

# APAMI 2006

Asia Pacific Association for Medical Informatics

## Proceedings

**Towards Global Interoperability for  
Electronic Health Records**

**亞太醫學資訊研討會論文集**

**October 27-29, 2006  
Grand Hyatt Taipei  
Taipei, Taiwan**



**2006 APAMI Triennial Conference**

Published by Taiwan Association for Medical Informatics

Copyright © 2006 by Asia Pacific Association for Medical Informatics & Taiwan Association for Medical Informatics

Printed in Taipei, Taiwan

Copyright and Reprint Permissions: Abstracting is permitted with credit to the source.

Libraries are permitted to photocopy beyond the limits of copyright law for private use of patrons those articles in this volume. Instructors are permitted to photocopy isolated articles for noncommercial classroom use.

ISBN:

This book was prepared camera-ready.

# Contents

|                                            |   |
|--------------------------------------------|---|
| ■ APAMI 2006 Committee & Information ..... | i |
|--------------------------------------------|---|

|                                          |    |
|------------------------------------------|----|
| ■ Oral Paper Presentation Schedule ..... | iv |
|------------------------------------------|----|

|                                                         |        |
|---------------------------------------------------------|--------|
| ■ Special Contributions & Asia-Pacific Regional Reports |        |
| .....                                                   | S1~S45 |

|            |                                                                               |    |
|------------|-------------------------------------------------------------------------------|----|
| <b>R01</b> | U-Health Approach to the National Health Information Systems in Korea (KOREA) | S1 |
|------------|-------------------------------------------------------------------------------|----|

*Young Moon Chae*

|            |                                                                  |    |
|------------|------------------------------------------------------------------|----|
| <b>R02</b> | National Health Information Network Development in Korea (KOREA) | S9 |
|------------|------------------------------------------------------------------|----|

*Yoon Kim*

|            |                                                    |     |
|------------|----------------------------------------------------|-----|
| <b>R03</b> | HIS Application status and trend in China. (CHINA) | S14 |
|------------|----------------------------------------------------|-----|

*Li Baoluo, Shen Tao, Chen Gi, Meng Xiaoyang*

|            |                                                     |     |
|------------|-----------------------------------------------------|-----|
| <b>R04</b> | Medical Informatics development in Taiwan. (TAIWAN) | S15 |
|------------|-----------------------------------------------------|-----|

*Jack Li*

|            |                                                                                                         |     |
|------------|---------------------------------------------------------------------------------------------------------|-----|
| <b>R05</b> | Development of Medical Informatics in Singapore -- Keeping Pace with Healthcare Challenges. (SINGAPORE) | S16 |
|------------|---------------------------------------------------------------------------------------------------------|-----|

*Walter T.L. Lim*

|            |                                                      |     |
|------------|------------------------------------------------------|-----|
| <b>R06</b> | EHR in Japan, -Recent Government Activities. (JAPAN) | S20 |
|------------|------------------------------------------------------|-----|

*Michio Kimura*

|            |                                                 |     |
|------------|-------------------------------------------------|-----|
| <b>R07</b> | eHealth trends in Pacific Islands 2006. (JAPAN) | S22 |
|------------|-------------------------------------------------|-----|

*Norio KIMURA, Isao NAKAJIMA, Muhammad Athar SADIQ, Yasumitsu TOMIOKA, Kenji TANAKA, Kiyoshi IGARASHI*

|            |                                                                                                                            |     |
|------------|----------------------------------------------------------------------------------------------------------------------------|-----|
| <b>R08</b> | Health Informatics in the Philippines. (PHILIPPINES)                                                                       | S25 |
|            | <i>Alvin B. Marcelo</i>                                                                                                    |     |
| <b>R09</b> | Thai Medical Informatics Association and medical informatics activity in Thailand. (THAILAND)                              | S29 |
|            | <i>Wansa Paojin</i>                                                                                                        |     |
| <b>R10</b> | Health Information Management: Malaysian Experience. (MALAYSIA)                                                            | S31 |
|            | <i>Datin Dr.S.Selvaraju</i>                                                                                                |     |
| <b>R11</b> | Recent Development of Health Informatics in Hong Kong. (HONG KONG)                                                         | S39 |
|            | <i>CP Wong</i>                                                                                                             |     |
| <b>R12</b> | Hospital-Wide Implementation of a Multi-Platform Pharmacological Decision Support System at the Point-of-Care. (SINGAPORE) | S41 |
|            | <i>Hui Nar Quek, Yuen Ho Chow</i>                                                                                          |     |

## ■ Full Papers ..... 1~265

|            |                                                                                                                       |    |
|------------|-----------------------------------------------------------------------------------------------------------------------|----|
| <b>A01</b> | Genomic Sequence Variation Markup Language (GSVML) for Global Interoperability of Clinical Genomics Data. (JAPAN)     | 1  |
|            | <i>Jun Nakaya, Michio Kimura</i>                                                                                      |    |
| <b>A02</b> | Design a pathway/genome expert system using a prolog machine incorporated with a parallel hardware searcher. (TAIWAN) | 9  |
|            | <i>Wenlung Shu</i>                                                                                                    |    |
| <b>A03</b> | Reconstructing protein-protein interaction networks from domain-domain interactions. (TAIWAN)                         | 15 |
|            | <i>Ka-Lok Ng</i>                                                                                                      |    |
| <b>A04</b> | Establishment of Disease Pathway Discovery System. (TAIWAN)                                                           | 19 |
|            | <i>Hao-Chung Hsu, Yuan-Chii G Lee</i>                                                                                 |    |



|            |                                                                                                                                                   |    |
|------------|---------------------------------------------------------------------------------------------------------------------------------------------------|----|
| <b>A05</b> | Constructing the Methionine Cycle and Its Related Disease Model with Hybrid Functional Petri Nets. (TAIWAN)                                       | 24 |
|            | <i>Austin H Chen, Zhi En Chen</i>                                                                                                                 |    |
| <b>A06</b> | Computational Methods to Uncover the Protein-Protein Interactions. (TAIWAN)                                                                       | 31 |
|            | <i>Fei-Hung Hung, Hung-Wen Chiu</i>                                                                                                               |    |
| <b>A07</b> | Mutual Information correction for eddy current-induced Geometric artefacts In Diffusion-Tensor Tractography. (TAIWAN)                             | 36 |
|            | <i>Kao-Lun Wang, Shiou-Ping Lee, Jack Li, Wen-Yih Tseng</i>                                                                                       |    |
| <b>A08</b> | A Virtual Bronchoscopy with Color-mapped Wall Thickness. (KOREA)                                                                                  | 39 |
|            | <i>Sang-joon Park, Jong Hyo Kim, Jin Mo Goo, Kwang Gi Kim, Sang Ho Lee</i>                                                                        |    |
| <b>A09</b> | K-means clustering approach for kinetic pattern analysis of dynamic contrast enhancement breast MRI. (KOREA)                                      | 45 |
|            | <i>Sang Ho Lee, Jong Hyo Kim, Kwang Gi Kim, Sang-joon Park, Woo Kyung Moon</i>                                                                    |    |
| <b>A10</b> | A study on the merits of Biometrics to perform e-Health. (JAPAN)                                                                                  | 49 |
|            | <i>Yasumitsu Tomioka, Toshihiko Kitano, Naoki Aida, Kei Kakehi, Athar Sadiq Muhammad, Hiroshi Juzoji, Isao Nakajima</i>                           |    |
| <b>A11</b> | Development of Portable Real-time Instrument for Three-Dimensional Measurement of Scoliosis. (TAIWAN)                                             | 54 |
|            | <i>Chung Pao Lin</i>                                                                                                                              |    |
| <b>A12</b> | Diagnoses and Surgical Planning of Lung Cancer in Color-Enhanced 3D Visualization System. (TAIWAN)                                                | 60 |
|            | <i>Horng-Shyang Liao</i>                                                                                                                          |    |
| <b>A13</b> | The use of geographic information systems(GIS) for rapid assessment of health facilities following tsunami disaster in Aceh province. (INDONESIA) | 68 |
|            | <i>Anis Fuad, Hari Kusnanto, Adi Utarini, Johan Groothoff</i>                                                                                     |    |

|            |                                                                                                                                                 |     |
|------------|-------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| <b>A14</b> | Dynamic PubMed: Automating bibliometric research using PubMed. (AUSTRALIA)                                                                      | 72  |
|            | <i>Kumara Mendis, Rasika Amarasiri</i>                                                                                                          |     |
| <b>A15</b> | Visualizing the Health Resources of Taiwan: The application of Web_GIS. (TAIWAN)                                                                | 78  |
|            | <i>Po-Huang Chiang, Chi-Pang Wen, Yaw-Tang Shih</i>                                                                                             |     |
| <b>A16</b> | A FORECAST MODLING AND DECSICION SUPPORT SYSTEM FOR VACCINE DEMAND DECISION MAKING. (TAIWAN)                                                    | 83  |
|            | <i>Chen k. Feng, Ruey k. Chiu</i>                                                                                                               |     |
| <b>A17</b> | Rasch Analysis Assists BNHI with Revealing Most Misfitting Strings of Hospital Medical Service Information. (TAIWAN)                            | 90  |
|            | <i>Tsair Wei Chien</i>                                                                                                                          |     |
| <b>B01</b> | The Design and Development of a Physician-Oriented Information Infrastructure for the Enhancement of e-Hospital Facilities. (TAIWAN)            | 94  |
|            | <i>Hsu-Hua Ho, Tzu-Yi Yu, Hung-Pin Hou, Wen-Yung Wu</i>                                                                                         |     |
| <b>B02</b> | Economic Evaluation of Japanese e-Health Systems: A Cost-Benefit Analysis. (JAPAN)                                                              | 102 |
|            | <i>Masatsugu - Tsuji, Masaaki Teshima, Fumio Taoka</i>                                                                                          |     |
| <b>B03</b> | Health Information on Cellphones over Wireless Application Protocol (WAP): Issues, Obstacles, and Strategies for the Philippines. (PHILIPPINES) | 110 |
|            | <i>Alvin B. Marcelo</i>                                                                                                                         |     |
| <b>B04</b> | A design of ubiquitous environment for medical problem-based learning with Web-technology. (KOREA)                                              | 118 |
|            | <i>Cho Hune</i>                                                                                                                                 |     |
| <b>B05</b> | Virtual Site Visit Abroad for EMR/PACS system by Broad Banded Teleconference system. (JAPAN)                                                    | 123 |
|            | <i>Naoki Nakashima</i>                                                                                                                          |     |

|            |                                                                                                                               |     |
|------------|-------------------------------------------------------------------------------------------------------------------------------|-----|
| <b>B06</b> | A Collaborative Workspace for Telemedicine System. (TAIWAN)                                                                   | 128 |
|            | <i>Hung-Wen Tung, Augustine Tsai</i>                                                                                          |     |
| <b>B07</b> | Using TAMIS Hospital Information System to Construct a Base for Clinical Decision Support in Schizophrenia Patients. (TAIWAN) | 132 |
|            | <i>CHIUNG HSU CHEN</i>                                                                                                        |     |
| <b>B08</b> | Problems on a Transborder EHR System in a Disaster in Asia-Pacific Region. (JAPAN)                                            | 138 |
|            | <i>Tanaka Takeshi</i>                                                                                                         |     |
| <b>B09</b> | Exercise Delivery Program for elder people in home. (JAPAN)                                                                   | 142 |
|            | <i>Nobuyuki Ashida, Taemi Hirai, Terumasa Higasi, Makoto Fukui, Youko Ohashi</i>                                              |     |
| <b>B10</b> | Integral Medical Information Interchange with HL7 CDA. (TAIWAN)                                                               | 146 |
|            | <i>Kuan Chih Lin, Ruey k. Chiu, Kuo Chin Tsai, Chia Ching Huang</i>                                                           |     |
| <b>B11</b> | Electronic Health Record System Evaluation Based on Patient Safety. (AUSTRALIA)                                               | 154 |
|            | <i>Yung-Yu Su, Khin Than Win, John FULCHER</i>                                                                                |     |
| <b>B12</b> | Construction and Evaluation of EHR which Supports Team Practice. (JAPAN)                                                      | 160 |
|            | <i>Hidehiko Tsukuma</i>                                                                                                       |     |
| <b>B13</b> | System Interoperability Study for Healthcare Information System with Web Services. (UNITED KINGDOM)                           | 165 |
|            | <i>Jing Kai Zhang, Wei Xu</i>                                                                                                 |     |
| <b>B14</b> | Features of Standardized Nursing Terminology Sets in Japan. (JAPAN)                                                           | 171 |
|            | <i>Hiromi itoh Ozaku, Kaoru Sagara, Akinori Abe, Futoshi Naya, Noriaki Kuwahara, Kiyoshi Kogure</i>                           |     |
| <b>B15</b> | Context-Based Retrieval System for Similar Medical Documents. (JAPAN)                                                         | 177 |
|            | <i>Kazuya Okamoto, Tadamasa Takemura, Tomohiro Kuroda, Hiroyuki Yoshihara</i>                                                 |     |

|            |                                                                                                                                  |     |
|------------|----------------------------------------------------------------------------------------------------------------------------------|-----|
|            | The Influence that Person's Sensibility Receives from                                                                            |     |
| <b>B16</b> | Mixed-Reality based on a Tanka (a traditional Japanese poem of 31 syllables). (JAPAN)                                            | 184 |
|            | <i>Hiroshi Juzoji, Athar Sadiq Muhammad, Isao Nakajima, Yasumitsu Tomioka</i>                                                    |     |
| <b>B17</b> | Evidence-Based Medicine Study between Bicycle Helmet and Head injury. (TAIWAN)                                                   | 188 |
|            | <i>LungWen Tsai, Ten Fang Yang</i>                                                                                               |     |
| <b>B18</b> | The effect of different height of posterior walker on gait parameters in children with spastic diplegic cerebral palsy. (TAIWAN) | 194 |
|            | <i>Yu Ju Chih</i>                                                                                                                |     |
| <b>C01</b> | Adaptable Access Control for Electronic Medical Records. (TAIWAN)                                                                | 200 |
|            | <i>Kung Chen, Yuan-Chun Chang, Da-Wei Wang</i>                                                                                   |     |
| <b>C02</b> | MAC Address-Based Position Detection of Illegal Access Terminals on Private Network for Hospital Asset Management. (JAPAN)       | 208 |
|            | <i>Tomohiro Kuroda, Hiroyuki Yoshihara, Naoto Kume</i>                                                                           |     |
| <b>C03</b> | Safeguarding clinical software – A managerial case study about project management and oversight. (GERMANY)                       | 214 |
|            | <i>Wetter Thomas</i>                                                                                                             |     |
| <b>C05</b> | Expanding the role of Universal Access Fund to Rural Telemedicine in Developing Countries. (JAPAN)                               | 220 |
|            | <i>Athar Sadiq Muhammad</i>                                                                                                      |     |
| <b>C06</b> | Safety Management and Crime Prevention by IC Tags – Cases of practical use of IC tag in medical and welfare fields (JAPAN)       | 224 |
|            | <i>Nobuyuki Ashida, Yoshie Ashida</i>                                                                                            |     |
| <b>C07</b> | Applying Grid Computation for Evaluating Potentially Inappropriate Medication Use in Elderly Patients. (TAIWAN)                  | 227 |
|            | <i>Yung-Tai Yen, Yu-Lun Liu, Yu-Chuan Li, Chien-Yeh Hsu</i>                                                                      |     |

|            |                                                                                                  |     |
|------------|--------------------------------------------------------------------------------------------------|-----|
| <b>C08</b> | Survey on Medical Error Reporting Systems in Korean Hospitals.<br>(KOREA)                        | 235 |
|            | <i>Jeongeun Kim, Sukwha Kim, David W Bates</i>                                                   |     |
| <b>C09</b> | Applying Contextual Analysis on Investigating Potential. (TAIWAN)                                | 236 |
|            | <i>Wei Yi Fan, Chien-Yeh Hsu</i>                                                                 |     |
| <b>C10</b> | Handheld RFID Patient Safety Improvement – Case Study in Wan Fang Hospital. (TAIWAN)             | 241 |
|            | <i>Li Xin Xu</i>                                                                                 |     |
| <b>C11</b> | The Requirement of Health Record to Satisfy the Patients' Safety Management. (JAPAN)             | 249 |
|            | <i>Kiyomu Ishikawa</i>                                                                           |     |
| <b>C12</b> | A Study on a Nationwide Ambulance Communication System via HEOs to Support Telemedicine. (JAPAN) | 257 |
|            | <i>Isao NAKAJIMA</i>                                                                             |     |
| <b>C13</b> | A Pilot Learning/Learning/Teaching System of Mammography Screening for Radiologists. (TAIWAN)    | 260 |
|            | <i>Mann-Li Chang</i>                                                                             |     |

## ■ Supplements: Poster Abstracts .....266~279

|            |                                                                                                                                                             |                                                       |        |
|------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------|--------|
| <b>P01</b> | Artificial Neural Networks: A Supporting Role In Improving Prediction Of Benign Prostatic Obstruction Evaluated By The International Prostate Symptom Score | <i>Chih-Cheng Lu<br/>Jinsheng Roan<br/>Yu-jen Lee</i> | TAIWAN |
| <b>P02</b> | A study on the Metaheuristics - apply biologically inspired dynamic formation to telecommunication network -                                                | <i>Isao NAKAJIMA</i>                                  | JAPAN  |
| <b>P03</b> | The Application and the Realization Of XML Transform Middleware in Electronic Medical Record                                                                | <i>Yi Chen<br/>Yong Feng Huang<br/>Si Fa Zhang</i>    | CHINA  |

|             |                                                                                                                                                |                                                            |        |
|-------------|------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------|--------|
| <b>P04</b>  | Development of Teleconsultation System                                                                                                         | Yuichi Ishibashi<br>Isao NAKAJIMA<br>Atsuko Hara           | JAPAN  |
| <b>P05</b>  | Correlation of EEG power ratio and severity of Alzheimer's disease                                                                             | Chih-Chung Chen<br>Chien-Yeh Hsu<br>Hung-Wen Chiu          | TAIWAN |
| <b>P06</b>  | Development of the Home-Based Multimedia System for Mental Health Management in adolescents via Information Superhighway                       | Jeongyee Bae                                               | KOREA  |
| <b>P07</b>  | Validating the role of "Mobile Telemedicine Package" in Marine Time Teleconsultations.                                                         | Athar Sadiq<br>Muhammad                                    | JAPAN  |
| <b>P08</b>  | A Rough Set-Based Prediction Approach on Medical Application                                                                                   | PEI CHUN<br>Jia-Yuarn Guo<br>Yu-Tung Dai<br>Chung-hsiu Kuo | TAIWAN |
| <b>P09</b>  | Practical operation of a document delivery system for medicine information, which used a hospital web-based system in risk management side     | Hirohiko Mishima                                           | JAPAN  |
| <b>P10</b>  | A study of formation and storage of laboratory results based on HL7 Clinical Document Architecture Standard                                    | Cho Hune                                                   | KOREA  |
| <b>P11</b>  | The development of additional automatic mapping protocols for local LIS system which do not make one-to-one correspondence with LOINC code set | Cho Hune                                                   | KOREA  |
| <b>P12*</b> | Design an assistive device for Parkinson's patients with impaired balance to improve the stride length                                         | Tz Hau Tsuei                                               | TAIWAN |
| <b>P13*</b> | Development of Active Gait Trainer : System Verification                                                                                       | Che Wen Yu                                                 | TAIWAN |
| <b>P14*</b> | An epidemiological study on psychiatric disorders based on the medical records of a tertiary care hospital                                     | Sabu K M                                                   | INDIA  |
| <b>P15*</b> | Building and Implementing a Comprehensive Care Healthcare Information System – for Mackay Memorial Hospital Medical Center                     | Chang Po-Ren<br>Hsu Shih-Xin<br>Liu Chien-Tsai             | TAIWAN |

\*Abstracts not available

## APAMI2006 Organizers

---

Asia-Pacific Association for Medical Informatics (APAMI)

Taiwan Association for Medical Informatics (TAMI)

Graduate Institute of Medical Informatics, Taipei Medical University (GIMI, TMU)

Center for Biomedical Informatics, Taipei Medical University (CBI, TMU)

Health Level 7, Taiwan (HL7 Taiwan)

## APAMI2006 Organizing Committee

---

**Honorary Presidents** Yun-Sik Kwak  
Sheng-Mou Hou  
Henry Chang  
Chung Yi Hsu

**Conference President** Yu-Chuan (Jack) Li

**Program Director** Ken N. Kuo  
Charng-Er Shyu

**Exhibition Director** CP Wong (HK)  
Ronlon Tsai

**Website Director** Hsien Ming Goh (MY)  
Yen-Cheng Chen

**Tour Director** Wen-Shan Jian

### **Secretariat**

Secretary General Hsu-Tien (Marian) Wan

CIS Designer Mei-Ling Huang

Treasurers Chia-Hui (Sophie) Chien

## APAMI2006 Science Program Committee

---

**Chair** Young Moon Chae (KR)  
Da-Wei Wang (TW)

**Co-Chair** Chun-Por Wong (HK)  
Michio Kimura (JP)  
Steven Chu (NZ)  
Jin-Shin Lai (TW)  
Chien Yeh Hsu (TW)  
JiaJie Zhang(US)

|                                  |                                 |
|----------------------------------|---------------------------------|
| <b>Members</b> Klaus D Veil (AU) | Hui Nar Quek (SG)               |
| Yang Zhong (CN)                  | Kwok Chan Lun (SG)              |
| Andre Greyling (HK)              | Tin Wee Tan (SG)                |
| Chun-Por Wong (HK)               | Tze Yun Leong (SG)              |
| Ngai-Tseung Cheung (HK)          | Yuen Ho Chow (SG)               |
| Saroj K. Mishra (IN)             | Chamaree Chuapetcharasopon (TH) |
| Michio Kimura (JP)               | Nipit Pirawej (TH)              |
| Jeongeun Kim (KR)                | Wansa Paoiin (TH)               |
| Sukwha Kim (KR)                  | Da-Dan Tang (TW)                |
| Young Moon Chae (KR)             | Da-Wei Wang (TW)                |
| Yun Sik Kwak (KR)                | Der-Ming Liou (TW)              |
| Hsien Ming Goh (MA)              | Ean-Wen Huang (TW)              |
| Jai Mohan (MA)                   | Jin-Shin Lai (TW)               |
| Stephen Chu (NZ)                 | Li Liou (TW)                    |
| Alvin B. Marcelo (PH)            | Phei-Lang Chang (TW)            |
| Herman Tolentino (PH)            | Polun Chang (TW)                |
| Rommel Palma Feria (PH)          | Shien-Young Chang (TW)          |
|                                  | Wan-Hsuan Lin (TW)              |
|                                  | Yu-Chuan Li (TW)                |



## APAMI Executive Committee

---

|                          |                                                                                                                                      |
|--------------------------|--------------------------------------------------------------------------------------------------------------------------------------|
| <b>President</b>         | Prof. Yun Sik Kwak (KR)                                                                                                              |
| <b>Vice-President</b>    | Prof. Yu-Chuan (Jack) Li (TW)                                                                                                        |
| <b>Secretary-General</b> | Dr. HM Goh (MY)                                                                                                                      |
| <b>Treasurer</b>         | Dr. CP Wong (HK)                                                                                                                     |
| <b>Member Society</b>    | Australia, China (PR), Hong Kong SAR, Japan, Korea, Malaysia,<br>New Zealand, Philippines, Singapore, Sri Lanka, Thailand,<br>Taiwan |

## NPO/Government Sponsors (Taiwan)

---

Department of Health, Executive Yuan  
National Science Council  
National Health Research Institutes (NHRI)  
Centers for Disease Control  
Institute of Information Science, Academia Sinica  
Ministry of Foreign Affairs  
Taipei City Government

## Corporate Sponsors

---

- Phillips Taiwan (台灣飛利浦)
- FlySheet Group (飛資得資訊有限公司)
- Hewlett-Packard Development Company (台灣惠普)
- Maya International (瑪雅資訊)
- Northrop Grumman Corp.
- HaBook Information Technology INC (網奕資訊)
- Data Systems Consulting Co., Ltd (鼎新電腦)
- Lenovo
- TMT Project
- Institute for Information Industry, Taiwan (資策會)

## Timetable of APAMI2006 Parallel Programs\*\*

|                   | Starting Time                 | Room A | Room B         | Room C | Room D<br>(MIST) | Room E                                                    |
|-------------------|-------------------------------|--------|----------------|--------|------------------|-----------------------------------------------------------|
| 2006/10/28 (Sat.) | Parallel Paper Sessions (I)   |        |                |        |                  | AP<br>Regional<br>Reports<br>(I)<br><br>R03~06,<br>R10~11 |
|                   | 13:10                         | A1     | B1             |        |                  |                                                           |
|                   | 13:25                         | A2     | B2             | C1     | M1               |                                                           |
|                   | 13:40                         | A3     | B3             | C2     | M2               |                                                           |
|                   | 14:55                         | A4     | B4             | C3     | M3               |                                                           |
|                   | 14:10                         | A5     | B5             | C5     | M4               |                                                           |
|                   | 14:25                         | A6     | B6             | C6     | M5               |                                                           |
|                   | 14:40-15:00                   |        | Coffee Break   |        |                  |                                                           |
|                   | Parallel Paper Sessions (II)  |        |                |        |                  |                                                           |
|                   | 15:00                         | A7     | B7             | C7     | M6               |                                                           |
|                   | 15:15                         | A8     | B8             | C8     | M7               |                                                           |
|                   | 15:30                         | A9     | B9             | C9     | M8               |                                                           |
|                   | 15:45                         | A10    | B10            | C10    | M9               |                                                           |
|                   | 16:00                         | A11    | B11            | C11    | M10              |                                                           |
|                   | 16:15                         | A12    | B12            | C12    |                  |                                                           |
|                   | 16:30-18:00                   |        | Poster Session |        |                  |                                                           |
| 2006/10/29 (Sun.) | Parallel Paper Sessions (III) |        |                |        |                  | AP<br>Regional<br>Reports (II)<br><br>R01, R02, R07~09    |
|                   | 09:00                         | A13    | B13            |        | M11              |                                                           |
|                   | 09:15                         | A14    | B14            |        | M12              |                                                           |
|                   | 09:30                         | A15    | B15            |        | M13              |                                                           |
|                   | 09:45                         | A16    | B16            |        | M14              |                                                           |
|                   | 10:00                         | A17    | B17            |        | M15              |                                                           |
|                   | 10:15                         | A18    | B18            |        |                  |                                                           |
|                   | 10:30-10:40                   |        | Coffee Break   |        |                  |                                                           |

\*\*Timetable shown here is subjected to change without notice, please see Conference Directory for the most accurate information.

## U-Health Approach to the National Health Information Systems in Korea

Young Moon Chae, Ph.D.

Graduate School of Public Health, Yonsei University, Seoul, Korea

Email: ymchae@yumc.yonsei.ac.kr

### Abstract

*This study describes the current status of computerization for hospitals and makes various recommendations to improve the level of computerization and information sharing between hospitals in order to strengthen the National Health Information System (NHIS). According to a survey of 314 out of 1,289 hospitals in August 2005, the computerization level was higher for the specialty tertiary hospitals and general hospitals than for smaller hospitals. The computerization rate by applications was 75.6% for outpatient CPOE (Computerized Physician Order Entry System), 70.6% for inpatient CPOE, 47.1% for PACS (Picture Archiving and Communication System), and 20.7% for outpatient EMR (Electronic Medical Record).*

*In addition, this paper also presents the NHIS plan in Korea and the survey results that examined to identify the user requirements for the electronic health record (EHR), which is a core element of NHIS. According to a survey*

*conducted in May 2005 of 317 people, ranging from community residents, health workers, and information systems people, most people responded that development of EHR was necessary and that they were willing to use it actively. With regard to key information included in EHR, they made the following recommendations in order: test results, health program, medical examination, immunization, allergy, information about medication and medical treatment, health behavior, genetic information, cause of death, and medical charge information. Most of the respondents wanted to apply EHR to the management of chronic diseases, but they were concerned about the security of private medical information. Based on the results, recommendations for development and implementation strategies were presented.*

**Keyword:** national health information system, standards, electronic health records

## 1. Introduction

As information technologies (IT) are rapidly improved, healthcare institutions in Korea are increasingly introducing information systems in order to improve quality of care and reduce administrative costs. There are several factors related to the increasing trends in computerization of health institutions and the increasing demand for the national health information system (NHIS). First, the proportion of the elderly population has been rapidly increasing for the last several years. The elderly population has been increasing at an annual rate of 5% and the elderly population is expected to reach about 23.9% of the total population in 2030 [1]. This rate of increase is higher than that of most OECD countries. Since elderly people place greater demand on healthcare services and resources, there is a need for information systems to support these health services and improve the efficiency in administrative activities at healthcare institutions.

Second, the legal system in Korea has been instrumental in the rapid computerization of the healthcare sector. Three application areas on e-health have been legalized in 2003: telemedicine, electronic medical record (EMR), and e-prescription system for pharmacies, which is a computerized physician order entry (CPOE) system between hospitals and

pharmacies.

Third, a major change in national hospital policy has also affected the information trend. Under the DDA (Doha Development Agenda), the Korean hospital market will be opened to foreign countries in 2007. Many Korean hospitals fear that world-class U.S. hospitals will enter the Korean market by targeting rich patients and using advanced information systems such as telemedicine and EMR. In a response to this, many Korean hospitals are developing more specialized medical services and strengthening their information systems.

Finally, advanced internet infrastructure and other IT in Korea have been a driving force for the rapid growth in computerization. The broadband uptake rate is in the top level in the world (Figure 1). In addition, Korea is also one of the top countries in the manufacture and export of PCs, computer monitors, hard disks, and other hardware.

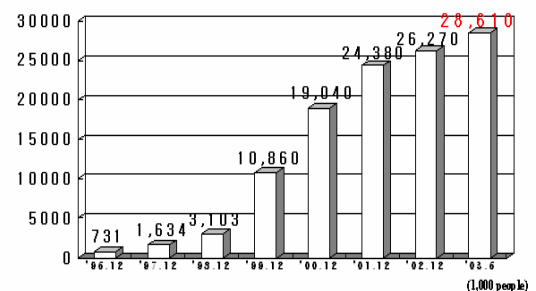


Figure 1. Number of subscribers of high-speed internet users in Korea

Many developed countries are actively developing their NHIS according to their health delivery systems. Most countries are developing electronic health record (EHR) as the core element of their national systems [2]. EHR has been defined as a systematic and continuous aggregation of computerized information about medical treatment and individual health. The EHR system includes the information about medical treatment and health collected throughout life from birth to death, as well as family genetic information and personal information on health wellness. In the United States, hospitals at the state level are integrated to form the Local Health Information Infrastructure (LHII), which is further integrated at the national level to form the National Health Information Infrastructure (NHII). Through NHII, an integrated EHR system was developed incorporating public and private insurance. This system was managed by a committee comprised of government officials, health providers, and consumers. The EHR system is being applied to patient care, electronic prescription service, and review of insurance claims [3]. In England, the public health system was divided into 5 regions and the nationwide, unified EHR system was developed and used in all regions to support e-medical treatment reservation service, telephone consultation service, and clinic research, etc.

The Korean government developed its first

10-year national ISP (Information Strategy Plan) for NHIS in 1990, followed by the second version in 2001. The scope of the ISP was broad covering almost all activities of the Ministry of Health and Welfare (MOHW): community health information system for health centers and subcenters, hospital information systems for emergency care and blood management, national health insurance system for processing insurance claims data with a focus on EDI (Electronic Data Interchange) for data transmission, national pension system, quarantine system for imported foods, pharmaceutical information system for processing administrative drug data for KFDA (Korea Food and Drug Administration), infectious disease surveillance system and vaccination system for CDC (Center for Disease Control), national registration system for the handicapped people, and national health statistics system.

The national ISP has been a driving force for upgrading information systems for MOHW. While these information systems have greatly contributed to achieving our national goals of reducing health care costs and improving quality of healthcare, our national systems are still years behind some western countries with respect to standards, e-health related laws, EHR, and telemedicine. This paper presents some key infrastructure (standards, law, and organization), national ISP, and the survey results on EHR.

## 2. Current Status of Computerization in Korean Healthcare Institutions

### 2.1 Subjects

Chae et al. [4] have conducted a nationwide survey on the computerization of healthcare institutions. 314 out of 1,289 hospitals were surveyed in August 2005.

### 2.2 Results

#### 2.2.1 Hospitals

As seen in Table 1, almost all specialty tertiary hospitals implement both outpatient and inpatient CPOE, PACS (Picture Archiving and Communication System), insurance claims processing, including transmitting claims data by EDI, and ADT (Admission, Discharge, and Transfer). However, only about 20% of hospitals were implementing EMR, perhaps because EMR was legalized only in 2003. It is expected that most tertiary hospitals will implement EMR by the end of 2007.

#### 2.2.2 Clinics and Pharmacies

Most private clinics and pharmacies are implementing some form of information system. Over 90% of all clinics and pharmacies are transmitting their insurance claims via EDI network. Even before the government legalized EMR, many clinics had introduced EMR and about 85% of clinics

were implementing EMR in 2005.

Table 1. Status of Computerization in Hospitals

( ): percentages

|                         | Tertiary Hospital | General Hospital | Hospital | Total |
|-------------------------|-------------------|------------------|----------|-------|
| Outpatient CPOE         | 97.6              | 84.2             | 66.9     | 75.6  |
| Inpatient CPOE          | 97.6              | 84.2             | 58.4     | 70.6  |
| Pharmacy                | 100               | 88.2             | 53.4     | 69.1  |
| Laboratory              | 97.6              | 86.7             | 54.1     | 68.5  |
| Radiology               | 97.6              | 88.2             | 53.4     | 68.7  |
| Specialized tests       | 97.6              | 82.2             | 43.1     | 60.9  |
| ADT                     | 100               | 98.7             | 95.4     | 96.8  |
| Administration          | 100               | 63.1             | 62.6     | 73.0  |
| Insurance claim         | 100               | 100              | 89.4     | 93.5  |
| Referral                | 97.6              | 69.1             | 36.4     | 53.8  |
| PACS                    | 90.5              | 78.6             | 22.6     | 47.1  |
| Inpatient EMR           | 21.4              | 14.5             | 21.0     | 19.6  |
| Outpatient EMR          | 19.1              | 14.8             | 23.4     | 20.7  |
| Insurance claims by EDI | 100               | 90.5             | 92.4     | 94.8  |

#### 2.2.3 Public Healthcare Institutions

Similarly, most public healthcare institutions were implementing information systems. All community health centers were implementing CPOE and the reporting systems on infectious diseases and vaccination. The Korean government is in the process of changing client-server based public health information to web-based systems using the ASP (Application Service Provider) approach.

#### **2.2.4 National Health Information System**

This survey also included items on government policies on NHIS. Most hospitals did not introduce any form of standards on clinical terminology, perhaps due to the absence of both national standards and incentives for using international standards. Despite the government policy on information sharing among health institutions, only 15% of institutions were exchanging information for the purpose of patient referrals. Over 80% of clinics expressed negative opinion on telemedicine due to the high risk of misdiagnosis and loss of patients to large tertiary hospitals. These results have significant implications for developing and disseminating the NHIS.

### **3. National Health Information Systems (NHIS ) in Korea**

#### **3.1 Information Strategy Plan (ISP)**

In 2004, MOHW revised its ISP, which was comprised of three phases. In Phase 1 (2004-2005), ISP for NHIS was developed and standards and e-health related laws were developed. In Phase 2 (2006-2007), EHR and public health information systems are being developed using new standards. In Phase 3 (2008-2009), standards and EHR for health centers will be redeveloped to meet the needs

for the private health institutions.

#### **3.2 Organization**

In order to effectively carry out ISP, the MOHW created the organization. Several advisory committees were created for providing policy directions and expert opinions. Two centers were created for carrying out system development and the related researches: Center for NHIS Development and R & D Center for EHR. In addition, several working groups were created for developing the detailed activity plans for each application: telemedicine, consumer health information, and so on.

#### **3.3 Standards**

Standard is an important infrastructure for the NHIS because it allows hospitals to exchange and share patient information. In order to develop the national standards, 14 subcommittees were created in 2004: clinical terminology, dental terminology, healthcare terminology, herbal medicine terminology, medical devices, image processing, security, etc.

The committees chose the UMLS (Unified Medical Language System) developed by the NLM (National Library of Medicine) in the United States as a framework for building the Korean standards because it is internationally

accepted standards and there is no charge, which can be a big financial burden for the small hospitals in the future. Matching tables for clinical and other terminologies were created under the UMLS framework in order to develop the Korean national standards.

### **3.4 e-Health Law**

Three revisions in the medical law enacted in 2003 greatly upgraded healthcare information systems in Korea: legalizing telemedicine, legalizing electronic medical record, and legalizing e-prescription system for outside pharmacies. However, there are needs for revising this law again. First, there is a need for providing a legal ground for NHIS. Second, there is a need for strengthening security and confidentiality of patient information. Third, there is a need for modifying some clauses related to telemedicine in order to increase its use in healthcare institutions.

Accordingly, government developed a draft of new e-Health law that includes the following elements: authorizing a creation of the Center for NHIS, including a separate Act for security and confidentiality of patient information, authorizing maintenance of EMR outside of the hospital that holds the ownership in order to allow ASP for public health information systems, and allowing tele-homecare for revisit patients.

### **3.5 Public Health Information Systems**

The public health system in Korea is comprised of a three tier system: 250 community health centers, 1200 health subcenters, and 2000 health posts. They provide primary health services and vaccinations to community residents and collect health statistics at the grass-root level. In an effort to improve their operation and quality of services, health centers started to introduce information systems in the early 90s and now entire health centers are implementing information systems. According to the evaluation survey [5], health center information systems were effective in improving patient satisfaction and increasing the productivity of health workers.

However, updating software and providing timely hardware maintenance have been serious problems for health centers because most health centers are located in remote rural areas and suffer a lack of skilled maintenance people. Since the public health information systems are the backbone of the NHIS, the government decided to strengthen the system by converting current client-server based systems to a web-based system using the ASP approach. In addition, newly developed national standards will be applied to develop community EHR for the new public health information system.



### 3.6 Telemedicine

First pilot project on telemedicine was carried out from 1995 to 1998 between the Wooljin Health center and the Kyungbook University hospital using a high speed T1 telecommunication network. While this project demonstrated a possibility of providing medical consultation from a distance through network, it was not widely used as much as people anticipated because of inadequate telecommunication network and lack of financial incentives for medical doctors. While several pilot telemedicine projects have been carried out mainly at community health centers since then, most of them had a limited success in community acceptance even after telemedicine was legally allowed in 2003.

### 3.7 Survey on EHR

According to a survey conducted in May 2005 of 317 people, ranging from community residents, health workers, and information system people, most people responded that development of EHR was necessary and that they were willing to use it actively [6]. As a prerequisite for the development of EHR, the respondents recommended the development of a security system, standards, budgets, restructuring the legal system, obtaining community support, and telecommunication network (Table 2).

Table 2. Prerequisites for Developing EHR

( ): percentage

| Item     | Health worker | Gov        | Aca        | Bus        | Res        | Tot         |
|----------|---------------|------------|------------|------------|------------|-------------|
| Security | 67<br>(42)    | 11<br>(31) | 10<br>(37) | 9<br>(26)  | 20<br>(48) | 117<br>(39) |
| Standard | 62<br>(39)    | 20<br>(57) | 10<br>(37) | 17<br>(50) | 16<br>(38) | 125<br>(42) |
| Budget   | 67<br>(42)    | 10<br>(29) | 6<br>(22)  | 15<br>(44) | 18<br>(43) | 116<br>(39) |
| Legal    | 50<br>(31)    | 15<br>(43) | 12<br>(44) | 17<br>(50) | 14<br>(33) | 108<br>(36) |
| Support  | 39<br>(26)    | 10<br>(29) | 10<br>(37) | 7<br>(21)  | 10<br>(24) | 76<br>(26)  |
| Network  | 28<br>(18)    | 4<br>(11)  | 4<br>(15)  | 3<br>(9)   | 6<br>(14)  | 45<br>(15)  |
| Total    | 159           | 35         | 27         | 34         | 42         | 297         |

(Gov: Government employee, Aca: Academics, Bus: Business, Res: Community residents)

The degree of participation was the highest for those in business, followed in order by government employees, academics, health workers, community residents, and health providers. With regard to key information included in EHR, they made the following recommendations in order: test results, health program, medical examination, immunization, allergy, information about medication and medical treatment, health behavior, genetic information, cause of death, and medical charge information. Most of the respondents wanted to apply EHR to the management of

chronic diseases, but they were concerned about the security of private medical information. The majority wanted to access their own personal information through Internet after security validation.

In conclusion, the unified England EHR model was proposed for the EHR system for public health institutions in Korea based on the similarities in the public health delivery system in the two countries and the Korean nationwide registration system. In order to protect the security of private health information, we recommend using an IC security card to access EHR. In addition, we recommend organizing an EHR committee, similar to that of the United States, in order to protect against an infringement of personal information due to the monopoly of information by a single organization.

#### 4. Conclusion

The Korean government is pushing hard with its ambitious plan for the NHIS. The factors which are critical for the success of the plan may include reliable and cost-effective IT, user-friendly application systems, standards, law, budgets, and strong support from various affected groups such as the Korean Medical Association, citizen's group, etc. Of these, the latter four factors are rather weak areas in

Korea, compared with the technical factors, and therefore they need to be strengthened. Especially, getting support from medical community is very important factor for the success of telemedicine.

#### 5. Reference

- [1] Ministry of Health and Welfare. "National Demographic and Health Statistics", 2003.12
- [2] E. Middleton, E. Hammond, P.F. Brennan, "Accelerating U.S. Adoption", *JAMIA*, 2005; 12, pp. 13-19
- [3] W.A. Yasnoff, B.L. Humphry, D.E. Detmer, "A Consensus Action Agenda for Achieving the National Health Information Infrastructure", *JAMIA*, 2004; 11, pp. 332-338
- [4] Y.M. Chae, et al. Status of Computerization in Korean Health Institutions", Health Insurance Review Agency, 2005.12
- [5] Y.M. Chae, S.I. Kim, S.H. Choi, I.S. Kim, "Implementing Health Information System: Measuring Success in Korea's Health Centers", *International Journal of Health Planning and Management*, 1994; 9, pp. 341-348
- [6] G.H. Kim, Y.M. Chae, S.I. Kim, Y.E. Kim, "Study on the Korean Model for Community-based EHR", *Journal of Korea Society of Medical Informatics*, 2005; 11(2), pp. S18-S23

## National Health Information Network Development in Korea

Yoon Kim MD, Ph.D., MS<sup>(1)</sup>

Department of Health Policy and Management,  
Seoul National University College of Medicine, Seoul, Korea<sup>(1)</sup>  
yoonkim@snu.ac.kr

### Abstract

*National level efforts have been made to adopt information and communication technologies in health care to improve the quality and safety and to reduce the costs and errors of health care in many countries. A national strategic plan, National Health Information Network, is prepared for the successful implementation of the interoperable health information technology in health care by the Ministry of Health and Welfare, Korea.*

*The purpose of this study is to describe how the National Health Information Network programs are developed and working in Korea. To ensure the anywhere, anytime access for anyone who wants to the lifetime electronic health record (EHR) and decision supports in secure way to improve the quality, safety and efficiency of health care by 2010, national health information standards, national health information infrastructure, lifetime electronic health record, public health information system, building consensus were the selected as NHIN programs. And the challenges faced in the process of NHIN development are discussed later.*

### 1. Introduction

Recently the national effort is going underway to adopt information and communication technologies to improve the quality and efficiency of health care and to reduce the cost of health care in all over the world. Ever since the series of IOM reports [1, 2] were published, many countries have made an effort to improve patient safety and quality of healthcare and to reduce the healthcare costs using information technology. According to the experience of countries including UK, US, Australia and Canada, which are leading the adoption of information technology in health care, a roadmap developed with a national consensus is the most necessary work to start for the successful health information technology (HIT) adoption.

Improving quality and safety of health care and reducing the health care costs are issues in Korea as well as in others. Since the Korean government felt needs for the HIT adoption and its efficient implementation, a national strategic plan for National Health Information Network (NHIN) has been prepared by the Ministry of Health and Welfare (MOHW), Korea. The health system of Korea is required to be more efficient and effective due to the consumer needs for the good quality, safe

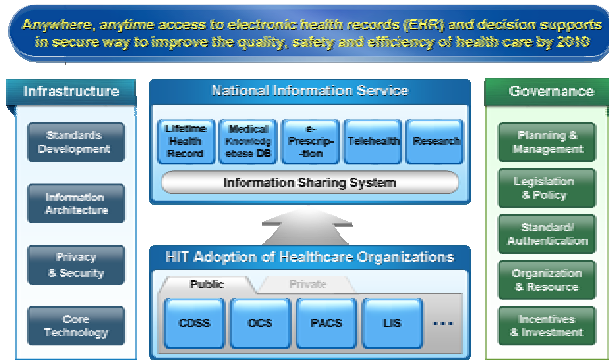
health care and the burden of the increasing cost of health care.

According to the recent national survey for health care consumers, 13% of people had experienced the needs for health information sharing of the past medical tests [3]. Another study showed that almost 20% of people experienced the need for laboratory information sharing from other healthcare institutions [4]. Health information sharing improves the quality of health care from the continuity of care as well as causes the decrease of cost from the tests duplications. Lab tests duplications, prescription errors, and preventable medical errors cause the high cost and low quality of health care. Another noticeable factor requiring the modernization of Korea health system is the rapidly aging society and the increasing health care costs. Aging society means the increase of chronic diseases, which are known that preventable services and self management are important.

The increasing cost of health care due to poor quality and safety, performance of health care system, and aging society are not the only problems for Korea. The benefit of national HIT adoption is estimated about 5% of national health expenditure savings and expected more [4]. The benefits of national HIT adoption may only be the estimation; however, many studies and reports prove the improvement of the performance and the quality of health care and many countries pursues the HIT adoption. National HIT adoption started already, now we are at the point to find the key factors to drive the national HIT adoption successfully.

### 2. National Health Information Network Framework [5]

In 2005, the MOHW announced the vision of future healthcare, “anywhere, anytime access to electronic health records and decision supports in secure way to improve the quality, safety and efficiency of health care by 2010”. For the sake of the vision achievement, the Korean government developed a 5-year national plan for NHIN and started to carry into action. The 5-year, NHIN plan includes implementing interoperable EMRs for public hospitals and public centers and facilitating EMR adoption among private health care organizations, developing national services through the infrastructure for health information sharing, and developing the infrastructure and governing structure for NHIN as shown in Figure 1.



## 2.1. NHIN Programs

The development of NHIN programs is meant to be as flexible and comprehensive as possible allowing for the future changes of techniques and circumstances. The developed standards for NHIN programs define the functions enabling health care institutions following the national standards to exchange the health information and .be part of NHIN.

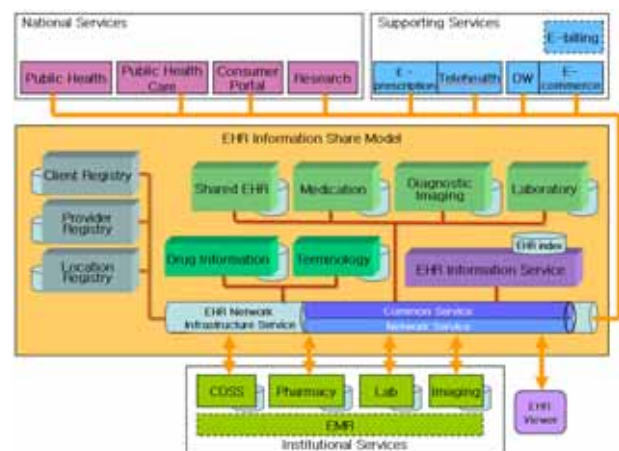
**2.1.1. HIT adoption of public hospitals.** The public hospital covers only 10% of the total number of beds in Korea. The performance and the quality of care in public hospitals are far behind the private hospitals. The HIT adoption in public hospitals is slow and inefficient. The utilization and satisfaction rate of public hospitals and health centers is low. According to the national plan for strengthening public health system, HIT adoption in public hospitals is considered as a way for the modernization of the public health care system.

Many public hospitals are planning to adopt HIT by their own; however, most of public hospitals are faced with the lack of budget, experience and expert personnel. Not only the HIT adoption effort of public hospitals costs high, but its implementation of individual institution makes it difficult the health information share. HIT adoption effort of public hospitals will be pursued according to the national NHIN plan for the efficient and effective implementation and investment.

The one of the NHIN program for public hospital HIT adoption is to develop and deploy a standardized hospital information system, which is flexible and standardized interoperable component-based hospital information system. Standardized hospital information system will be developed following the functions enabling health information share between health institutions following the national standards. A well developed standardized hospital information system and its core technologies will be provided to private hospitals and we are expecting that HIT adoption of public hospitals and standardized hospital information implementation can boost the HIT adoption in nation wide.

### 2.1.2. National health information share model

Neither the data sets for health information sharing, nor the NHIN data sharing model did not earn the consensus of the public yet. However, a data share model is necessary and inevitable to support the information exchange between health systems. The national health information share model plays the role for any health care institutions to be connected to the national services and to exchange the health information only if they follow the national standards. Though the public health system is taking up only 10% of total health care system, the national health information share model is not only aimed for the information exchange of public health care systems. At the present time Center for interoperable Electronic Health Record (CiEHR) is responsible for the development of the national health information share model and to get the public consensus on the model.



### Figure 2 National health information share model

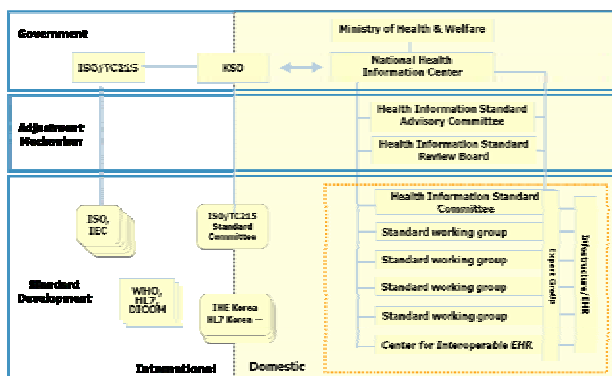
### 2.1.3. National health information services

If the national health information share model is the technical information exchange model such as networks, national health information services are made of the services enabling the health information sharing at the national level beyond only one health care institution such as e-prescriptions. The national health information service are e-prescriptions, lifetime health records, Telehealth, image information sharing, public health, data warehouse, e-procurement, and medical knowledge base sharing. E-prescription is known effective for improving the convenience of patients and reducing errors from the handwriting prescriptions; lifetime health records for the quality of care with the continuity of care, and the self management of chronic diseases; data warehouse for health care researchers; medical knowledge database enabling the data sharing for evidence based practice; e-procurement for efficient management of health care organizations, and so on. The development of national health information services is still undergoing and need to earn the consensus from stakeholders.

#### 2.1.4. Infrastructure for NHIN

**National health information standards.** A national health information standard is the most critical factor for the interoperable national health information system and the health information exchange. Since National health information standards committee was organized first in 2004, now the phase three of national health information standards development has been started. 10 subcommittees started at the phase one and became 14 subcommittees at the phase three now. 14 health domains were selected for the national health information standards development and the domain experts including physicians, nurses, health information administrators, medical informaticians, etc have been participating in the standards development process.

National health information standards committee adopted the Unified Medical Language System (UMLS) to create, to process, to retrieve and to integrate the national health data and information in Korea. All health data concepts collected in the first and second phases were mapped into UMLS concept. The basic principles of the health information standardization in Korea were to localize the international standards, to add the required local concepts, and to take a gradual approach. Besides the National Health Information Standards Committee, Korea Standard Organization (KSO), ISO/TC215 working groups, and other international health information standard organizations are working. The cooperation and coordination effort among health information standards organizations has been made for the efficient and successful NHIN implementation. The Figure 3 described the relationship of health information standards organizations and the strategy of collaboration.



**Figure 2 Health Information standard organizations**

**NHIN architecture.** NHIN architecture is a business and technical framework of NHIN, and the specification of functions enabling to exchange health information such as the Blueprint of Canada. NHIN architecture describes how NHIN programs are collaborating and working for the health information exchange and electronic health records. Health care organizations, health information system vendors following NHIN architecture will be able to join NHIN and to access the national health information services. This effort is made for the successful HIT adoption in Korea with the government will to provide fair opportunities to any

organizations who want to participate and to foster the many health information system vendors.

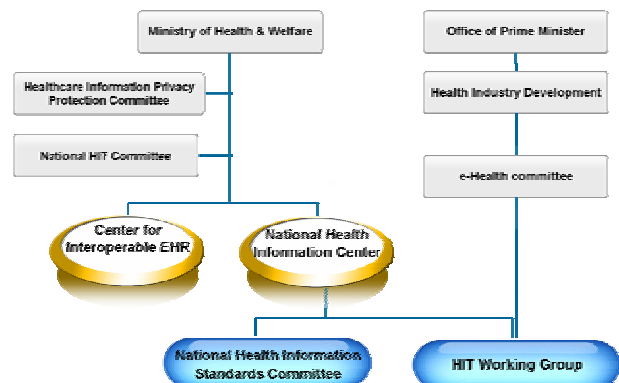
**Privacy and Security.** Privacy and security is the most critical issues not only in building consensus, also in operating NHIN in real. When NHIN plan was announced by the government, privacy and security was the first concern from the public, and many organizations. Privacy and security standards are developed with a great concern by many privacy and security experts in the subcommittee of a national health information committee.

### 2.1.5. Governing structure

The governing structure of NHIN is developed for the purpose of efficient and effective implementation of NHIN plan. Strategic planning, law, regulation and policies, standard authentication process, organizations, and incentives and strategic investment are considered for the management purpose of NHIN. Governing structure is important as much as other components of NHIN. Currently legislations including the hospital information management, individual consent on the health information use are prepared and undergoing. National strategic plan for NHIN is prepared by the government. However, standard management process and incentives and strategic investment plans are still remained for the further considerations.

## 2.2. Organizations

There are several organizations engaged in the NHIN plan to translate into action: MOHW, Health Industry Development Commission under the Office of Prime Minister, e-Health committee, Health Information Working Groups, National Health Information Standards Committees, the Center for Interoperable EHR, and etc. Some are newly established and others are to be. The coordination and cooperation among organizations are important for the efficient NHIN implementation. These are the key organizations to be described for the successful implementation of NHIN in near future.



**Figure 4 Organizations for NHIN development**

**2.2.1. The center for interoperable EHR (CiEHR).** CiEHR is a research and development institution for the



development and dissemination of the national architecture of NHIN and the core technologies of the standard based health information systems. It is established in the late 2005, and now actively engaged in planning NHIN and developing core technologies of standard based health information systems.

**2.2.2. National health information standard committee.** The national health information standard committee was established in 2004 and has been responsible for the development, maintenance of the national health information standards. 14 domains of the vocabulary, document and technology standards were selected for the development of health information standards. The standard development process will be belonging to National health information center, and the evaluation, verification and authentication process will be the responsibilities of the National health information standards committee.

**2.2.3. National health information center (NHIC).** National health information task force team, now, is the former self of NHIC. NHIC is tentative to be the main body to promote the NHIN plan and to be responsible for the development, maintenance and management, and deployment of national health information standards, and the training of skilled personnel. NHIC is in charge of all the technical supports of NHIN implementation.

### 3. Challenges

**To build consensus.** The one of the biggest challenges for NHIN development and implementation is to build public consensus by clearing up the common misunderstandings and explaining the benefit of NHIN. There have been several projects proposed for national health information implementation for the past decade. The main reason for the program failure was again, the lack of consensus by public, health care providers, payers, policy makers, health information industries and etc. Not much difference can be found in the former projects or in the concept of health information exchange, but the circumstances shown in a global trend and the strong government will for health information technology adoption at the national level have been changed. It is the very critical time for Korea to drive strongly forward the health information technology adoption at the national level to ensure the more efficient and higher level of health care system facing the aging society. We need to focus on the development of strategies for building consensus of the society.

**To increase the needs for health information standards.** Another difficulty in adopting the national health information technologies is to make people realize the needs for a national health information standard. Most of people are not aware of the importance of the standard adoption in implementing interoperable life-time electronic health records. One of

the NHIN programs, national health information standards development is not evaluated and verified the usability and feasibility of them. The pilot programs are needed to be tried as soon as possible to apply them in real life system.

**To encourage the participations of consumers and health professions.** The engagement of health professions is another key factor for the successful NHIN implementation. However, the channel of participation to NHIN plan is limited to the few who are actively engaging already from the lack of general awareness. Besides the lack of health professions' participation will, public indifference for the idea of having EHR for their convenience and consumer rights causes the ignorance of healthcare providers. There are many communication channels including public relations, seminars, public hearings, workshops, etc. Better engagement of physicians and other health professions working in real system and the public interests will lead the NHIN program to be more successful by reducing the resistance of the system users and increasing the compliance of the program.

**To assure the evidence of NHIN benefits.** The evidence which NHIN implementation could bring benefits to not only for the public also for the healthcare providers is the most effective way to persuade them. The analysis of cost and benefit of NHIN implementation are rarely reported in the past. Again, the most effective method to increase the participation rate of stakeholders is to show the benefits what they can expect. More evidence studies of NHIN implementation at the national level, EMR adoptions, health information standard adoptions, and lifetime electronic health records use are required in near future.

### 4. Conclusion

A national strategic plan for the HIT adoption has been developed by the government initiations. With the advanced information and communication technologies of Korea and the flexible attitudes toward new technology, we can expect the synergy effect from promoting the HIT adoption by the government for health care markets.

The value of the health information exchange could be much higher than we can estimate at the current status since there are many factors we hadn't had counted yet. However, the value of NHIN could be achieved only with the high level of HIT adoption. In addition, a nation-wide effort for the health information technology adoption is even necessary than partial or local effort by individuals. The successful HIT adoption can only be achieved when the various needs of stakeholders are considered with a comprehensive national strategic plan of NHIN and the benefit of NHIN is agreed.

### 5. References

[1] L.T. Kohn, J.M. Corrigan, M.S. Donaldson, and Editors, *To error is human: Building a safer health system*, National Academy of Sciences, Washington, D.C., 1999

[2] Institute of Medicine, *Crossing the quality chasm: A new health system for the 21<sup>st</sup> century*, National Academy of Sciences, Washington, D.C., 2001

[3] J. Kim, *2006 National Consumer Survey for NHIN*, Seoul, Korea, 2006.

[4] Center for Information Technology Leadership, *The value of healthcare information exchange and interoperability*, HIMSS, MA, US, 2004

[5] The center of interoperable EHR, *The National health information network development in Korea*, MOHW, Seoul, Korea, 2006

## **HIS Application status and trend in China**

-- Summary Analyses Report based on the 2005 Survey

Li Baoluo, Shen tao, Chen qi, Meng Xiaoyang

### **Abstract:**

Healthcare providing system reforming is puzzling the Chinese healthcare stockholder. HIS is a powerful tool to solve the “High cost of medication and low accessibility of healthcare” problem complained by public.

Chinese Hospital Information Management Association (CHIMA) launched a national survey project from February to September 2006 in order to analyses HIS application status and trend in China. They already collected 512 eligible responses provided by chief information officers (CIOs)/directors of information systems at healthcare facilities who were invited to participate via meeting, phone call and e-mail across the China.

This presentation will describe the Chinese general situation of healthcare system, the survey purpose, methodology and several analysis profiles: Description Statistics, Survey Respondents, IT Priorities, IT Barriers, IT Applications, Technology Adoption, IT Budget and Staff.

This summary report should be very helpful for both the domestic healthcare administrative officers/vendors and oversea HIT vender to understand the Chinese HIS current status, future trends, to make their stratagem development plan and to master the HIT market requirements.



## The Health Smart Cards and the National Health Informatics Project in Taiwan

Yu-Chuan (Jack) Li, M.D., Ph.D.<sup>1,2</sup>

1. Vice President, Asia Pacific Assoc for Med Info (APAMI)

2. President Emeritus, Taiwan Association for Medical Informatics (TAMI)

2. Professor and Chair, Institute of Biomedical Informatics

National Yang-Ming University, Taipei, Taiwan

Correspondence: [jackli@ym.edu.tw](mailto:jackli@ym.edu.tw)

The issuance of the first Health Smart Cards (HSC) from the Bureau of National Health Insurance (NHI) in the July of 2002 started a new era for personal health information exchange in Taiwan. Connected by a ubiquitous "Health Information Virtual Private Network", all the 600 hospitals, 5000 pharmacies and 17,000 clinics in Taiwan are now submitting health information to the IDC (Internet Data Center) of NHI from 40,000 card readers everyday. The HSCs are swiped through the readers 360,000,000 times a year. With visits and hospitalization information like vaccinations, diagnoses, medications and major procedures registered, the HSC has become a precursor for a nation-wide health information infrastructure.

In the end of 2005, the Department of Health initiated a 5-year plan called the National Health Informatics Project (NHIP). The NHIP is formulated in a three-layer structure, namely, the Core Architecture, the Key Applications and the Value-Added Applications.

In the Core Architecture layer, infrastructure and common platforms are laid out in 5 different component projects:

1. Information standards for Electronic Health Record and Personal Health Record
2. Ethics, legal and social impact
3. Standard platforms for health information exchange
4. Incentives and Innovations for the Health Information Technology (HIT) industry.
5. Training and education of HIT for health professionals.

The 2nd layer of the NHIP is for the Key Applications. There are four key applications planned in this layer:

1. Evidence-based medicine and clinical decision support systems for patient safety and quality of care
2. Personal health information management system for the public
3. Integrated public health information systems
4. Consumer health knowledge repository

The 3rd layer of the NHIP is for the Value-Added Applications. There are two value-added applications planned in this layer:

1. Tele-care and U-care for chronic and home care patients
2. Establishing value-added secondary medical databases from de-identified electronic health records and national health insurance research databases

With its goals set on population health, personal health, patient safety, quality of care and HIT industry, this 5-year National Health Informatics Project will hopefully become a foundation for a future health care system with better efficiency, efficacy, safety, patient-centeredness, timeliness and equitability.

## Development of Medical Informatics in Singapore -- Keeping Pace with Healthcare Challenges

Walter T.L. Lim, MBBS

*Association for Medical and Bio-Informatics, Singapore*

*Correspondence: walter.t.l.lim@singhealth.com.sg*

### Abstract

*The development of Medical Informatics in Singapore is closely intertwined with that of the overall healthcare system and has moved in concert with changing healthcare needs. This paper provides an overview of Medical informatics development in relation to key strategic directions in Singapore healthcare.*

### 1. Introduction

In recent years, the Singapore healthcare system has had to react rapidly to the changing needs of the nation and its population. There have been several key drivers of change: fundamental shifts in national demographics and the burden of disease, the threat of emerging infectious diseases and the globalization of healthcare services.

The development of Medical Informatics in Singapore is closely intertwined with that of the overall healthcare system, and thus far has moved in concert with changing healthcare needs.

This paper provides an overview of Medical Informatics in Singapore and its development in relation to some of the key strategic directions in Singapore healthcare.

### 2. Overview of Singapore Health System

As with most other developed countries, the burden of disease has shifted towards non-communicable chronic illnesses: Cancer, stroke and cardiovascular diseases together account for approximately 65% of the total causes of death.

In Singapore, there is a dual system of healthcare delivery. The public system is managed by the Government while the private system is provided by the private hospitals and general practitioners. The majority of primary healthcare is provided by private practitioners, whereas the reverse is true for hospital care. Public healthcare delivery is structured into two vertically integrated healthcare clusters, namely Singapore Health Services (SHS) and National Healthcare Group (NHG).

Medical savings accounts are the cornerstone of healthcare financing. These are supplemented by Government as well as private insurance schemes. Until recently, use of funds in medical savings accounts was restricted to in-patient stays.

### 3. Electronic Medical Records Development

#### 3.1. First Steps

The first hospital information systems in Singapore were primarily administrative systems supporting admissions/discharge and billing processes. Users of these systems were largely administrative staff and nurses.

The process of growing these systems into comprehensive electronic medical records (EMRs) has been an organic one. Initial functionality was a reflection of the system's administrative origins: patient demographics, medical alerts and electronic discharge summaries.

In para-clinical areas, pockets of computerization have arisen independently. The implementation of such para-clinical systems (Laboratory Information Systems, Radiology Information Systems) have been facilitated by the relatively compartmentalized nature and specific scope of their related work-processes. This has provided avenues for organic growth of core EMR systems, by providing additional sources of electronic clinical data.

As such, current hospital-based EMRs have grown into the role of umbrella systems that aggregate clinically relevant electronic data from various sources.

#### 3.2. Beyond Organic Growth

The organic model still offers room for expansion. New systems such as operating theatre management systems, e-Archival for paper-based records, web-based radiological image and ECG repositories continue to emerge, and can be integrated under the EMR umbrella.

However, with regards to clinical documentation, functionality has been relatively slow in developing. Fully electronic clinical documentation has been limited to specific clinical scenarios i.e. Emergency Medicine Departments, Armed Forces Outpatient Services.

To address this, the two healthcare clusters have taken steps to engage commercial vendors in the development of the next generation of EMRs. The distinguishing feature of these next generation systems is a ground-up attempt to incorporate clinical workflows such as computerized physician order entry (CPOE), clinical results tracking and nursing records. This has necessitated deeper involvement on the part of healthcare users, which could be seen as a co-development role.

These large scale efforts to develop and implement next generation EMRs are similar to efforts in other healthcare systems such as UK's NHS.

### 3.3. Towards a Life-time Health Record

The shift in the burden of disease to chronic illnesses such as diabetes, hypertension, lipid disorders and stroke has changed the mode of healthcare delivery from one that is episodic and acute to one that is life-long and involves a spectrum of healthcare providers.

The implication for medical records is that institution or provider-centric silos of information would not suffice. What is needed is a record that brings together relevant clinical information irrespective of timeframe and provider, – a life-time health record.

In this respect, there have been efforts to bring a level of inter-operability to EMRs belonging to various healthcare system stakeholders. However, promotion of data standards has been hampered by the heterogeneity of the overall EMR environment. There is no dominant EMR system and there is a great deal of variation in the stage of EMR implementation amongst providers. Part of the reason for this is the organic fashion in which EMRs have developed.

To circumvent the thorny issue of data standards, the Ministry of Health (MOH) as well as the two clusters launched the EMRX (Electronic Medical Record Exchange) initiative in 2004. The scope of the initiative was to provide an electronic platform for the sharing of medical documents such as discharge summaries, laboratory/radiological reports and prescriptions. Although this approach does not provide for full inter-operability i.e. exchange of machine-readable information, EMRX has functioned as credible bridging solution within a relatively short implementation time-frame.

The main public sector stakeholders do recognize data standards as a desirable longer term objective. This is being approached in a targeted fashion, focussing on critical medical information via the CMIS (Critical Medical Information Store) initiative. CMIS serves as a shared electronic repository for structured medical alerts and allergy data. In requiring participating organizations to adopt standards for exchange of such information, CMIS is one possible avenue for greater promotion of data standards.

## 4. Development of Integrated Care

### 4.1. Building Connections

As chronic disease management requires the involvement of a spectrum of providers and services, building connections and improving co-ordination amongst these parties is crucial in avoiding “disconnects” or lapses in the delivery of care.

There have been recent efforts to address this on a community basis. For example, the Jurong HealthConnect project [1] was launched in 2004 with the objective of connecting residents of Jurong Town with healthcare providers, services and resources in their community. One of the features of the project is a

community web portal that provides information on healthcare providers, services and resources in the Jurong area. Other than directory services, the portal also offers self-administered questionnaires that are able to suggest resources that best match a resident's medical conditions and daily needs. The web portal is complemented by a bricks and mortar information counter and a call centre hotline for residents who are less IT savvy.

On a national level, the ICS (Integrated Care Services) web portal [2] provides similar tools to connect acute hospitals, primary care providers and the public with step-down services such as nursing homes and chronic sick facilities. In addition, the previously manual process for making referrals to these facilities has been replaced by online referral forms.

### 4.2. Strengthening Provider Networks

To further develop integrated care, some of the key connections will have to be formalized. The main limitation of existing managed care schemes is that their provider networks are primarily in the private sector which is largely oriented towards primary health care, whereas the public sector accounts for the bulk of hospital care.

There is a need to form healthcare networks that include both public and private providers, so as to offer a balanced range of services from primary care to quaternary hospitals.

Within these networks, healthcare processes, in particular transfer of care processes (step-down, escalation), need to be standardized and made explicit. Although, existing shared care and disease management programmes do achieve some of these goals in terms of specifying and documenting processes, the medical informatics challenge is to streamline these processes.

This may take the form of standardized communication models, minimum datasets or electronic platforms for collaboration.

### 4.3. Patient Empowerment

The key to integrated care is still the patient. The patient has to want to connect with the network of healthcare providers and services.

One aspect of this is education. This may take the form of general education on chronic illness and self-care and is currently readily available via community and national health portals.

Specific to the individual, there needs to be individual ownership of health information. The personal health record (PHR) is one way of empowering the patient to educate himself on his individual condition.

Currently, these take the form of hardcopy personal health books that are available from public primary care facilities. Recent efforts have been to re-focus these hardcopy PHRs on key chronic illnesses such as diabetes, hypertension, lipid disorders and stroke. The key goal in this regard is not technological but rather to effect a change in information ownership attitudes. i.e. to have patients “own their condition”. However, technology such as online PHRs or portable flash memory PHRs can facilitate this.

## 5. Developing Health Intelligence

### 5.1. Infectious Disease Threats

Recent years have seen the emergence of infectious disease threats such as SARS and Avian Influenza. In responding to these threats, the requirement is not individual but aggregate information.

Traditionally one of the main sources of infectious disease intelligence has been the pool of practising medical professionals. Under the Infectious Diseases Act [3], there is a statutory requirement to report any of a schedule of 29 notifiable diseases. Currently, aggregation of disease notification data is facilitated by an online portal for electronic disease notification.

There are on-going efforts to expand the scope of disease surveillance beyond disease notification. These include the use of intelligent agents to trawl news feeds and internet sites for relevant information. The Defence Science Organization (DSO) together with the Armed Forces Medical Corps has successfully built a prototype system of this nature, Medical Large Scale Integrated Search and Analysis System (MELISA), that uses natural language processing to process large amounts of textual data.

In hospital settings, wireless positioning and tracking based on RFID and WiFi technologies have been piloted in both healthcare clusters. The aim of such systems is to provide an alternative to tedious and error-prone manual contact tracing. Some of the newer wireless patient tags have incorporated sensors for wireless monitoring of vital signs such as temperature.

As automated processes for aggregating infectious disease information grow in use, the problem shifts to making sense of large volumes of information. New tools such as Geographic Information Systems (GIS) are being used to help visualize disease information with a spatial component. The National Environment Agency (NEA) uses a GIS to map vector distribution and plan eradication efforts. The Ministry of Health is also in the process of applying GIS technologies to mapping of other diseases, for example in portraying fever clusters.

One of the possible areas for exploration is the development of similar information visualization tools to generate contact tracing “maps”.

### 5.2. Population-based Disease Management

Disease registries are the means by which chronic disease information is aggregated. Registries have proliferated both at hospital department as well as national levels.

Although national registries exist for a considerable number of chronic disease e.g. cancer, renal, cardiac, myopia, stroke, thalassemia etc., the number of hospital departments attempting to maintain their own intra-departmental registries indicates that beyond administrative or epidemiological requirements, there is an operational requirement for disease registry data.

It has not been possible to implement national registries that cater to the diverse operational requirements of a multitude of hospital departments. The future of

registry systems may be one in which registries are based at both institution and national levels, and designed to co-exist and inter-operate seamlessly.

Institutional registries could then serve an important role in bringing information back to the point of care by:

- Identifying patients overdue for assessment/treatment or not meeting treatment goals
- Stratifying patients into various categories of risk in order to target interventions at high-risk populations
- Providing information on how well individual care teams are doing in the delivery of care

### 5.3. Healthcare Performance

The concept of a standardized approach to measure healthcare quality and performance is relatively new. In spite of this, there has been a strong trend towards healthcare accreditation and rating. Part of the reason for this is that quality marks allow providers to demonstrate their quality of patient care to an evermore discriminating pool of domestic and international patients. Singapore is the only country in the world where all public hospitals have achieved Joint Commission International (JCI) [4] accreditation.

Medical information systems can serve an important role in providing the raw data for the measurement of clinical as well as business process performance, from turnaround times for reporting of X-rays to average length of stays.

Much of this data is currently extracted on an ad-hoc basis from operational systems. In order to support the on-going measurement of healthcare and business process performance, there is a need to enhance these ad-hoc processes. One possible avenue is through the implementation of the business intelligence equivalent of disease registries. In fact, there may be a degree of overlap between business intelligence systems and disease registries, as clinical performance indicators are also an integral part of the disease registry dataset.

## 6. Supporting New Models of Healthcare Financing

The denominator of performance is cost. In the financing of healthcare, Singapore has always been on the cutting edge. The government has consistently introduced new financing models to ensure continued healthcare affordability.

Medical Informatics can help streamline financial and administrative processes. This is likely to be a mutually beneficial arrangement.

For example, government subvention to public hospitals was previously based on the Casemix classification system. Although this is no longer in use, it had a beneficial effect in spurring the adoption of standardized diagnoses coding (ICD coding) amongst public hospitals.

More recently, the Ministry of Health has decided to extend the use of Medical Savings Accounts to outpatient consultations for key chronic illnesses. (previously deductions from Medical Savings Accounts

were allowed only for in-patient care) As participating primary care providers will have to provide standardized claims information, this represents an opportunity to further promote the use of data standards for the coding of diagnoses, investigations, medications and procedures. Downstream, standardized coding would facilitate the monitoring of provider performance, with the possibility of adopting a pay-for-performance approach.

Data-standards aside, the administrative cost savings from electronic claims processing may be sufficient to self-fund the IT infrastructure needed for healthcare inter-operability. This IT infrastructure could then be used for electronic exchange of clinical information.

## 7. Improving Global Competitiveness

Economic pressures are driving the globalization of Medicine.

As a consumer of healthcare services, Singapore can benefit from outsourcing components of medical care to more cost effective foreign providers through telemedicine.

Currently, teleradiology is being piloted in three public primary care clinics. X-rays are now read in India, with an average turnaround time of one hour.

Conversely, telemedicine can be used to position Singapore as a healthcare service provider in the global marketplace.

## 8. Conclusion – Enhancing Medical Informatics Education

Medical Informatics initiatives should be approached in the same empirical manner as clinical interventions. Ideally, with a strong body of supporting evidence.

However, given the sheer number of initiatives, it is unlikely that all will benefit from the presence of an experienced Medical Informatician.

One of the ways in which experienced Informaticians can maximize their value is by imparting key informatics concepts and principles to other healthcare team members. This can be done via a role as internal consultant, or in providing continuing education to other healthcare team members. This will help provide some uniformity and structure in the approach to new initiatives.

Ultimately, it is important to develop Medical Informatics as an academic discipline. Although local universities do not currently offer specific programmes in Medical Informatics, collaborations with established international centres may provide an avenue for them to do so in the near term.

## References

[1] Jurong Healthconnect Accessed at <http://www.juronghealthconnect.com.sg/eng/index2.aspx> on 20 May 2006

[2] Integrated Care Services Accessed at <http://www.ics.com.sg/ICS2006/Main/default.aspx> on 20 May 2006

[3] Infectious Diseases Act (Chapter 137) Accessed at <http://www.moh.gov.sg/corp/systems/acts/infectious.do> on 21 May 2006

[4] Joint Commission International Accessed at <http://www.jointcommissioninternational.com/> on 21 May 2006



## EHR in Japan, -Recent Government Activities

Michio Kimura

Hamamatsu University, School of Medicine, Hamamatsu, Japan

Correspondence: kimura@mi.hama-med.ac.jp

### Abstract

*This paper figures Japan's activities concerning EHR since 1993, according mainly to related government activities.*

*This also includes current status of EMR and CPOE in Japan, relatively high rate of installation in the world.*

### 1. Healthcare Delivery System in Japan

In Japan, "hospitals" are defined as healthcare provider with more than 20 beds. In 2005, we have 9333 hospitals, and about 90000 clinics. Averaged revenue is 137Myen per 100beds, and averaged employees are 112 per 100beds.

Medical care insurance system is based on "fee for service" payment system, while government is looking to prospective payment system, which is already applied to some 200 large hospitals. All citizens are covered by either employees healthcare insurance or community-based health insurance.

### 2. EMR Strategies in Japan[1]

Past major government activities for EMR are as follows;

Healthcare Information Systems Advisory Committee ...10/93

"Healthcare Information System Strategy 21" ..7/94

New Healthcare Information System Joint Committee ...11/94

Research and Development for EMR (purse 800Myen) ...6/95

EMR was authorized as Formal Document ...4/99

HELICS(Standard Board) was established ...5/'00

IT Grand Design for Healthcare System ...12/'01

Aid Money for EMR Installation (half aid, total purse 20Byen) by MHLW ...'01/02

Aid Money for Regional EHR (total purse 25Byen) by METI ...'02/03

Privacy Law ...5/'03 (effective 4/'05)

Project on Interoperability for Health IT ...06/'04.

Among them, IT Grand Design for Healthcare noted in 2001, as;

To provide practical strategies and objectives to achieve annual and numerical targets in order to promote information system

To present measures to achieve the objectives set for each phase of information system, and lay down an action plan that provides roles and targets for public and

private sectors

And its action plan says;

1. Standardization in healthcare[2]

2. Infrastructure for information system

3. Implementation of model projects

4. Subsidy for introduction and maintenance of information system

5. Raising awareness

Its target stated;

By the end of the 2004 fiscal year

At least one institution in every secondary medical area across Japan should have an electronic medical record system. At least 60% of the hospitals should have an e-claim system.

By the end of the 2006 fiscal year

On a national basis, at least **60% of the hospitals with more than 400 beds** and 60% of the clinics should have an electronic medical record system. At least 70% of the hospitals should have an e-claim system.

### 3. Current Status of Medical Information System in Japan

Order entry system

**65%** or more of the hospitals with more than 400 beds, in 2004.

Japan is one of the highest rate of order entry installation in the world.

Electronic medical record system

**12%** of the hospitals with more than 400 beds (as of the end of 2004 fiscal year) and 3% of clinic.

This is based on definition of EMR as "paperless". Author's study on Shizuoka prefecture hospitals shows this rate as 48%, which is based on the definition of "quick and thorough use of order entry database which serves information to patient, not necessarily paperless".

### 4. Recent Government Activities

#### 4.1. Interoperability of Health Information Systems Project by METI ...'04-'07

Its total budget is 1.5Byen. It comprises

1. IHE-J activity support, which is Connect-a-thon support, and installation support as showcase at Saitama Medical College Hospital.

2. Inter-EMR fundamental data set for system update/vendor change by JAHIS(vendor assoc.)

3. HL7 tools development

4. Health IT CIO training course.

#### **4.2. Standardized EMR Promotion Committee Report by MHLW...05/'05**

Its report included;

1. redefinition of EMR dreams,
2. promotion of standards, i.e. HL7, DICOM, HL7 CDA, and other codes & terminology are recommended,
3. Interoperability Promotion
4. Promotion Strategy.

#### **4.3. MHLW promotes Shizuoka Prefecture EHR for Nationwide Use ...04/'06**

Shizuoka Prefecture EMR project produced many EMR components, such as progress notes, nursing observations, referral documents, clinical research database, PACSystem, and formatted document system[3],

These components are free for Shizuoka prefecture hospitals (software package only, hardware and maintenance not included). They are operable by information from order entry system in HL7 standard. Ministry invested 88Myen to Shizuoka prefecture to improve its components and let new these components freely available nationwide.

On the other hand, Ministry approved for healthcare providers to collect 3000 yen (example) for handing out clinical data CD to patients, provided that it is in the recommended standard (HL7 and DICOM).

#### **5. IT Strategy 2006 by Cabinet's Office, ...06/'06, Primary Accent is on Healthcare System Restructure by IT**

Its action plans are;

1. New Grand Design for IT in Health ...by 2006
2. Prevention Medicine and EMR Promotion by Interoperable Healthcare & Health-checkup Information Healthcare Public Key Infrastructure ..by 2006, which comprises;
  - Safe and inexpensive broad network,
  - IC card feasibility study, standardization promotion,
  - Measurement for IT in healthcare providers,
  - Interoperable EHR for large hospitals (By 2008 for 400+ beds, by 2010 for smaller),
  - Start implementing standardized exchange from 2006 (Nationwide Shizuoka Software)
  - Standard codes recommendation by 2007
  - Interoperability test (Connect-a-thon) for EHR from 2007
  - Support for IT in smaller healthcare providers (Nationwide Shizuoka Software)
  - Incubation of CIO in healthcare providers,
  - Ontology development from 2006

Development of health information gathering and utilization in EHR era

Health checkup data utilization study from 2007  
 Health checkup mandated from 2008 for all 40+ ages  
 Reimbursement claim data utilization study from 2007  
 Handling health data management by patients study by 2008 (Shizuoka Style CD).

3. Online Reimbursement Claim by 2011  
 e-Document law applied to healthcare  
 Incentive promotion for e-claim from 2006  
 Standard codes mandated by 2010  
 Simplification of reimbursement scheme  
 Electronic reimbursement tariff from 2008  
 Online real time verification of insurance from 2011

4. Communications  
 Telemedicine,  
 Surface television broadcast,  
 IC tags in materials and medicines.

#### **6. Final Remarks -- After All,,**

Who is paying for EHR?  
 (Within Healthcare Providers, Network and Registries)  
 What is going to be the Terminology?  
 (SNOMED-CT talk with WHO ICD11, Global vs Local Culture).

#### **7. References**

- [1] Kimura M.: Roles of Health Informatics Standards in supporting seamless healthcare and improving health outcomes, eHealth Asia 2004 Conference, Kuala Lumpur, Malaysia, April 6, 2004.
- [2] Kimura M.: Standard of Health Information – What is Going on and What is Helps, Medinfo 2004, Sanfrancisco, USA, September 7.2004.
- [3] Kimura M.: General Purpose Portable Data: MERIT-9 Referral document conforming both HL7 CDA R2 and IHE PDI (Portable Data for Image) in Shizuoka Prefecture EHR project, The 6<sup>th</sup> HL7 International Affiliates Meeting & The 4<sup>th</sup> Asia-Pacific HL7 Conference on Healthcare Informatin Standards, Taipei, Taiwan, July 21, 2005.

## eHealth trends in Pacific Islands 2006

Norio KIMURA<sup>1)</sup>, Isao NAKAJIMA<sup>1)</sup>, Muhammad Athar SADIQ<sup>1)</sup>, Yasumitsu TOMIOKA<sup>1)</sup>,  
Kenji TANAKA<sup>2)</sup>, Kiyoshi IGARASHI<sup>2)</sup>

1: Tokai University School of Medicine, Japan

2: the National Institute of Information and Communications Technology, Japan

### 1. Purpose

In the past APAMI 2000 held in Hong Kong, we have reported eHealth trends in Pacific island. In this 2006 APAMI, we also intend to report our study and investigation concerning about their recent activities. Because there are trends to help and support the public health and/or medical care in developing countries, in particularly the Government of Japan, ITU, WHO, UNDP, and other UN organizations. In 2006, around the Pacific, there are three major projects running that covered wide area, while small populations.

### 2. Investigation

Health professionals in various Pacific Island countries, particularly those working in rural and remote areas, have less opportunity than their peers in developed and even some developing countries to keep abreast with both technical and technological advancements in medicine and healthcare.

### **POLHN ( Pacific Open Learning Health Network )**

For their continuing education and professional development, distance education using a variety of media is an option that has been support via funding from the Japanese Government.

In general, POLHN is a successful initiative based on actual web-based operation via internet with package media, like CD-ROM, the evaluation report, and the meeting discussions. The pilot open learning courses are well received by health professionals and the learning centers in each country, equipped with information, communication, technology (ICT) for open learning and health educational resources, are being used by countries for their health professional training and continuing education, including ICT skills trainings. There was a strong interest among the participating countries to contribute to the long-term sustainability of the POLHN and a desire for greater ownership by the countries, of the open learning centers. There was also consensus that it would be important to integrate POLHN as a component of national human resources for health programmes of the ministries of health, and that the network's coverage should be expanded to include other countries and to make it a Pacific regional learning network. There are two phase projects in the POLHN. Phase One: the establishment of computer laboratories as resource centers for the health professionals (HPs) in ten countries in the Pacific (Cook Islands, Fiji, Kiribati, Marshalls, Federated States of Micronesia, Palau, Samoa, Solomons, Tonga and Vanuatu) has been completed. Computer laboratory equipped with computers, servers, networked printers, data projectors, web cams and Internet connectivity. Various learning and reference materials are also available in the resource centers including the blue trunk library kit.

Phase Two: using the computer laboratories as resource centers for the continuing education of professional health workers of the Pacific started in later 2003.





**Pacific Basin Telehealth Initiative**

The U.S.-affiliated Pacific Islands jurisdictions represent over 2,200 small volcanic and coral islands stretching across 4,500,000 square miles, with about 400,000 inhabitants living on a landmass smaller than the state of Rhode Island. Due to their geography and isolation, coupled with limited finances, poorly developed infrastructures and lack of skilled staff, these islands provide an ideal environment for developing and implementing telehealth applications.

In 1999, the Health Resources and Services Administration (HRSA) launched a multi-year Pacific Basin health initiative to bring telemedicine to the region. This initiative came out of recommendations from two events the previous year: a HRSA-sponsored summit meeting in Rockville, Maryland with the health ministers from the six U.S.-affiliated Pacific Basin jurisdictions and the publication of the Institute of Medicine's report "Pacific Partnerships for Health: Charting a New Course." Both the report and participants at the summit identified telehealth (i.e., telemedicine and distance education) as a key strategy for improving primary care delivery, enhancing prevention activities and supporting the training of health personnel. This telehealthnet work has been funded by the OAT( Office for the Advancement Telehealth) of the U.S. Government.

**PPHSN ( Pacific Public Health Surveillance Network)**

"Training in applied epidemiology and public health surveillance" is one of the five strategies of Pacific Public Health Surveillance Network (PPHSN). It aims to develop a regional pool of experts in public health surveillance and response for the Pacific Islands. The PPHSN was created in 1996 under the joint auspices of SPC (South Pacific ) and WHO. The Goal is to improve public health surveillance and response in the Pacific Islands in a sustainable way. In 1996, the idea of a modular training programme with field-based components, involving universities, health development agencies and professional associations came up. This programme should be accredited by a training institution. From 1998 to 2001, SPC made a first move in that direction and ran two different series of subregional training sessions in surveillance, outbreak investigation and the use of Epi Info 6 software for surveillance activities. The overall objective of these training sessions was to build a critical mass of health professionals who share a common set of tools and methods for public health surveillance. SPC can also provide attachment and hands-on field training opportunities for trainees in public health practice. A memorandum of understanding was signed between SPC and FSM in 1999. The article suggests the areas of future FSM/SPC collaboration in public health surveillance: the accreditation of SPC courses by FSM; the addition of a microbiology component; the identification of opportunities for field training; and the evaluation and harmonisation of the training programme(s).



### 3. Considerations and our action

The objectives of our cooperation related Pacific Island through ITU, APT, NICT, Sasakawa Fund, and other private sectors in developed nations is to bring research teams in health care field, governments sectors in island, and other stakeholders in the region to discuss the different applications of telecommunication and information technology in the health care sector. Based on our investigations, we intend to invite some researchers from Pacific Islands to provide an opportunity to study more advanced telehealth at the ITU Telemedicine Expert Training Course in the Shonan Campus, Hiratsuka Japan from 2006 to 2009.

To familiarize themselves with the specifications and implementation details of telehealth;

To assist in the management of the transition process from IPv4 to healthcare specialist to end user(client); and

To share information on the development and deployment of telehealth to other countries and regions.

### Acknowledgments

This study was supported by the APT-HRD, NICT and PIDO (a self-funded initiative of the Sasakawa Pacific Islands Nation Fund)

### References

01. I. Nakajima, S. Chida: Telehealth in the Pacific: Current Status and Analysis Report(1999-2000), J. of Medical Systems, 24, 6:pp321-331, 2000.

## Health Informatics in the Philippines

Alvin B. Marcelo, MD, FPCS

*National Telehealth Center, National Institutes of Health, University of the Philippines Manila*

*Correspondence: [alvin.marcelo@telehealth.ph](mailto:alvin.marcelo@telehealth.ph)*

### Abstract

*The progress of biomedical informatics in the Philippines has been fraught with many highs and lows. Called 'health informatics' locally, the field has been in informal and formal development for the past ten years since the first professionals commenced investing time and energy to pursue the field as a distinct discipline. Although local efforts have been hampered with infrastructure issues, there have been many activities that have provided strategic foundation for the implementation of future activities. Foremost among these is the Master of Science in Health Informatics offered by the University of the Philippines Manila. The program offers a unique approach to learning health informatics by putting emphasis on community-based and community-managed health information systems that are appropriate for resource constrained environments. It is also unique for its heavy adoption of the principles of primary health care (as manifested by the Declaration of Alma Ata) in its curriculum design and implementation.*

### 1. Introduction

As a science, biomedical informatics had been loosely practiced in the Philippines as early as the nineteen eighties. Residents in tertiary care facilities who had access to IBM compatible machines were already using word processors to store patient information. In other areas, anecdotal evidence of database management systems was used for storing patient information. Epidemiologists from the Department of Health who took the Field Epidemiology Training Program (DOH-FETP) used Epi-Info<sup>1</sup> extensively in their practice. In fact a professional association, the Philippine Association for Medical Informatics (PAMI) led by Dr. Benjamin Marte was formed with membership mostly composed of Department of Health staff. The PAMI was able to represent the country in initial meetings of the Asia Pacific Association for Medical Informatics and was the de facto representative for the country in APAMI for the years 1994 to 1995.

It was however during the late nineties when a more structured approach to 'medical informatics' (as it was called then) appeared. This was through the establishment of the Medical Informatics Unit in the

University of the Philippines Manila. Preceding the MIU however were several important events. First there was the incorporation of a non-governmental organization called the Philippine Medical Informatics Society, Inc composed mostly of academicians from the University of the Philippines Manila (UPM). Then the PMIS initiated many of the scientific activities related to the field such as the First Symposium on Medical Informatics and Seminars on Telepathology (Dr. Paul Fontelo, Armed Forces Institute of Pathology), Medical Records Management (Dr. Michael Yang, OACIS), and Integrating Technology into Medical Education (Dr. Emmanuel Besa, Medical College of Pennsylvania). However, although these activities served to keep the interest and awareness in medical informatics, they were too far in between to actually push the science forward in the country.

### Academics

By 1998, faculty members from the UP College of Medicine began taking formal courses in medical informatics in different institutions around the world; in particular, at the University of Washington, the National Library of Medicine in Bethesda, Maryland, and University of Warwick in Coventry, England. These faculty members served as the core group of what will be the Medical Informatics Unit at UP Manila.

### Research

Upon return, the faculty core group started working on health information systems projects and applied many of the principles they learned from their training. Most of these researches were conducted at the Medical Informatics Unit of the UP College of Medicine while others at the World Health Organization and the Department of Health. With the experience from these projects, faculty members were able to refine their techniques and protocols to address the growing informatics needs of the country. One such project was the Community Health Information Tracking System<sup>2</sup>, a Linux, Apache, MySQL, PHP-based system released under the general public license (GPL). CHITS was named finalist at the Stockholm Challenge 2006 and one of top three e-government projects in the Philippines by the Asia Pacific Economic Cooperation Digital Opportunity Center. Several other projects with the

United States Agency for International Development and International Development Research of Canada provided the experience for the faculty and staff of the MIU to handle increasingly complex projects in medical informatics.

By 2003, the faculty members realized the need to generate new blood and to formulate a structured learning program for medical informatics. A proposal for a Master of Science in health informatics was crafted together with the College of Arts and Sciences (who had a program in bioinformatics). This combined program, now called Master of Science in Health Informatics (MSHI) was finally approved in 2004 and implemented in academic year 2005-2006.

## Service

A key component of the MSHI was service as well as research. As designed, students of the MSHI are expected to implement the principles and concepts of health information systems development within an existing environment relevant to their practice. As seen with the first batch of students, most of these projects are for databases for their professional organizations or specialty affiliations.

## Context

A unique spin in medical informatics in the Philippines is the bias of pioneering faculty on community health informatics. This is due to the fact that most of the initial efforts began at the UP College of Medicine where a community-oriented curriculum serves as the central recurring theme in all institutional activities. This community-orientation permeates the design and delivery of the MSHI, and in the program students are asked to immerse themselves and to integrate with local health facilities to understand the issues in health information management at the grassroots level. This is very different from the approach of first world countries who expose their students to complex, sophisticated, and resource-intensive health facilities, many of them at the secondary or tertiary level. In fact, the MSHI program is unique in that it revolves around patient level data within a public health system infrastructure forming the staging area for the development of public health informatics in the future.

A clear manifestation of this philosophy is the staunch advocacy of the UPM for the use and promotion of free and open source software (FOSS). In fact, the National Telehealth Center (an institutional product of the MIU) has been named as UNDP Regional Center of Excellence for Free and Open Source Software and node of the International Open Source Network<sup>3</sup> for ASEAN+3.

## Issues

Despite the many developments in health informatics in the Philippines, the nation still suffers from several issues that impede progress. Foremost is the lack of health human resource interested in the field. Most of the initial enthusiasts were clinician specialists who were engaged in health informatics more as a novelty rather than as a profession. When the pull of economic and professional constraints are felt, priorities shift towards clinical responsibilities to the detriment of the health informatics discipline.

A second problem is the network infrastructure (which also involves IT human resource). While connectivity is an important component of a health information systems, the availability of affordable IT human resource and their retention in the age of globalization remains an issue.

Third, the benefits of information technology have not yet dawned to many decision-makers in the health sector. The huge capital outlay for a health information system remains a stumbling block to the integration of IT in health operations.

## Opportunities

An exciting development in the country is the rapidly maturing wireless cellular network with value-added services. The Philippines is the texting capital of the world and is home to 300 million messages a day. It is also a large laboratory for SMS based applications that range from games to banking. Definitely, with a wide penetration nearing fifty percent<sup>4</sup>, SMS is a protocol waiting for a health informatics application.

A joint cooperation of the Department of Health, Department of Science and Technology, the University of the Philippines Manila, and the Philippine Health Insurance Corporation to formulate standard secure health messaging across health facilities is opening up a lot of business opportunities both for health informatics practitioners and for IT vendors as well. Called the Philippine National Health Information Infrastructure (PNHII), it will form the framework for the automation of many health transactions and may eventually pave the way for the justification of IT investments in health.

The BuddyWorks Telehealth project of the National Telehealth Center, an e-government project, is nearing completion and offers the possibility of managing electronic patient records needs of the government health bureaucracy. BuddyWorks is based on open source and is flexible enough to plug into any system that will be chosen by DOH. It can also serve as the reference implementation of the Philippine National health Information Infrastructure.

## Next Steps

At center stage in health informatics in the Philippines is the Master of Science in Health Informatics program. Designed to produce the leaders who will push for the needed organizational and technological changes in the health sector, the MSHI program aims to participate actively in the efforts of the PNHII and of BuddyWorks.

The cellular infrastructure will play a major role in the delivery of health messages because of its wide coverage in the country.

The Philippines will continue to focus its informatics to the real needs of the Filipino people, especially those who are marginalized and underserved.

<sup>1</sup> Epi-Info. <http://www.cdc.gov/epiinfo/>

<sup>2</sup> Tolentino H, Marcelo A, Marcelo P, Maramba I. Linking primary care information systems and public health information networks: lessons from the Philippines. *Stud Health Technol Inform*. 2005;116:955-60.

<sup>3</sup> International Open Source Network. <http://new.iosn.net>

<sup>4</sup> Lucas D. Cell phone penetration rate seen to peak at 50%. [http://news.inq7.net/infotech/index.php?index=1&story\\_id=78818](http://news.inq7.net/infotech/index.php?index=1&story_id=78818)

## Thai Medical Informatics Association and medical informatics activity in Thailand

Wansa Paoin MD.

[wansa@inet.co.th](mailto:wansa@inet.co.th)

### Abstract

*Thai Medical Informatics Association(TMI) had established in 1991. Until now it had devoted most of their efforts on many works to make Thailand fully implemented medical informatics sciences and technologies for enhance and promote health status of the Thai people. And to make the Thai medical alliance workers fully literate on medical informatics. During 2004-2006 Medical informatics activities in Thailand focused on standardization of health data, coding and classifications, Medical Imaging and PACS(Picture Archiving and Communication System) and open source applications in hospital information systems. In the near future, Thai Medical Informatics Association will expand their activity further into various activity like conduct short course training on some issues of medical informatics, research further on HL-7, SNOMED modification and implementation in Thailand. Enhancing organization structure by recruit full time staff, upgrade hardware and web server. Increase activity with the Asia Pacific Association of Medical Informatics(APAMI).*

### 1. Introduction

Thai Medical Informatics Association(TMI), formerly known as Thai Medical informatics Society, was established in the year 1991. On vision of senior doctors of The Consortium of Medical Schools at that time when computer usage in medical facilities began in Thailand. Started with 300 members who work in medical fields and computer sciences like medical doctors, nurses, clinical coders, programmers, system analysis etc. Now the Thai Medical Informatics Association had more than 600 members. Main activities are Annual conference, Special Interest Groups seminars, affiliated with the Ministry of Public Health in promotion and implementation of standards, concepts and new technology of medical informatics in Thailand.

### 2. Brief History

When first established in 1991 most of the founder members came from existing medical schools(in that time there are twelve medical schools in Thailand. Main objectives of Thai Medical Informatics Society were 1. To be a center for coordination and distribution of medical informatics knowledge. 2.To develop models of administration, education and research in medical informatics. 3.To be a center for knowledge sharing in

medical informatics. 4.To encourage continuous education, studying and practicing for medical informatics staffs. 5. To be a resource centers in medical informatics for the Health institutes both in public and privates sector. 6. Not to be involved in commercial and politician activities

During 1991 – 1999, Main activities of the society are annual conferences that was held in many provinces around Thailand. Those provinces are Khonkaen(1991), Songkhla(1992), Chiangmai(1993), Bangkok(1994), Petchburi(1995-1996), Bangkok(1998), Choburi(1999). Which total attendants around 600- 1300 persons each year. Main theme of each conference was a current intereriting issue or new technology of medical informatics in that year.

Since 2000, 4 Special interest groups were established in the society. New activities from special interest groups emerged including seminars, researches and small group meetings. Main annual conference were conducted in Khonkaen(2000), Bangkok(2001-2003), Pathumthani (2004), Chiangmai(2005). In 2005 the society was changed to Thai Medical Informatics Association.

During the same period medical informatics gain medium speed of acceptance in Thailand. Started for computer usage in medical records, upgraded to hospital accounting system until now most computer usage in hospital of Thailand are hospital informations system. Some hospitals can developed further to clinical information system, but most hospitals could not achieve that state.

Medical informatics were studying as a subject in 3-4 medical school. There was 1 course of medical informatics developed and conducted in Ramkhamhaeng University, but no additional course emerged from another university until now.

Researches on medical informatics in Thailand most happened in the Ministry of Public Health projects and some medical schools. Research topics included standardization of data, ICD-10-TM(Thai modification), Thai DRG(Diagnosis Related Groups), Telemedicine, Medical Imaging and PACS(Picture Archiving and Communication System), Open source applications in hospital information systems.

### 3. Situation Analysis

SWOT Analysis done by Thai Medical Informatics Society before changing to association revealed that Thai Medical Informatics Association has strengths on flexible organization structure, multiple stakeholders coordinations and teamwork, knowledge sharing



empowerment and acceptance from both public and private health sectors. But had weaknesses on inadequate full time staffs, organization planning and management, lack of long term funding and most central committees had another full time jobs so those key persons can not devoted much time on the association activities.

Opportunities for TMI are increasing demands on medical informatics knowledge and implementations, new young generations people interesting in the fields and coordinative works with the Asia Pacific Association of Medical Informatics(APAMI). Threats include deficiency of academic course and curriculums for medical informatics in Thailand, deficiency of resource persons in medical informatics and lack of new knowledge and innovations.

Overall Thailand situation of medical informatics activity could be separated into 2 groups of activities. Implementation of new technologies of medical informatics mostly occur in private healthcare sectors. But most of the research activities in medical informatics occur in the Ministry of Public Health and medical schools.

#### 4. Recent and current activities

During 2004-2006 Medical informatics activities in Thailand focused on standardization of health data, coding and classifications, Medical Imaging and PACS(Picture Archiving and Communication System) and open source applications in hospital information systems. These activities would be discussed below.

##### 4.1 ICD-10-TM(Thai Modification)

ICD-10(International classification of Diseases and related health problems, 10<sup>th</sup> Revision) was implemented as a tool for encoding morbidity and mortality data in Thailand since 1994. After 1997 it was used extensively in all hospital and health facilities across the country both in public and private health sectors. Users of ICD-10 found that the classification comprehensively include all codes for the diseases occur in the western country, but lack some codes for tropical area for example, ICD-10 had no code for necrotizing fasciitis, dengue shock syndrome, wheezing associated respiratory illness, cobra and green pit viper snake bite etc. So it was apparently known that if a country need to collect data on diseases frequently occur in this area but no specific code found in ICD-10, that country should add some more codes under original framework of ICD-10 to match their requirement of code. This phenomenon had occurred in several countries around the world like USA(ICD-10-CM), Australia(ICD-10-AM), Canada(ICD-10-CA).

In the year 2000, working groups were formed in the ICD-10-TM research project conducted by Bureau of Policy and Planning in Health, Ministry of Public Health. Also affiliated by 13 specialist Royal colleges(groups of medicine experts ie. internal medicine, surgery, OB-Gyn, pediatrics etc.), Thai

medical council, 4 Association of medical specialist in oncology, dermatology, family medicine, emergency medicines, modification of codes was finished in the year 2003. About 1,500 new codes were add to original 14,200 codes of ICD-10. These working groups also developed new procedural coding system to be use as a doctor fee coding schema.

Pilot study for implementation of codes was conducted in 25 hospitals across the country. The results show that new codes were used in 5-10% of all codes in morbidity data. Preliminary research on cost analysis for diseases also found that new code can be used to discriminated high cost diseases and result in better reimbursement to both the hospital and the universal coverage agency.

In the year 2005, the World Health Organization revised original ICD-10 to be ICD-10 2<sup>nd</sup> edition. New codes from ICD-10 2<sup>nd</sup> edition also were added into ICD-10-TM. Now it became ICD-10-TM 2<sup>nd</sup> edition.

##### 4.2 Thai DRG(Diagnosis Related Groups)

Universal coverage in health care system were studied by health system researchers since 1990. From that time, a working group on casemix system studied various casemix system in the world and conducted several researches to find suitable model to be used in reimbursement process under universal coverage schema. That model was known as Thai DRG version 1 with 500 patient groups developed in the model.

In the year 2001, new government leaded by Thairakthai party fully implemented the universal coverage in the public healthcare sectors(the 30 bahts for every disease campaign). The old Thai DRG model was developed further including about 1,000 groups into the system. That became Thai DRG version 2. It was used extensively as a tool for calculate the reimbursement rate for every hospital provided healthcare under the universal coverage system.

2 years later after trial and error period, errors were reported on some grouping logic. Working group then developed Thai DRG further into version 3 and 3.5 until now the Thai DRG included about 1,400 groups and be used as a major tool in reimbursement process of the universal coverage health care system.

##### 4.3 Picture Archiving and Communication System (PACS)

Picture archiving and Communication System(PACS) implementation in hospital computer system Thailand in the past were increasing at slow speed. The major challenge was high costs of the system both hardware and software. Main activity now are studying of the way to implement PACS at low cost but achieving it's effectiveness and efficiency.

Special interest group on PACS was formed in the year 2000 with 2 subgroups. The first group was studying the storage, networking, retrieving and display



technology to find the most cost effective technology that should be implemented in Thailand. The second group was studying open source freeware and shareware on PACS/DICOM servers and clients, tools, libraries.

Now the first group can implement low cost PACS system in one of the medical school as a pilot model. The low cost PACS pilot model was implemented at the Ramathibodi hospital, Bangkok, Thailand since 2002. While the second working group developed open source PACS/DICOM libraries software, named *dcm4che*.

### 4.3 Open source applications in hospital information systems

Hospital information system in Thailand development in Thailand began around the year 1991 start from electronic medical records module. In that beginning periods most of programming processes were performed by a group of medical doctors who devoted their efforts to created hospital information systems to be used in their hospital. Some doctors share their program to other hospital to be use free of charge.

Commercial hospital information systems came around the year 1995, mostly implemented in the private hospitals and acquired popularity not so rapidly. Until now both open source application and commercial hospital information system gain equal popularity. Most of the public hospital in the Ministry of Health implement open source applications, while most of the private hospital used commercial hospital information systems.

Special interest group on open source applications in hospital information systems was formed in 2000. The group developed their open source hospital information system named *HospitalOS*. The second group appeared in 2003, developed another open source hospital information system called *HospitalXP*. Both groups distributed their software across the country to be used in more than 100 hospitals now.

### 5. Future Plan

In the near future, Thai Medical Informatics Association will expand their activity further into various activity like conduct short course training on some issues of medical informatics, research further on HL-7, SNOMED modification and implementation in Thailand. Enhancing organization structure by recruit full time staff, upgrade hardware and web server.

Increase activity with the Asia Pacific Association of Medical Informatics(APAMI).

The Thai Medical Informatics Association also have a plan to enhance implementation of medical informatics sciences in Thailand , plan to co-ordinate with universities to developed curriculum and course on medical informatics study in various degree of educations.

### 6. Conclusion

While we had done our works and activities for 15 years, the Thai Medical Informatics Association have a long way to go to accomplished our final goal. That is to make Thailand fully implemented medical informatics sciences and technologies for enhance and promote health status of the Thai people. And to make the Thai medical alliance workers fully literate on medical informatics.

### 7. References

- [1] Rames Wacharasin, Preecha Rukpolamueang, "History of the Thai Medical Informatics Society", *Proceeding of TMI annual conference 2001*, Thai Medical Informatics Society, Bangkok, Nov. 2001.
- [2] Choosna Makarasarn, "TMI strategic plan", *TMI Newsletter*, 2005; Volume1, June 2004, pp. 2-3.
- [3] Wansa Paoon, "ICD-10-TM development and implementation in Thailand", *Proceeding of TMI annual conference 2004*, Thai Medical Informatics Society, Bangkok, Dec. 2004.
- [4] Pannarunothai S, "Thailand Casemix System", *Proceeding of TMI annual conference 2004*, Thai Medical Informatics Society, Bangkok, Dec. 2004.
- [5] Nithipong Ballawanich, Suthee Poojareonchanachai, "Technology for the third generation PACS", *Proceeding of TMI annual conference 2004*, Thai Medical Informatics Society, Bangkok, Dec. 2004.

## Health Information Management: Malaysian Experience

Datin Dr.S.Selvaraju  
Health Informatics Center  
Ministry of Health, Malaysia  
[selvaraju@moh.gov.my](mailto:selvaraju@moh.gov.my)

### ABSTRACT

*The evolution of Health Information Management in Malaysia has started from a basic statistical reporting system through papers to an ICT enabled Health Information Management System. The building blocks to facilitate the implementation of such system are through the development of Health Informatics standards and a mechanism for the adoption of such standards. The legislative support required to enable a wider adoption by all stakeholders is vital for successful implementation. This paper provides the historical perspective of the transformation of Health Management Information System to Health Information Management System, the challenges and the way forward in realizing a National Integrated Health Information System.*

1.2 Health care in Malaysia is provided by the public and private sectors and non-governmental organizations. The major provider and financier of health services is the Ministry of Health (MOH). Basic health care through health facilities is currently available to and accessible to more than 95 % of the population of Peninsular Malaysia and about 70 % of the population in Sabah and Sarawak. Secondary and tertiary care services are now widely available in government hospitals and private hospitals across the country. In 2005, there were 2,877 primary health care clinics, 125 public hospitals (including institutions) under the MOH and 9,410 doctors working in public hospitals.

### 1. INTRODUCTION

1.1. Malaysia is composed of Peninsular Malaysia and the states of Sabah and Sarawak on the island of Borneo. The total land area is 330 252 square kilometers. Malaysia, a constitutional monarchy with a non-elected upper house and an elected lower house, is comprised of 13 states and three federal territories. In 2005 Malaysia has a population of 26, 127,700 million, more than 60% of who live in urban areas. The population is relatively young with 42.3% being below the age of 20years & 6.6% aged 60years and above. Its ethnically diverse people are comprised of 65% Bumiputera, 26.1% Chinese, and 7.6% Indian. Life Expectancy (LE) at birth for both genders has increased over the years. Malaysians enjoy a relatively high overall standard of health. In 1990, the LE for males and females were 68.8 years and 73.5 years respectively. The LE for males and females in 2005 had improved further to 70.6 years and 76.4 years respectively

1.3. Over the past few decades there has been an increasing role of the private sector in the provision of health care for the country. However, there is inadequate integration between public and private health services. There were 225 private hospitals in 2005 compared with 50 hospitals in 1980, mostly located in urban areas. Doctor to population ratios in 2005 ranged from 1 doctor per 1,268 populations in Peninsular Malaysia to 1 doctor per 2,765 populations in Sabah and 1 doctor per 2,115 populations in Sarawak. In 2004 there were 8,836 doctors in private hospitals accounting for 46% of total number of medical practitioners but private hospitals only accounted for 21% of the total hospital beds in the country.

1.4. The Healthcare delivery System that is strategically designed to improve the health status of the population is one that focuses on wellness and empowerment. It provides focus on the individual and community responsibility for health. Health system will be able to support the access to right information at the right time to make the right life style choices. Access to an individual's, longitudinal medical and health record will support continuity in care. The use of ICT as an enabler to further improve health care delivery is being planned within the following context:

- 1.4.1. Informed and knowledge individual and population that will give consumers more access and involvement in healthcare decision making and be more responsible in taking care of their own health
- 1.4.2. A proactive wellness and illness management through competent and knowledgeable healthcare providers
- 1.4.3. The delivery of service in a networked environment based on care network concept. Sharing of information and resources shall be the thrust for efficient and effective delivery of services.
- 1.4.4. There will be a provision for services to be offered from home where applicable where patients can have access to health information and being able to interact with healthcare providers through multimedia network.

1.5. During the Eighth Malaysia Plan (2000-2005) concerted efforts were undertaken to provide a stronger platform for the country's transition towards knowledge based economy. There was increasing effort on the promotion and use of ICT in various sector including health sector .For the Ninth Malaysia Plan ( 2006-2010) efforts will be intensified for the wider adoption and use of ICT .The government is committed to enhance ICT related skills and competencies as well as info structure expansion.

## 2. EVOLUTION OF HEALTH INFORMATION MANAGEMENT IN MALAYSIA

### HISTORICAL PERSPECTIVE

2.1. The collection and collation of health information for routine statistics reports was developed since colonial days and which was further strengthened to a Health Management Information System. During the British Civil Administration days (1874-1941), health information and events

Were recorded and documented as a by-product of the services provided by the Medical Department. Data were generated at the services outlet, namely at hospitals, static (outdoor) dispensaries, traveling dispensaries maternal and child welfare centers, municipal clinics, leper settlements and mental institutions. Some form of procedure existed for recording into registers and collection of medical and health data on some predetermined formats. Sanitary inspectors and public health nurses in the course of their work carried out the collection of data for the public health activities.

While in the hospitals, static (outdoor) dispensaries and institutions for special diseases this was done by hospital assistants and nurses. The data collection was coordinated by Medical Branch of the Medical Departments in each State. The annual State Health Situation Assessment Report was given to the Medical Department Headquarters (one for Straits Settlement States and the other for the Federated Malay States) to prepare an annual report on the health situation of the country as a whole.

2.2 During the period from 1948 to 1962, the gathering of health information rested with the Medical Services under the Ministry of Social Welfare. There was minimal effort at coordination of health information although the quantity and variety of data had increased as a result of more diversified health activities. In 1963, with the formation of Malaysia, the Ministry of Health and Social Welfare becomes the Ministry of Health. The responsibility for health information was still with the



Medical Services Division which continued to be involved in the collection of workload statistics for report writing purposes.

2.3. By the second half of the 1960's, a Medical Records and Health Statistics Unit were set up. It functioned as a data coordination unit for the Ministry of Health (MOH). In 1972, the Operation Research Unit was also set up. The primary function of this unit was to monitor the development of the National Health Management Information System (NHMIS). To further improve the health information system, the World Health Organization (WHO) was requested in 1975 to study and recommend improvements to the existing system. The study recommendation included the establishment of a comprehensive information system through a Health Management Information System Development Project (NHMISDP). The Information and Documentation System (IDS) Unit in the Ministry of Health was established in 1981 to replace the Medical Records and Health Statistics Unit and the Operations Research Unit.

2.4 With the establishment of the HIMS systems medical records officers were trained and posted in all large hospitals to responsible for the compilation of the hospitals statistics. The statistics for public health was continued to be compiled by public health nurses, medical assistants and public health inspectors. The IDS unit at Ministry of Health continued to produce annual reports for use at the national and international level

2.5. The last two decades saw a very rapid growth in the use of ICT in health care. The Tele-health Project launched in 2000 paved the way for the need to provide standards for data collection to provide for interoperability and to produce reports with high degree of data quality and integrity. The subsequent deployment of HIS in public and private hospitals enabled health care providers to produce reports of better quality. However this did not solve the interoperability issue as the disparate systems were using their terminology and data definitions. This again resulted in problems related to data mining and analysis at the national level

2.6 Strengthening of the Health Information Management was one of the supporting goals determined to achieve the objective of the 9MP.to ensure the Integrated Health information Management for the country. The Health Informatics Center established in 2006 at the Ministry of Health will be the custodian of all health and health related information in the country.

### 3. CHALLENGES

#### 3.1 DATA QUALITY, ACCURACY, TIMELINESS

In general, health information required for national and international use is not adequate and often, is not timely for effective and efficient management. The information collected is from MOH facilities only and hence incomplete for analysis for population health status from the whole country perspective.

The current HMIS data collected is based on formats developed for a predominantly annual system of data collection. Over time, additional information requirements to monitor the new activities have resulted in adhoc collection of information. This has resulted in use of data sets which are non-standardized, where the data definition used were different, values varied. Data for research is collected by different research organization for specific purposes

The private sector hospitals are required to submit returns as per Private hospitals Act regulation .however the quality and timeliness of the data is so varied between hospitals thus causing problems for meaningful data analysis. Currently there is very little provision for information collection from private clinics and NGOs.

#### 3.2 ICT IN HEALTH CARE

Healthcare is an information intensive service. It is inevitable that as long as an individual has had an encounter with the health care system, the health information ought to be stored in some kind of "record", probably kept in the provider's memory, paper and computer media for recording health information. Taking cognizance that gathering and reporting healthcare



information have to be shared within a care network concept in an enterprise wide environment, all stakeholders have to take concrete steps to break the deadlock and overcome the grappling issue of data sharing and information flows that we are currently facing. One of the major impacts of ICT in healthcare is in the field of health and disease surveillance. By leveraging the technology, on-line notification of disease and incidents is made possible to support epidemiologists in their capability of pursuing early warning and alert to potential disease outbreak. Integrating IT and advancement in technology into healthcare will change the way we do things. The key issue is in addressing how IT influences the way we work and manage our patients taking into consideration the business process reengineering which is driven by the vision to achieve value-added objectives of our business functions.

### **3.3. EVIDENCE BASED PLANNING / PRACTICE**

The advent of Clinical Practice Guidelines and its use has facilitated the use of evidence based clinical practice in patient care. However the use of information for evidence based planning in measuring health outcomes requires the development of National Data Sets that allows for comparison between different health care providers. The formulation of data sets for specific diseases and conditions through a consensus driven approach is critical for wider acceptance and compliance within the clinical fraternity

### **3.4. MEDICAL INFORMATICS**

The growth of medical/ health informatics has provided a new impetus in the formulation of policies and strategies towards Health Information Management in Malaysia. This field of specialty that is rather new started with the launching of the Tele Health project. Since then there has been a considerable interest in the field both in the public and the private sector. The establishment of the Malaysian Health Informatics Association and the HL 7 Malaysia International affiliate status has provided the platform for experts to share experiences and get involved in some of the works. The government driven projects such as the Tele Health project including the Lifetime Health Record (LHR) provides the ground for all stakeholders to participate. The challenge is the extent of participation from the private sector and industry partners

### **3.5 RESOURCES**

An investment in health informatics infrastructure requires a corresponding investment in human resources to ensure that the infrastructure is appropriately developed, implemented, maintained and evaluated. It is important to have adequate numbers of skilled and competent IT staff. In view of the high cost of skilled IT personnel, the big question seems to be not only whether there will be enough ICT funds to hire them, but rather whether there is enough knowledgeable expertise and skills resource in the country. There is a need for more people, better training and development and career prospects as professionalism is fundamentally important to making ICT more attractive. Thus, the challenge is being able to identify and assign train staff on core IT competencies. Getting people on board in leading the project planning is also vital as well as retaining this “top-notch” people whose expertise is very valuable especially for future undertaking.

It appears that the vast deployment of IT in our day today business may not be still possible in view of lack of infrastructure capability and capacity. Thus investment in ICT infrastructure in terms of adequate storage space, high bandwidth LANS and WANS, wireless technology and planned replacement of workstations, need to be addressed.

## **4. DISCUSSIONS**

### **4.1. POLICY AND GOVERNANCE**

The thrust for the government to strengthen the Health information Management system in the country in order to achieve the goals set in the plan period provides an impetus in moving forward towards an Integrated Health Information Management System. In order to ensure greater compliance by all stake holders the National Health Policy on Health Information Management is absolutely essential. The issues that need to be addressed includes

- Confidentiality & Privacy
- Security & data protection
- Consent
- User access
- Role of stakeholders
- Infrastructure support
- Health informatics standards
- Capacity and Capability building

Such issues can only be addressed through a National Coordinating Body responsible for the collection and storage of all health and health related information. The membership should be drawn from all stakeholders to allow for ownership and accountability amongst all interested parties. Such a mechanism will also provide a forum to accommodate the specific needs of all stakeholders. The information needed for research, policy planning and outcomes measurement should be built integral to all operating systems developed /deployed in the country. A national approach for the coordination of health information in the country will allow for maximizing the potential of various initiatives implemented in the country and therefore will allow to maximize the ICT investments for purposes of making available accurate quality information meant for all types of user.

## 4.2 LEGISLATIONS & REGULATIONS

Currently there is no single legislation that mandates the provision of health and health related information for compilation and analysis of medical statistics in the country. There are several Acts and regulation such as the Tele Medicine Act, Private Health Care Facilities Act, CDC Act, Medical Act, etc which regulates the information required for the enforcement of the Act. Such Acts are very much service driven and serves the specific purpose. As such much of the information collected currently is very much dependent on the information that is gazette for collection. Any additional information required cannot be collected due to lack of provision in the Act. There is no provision for incentives for care providers to deploy IT to facilitate information provision. In addition the issue of consent and confidentiality in relation to information sharing within an enterprise has to be adequately addressed through the Data Protection Act. Experiences from other countries show that legislative support facilitative the

Implementation of ICT initiatives for e.g. which passed several legislations prior to the implementation of the EHR project. As much of the information is collected from patient encounters by various stakeholders the need for a national minimum data set to made mandatory will assist in ensuring appropriate information provided for continuity in care and measuring outcomes analysis. The advent of HIS hospitals has created a new generation of electronic medical records for easy collection, maintenance and storage. the medico-legal implications and legal requirements on the use of electronic records as against paper medical records has to be adequately addressed .this becomes increasingly important inn care network sharing information between private and public sector .therefore privacy protection arrangements have to be made to provide assurance on the consistency in which records are handled at the public and private sector

## 4.3 HEALTH INFORMATICS STANDARDS

Standards for health sector data and communication technology is critical for the implementation of interoperable services. A substantive amount of work has been done in coming up with health informatics standards required for integrating health information over the past 5 years. However, there is much more to be done in view of the sheer volume of the standards. Immense effort is made to set up different workgroups to look into the various field of health informatics standards development, in which the National Health Data Dictionary is one of them. What is more challenging is the maintenance and updating of these standards, and necessitating constant review and changes to those standards which have been developed so that they are current. The setting up of the National Health Informatics Committee should be able to facilitate the enhancement of this development of standards, protocols and other governance issues related to health informatics standards and its utilization. Further more, affiliation and collaboration with the International Standards Bodies and IT industries have to be promoted so as to enhance further the widespread adoption of the data guidelines and rules for global information sharing.



The approach that Malaysia took in realizing the value of health informatics standards as the vital element of integration is through the establishment of a working committee on the Health Informatics Standards. The preliminary effort undertaken by this committee is to establish Special Interest Groups (SIGs) on specific areas to develop the Business Functional Model. A National Health Data Dictionary (NHDD) has been developed to provide standard definition of terms for health care industry. The Malaysian Drug Codes (MDC) has been formulated to make available standardized codes for medication orders for use in the health care facilities. The clinical procedures have been listed and coded by the stake holders who are involved in health care business. All standards have been approved through consensus building held amongst relevant stakeholders. Currently work is underway to convert these standards into national standards through Malaysian standards Department that is affiliated to ISO.

#### 4.4. HEALTH INFORMATION MANAGEMENT – Way Forward

A Health information Management blueprint has been developed that outlines the business case and provides an action plan on which all future strategies and action plans will be based. The Telemedicine Blueprint 1996 has provided a conceptual framework on the health care delivery system of the future and the use of ICT in achieving it. The LHR project will provide the overriding concept for info structure framework for implementation in the country to allow for integration and interoperability. The Information Technology Strategic plan (ITSP) will provide for strategies and action plan to support the implementation of the Health Information Management Action Plan.

The National Health Information Action Plan is being developed in consultation with all health information stakeholders in the public and private sector. Thus it will provide the thrust for all health information initiatives in the country. Amongst the key areas that will be included in the action plan are to ;

- Establish a mechanism for National Coordination and Partnership
- Ensure the building blocks such as the Health Informatics standards, security standards, privacy protection and infrastructure readiness
- Implement key national initiatives that addresses the needs of all stake holders and includes

- LHR services
- Electronic Reporting system :HIMS e
- National Health Data Warehouse

- Ensure the legal, data protection and security framework is in place to ensure secured transfer of all health and health related information

## 5. CONCLUSION

In Malaysia's commitment towards a paradigm shift in the health care delivery system and with the increasing use of ICT health information management becomes a crucial and compelling issue. The role of other stakeholders in providing care becomes critical for the development of Integrated Health Information system. A strong political will and governance through the establishment of Health Informatics Center will pave the way for the national approach in the management of health and health related information. The establishment of the National Health Data Warehouse which manages information in compliance to health informatics standards will allow for evidence based health planning in the country. However the human capital required for planning, operations of this national collaborating

## REFERENCES

1. Economic Planning Unit, Prime Minister's Department, Ninth Malaysia plan 2006- 2010, chapter5 "Mainstreaming Information & communications Technology & , Chapter 20 "Achieving Better Health Published March 2006
2. Ministry of Health; Telemedicine Blue Print "Leading Health Care into information Age "published 1997
3. Ministry of Health Malaysia "My health; Health status of Malaysia 2005 published 2006
4. Ministry of Health Malaysia; Ninth Malaysia Plan TWG paper "Transforming Health Services through ICT "unpublished 2005
5. Ministry of Health Malaysia "Health Information Management system Blueprint" Towards Excellence in Health Information Management "unpublished 2005
6. Ministry of Health Malaysia "Director Generals Report 2006 "Health Information Management "–under publication



## Recent Development of Health Informatics in Hong Kong

Dr CP Wong

Chairman, Hong Kong Society of Medical Informatics

Correspondence: cpwong@hksmi.org

### Abstract

*Hong Kong started late in the development of health informatics but has achieved a great deal in the last 15 years under the auspice of the Hospital Authority. At present, all 40 hospitals and 120 clinics within the public sectors are using the same information system CMS which contained 7.2 million citizen records sharable online. Using innovative approaches, careful implementation policies and enduser-oriented design, the health information system in Hong Kong is gaining very good acceptance by frontline doctors and achieving a very high utilization rate. Future development lies in revamping the system to a new platform, more public health information system and the new initiatives of eHealth by the newly formed eHealth Consortium.*

### 1. Introduction

Although Hong Kong started quite late in the initial development of health informatics, we do accomplish a lot since 1991 after the establishment of the Hospital Authority. By now, all 40 public hospitals and 120 clinics are connected to the same system (the Clinical Management System CMS) with a data ware-house consisting of 7.2 million records, which are freely interchangeable within all institutions in the public sector.

The Hong Kong Society of Medical Informatics was founded in 1986 by a group of medical practitioners and has now evolved into a 300-membered Society with different grades of healthcare workers, IT professionals and academics. It has been very active both in the local and international scene. We are one of the founder members of the Asia Pacific Association for Medical Informatics APAMI, and the national member of the International Medical Informatics Association IMIA.

### 2. Electronic Health Records

Since 2004, the Electronic Health Records System ePR in Hong Kong incorporated also image distribution of all digitally-captured radiological images and made them available online within the public sector of healthcare institutions, which serves 90% of the hospital market in Hong Kong. In this manner, a citizen

electronic longitudinal health record system is in shape, whereby the recent 10 to 15 years of health records of individual citizens has been built and made sharable.

A recent pilot program was started in April 2006 to allow access by the private medical practitioners into this ePR system to facilitate the public-private interfacing of healthcare. An evaluation report will be published after a year.

### 3. eSARS

During 2003, Hong Kong was hit hard by the SARS epidemic. Within a very short time, a special module called eSARS was put online within the CMS system to capture realtime situations, reporting and surveillance data and shared with the Department of Health and the Police Department Geographic Information Systems. This is probably the largest scale IT project for public health and was rewarded as the winner of the health section of the Stockholm Challenge 2004.

We continue to utilize IT for more public health surveillance by incorporating the Notifiable Disease and Outbreak Reporting System NDORS and setting up the future Communicable Disease Information System CDIS.

### 4. Joint venture with vendors

As the old CMS of the Hospital Authority IT system became matured, a new strategy is adopted to revamp the whole system when many of the tools and components will become obsolete. A proposed model will lead to a probable joint venture with software and system vendors for the future redevelopment of the CMS system.

A successful conference on the theme of "The Next Generation CMS" was held in December 2004 when 400 participants and more than 20 vendors attended. Useful experiences were shared and new insights were captured.

### 5. eHealth Consortium

The Hong Kong Society of Medical Informatics joined hands with the Internet Professional Association iProA, the Hospital Authority, the Department of Health

and the Office of the Government Chief Information Officer in September 2005 to establish the eHealth Consortium of Hong Kong. This is in line with the appeal from the World Health Organization to initiate more activities in eHealth for all countries.

The eHealth Consortium now has more than 30 supporting organizations and academic institutions. The three main priorities this year will be to set up a task force on data sharing and health information data standards, another task force for education and capacity building in health informatics and to host the inaugural

conference on 15-16 September 2006 (the eHealth Conference).

## **6. Conclusion**

We look forward to more collaborations with APAMI members for a better development of Health Informatics within the Asia Pacific region.

# Hospital-Wide Implementation of a Multi-Platform Pharmacological Decision Support System at the Point-of-Care

Hui Nar Quek, M.A., Yuen Ho Chow, M.B.B.S.

*Department of Medical Informatics  
Singapore General Hospital, Singapore*

## ABSTRACT

*Providing accurate, relevant pharmacological information at the point of care is the basic function of Pharmacological Decision Support Systems (PDSS). Benefits of such systems at the point-of-care include reduction of medication errors and better patient care. In the recent years, point-of-care technology has become more widely available and more practical with the advent of mobile devices. To fully exploit technology to facilitate decision support at the point-of-care, we integrated our 1600-bed hospital formulary with an electronic pharmacopoeia to create a fully customized site-relevant knowledge base. In addition to existing basic pharmacological information, this also included highly institution-specific information such as local drug-pricing, subsidised selling prices, locally available alternatives within the same therapeutic category and local availability. This information was made available via handheld devices, mobile carts and satellite PC-based clients. We selected 30 clinicians and pharmacists to evaluate this in their respective care areas for 4 months and they found the product useful in providing accurate and comprehensive drug information at the point-of-care, thereby reducing ambiguity and reliance on the human memory and potentially reducing preventable adverse drug events.*

*Keywords: PDAs, handheld devices, pharmacopoeia, formulary, adverse drug events, medication error*

## INTRODUCTION

Medical knowledge is growing rapidly, especially in the areas of pharmacology and therapeutics[1]. Clinicians are expected to know the ever-increasing variety of medications and their complex dosing regimens, changing indications, unique contraindications and possible drug-drug or drug-food interactions. A lack of knowledge or oversight in any of these areas can result in serious adverse drug reactions.

Inquiries relating to medications are one of the most frequent questions that clinicians have[2].

Studies have shown that 22% of the questions doctors have when they see patients relate to medications[3]. Drug information is often not available when required at the point-of-care or when questions pertaining to medications arise. This makes point-of-care technology – which can provide prescribing information at the bedside, in the hospital ward, and in the examining room – especially attractive[4]. Bringing accurate and complete drug information to the point-of-care in an easily accessible manner enables the clinician to quickly obtain the information he needs to provide accurate prescribing, thereby optimizing patient safety by reducing preventable medication errors.

## LITERATURE REVIEW

There have been a number of studies pertaining to the use of drug information software on Personal Digital Assistants (PDAs)[5,6,7,8,9]. Most of these studies do comparisons of the different software available in the market. However, to date the authors did not find any evaluation study on the design, usability and feasibility of a drug information software coupled with the hospital's formulary in the hospital.

## MATERIALS & METHODS

### Background

Singapore General Hospital (SGH) is a 1600-bed tertiary referral centre. Prior to the implementation of a PDSS, drug information and the hospital's formulary are not readily available at the point-of-care. Drug reference books and formulary booklets are available in the wards and clinics in very few quantities and they are updated only bi-annually. The formulary classification of a drug affects the costs borne by the patients and the hospital. This includes information on whether patients can enjoy government subsidies on prescribed 'standard' drugs, or if they will have to pay full retail prices on 'non-standard' or 'non-formulary' drugs. Clinicians often do not have this information readily during prescribing.

**Table 1: Top 3 Most Important Types of Drug Information at Point-Of-Care**

| Information Category              | Information Components                                                                  | % Ranked as Top 3 Most Important (n=30) |
|-----------------------------------|-----------------------------------------------------------------------------------------|-----------------------------------------|
| Dosing                            | Dietary considerations, dose adjustments, paediatric/usual dosing, IV solution capacity | 80%                                     |
| Drug effects                      | Adverse/Common/Serious reactions, contraindications, warnings                           | 77%                                     |
| General drug information          | Administrative route, class/therapeutic category, indications/use, pregnancy issues     | 50%                                     |
| Hospital-specific information     | Dosage form, formulary status, cost                                                     | 30%                                     |
| Pharmacodynamics/Pharmacokinetics | Kinetics, mechanics of action, metabolism, monitoring parameters                        | 30%                                     |
| Drug identification               | Brand names, drug interactions, generic equivalents, pharmaceutical company             | 27%                                     |
| User specifications               | Ability to add drugs, notes by users, user-customized lists                             | 10%                                     |
| Information types                 | Clinical guidelines, overdose/toxicology data, references                               | 10%                                     |

### Development of ePharmacopoeia

Our PDSS system, named “ePharmacopoeia” is a multi-platform capable pharmacopoeia that contains generic drug information as well as the hospital’s formulary information. It enables clinicians and pharmacists to access drug information easily at the point-of-care.

A small team consisting of clinicians, pharmacists and IT personnel was formed to look into the design and usability of ePharmacopoeia. The following design considerations were taken into account:

- (1) Ability to show comprehensive drug information at the point-of-care without information overload.
- (2) Ability to search and retrieve by generic and trade names.
- (3) Ability to show hospital-specific information for the drug, e.g. formulary classification, price, dosage forms, other comments.
- (4) Ability to easily find similar drugs that are cheaper or more suitable for the patient.
- (5) Ability to navigate quickly to desired sections within a drug monograph.

To ensure that the drug information provided will be useful to the clinicians and pharmacists, a survey was done amongst 30 doctors, pharmacists and nurse clinicians to determine the 3 most important types of drug information they need at the point-of-care based on their practice. As shown in Table 1, 80% ranked dosing information as the most important, followed by information on drug effects (77%) and general drug information (50%).

Lexi-Drugs was chosen as the drug information source for ePharmacopoeia based on positive reviews

from peer review case studies and articles[5,6,7,8,9]. The international version is used as it includes trade names of drugs used outside the USA. To ensure that the drug information on ePharmacopoeia is adequate and of sufficient depth yet not overloaded with information, we have taken into consideration Table 1 and removed some of the information components under the categories of “User specifications” and “Information types”. The hospital formulary is then reviewed internally and included with Lexi-Drugs to become a one-point drug and formulary information source.

ePharmacopoeia runs on several platforms. The PDA version runs on either a PocketPC device or a PalmOS device. The PC-based client runs on any browser-enabled personal computer (PC) or compatible device, and they are deployed to wireless mobile carts used on the ward rounds and on stationary PCs in the wards. The application can be launched from the institution’s intranet, as well as from a toolbar within the hospital’s electronic medical records (EMR) system. Updates on Lexi-Drugs are available to users on a daily basis by synchronizing the PDA with an online website. The browser-based client is automatically updated via a central server hosted by the vendor. Updates on the hospital formulary are available to users on a regular basis via the same method.

### Users

30 doctors, pharmacists and nurse clinicians took part in this pilot study. They were chosen for this pilot due to the factors of availability and interest. Special consideration was taken to select doctors from different specialties and with varying amount of clinical experience so as to obtain a more representative subject group. Of the clinicians, 25%

are Senior Consultants, 30% are Consultants, and the remaining 45% are Associate Consultants, Registrars and Medical Officers. Their specialties include Emergency Medicine, Endocrinology, Renal, Gastroenterology, Otolaryngology, Orthopaedic Surgery, Cardiothoracic Surgery, Respiratory & Critical Care. Each of them owns a personal handheld device which is either a PocketPC device or a Palm device and all were issued with the corresponding ePharmacopoeia version for evaluation.

### Questionnaire Design and Analysis

A questionnaire was designed to aid in the evaluation of ePharmacopoeia, focusing on its content and usability, layout and design, and its comparison with other pharmacopoeias. A 5-point Likert scale was used for most questions in which 1 meant “disagree strongly” and 5 “agree strongly”.

In the study, all users used ePharmacopoeia for 4 months in their respective areas of care which included the outpatient clinics, satellite pharmacies and inpatient wards. During this period, they were strongly encouraged to use the program as much as possible, e.g. during ward rounds, at the bed-side, during prescribing. At the end of the study period, all users were required to complete the questionnaire.

## RESULTS

### The ePharmacopoeia Program

Some sample screen shots of ePharmacopoeia are shown in Figure 1. The Main Index shows the list of drugs, together with the formulary status of the drug and the price indicator. [NA] indicates that the drug is not in the hospital's formulary. As such, there is no corresponding price and status information, but users are still able to access the drug monograph.

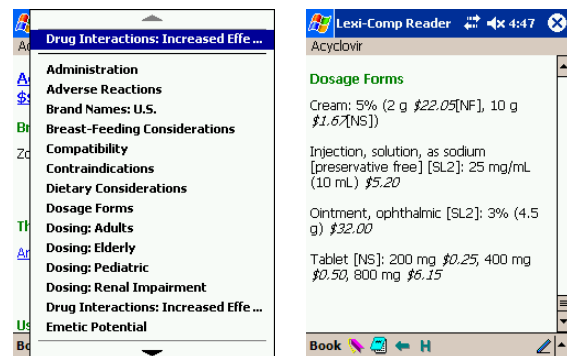
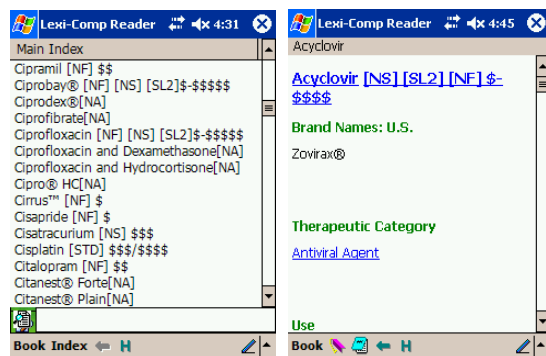


Figure 1: PDA Screen Shots of ePharmacopoeia

Information of each drug from Lexi-Drugs is integrated with the hospital's formulary information. At a quick glance in the index one can see the list of drugs, together with its formulary status and price indicator. The detailed information for each drug can be retrieved by simply tapping on the drug name in the index. For the drug 'Acyclovir', one can see its drug monograph, together with the hospital's pricing information, dosage forms and formulary status. If users are interested in similar drugs that may be cheaper, they can tap on the therapeutic category link with the respective price indications. For example, selecting Therapeutic Category in the 'Candesartan' monograph will bring up Angiotensin II receptor blockers within the SGH formulary and reveal that Losartan is the cheaper alternative to Candesartan. The online interface is similar in terms of features, but allows more detail and is more convenient to read.

### Survey Results

The survey results are shown in Tables 2 and 3. In terms of content and use, more than 90% of the users found ePharmacopoeia easy to use, reliable, useful and beneficial in the practice setting. 80% found it made their practice more efficient. Only 57% of the users said the program easy to install, and 63% found it error-free.

In the program's layout and design, more than 70% of the users found it easy to understand and to find information. However, only 40%-57% found it appealing and easy to read and navigate.

**Table 2: Doctors, Nurse Clinicians and Pharmacists' Perceptions on the Content, Use, Layout & Design**

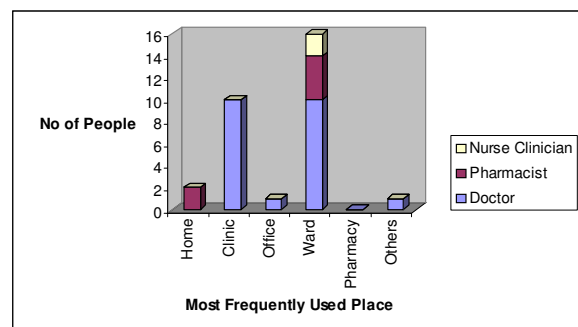
| Questions                       | Users "Agree or Agree Strongly" with the Questions (%) |                       |                  |              |
|---------------------------------|--------------------------------------------------------|-----------------------|------------------|--------------|
|                                 | Doctor (n=22)                                          | Nurse Clinician (n=2) | Pharmacist (n=6) | Total (n=30) |
| <b>Content &amp; Use</b>        |                                                        |                       |                  |              |
| Easy to install                 | 59%                                                    | 100%                  | 33%              | 57%          |
| Easy to learn                   | 95%                                                    | 100%                  | 100%             | 97%          |
| Easy to use                     | 91%                                                    | 100%                  | 100%             | 93%          |
| Useful in my practice setting   | 95%                                                    | 100%                  | 100%             | 97%          |
| Made my practice more efficient | 77%                                                    | 100%                  | 83%              | 80%          |
| Provides benefit to my practice | 91%                                                    | 100%                  | 83%              | 90%          |
| Provides rich content           | 91%                                                    | 100%                  | 100%             | 93%          |
| Reliable                        | 86%                                                    | 100%                  | 100%             | 90%          |
| Error-free                      | 55%                                                    | 100%                  | 83%              | 63%          |
| <b>Layout &amp; Design</b>      |                                                        |                       |                  |              |
| Aesthetically appealing         | 32%                                                    | 50%                   | 67%              | 40%          |
| Easy to read                    | 45%                                                    | 50%                   | 83%              | 53%          |
| Easy to understand              | 77%                                                    | 100%                  | 100%             | 83%          |
| Easy to navigate                | 55%                                                    | 50%                   | 67%              | 57%          |
| Easy to find information        | 64%                                                    | 50%                   | 100%             | 70%          |

**Table 3: Comparison of ePharmacopoeia with Other Pharmacopoeias**

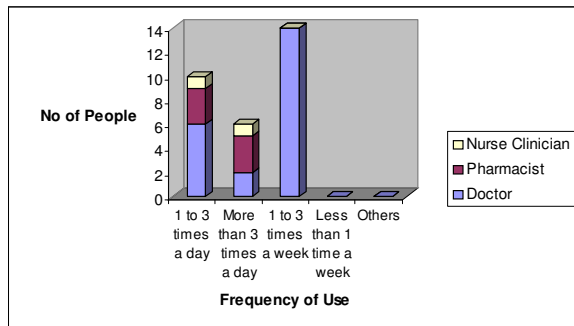
| Other Pharmacopoeia             | Users Find ePharmacopoeia Comparable or Better<br>n = 30 (100%) |                                      |                                           |                                        |                                    |
|---------------------------------|-----------------------------------------------------------------|--------------------------------------|-------------------------------------------|----------------------------------------|------------------------------------|
|                                 | In terms of Ease of Use & Convenience                           | In terms of Relevance of Information | In terms of Information Comprehensiveness | In terms of Information Up-to-Dateness | In terms of Patient Safety Profile |
| Book-Based Reference            | 100%                                                            | 96%                                  | 74%                                       | 96%                                    | 91%                                |
| Computer-based/Online Reference | 100%                                                            | 100%                                 | 77%                                       | 100%                                   | 92%                                |
| Handheld Reference              | 74%                                                             | 95%                                  | 84%                                       | 89%                                    | 95%                                |

ePharmacopoeia compared well with other forms of pharmacopoeia. In terms of ease of use and convenience and information updatedness, more than 95% found it similar or better compared to book and online references and 74%-89% found it similar or better compared to other handheld references. ePharmacopoeia was found to be similar or better compared to all 3 other forms of pharmacopoeias by more than 90% of the users in terms of relevance of information and patient safety profile, and 74%-84% in terms of information comprehensiveness.

The frequencies of usage are shown in Figures 2 and 3. More than 50% of the users use ePharmacopoeia most often in the wards. This includes 10 doctors, 4 pharmacists & 2 nurse clinicians as shown in Figure 2. Figure 3 shows that 100% of the users use ePharmacopoeia at least 1 to 3 times a week, and of which 20% of them use it more than 3 times a day.


**Figure 2: Most Frequency Used Place**





**Figure 3: Frequency of Use**

### General Feedback

Users were asked for comments on areas which satisfied and frustrated them the most, as well as areas where the product exceeded and fell short of their expectations. The comments are summarized in Table 4.

Users also commented that ePharmacopoeia generally saves time during information retrieval, is easily incorporated into their usual workflow, and improves drug-related decision making. Doubts and queries that arise at the point-of-care pertaining to medications can be answered through quick referral to the program, and this greatly increases the confidence level of the clinicians whilst caring for their patients. They agreed that the program helps to reduce the rate of preventable adverse drug events.

### DISCUSSION

The results show a potential in having a PDSS at the point-of-care. As seen from the results and feedback, most users find it useful and beneficial. More importantly, it can potentially help to reduce medication errors leading to improved patient safety.

It should be highlighted that ePharmacopoeia is not intended to replace other forms of pharmacopoeias. Depending on the clinical setting, purpose of use and other factors, different pharmacopoeias may be preferable. Thus in the comparison, the benchmark is to ensure that ePharmacopoeia generally does not and should not perform any worse than other pharmacopoeias.

Further investigation was done to ascertain the accuracy of the information provided in the program as only 63% found it error-free. It was found that the comments field of 26 drugs were not inserted correctly during the development. The comments field forms part of the hospital formulary information and indicates if the drug is a controlled drug or if it is restricted to specific specialties only. This was confirmed to be the only source of error and was promptly rectified. There was no significant impact on the decision support at the point-of-care.

**Table 4: General Feedback from Users**

|                                                                                                                                                                                                                                                                                                                                                             |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p><b>Which parts of the program did you enjoy the most?</b></p> <ul style="list-style-type: none"> <li>• Availability of hospital formulary with prices</li> <li>• Availability of international brands (e.g. Panadol)</li> <li>• Ease of use &amp; easy navigation</li> </ul>                                                                             |
| <p><b>Which parts of the program frustrated you the most?</b></p> <ul style="list-style-type: none"> <li>• Too wordy</li> <li>• Appearance of many unfamiliar brand names at the beginning of each drug monograph</li> <li>• Confusion over icons at the bottom</li> </ul>                                                                                  |
| <p><b>Which parts of the program exceeded your expectations?</b></p> <ul style="list-style-type: none"> <li>• Very comprehensive information</li> <li>• Has latest FDA-issued alerts &amp; cautionary reports</li> <li>• Information available for drugs that are not US-based</li> </ul>                                                                   |
| <p><b>Which parts of the program fell short of your expectations?</b></p> <ul style="list-style-type: none"> <li>• Some local medications not found (e.g. Saridon, Vasteral, Arcalion)</li> <li>• Information limited as compared to textbook references</li> <li>• The mechanism of drug is not included. This is very important to pharmacists</li> </ul> |



## CONCLUSION

This study has demonstrated the feasibility and usefulness of a PDSS integrating drug information and formulary information for clinicians at the point-of-care. Interesting to note is the comparison between ePharmacopoeia and other traditional methods of delivering pharmacopoeia information. A clear majority of users found it comparable to or better than the traditional means. Informal interviews with users related that importance had been placed on being able to find information on all the drugs carried by the hospital in the pharmacopoeia. This is a relatively simple request, but this has traditionally been difficult to achieve.

We are now looking into the extension of ePharmacopoeia to other mobile devices such as the smartphone and blackberry, as well as tighter integration with the hospital's Computerized Physician Order Entry (CPOE) system. For example, the ability to provide context-sensitive information to the requestor from a prescribing field in the CPOE system would be a valuable and convenient tool[10]. As we advance into the era of wireless communication and mobile technology, the clinical and practical value of using these devices in clinical settings will continue to grow further.

## REFERENCES

1. Wyatt J. Use and sources of medical knowledge. *Lancet* 1991;338:1368-73.
2. Ely JW, Osheroff JA, Ebell MH, et al. Analysis of questions asked by family doctors regarding patient care. *BMJ* 1999;319:358-61.
3. Ely JW, Osheroff JA, Gorman PN, et al. A taxonomy of generic clinical questions: classification study. *BMJ* 2000;321:429-32.
4. Rothschild JM, Lee TH, Bae T, Bates DW. Clinician use of a palmtop drug reference guide. *JAMIA* 2002;9:223-9.
5. Knollmann BC, Smyth BJ, Garnett CE, Salesiotis AN, Gvozdan DM, Berry NS, Lee H, Min FD. Personal digital assistant-based drug reference software as tools to improve rational prescribing: Benchmark criteria and performance. *Clin Pharmacol Ther* 2005;78:7-18.
6. Clauson KA, Seamon MJ, Clauson AS, Van TB. Evaluation of drug information databases for personal digital assistants. *Am J Health-Sys Pharm* 2004;61:1015-24.
7. Galt KA, Rule AM. PDA-based drug information sources: Potential to assure medication safety. American College of Clinical Pharmacy 2003 Spring Practice and Research Forum.
8. Enders SJ, Enders JM, Holstad SG. Drug-information software for palm operating system personal digital assistants: breadth, clinical dependability, and ease of use. *Pharmacotherapy* 2002;22:1036-40.
9. Keplar KE, Urbanski CJ. Personal digital assistant applications for the healthcare provider. *Ann Pharmacother* 2003 Feb;37(2):287-96.
10. Kaushal R, Shojania KG, Bates DW. Effects of computerized physician order entry and clinical decision support systems on medication safety: A systematic review. *Arch Intern Med* 2003 Jun 23;163(12):1409-16.

# Genomic Sequence Variation Markup Language (GSVML) for Global Interoperability of Clinical Genomics Data

Jun Nakaya<sup>a</sup>, Keisuke Ido<sup>a</sup>, Kaei Hiroi<sup>a</sup>, Woosung Yang<sup>a</sup>, Michio Kimura<sup>b</sup>

<sup>a</sup> Graduate School of Medicine, Kobe University, Kobe, Japan

<sup>b</sup> School of Medicine, Hamamatsu University, Hamamatsu, Japan  
junnaka@med.kobe-u.ac.jp

## Abstract

**Objective:** Aiming to make good use of internationally accumulated genomic sequence variation data that is a result of the explosive amount of recent genomic researches, the development of the interoperable data exchanging format and its international standardization are the demands. The GSVML is focusing on genomic sequence variation data and human health application domain such as the gene based medicine or the pharmacogenomics.

**Design and Method:** We developed the GSVML through eight steps having the use case analysis and the domain investigations. By focusing on the design scope to human health application and genomic sequence variation, we tried to eliminate the ambiguity and to give the practicability. Basically we intended to follow and to satisfy the requirements derived from the use case analysis of human based clinical genomic applications. By the database investigations, we tried to minimize the redundancy of the data format, while we tried to maximize the data covering range. We tried to have communication ability and interface ability to the other markup languages with intention to exchange various omics data among different kinds of omics researchers or facilities. The interface ability with the developing clinical standards such as the Health Level Seven Genotype information model were analyzed in detail.

**Results:** The Genomic Sequence Variation Markup Language (GSVML) is developed. The GSVML is human health oriented and has three categories as variation data, direct annotation, and indirect annotation. The variation data category is the required category, while the direct annotation category and the indirect annotation category are optional. The annotation categories contain the omics and clinical information and have internal relations. Six use cases and three criteria in human health application were examined. Eleven data elements and three criteria are examined as the data format for genomic sequence variation data exchange. The data format of five international SNP databases and six markup languages were investigated. An interface ability to Health Level Seven Genotype Model was investigated in terms of 317 items.

**Conclusion:** The development of the data exchanging format for the genomic sequence variation data exchanging with focusing on human health application including researches is the demand. The GSVML can be

the pinpointed answer for this demand. The international standardization of the GSVML is also the demand and is in progress. The GSVML has an ability to enhance the genomic sequence variation data utilization internationally by providing a standardized platform for both clinical and research applications.

## 1. Introduction

In current electronic world, we have been dealing with clinical data and image data as major data types for healthcare. Storming into the post genomic era, we are urged to handle the additional huge and internationally overswollen genomic data additionally besides clinical data and image data. The huge number of experimental data from the recent explosive amount of sequence variation researches especially SNP researches produced the overswollen data and many databases with various types of data formats world widely. Nowadays there are so many kinds of omics data having relation with the sequence variation data around the world and they are waiting for their effective utilization especially for human health. To utilize them truly, opening the way to exchange sequence variation data around the world independently from the types of data format is the entry point and is the first hurdle. Providing the globally interoperable data format is in urgent demand for managing, analyzing, and utilizing these genomic data.

This paper intends to show the development process and to provide the outline of the interoperable data exchanging format for genomic sequence variation data as a Markup Language in human health domain without forcing to change any existing database schema with an eye to the international standardization.

As the genomic data, we have genome sequence, genomic sequence variation, and other genome based data as expression data, proteomics data, molecular network, etc. In this paper, we are focusing on the genomic sequence variation. Among the genomic sequence variation, we started with SNP and gave priority to it because of the following three reasons.

1. SNP is the most researched sequence variation for human health.

2. In the current context, SNP data is overswollen around the world with various types of data formats.

3. SNP data already has a great impact for human applications as the Gene Based Medicine and the pharmacogenomics.



Considering that genomic sequence variation especially SNP has its significant meaning in the gene based medicine and the pharmacogenomics for human health domain, the sequence variation data exchanging format is the key to enhance the gene based clinical research and the gene based medicine. In the present circumstances, SNP is expected to be a key to understand human response to the external stimuli such as alien invasions, therapies, and the environmental interactions [1]. In case of alien invasion, the bacterial infection is an example and a response to an infection is different among individuals. In case of therapy, the degree of a side effect of a drug is different among patients. These responses are also different at various environments. Considering that SNP is the major and simple polymorphism in human health genomic research domain, setting a starting point to it and expanding the specification from it to the other sequence variation data will be reasonable. Under such circumstances, the informational approaches to the variation data are getting more important to support genomic research and genomic medicine [2]. Starting with SNP, we widened the covering range of GSVML from SNP to the other sequence variation data. GSVML is focusing on DNA sequence variation data in human health domain. To utilize the accumulated variation data among many facilities around the world, the point of the data exchanging is in the messaging and its interoperable data format. In the current context, Markup Language is the reasonable choice to make an answer for this need.

The Markup Language is a set of symbols and rules for their uses when doing a markup of a document [3]. As for Markup Language development, the first standardized markup language was SGML (standard generalized markup language) [4] which has strong similarities with troff and nroff text layout languages supplied with Unix systems. HTML (Hypertext Markup Language) was based on SGML [5]. XML (Extensible Markup Language) is a pared-down version of SGML, designed especially for Web documents [6]. XML acts as the basis for XHTML (Extensible HTML) [7] and WML (Wireless Markup Language) [8] and for standardized definitions of system interaction such as SOAP (Simple Object Access Protocol) [9]. By contrast, text layout or semantics is often defined in a purely machine-interpretable form, as in most word processor file formats [10]. Markup Language for the biomedical field based on XML is in developing for these decades to enhance the exchange data among researchers. BSML (Bioinformatic Sequence Markup Language) [11], SBML (Systems Biology Markup Language) [12], Cell ML (Cell Markup Language) [13], Neuro-ML (Neuro Markup Language) [14] are the examples. Polymorphism Mining and Annotation Programs (PolyMAPr) [15] is centric on SNP and tries to achieve mining, annotation, and functional analysis of public database as dbSNP [16], CGAP [17], and JSNP [18] through programming.

As a result of recent explosive amount of SNP researches, the huge number of experimental SNP data have been accumulated in many databases with various types of data formats and are waiting for utilization in drug discovery, clinical diagnosis, and clinical researches. We investigated the diversity of databases of dbSNP, JSNP, CGAP, and ALFRED [19]. The international standardization organizations like Health Level Seven (HL7), International Standardization Organization (ISO), DICOM and JPEG have been keeping efforts to standardize the medical data. GSVML project is in the process of the standardization at ISO TC215 [20]. GSVML project also has a relation with HL7 efforts as for genomic data message handling. Health Level Seven Clinical Genomics Special Interest Group (HL7 CG SIG) [21] is summarizing the clinical use cases for general genomic data. We investigated the interface ability of GSVML with HL7 genotype model [22] and the SNP use case of the Japanese millennium project [23] at the HL7 CG SIG

Based on these contexts and investigations, this paper elucidates the requirements, the issues, and the development process of the GSVML with an eye to the international standardization. Then this paper shows the design and the specification of the GSVML.

## 2. Design and Method

### 2.1. Eight Steps Development

We developed the GSVML through following eight steps:

Step 1: We set the required elements and the needs according to the use case analyses. Prior to the development, we elucidated the specification needs for genomic sequence variation data exchanging among the human health facilities. We set the elements and the issues according to the classified use cases. To clear up the elements and the issues that we should solve, we limited the use case of SNP data in the human health related fields. Here the human health related fields indicate the human applications including clinical practice, clinical trial, and translational research. We prepared six use cases for three typical criteria. Four use cases are concerning about the clinical practice, and one use case for each clinical trial and translational research.

Step 2: We designed the initial basic structure and DTD

Step 3: We investigated the applicability of existing Markup Languages to the sequence variation data derived requirements. Investigated markup languages were MAGE-ML, BSML, SBML, Cell ML, and RNAML [24] to the use case needs. We also investigated the PolyMAPr program that is not a markup language but a SNP centric analyzer. Comparing with existing Markup Languages, we set the position of GSVML in the domain of the Markup Language.



Step 4: We refined the initial structure and DTD.. And we designed the XML Schema as more validated information model.

Step 5: We investigated the existing SNP databases. By investigating the existing SNP databases data formats, we confirmed the specification needs for data exchanging format of SNP data. We extracted the sharable format based on the data format diversity investigation. By extracting the common format and their characteristics, we try to trim the fat of data format off. Investigated databases were JSNP, dbSNP, HGVBBase, ALFRED, and Human SNP Database.

Step 6: We checked the interface ability to the Health Level Seven Genotype Model. As Health Level Seven is one of the good choices of the clinical standardizations, we checked an interface ability of GSVML to HL7 genotype informational model. To use GSVML actually both in clinical application and in clinical research application, having an interface ability to HL7 will be the good user-friendly point.

Step 7: We redefined the needs to GSVML and its demanded elements. Based on the above investigations, we set the position of GSVML in the field of the Markup Language and also we cleared up its data specification needs.

Step 8: We refined the initial structure, DTD, and XML Schema.

We have done the design work in collaboration with HL7 CG SIG. Aspiring to become the ISO standard, "to and fro" processes between design work, measurement work, and the standardization process were needed.

## 2.2. Design Principles

By focusing on the design scope to human health application and genomic sequence variation, we tried to eliminate the ambiguity and to give the practicability. Basically we intended to follow and to satisfy the requirements derived from the use case analysis for clinical genomic applications. Based on this standing point, we designed the initial version of GSVML that was called as PML (Polymorphism Markup Language).

To satisfy the data exchange needs among the variation databases, we tried to minimize the redundancy of the data format while we tried to maximize the data covering range for the databases. This compelled to optimize the size and the specification of the sharable data format of GSVML.

To have communication with other markup languages in biomedical field with intention to exchange various omics data among different kinds of omics researchers or facilities, we tried to have communication ability and interface ability to the other markup languages. These efforts modified the GSVML data structure and improved the interface ability.

Considering that the use cases of the GSVML are mainly in clinical fields, the interface ability with the existing clinical standards such as the Health Level Seven is important. We tried to give GSVML an

interface ability to the Health Level Seven Clinical Genomics SIG Genotype information model.

## 3. Measurement

### 3.1. Use Case Analysis

We analyzed the use cases and the required elements, and the indicated elements derived from these use cases and requirements for the GSVML. Here some use cases are supposed. In case of SNP application, the SNP associated genes are in the SNP annotation. The clinical information and observations are included in the clinical annotation of indirect annotation. All kinds of omics data including proteomics data are included in the omics annotation of indirect annotation. The demands to these elements are different among the use cases. As an example, the Omics annotation as indirect annotation is necessary to the great extent possible for the Gene therapy among MD and other paramedicals.

In translational research, the SNP data is exchanged among hospitals, research institutes, MDs, researchers, and the pharmaceutical company. In this case, individual SNP data is sent/received with individual clinical data and the other additional data that are needed to specify the experiment and the patients. The number of the demanded data is almost several dozen, while the parameters for each individual are many.

In clinical trial, the number of the demanded data depends on the clinical phase. Early phases do not need many individuals but need many parameters, while late phases are in opposite.

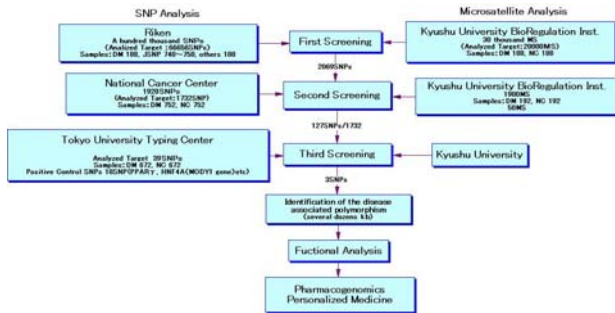
In clinical practice, individual SNP data is sent/received with individual clinical data. For more advanced diagnosis, individual genomic data including omics data are demanded. For the prescription derived from pharmacogenomics, the SNP data will not be exchanged in most cases. The exchanging data will be the prescription, reasons, and its annotations. For gene therapy, individual SNP data is sent/received with individual clinical data and individual genomic data. For disease prevention based on the individual polymorphism, individual SNP data is sent/received with individual clinical data.

In the general Use Case of GSVML in clinical scene, every actor such as MD, Hospital can exchange data smoothly without forcing to change their existing database schema through GSVML. In the same way, the researches can exchange their genomic sequence variation data without any pain. As an example, in case of genetic diagnosis, the individual SNP data is exchanged among the facilities as hospitals, Medical Laboratories. This data is also exchanged among the persons such as MDs, Laboratory Analysts, Counselors, and in some case patient him/herself. Here individual SNP data is encapsulated with the individual clinical data and his/her omics data in some cases for further examination. To analyze this individual SNP data, the individual SNP data need to be compared with the





Figure 1 is an example work flow of the SNP analysis [25]. This is the case of the "Japanese National Millennium Project". This project tries to find the dominant SNPs or genes for the 5 Life Style Diseases. This is the case for the Diabetes Mellitus. In this case SNP data and additional information are exchanged among the facilities. Here we have not only the SNP data but also the clinical data and SNP annotations as the clinical data, the Omics annotations, and the environmental data.



### 3.2. Diversity of SNP Databases

**Table 1. Diversity analysis of international databases.** (quoted from [25])

| Column: terms of comparison | JSNP                                                                                                | dbSNP                                                                                                                                                                                                                                                  | HQVBase                                                             | ALFRED                                                                                      | Human SNP Database                                                                              |
|-----------------------------|-----------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------|---------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------|
| URL                         | <a href="http://www.ncbi.nlm.nih.gov/snp/index.html">http://www.ncbi.nlm.nih.gov/snp/index.html</a> | <a href="http://www.ncbi.nlm.nih.gov/projects/SNP/">http://www.ncbi.nlm.nih.gov/projects/SNP/</a>                                                                                                                                                      | <a href="http://hqvbase.csh.bu.edu/">http://hqvbase.csh.bu.edu/</a> | <a href="http://alle.ed.ac.uk/alfred/index.html">http://alle.ed.ac.uk/alfred/index.html</a> | <a href="http://www.human.snl.se/human/index.html">http://www.human.snl.se/human/index.html</a> |
| Molecular Type              | NA                                                                                                  | genomic, cDNA                                                                                                                                                                                                                                          | DNA, RNA                                                            | NA                                                                                          | cDNA (Affymetric)                                                                               |
| Variation Type              | SNP<br>Deletion/Insertion<br>Microsatellite                                                         | SNP<br>Deletion/Insertion<br>Heterozygous sequence<br>Microsatellite<br>short tandem repeat<br>Named variant<br>No-variation<br>Mixed<br>Multi-Nucleotide<br>Polymorphism                                                                              | SNP<br>Deletion/Insertion<br>Short tandem repeat<br>Generic         | Allele<br>Frequency                                                                         | SNP                                                                                             |
| Population                  | Japanese only                                                                                       | approximately 700                                                                                                                                                                                                                                      | Plural                                                              | Plural                                                                                      | Plural                                                                                          |
| Organism                    | Human                                                                                               | Homo sapiens<br><i>Arabidopsis thaliana</i><br><i>Caenorhabditis elegans</i><br><i>Drosophila melanogaster</i><br><i>Escherichia coli</i><br><i>Gallus gallus</i><br><i>Macaca mulatta</i><br><i>Plasmodium falciparum</i><br><i>Rattus norvegicus</i> | Homo sapiens                                                        | Human                                                                                       | Human                                                                                           |

|                             | JSNP                                           | dbSNP                                                                     | HGVBase                                               | ALFRED                 | Human SNP Database                   |
|-----------------------------|------------------------------------------------|---------------------------------------------------------------------------|-------------------------------------------------------|------------------------|--------------------------------------|
| <b>5' Flanking Sequence</b> | <5_flank_seq><br>CAGGAAAC ...<br><5_flank_seq> | <NSE ss_flank5><br><NSE ss_flank5_E><br>CAGGAAAC ...<br><NSE ss_flank5_E> | <UpstreamSeq><br>CAGGAAAC ...<br><DownstreamSeq>      | 5'-ta-...              | NA<br>(Primer)                       |
| <b>3' Flanking Sequence</b> | <3_flank_seq><br>CAGGCAAC ...<br><3_flank_seq> | <NSE ss_flank3><br><NSE ss_flank3_E><br>CAGGCAAC ...<br><NSE ss_flank3_E> | <UpstreamSeq><br>CAGGCAAC ...<br><DownstreamSeq>      | ...at-3'               | NA<br>(Primer)                       |
| <b>Allele</b>               | <na_var><br>OT                                 | <NSE ss_observed>                                                         | <Allele>C</Allele><br><Allele>T</Allele>              | ... C ...<br>... T ... | C/T                                  |
| <b>Repetition</b>           | CACACA<br>CACACACA<br>CACACACACA               | <NSE ss_observed><br>Observed (CA) <sup>18</sup> /45                      | Allele CACACA<br>Allele CACACACA<br>Allele CACACACACA | 0/3                    | CACACA<br>CACACACACA<br>CACACACACACA |
| <b>Deletion</b>             | A/-                                            | Observed A/-                                                              | Allele A<br>Allele -                                  | A/N                    | A/N                                  |

**Table 3. Diversity of sequence variation data representation in variation databases.** (quoted from [25])

[illegible]

○...Terminology is exact    △...Terminology is not exact  
 ⊗...The database had existed before. But, it is not available now

### 3.3. Markup Language representation comparisons

4





| Sequence       | MAGE-ML                                                  | RNAML                                                    | BSML                                                     | ProML                                                    | SBML                                                     | CellML                                                   |
|----------------|----------------------------------------------------------|----------------------------------------------------------|----------------------------------------------------------|----------------------------------------------------------|----------------------------------------------------------|----------------------------------------------------------|
| Sequence       | Yes (Sequence, Position, Length, Orientation, etc.)      | Yes (Sequence, Position, Length, Orientation, etc.)      | Yes (Sequence, Position, Length, Orientation, etc.)      | Yes (Sequence, Position, Length, Orientation, etc.)      | Yes (Sequence, Position, Length, Orientation, etc.)      | Yes (Sequence, Position, Length, Orientation, etc.)      |
| Variation Data | Yes (Variation, Position, Length, Orientation, etc.)     | Yes (Variation, Position, Length, Orientation, etc.)     | Yes (Variation, Position, Length, Orientation, etc.)     | Yes (Variation, Position, Length, Orientation, etc.)     | Yes (Variation, Position, Length, Orientation, etc.)     | Yes (Variation, Position, Length, Orientation, etc.)     |
| Clinical Info  | Yes (Clinical Info, Position, Length, Orientation, etc.) | Yes (Clinical Info, Position, Length, Orientation, etc.) | Yes (Clinical Info, Position, Length, Orientation, etc.) | Yes (Clinical Info, Position, Length, Orientation, etc.) | Yes (Clinical Info, Position, Length, Orientation, etc.) | Yes (Clinical Info, Position, Length, Orientation, etc.) |
| Transcriptome  | Yes (Transcriptome, Position, Length, Orientation, etc.) | Yes (Transcriptome, Position, Length, Orientation, etc.) | Yes (Transcriptome, Position, Length, Orientation, etc.) | Yes (Transcriptome, Position, Length, Orientation, etc.) | Yes (Transcriptome, Position, Length, Orientation, etc.) | Yes (Transcriptome, Position, Length, Orientation, etc.) |
| Proteome       | Yes (Proteome, Position, Length, Orientation, etc.)      | Yes (Proteome, Position, Length, Orientation, etc.)      | Yes (Proteome, Position, Length, Orientation, etc.)      | Yes (Proteome, Position, Length, Orientation, etc.)      | Yes (Proteome, Position, Length, Orientation, etc.)      | Yes (Proteome, Position, Length, Orientation, etc.)      |
| Metabolome     | Yes (Metabolome, Position, Length, Orientation, etc.)    | Yes (Metabolome, Position, Length, Orientation, etc.)    | Yes (Metabolome, Position, Length, Orientation, etc.)    | Yes (Metabolome, Position, Length, Orientation, etc.)    | Yes (Metabolome, Position, Length, Orientation, etc.)    | Yes (Metabolome, Position, Length, Orientation, etc.)    |
| Signalome      | Yes (Signalome, Position, Length, Orientation, etc.)     | Yes (Signalome, Position, Length, Orientation, etc.)     | Yes (Signalome, Position, Length, Orientation, etc.)     | Yes (Signalome, Position, Length, Orientation, etc.)     | Yes (Signalome, Position, Length, Orientation, etc.)     | Yes (Signalome, Position, Length, Orientation, etc.)     |
| Other Omics    | Yes (Other Omics, Position, Length, Orientation, etc.)   | Yes (Other Omics, Position, Length, Orientation, etc.)   | Yes (Other Omics, Position, Length, Orientation, etc.)   | Yes (Other Omics, Position, Length, Orientation, etc.)   | Yes (Other Omics, Position, Length, Orientation, etc.)   | Yes (Other Omics, Position, Length, Orientation, etc.)   |

Table 4 shows the results of comparison among Markup Languages. The column lists the Markup Languages and the row shows the comparison terms. This time we investigated the markup languages as MAGE-ML, RNAML, BSML, ProML, SBML, and CellML. The compared terms are Sequence, Variation Data, Clinical Info, Transcriptome, Proteome, metabolome, signalome, and other Omics data. The results shown in Table 4 can be summarized as follows:

- All Markup Languages can describe the DNA sequence data, but the representations are different.
- Some can describe the variation data, but the details of it were difficult at every markup languages. Some does not have the descriptive ability of it.
- The proteomics information can be described by almost all Markup Languages, but the terms are different.
- The definitions of the basic vocabulary are different among Markup Languages. As an example, the word "Species" means chemical classification at SBML, while it means biological classification at the other Markup Languages.
- No Markup Languages have descript ability or expandability of clinical annotative data at this time.
- No Markup Languages have ability of interface to HL7 Genotype Model at this time.

### 3.4. GSVML requirement

The pit fall of genomic data handling is in the lack of the genomic sequence variation centric data exchanging format. GSVML is centric on genomic sequence variation, human, and clinical use. All of its needs and specifications are derived from these directions. Fundamentally GSVML should have the sharable representations for genomic sequence variation data such as allele, type, position, length, and region. These representations also should have expandability to the possible other sequence variation data. The annotations of variations such as variation associated genes, individual sequence, experimental assay are essential to understand the basis and the situation of the genomic sequence variation. To understand the clinical significance or to use in clinical situations, the peripheral

annotations of variations such as clinical observation, phenotype are necessary to determinate the meanings of variations.

### 3.5. Interface Analysis to Health Level Seven

The interface of GSVML to HL7 was examined through comparing the terms of the HL7v3 Ballot 10 Genotype model and the GSVML.

The difference is typically reflected in the difference of the entry point, structure and content.

The entry point to GSVML is the variation locus. In contrast, the entry point to HL7v3 Genotype Model is the genotype (genetic locus).

In GSVML, variation data and its annotations are hierarchically categorized in three criteria as variation data, direct annotation, and indirect annotation. The variation data is associated with genotype, alleles, and sequences in the variation data criterion. The annotations as omics and clinical concerns are described hierarchically in the direct annotation criterion or the indirect annotation criterion. In HL7 Genotype model, the main elements are genotype, allele, variation, expression, sequence, and phenotype. Genotype is associated with a pair of alleles.

GSVML has the variation centric information model, while HL7v3 has the genotype centric information model. Both HL7v3 and GSVML have the genetic information and their supportive annotations, while the scope is different. GSVML describes the annotative information with hierarchical classification in one information model, while HL7v3 describes it with the multiple information models that are described with relatively planar classifications.

## 4. Result

Genomic Sequence Variation Markup Language (GSVML) version 1 was created.

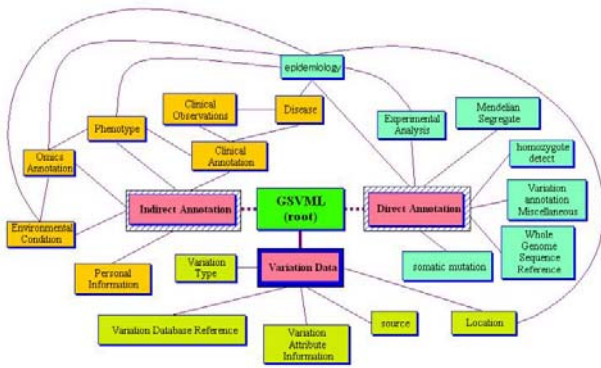
The main role of GSVML in markup language context is to offer the sharable data format for exchanging genomic sequence variation data among the facilities that have the various types of data formats.

The envisioned application of GSVML is in human health domain, this indicates that GSVML standardization need harmonization with the clinical information models and other omics information models.

### 4.1. GSVML structure

The outlined structure of GSVML is shown in Figure 2. GSVML is consist of three data criteria as variation data, direct annotation, and indirect annotation.

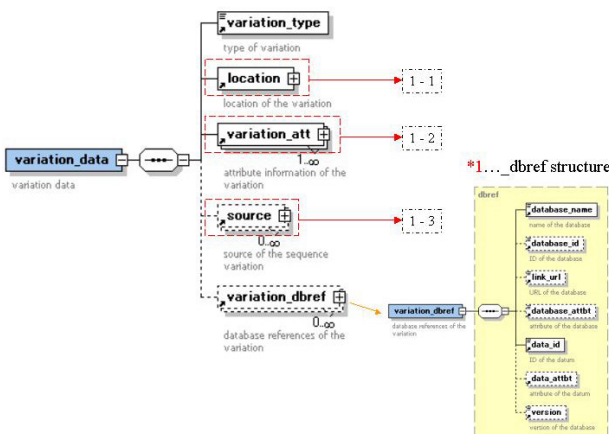




**Figure 2. Outlined structure of GSVML.** (quoted from [25])

The variation data criterion describes the straightforward variation data as allele, type, position, length, region, etc (Figure 3).

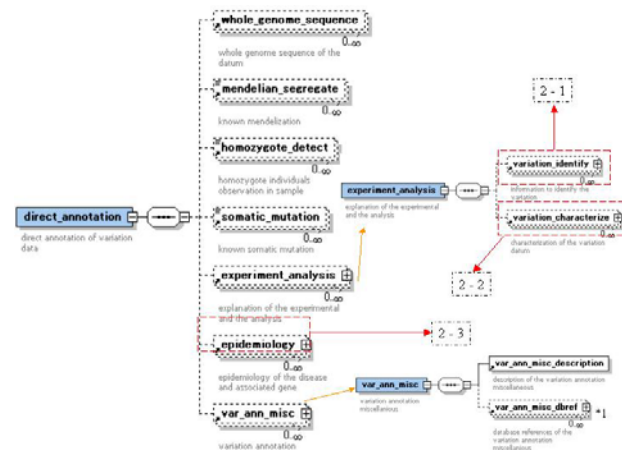
Some elements in the variation data criterion as variation type, location, variation attribute, and source are required, while the other elements are optional. Here database reference structure is defined as a dbref type that is a user defined complex type. The associated gene element is coupled with location element and the epidemiology element.



**Figure 3. Variation data criterion of GSVML** (quoted from [25])

The direct annotation criterion describes the attached data of variation data as whole genome sequence, mendelian segregate, homozygote detect, somatic mutation, experiment analysis, epidemiology, and miscellaneous (Figure 4). All of the direct annotation elements are optional. The experiment analysis element is consisted of the two categorized elements as variation-identify and variation-characterize. The variation-identify element has child elements to describe the experimental background to identify the variation data. The variation-identify element also has elements for publication and submitter, and recursive reference is allowed between the publication elements and submitter elements. The variation-characterize element has child elements to describe clinical statistics or genetic statistics. All kinds of statistical methodologies are

allowed, while the typical elements as p-value, linkage disequilibrium index, descendent index, and maximum lod score are defined as the isolated elements. The epidemiology category gathered the statistical elements from all of GSVML elements, and it describes the statistical data. This category includes the associated gene from variation data, the disease epidemiology from indirect annotation, population and frequency from direct annotation. Each element of the disease epidemiology element is defined in the indirect annotation criterion.



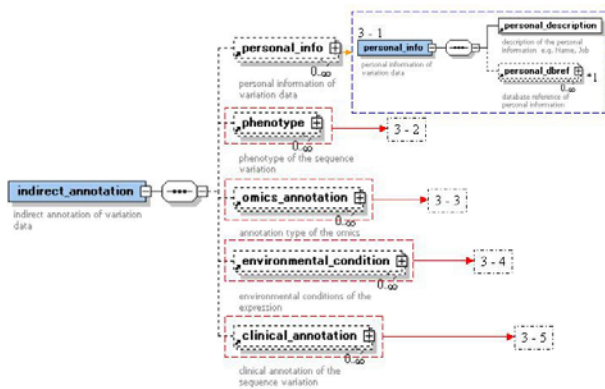
**Figure 4. Direct annotation criterion of GSVML** (quoted from [25])

The indirect annotation criterion describes the explanatory/higher-level information of variation data as the omics data, the clinical information, and the environmental data (Figure 5). The personal information is defined with personal description element and database reference. GSVML supposes all situations of describing the personal information, while at the most cases the personal information is encrypted or numbered. The phenotype element and the omics element allow broad data type to be able to describe it in any types of data formats. The clinical annotation category has disease element, clinical observation element, and database reference. The disease element is consisted of minimum set of disease descriptions, disease epidemiology, and database reference. Some elements in the disease description category are coupled with its expression probability. This probability elements are referred in the epidemiology data section. The database reference allows to describe with the other type of ontological description like SNOMED-CT. The clinical observation element is followed to SOAP description, while the almost all elements are diverted from the disease elements. The family history element is defined under the clinical observation element. The family history element for each family member is coupled with personal information, phenotype, and clinical annotation. To describe the character of each member, the recursive description is allowed.

These data criteria have relations to each other internally.







**Figure 5. Indirect annotation criterion of GSVML** (quoted from [25])

## 5. Discussion

In the current context, annotative information around the genomic sequence variation data is increasing and is getting to embed the information holes. The variation data themselves are also increasing and resulted in various databases. This trend is typical in SNP data. Here the pit fall of genomic sequence variation data handling is in the lack of the sequence variation centric data exchanging formats. Historically many markup languages and programs are developed to handle the genomic information. However, there have been no SNP centric or sequence variation centric markup languages so far. GSVML can be addressed as the first genomic sequence variation centric markup language.

From application side, GSVML is human health centric. Considering that SNP is the highly researched polymorphism and has the great impact especially in human health domain, we can say that GSVML has the greatest potential to be the pinpointed ML for human healthcare. On the other hand, setting the applications to practical human health means to handle the direct or indirect annotative information. Here direct annotation indicates general information such as SNP associated genes, and indirect annotation indicates all of omics information and clinical information. To understand the situation of each patient, we need these kinds of additional information. For this reason, the development of GSVML need harmonization with the clinical standardization organizations such as Health Level Seven, International Organization for Standardization (ISO). The development work of GSVML collaborated with the Health Level Seven Clinical Genomics SIG work. The standardization effort of GSVML in ISO is in process, the improvement raised from the standardization process of ISO activities will be fed back and the design will be changed if need at the next version of GSVML. The "to and fro" process between the design work and the standardization process will continue to reflect the demands in future.

GSVML intends to apply for exchanging messages related to human health. In development and standardization of GSVML in this application domain, we kept an eye on the patient safety, the clinical

efficiency, and the medical costs. For the patient safety, the conservation and the secrecy of patient information are important. The sharable data format and the standardization activity of GSVML can contribute to the data conservation of the domain field internationally, and the public key infrastructure will also need these kinds of the sharable data format. For the clinical efficiency, the simplicity and the easy understandability are important. The structure of the GSVML is hierarchically classified from end-user's view such as clinicians or researchers. The information model of the GSVML adopted the element-based-definition to simplify the usage of the GSVML. For the medical costs, the installation ability is important. Providing the GSVML with DTD and XML schema will be a good offer for installation at current context. The GSVML designed with intention to adopt the end-user understandable classification and the simplified information technology.

GSVML can be used for the clinical variation data exchanging among various facilities having various types of data formats. In the greater framework of clinical data standardization, GSVML will play a part of describing the variation data and its necessary information. At the version 1, we validated the annotative information such as clinical information or omics information with intentional roughness to accept the various representations of the end-user's descriptions. The many efforts to standardize the data format of these annotative information are going on. If these efforts reach a stage where one can rest, the more detailed validation of the annotative information will be our future work.

## 6. Conclusion

GSVML is in the demand for genomic sequence variation data exchanging. The GSVML is the sharable data exchanging format to exchange the genomic sequence variation data and the annotative information among the facilities having various types of data formats. The envisioned applications of GSVML are in human health domain, and the GSVML are demanded to equip a harmonization with clinical information and omics information as annotations of variation data. The GSVML can enhance the genomic sequence variation data utilization internationally by providing a sharable platform for data exchanging.

## 7. Acknowledgements

We thank the Health Level Seven Clinical Genomics Special Interest Group and the International Organization for Standardization TC 215 Working Group who gave us the valuable advises.

## 8. References





- [1]. Holden AL. 2002. The SNP consortium: summary of a private consortium effort to develop an applied map of the human genome. *Biotechniques Suppl*: 22-24, 26.
- [2]. Elias Zerhouni, "Medicine. The NIH Roadmap." *Science*. 2003 Oct 3;302(5642):63-72.
- [3]. Cognitive Science Princeton University, "Overview for Markup Language," internet article of <http://www.cogsci.princeton.edu/cgi-bin/webwn2.0?stage=1&word=Markup+Language>, 1998
- [4]. International Organization for Standardization, ISO 8879: Information processing --Text and office systems -- Standard Generalized Markup Language (SGML), ([Geneva]: ISO, 1986)
- [5]. T. Berners-Lee and Dan Connolly, "HyperText Markup Language Specification -- 2.0", RFC 1866. Proposed Standard, Nov. 1995.
- [6]. W3C recommendation, "Extensible Markup Language (XML) 1.0 (Second Edition)", internet article of <http://www.w3c.org/TR/2000/REC-xml-20001006>, 1998
- [7]. W3C recommendation, "XHTML™ 1.0 The Extensible HyperText Markup Language (Second Edition) A Reformulation of HTML 4 in XML 1.0", internet article of <http://www.w3.org/TR/xhtml1/>, Jan. 2000
- [8]. W3C recommendation, "WAP Forum - W3C Cooperation White Paper ", internet article of <http://www.w3.org/TR/1998/NOTE-WAP-19981030>, 1998
- [9]. W3C recommendation, "Simple Object Access Protocol (SOAP) 1.1", internet article of <http://www.w3.org/TR/2000/NOTE-SOAP-20000508/>, 2000
- [10]. Flying Boat Mobile Communications, "Glossary of Terms relevant to Mobile Communications," internet article of <http://homepages.nildram.co.uk/~jidlaw/pages/glossary.html>, 2004.11.
- [11]. Laurent SS. Biggar RJ. 1999 "Inside SMLDTDs: Scientific and Technical." Berkeley, CA: McGraw-Hill.
- [12]. Hucka M, Finney A, Sauro HM, Bolouri H, et al. "The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models." *Bioinformatics*. 2003 Mar 1;19(4):524-31.
- [13]. Hedley, W.J., Nelson, M.R., Bllivant, D.P. and Nielson, P.F. "A short introduction to CellML." *Phil. Trans. Roy. Soc. London A*, 359, 1073-1089, 2001.
- [14]. Goddard NH, Hucka M, Howell F, Cornelis H, Shankar K, Beeman D. "Towards NeuroML: model description methods for collaborative modelling in neuroscience." *Philos Trans R Soc Lond B Biol Sci*. 2001 Aug 29;356(1412):1209-28.
- [15]. Freimuth RR, Stormo GD, McLeod HL. "PolyMAPr: programs for polymorphism database mining, annotation, and functional analysis." *Hum Mutat*. 2005 Feb;25(2):110-7.
- [16]. Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, Sirotkin K. 2001. dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res* 29: 308-311.
- [17]. Buetow KH, Edmonson MN, Cassidy AB. 1999. Reliable identification of large numbers of candidate SNPs from public EST data. *Nat Genet* 21: 323-325.
- [18]. Ohnishi Y, Tanaka T, Yamada R, Suematsu K, Minami M, Fujii K, Hoki N, Kodama K, Nagata S, Hayashi T, Kinoshita N, Sato H, Kuzuya T, Takeda H, Hori M, Nakamura Y. 2000. Identification of 187 single nucleotide polymorphisms (SNPs) among 41 candidate genes for ischemic heart disease in the Japanese population. *Hum Genet* 106: 288-292.
- [19]. Cheung KH, Miller PL, Kidd JR, Kidd KK, Osier MV, Pakstis AJ. "ALFRED: a Web-accessible allele frequency database". *Pac Symp Biocomput* 2000.:639-50.
- [20]. International Organization for Standardization Technical Committee 215, <http://www.iso.org/iso/en/stdsdevelopment/tc/>
- [21]. Health Level Seven Clinical Genomics Special Interest Group, internet article of <http://www.hl7.org/Special/committees/>, Since Sep. 2002
- [22]. HL7 information model of genotype (HL7 POCG\_DM000023), <http://www.hl7.org/special/Committees/clingenomics/docs.cfm>
- [23]. Yoshida T. "[SNP project in the Millennium Genome Project, Japan]" *Gan To Kagaku Ryoho*. 2002 Jun;29(6):963-7.
- [24]. Waugh A, Gendron P, Altman R, Brown JW, Case D, Gautheret D, et al. "RNAML: a standard syntax for exchanging RNA information." *RNA*. 2002 Jun;8(6):707-17.
- [25] International Organization for Standardization, TC215. Genomic sequence variation markup language. Available from: <http://www.iso.org/iso/en/CatalogueDetailPage.CatalogueDetail?CSNUMBER=43182&scopelist=PROGRAMME> E. <Accessed July 11, 2006>





# Design a pathway/genome expert system using a prolog machine incorporated with a parallel hardware searcher<sup>§</sup>

Wenlung Shu<sup>1</sup> and Junming Lan<sup>2</sup>  
 Bioinformatics Department<sup>1</sup> and CSIE Department<sup>2</sup>  
 Chung Hua University, Hsin Chu, Taiwan  
[wlsu@chu.edu.tw](mailto:wlsu@chu.edu.tw)

## Abstract

It is well recognized that processing complex pathway/genome databases (PGDB) is a very time consuming task. A specific prolog machine which is incorporated with a parallel hardware searcher is proposed to process PGDB in this paper. These databases and the relevant programs are represented as prolog horn clauses (facts and rules) which are treated as objects, and stored in DRAM to increase access time. Search module is a parallel hardware searcher with 9 processor units and 9 corresponding DRAM modules. Index files of object databases can be efficiently processed in search module, and each record of index file contains a search value of an object and corresponding physical address of this object.

Java is used to integrate search module and SWI prolog interpreter. Search module finds all goal related objects by searching index files of object databases continuously. Therefore, only a small portion of necessary objects is transferred to prolog interpreter for further implementation. This specific searcher can readily solve the bottleneck of prolog machine. This expert system itself can have enough intelligence to answer pathway/genome problems, and can be adopted for biological process, drug discovery and medical diagnosis research. This type of prolog machine can also be applied to process huge knowledge bases rapidly for robotics and fifth generation computers.

## 1. Introduction

As genomic information becomes available for a growing number of organisms, it becomes essential to have efficient method to process knowledge resources for every organism. The biocyc collection of pathway/genome databases [1] contains more than 160 databases, including: ecocyc and metacyc. The pathway tools software [2] can be used to query, visualize, and analyze existing pathway/genome databases from the biocyc collection, and to create new databases for an annotated genome.

The main purpose of this paper is to develop an intelligent system for biological process, drug discovery

and medical diagnosis research through pathway/genome databases. A specific prolog machine which is incorporated with a parallel hardware searcher is proposed in this paper. The pathway/genome databases and the programs to process these databases are represented as rules and facts. These horn clauses are treated as objects and stored in DRAM to increase access time. Each record of index files for these objects contains search attribute and physical address. Search module is a parallel hardware searcher with 9 processor units and 9 corresponding DRAM modules. Index files of object databases can be efficiently processed in search module. Java is used to integrate search module and SWI prolog interpreter [3, 4]. Search module finds the goal related objects through index files. Since only a small portion of objects is transferred to prolog interpreter for further procession, this searcher can readily solve the bottleneck problem of prolog machine. This expert system itself can have enough intelligence to answer gene network problems.

In 1980's, Japanese proposed the concept of intelligent computers called fifth generation computers, since they possess the superior robotic technology. Prolog is expected to be utilized as major language in this project. This hardware searcher proposed in the paper can provide fast search among huge data volume. Such search capability is desperately needed in the deduce process of prolog machine. The most significance of this searcher is trying to provide the key technique which can overcome the bottleneck of next generation intelligence computer systems. Therefore, computers can be upgraded to "electronic brains".

Since earlier computer developing stage, many researchers had been devoted to study different search algorithms. These algorithms are quite mutual. The cost of B tree index structures [5, 6, 7] is depending on the height of tree. B+ tree [8, 9] is a dynamic multi-level index file which is designed for very large files. But B+ tree is a one-dimensional access method. Search performance can be improved by using parallel technique and tree maintaining cost is too high.

The distributed and parallel architecture becomes the trend of computer technology when double core processor is popular in the market. Parallel search algorithms, such as R tree [10, 11, 12] and other

<sup>§</sup> The design concept of this searcher had been utilized to apply the patent of Taiwan (application No: 095108044) and the patent of USA (application No: 11/416, 229) with financial support of Chung Hua University.





algorithms [13, 14, 15, 16], had been explored. But they still have overhead on maintaining tree and performance improvement is limited. Because, it is difficult to allow all processors involving on searching distributed data in each search step. The complexity and overall cost for this kind of hardware searcher is just too high to build.

In the proposed parallel hardware search machine, one of  $m$  processors has to enter into a rest state. The performance of this system is almost unaffected when  $m$  is becoming large. However, each processor can only process data in his memory location, and virtual search tree can be adopted in this machine. The cost of maintaining tree can be eliminated completely. The hardware of searcher can be easily designed by using FPGA chips. The performance of searcher can be dramatically increased at low production cost.

The remaining paper is divided into three main sections. Research projects of biocyc are introduced in Section 2. The pathway/genome expert system is designed in Section 3. The design of this parallel hardware searcher is given in Section 4.

## 2. Research projects of biocyc

The biocyc pathway/genome databases contain more than 160 databases, including: ecocyc, metacyc. The ecocyc is a model organism database for *Escherichia coli* K-12. The ecocyc project has integrated information on the *E. coli* genome, and on the *E. coli* metabolic and genetic networks from more than 11,000 publications. The metacyc is an encyclopedia of experimentally elucidated metabolic pathways and enzymes derived from 450 organisms.

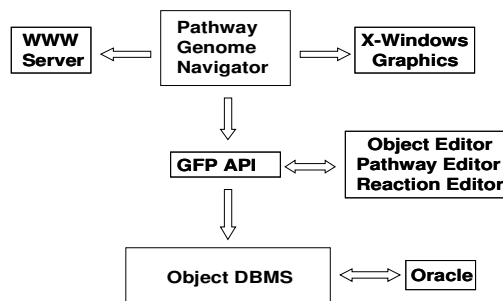


Figure 1. The overall architecture of biocyc.

Ocelot object database management system is utilized to manage pathway/genome objects which are named as frames. Common lisp is used as the native language of pathway tools. Bio-Loader is the interface program between lisp and databases. Many utility functions are developed by using lisp. Pathway tools provides querying capabilities, visualization tools such as for drawing pathways and genome maps, and interactive editing tools to allow users to update data such as modifying a metabolic pathway or defining a new DNA binding site for a transcription factor. The visualization and querying capabilities allow PGDBs to

be published on the Web, or to be accessed as a desktop application on a PC.

## 3. Design a gene network expert system using a specific prolog machine

The overall architecture of the expert system is described in Section 3.1. Knowledge representation for gene network databases and programs is discussed in Section 3.2. In Section 3.3, several commands are implemented in the search module. The advantages of prolog for implementing PGDB are given in Section 3.4.

### 3.1 The overall architecture of the expert system

The overall architecture of our specially designed expert system is exhibited in Figure 2. Java language is used as API to integrate SWI prolog interpreter and search module. The gene network databases are represented as prolog facts. All the programs that process gene network databases are represented as prolog rules and facts. These horn clauses are treated as objects. Since recent cost reduction of DRAM, it is reasonable to store all objects into DRAM in order to increase data access time. The main functions in each block are explained below:

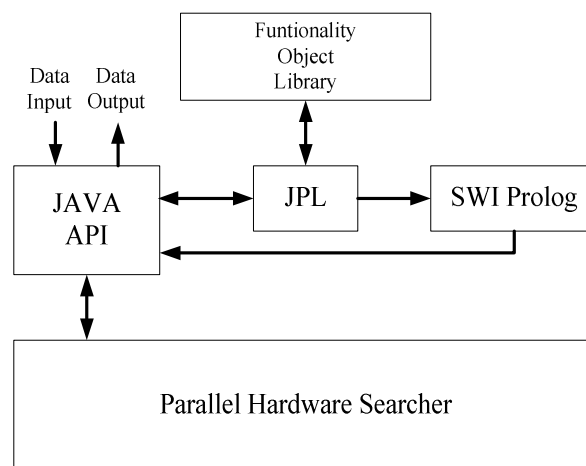


Figure 2. The overall architecture of our prolog machine.

1. **Java API:** Java language can be embedded with assembler language. Hence PC can directly store objects into or access objects from DRAM. Java can also input data files, goals or functionalities to hardware module, and display text output file or graphic data from prolog to GUI. PC can send a number of commands to searcher using interface addresses. These commands are: open pc and searcher connection, close pc and searcher connection, search, insertion, deletion, output result.



2. **SWI prolog:** SWI-Prolog has become a popular free software implementation of the Prolog language. It is formerly known as SWI at the University of Amsterdam. The programmer's environment has contributed most to its popularity. Many expert software developers adopt SWI prolog system and expend effort on large portable prolog applications where scalability, interfaces and networking are often important characteristics.
3. **JPI and functionality object library:** JPL is the interface program between search module and SWI prolog interpreter. All the prolog objects of horn clauses which are related to input goal are rapidly searched and collected from DRAM using hardware searcher. This set of horn clauses can be sent to SWI prolog interpreter to perform, can also be store in functionality object library. Therefore this functionality can be directly accessed from library and performed in SWI interpreter.
4. **Search module:** The architecture of parallel hardware searcher is exhibited in Figure3. There are 9 Processor Units (PUs) and 9 corresponding modules in the interface card of the search module. A number of connection lines are constructed for the communication among PUs. Each PU is corresponding with one giga byte DRAM. Only a small amount of qualified data needs to transfer to SWI prolog for further procession. The performance of the system can be dramatically increased to solve the time consuming problems in implementing huge and complex gene network databases.

### 3.2 Knowledge representation for PGDB and relevant programs

When PC opens connection with searcher, PC can store or retrieve gene network databases which are distributed in 9 DRAM modules. Gene network databases, including gene regulation network, signal transduction network, metabolic network and protein interacting network databases, can be stored in DRAM as prolog fact objects.

This object database system can contains databases from a number of organisms. The object databases of each organism contain several tables, and each table can have several index files which are convenient for data retrieval. The database dictionary contains the information of organism name, search attributes in each table, the size of each table (N), maximum loop number ( $\lceil \log_{m-1} N \rceil$ ) and starting address of index file. It is noted that index file is a binary relation that contains two attributes: search attribute and object identifiers (can be represented as physical address). Multiple value attribute and referential information can be represented as an array of object identifiers.

Many software programs can be written in prolog rules and facts, such as: documentation cluster, predicting pathway, predicting hole filler and operon, applying gene network for different research purposes. This type of objects is combined into a control database file. The index file for these databases which contains prolog function name and physical address can be constructed for fast retrieval.

### DRAM Interface Slot

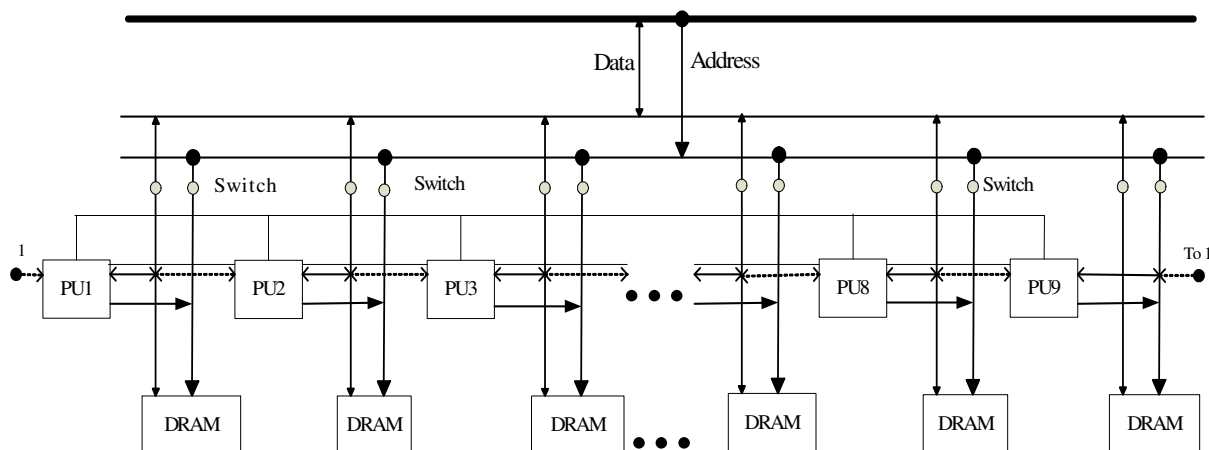


Figure 3. The design of a parallel hardware searcher.

### 3.3 Implement commands in search module

The search command implementation steps are described below:





1. PC opens the connection with searcher.
2. PC sends search command and three consecutive 64-bit data to searcher. Two 64-bit data for search criterion, and another 64-bit represents starting address of index file, maximum loop number.
3. PC closes the connection with searcher, and start to enter search loop.
4. The hardware of each PU performs the algorithm discussed in next section. Search operation can be completed within maximum loop number.
5. Searcher interrupts PC after completing the search operation.
6. PC sends result command, then PU that finds the required search value will transferred object physical address, object size and qualified object number to data bus. If no object is found, PU outputs the address for insertion.

To implement insertion, search the insertion value first. If criterion is found, physical address to insert will be calculated and reported by using count value of same function name. Otherwise, a PU will reports insert address.

### 3.4 The advantages of prolog language

Several advantages of prolog program are:

1. Prolog has strong capability to represent knowledge contained in natural language. Document clustering can be well performed by prolog.
2. All the other language use interface program to access databases. Databases are represented as prolog facts. Hence databases are a part of program and data access and processing is much easier.
3. Instead of writing all the detailed program steps, only logic rules are required in prolog. It will reduce the program developing cost, and has good capability for implementing complex programs.
4. Program size is greatly reduced, comparing with program size of other language.
5. The prolog program is considered as knowledge bases of expert system. The system using prolog has better intelligence than the system developed using other language.

The major disadvantage of prolog program is that this program requires fast searching capability over huge amount of knowledge bases. Parallel hardware searcher can be specially designed to overcome such disadvantage.

### 4. Design a parallel hardware searcher

Search is implemented by using an example in Section 4.1. Insertion and deletion is processed in Section 4.2. Search algorithm is given in Section 4.3.

#### 4.1 Implement search operation

Let  $N$  represent the data size and  $m$  represent total processor number.  $N$  data is ordered according to search key values and distributed to  $m$  processors' memories.  $N$  is 64 and  $m$  is 5 in the example shown in the Figure 4. It is noted that each Location  $k$  (where  $k$  is between 1 and 64) contains a search key value and  $k$  is merely sequence ordered number.  $P_i$  is the  $i$ -th processor where  $i$  is between 1 and  $m$ . Therefore, the relationship among  $k$ ,  $m$  and  $i$  is found:  $k \text{ MOD } m = i$ . In another words, if we want to know the search key value of location  $k$ , then this value can be found in the memory of processor  $P_i$ .

The search key values of data can be sorted and distributed by using a hardware sorter. At initial, host sends  $\text{BlockSize} = 64$ ,  $\text{UpperBound} = 64$ ,  $m = 5$ ,  $\text{height} = \lceil \log_{m-1} N \rceil = 3$  and  $\text{path} = 4$  to all processors. This search example can be completed in 3 levels with 3 comparisons in worst case. If search value is found earlier by a processor, this processor will broadcast stop signal to all processor. Assume the search criterion can be found at  $\text{Location} = 38$ . The proposed algorithm can be described on detail below:

| p1 | p2 | p3 | p4 | p5 |
|----|----|----|----|----|
| 1  | 2  | 3  | 4  | 5  |
| 6  | 7  | 8  | 9  | 10 |
| 11 | 12 | 13 | 14 | 15 |
| 16 | 17 | 18 | 19 | 20 |
| 21 | 22 | 23 | 24 | 25 |
| 26 | 27 | 28 | 29 | 30 |
| 31 | 32 | 33 | 34 | 35 |
| 36 | 37 | 38 | 39 | 40 |
| 41 | 42 | 43 | 44 | 45 |
| 46 | 47 | 48 | 49 | 50 |
| 51 | 52 | 53 | 54 | 55 |
| 56 | 57 | 58 | 59 | 60 |
| 61 | 62 | 63 | 64 |    |

Figure 4. Distributing ordered data into multiple processor units for searching.

1. At level 1, all processor will process  $\text{BlockSize} = \text{BlockSize} / (m-1) = 16$ ,  $\text{UpperBound} = 64$  and  $\text{UpperBound} \text{ MOD } m = 4$ . Therefore, all processor know  $P_4$  represents path = 4 with location = 64. Since level number is odd, processor numbers must be increasing when corresponding path numbers are increasing. Each processor can calculate his representing path and the location to retrieve data. In Figure 5,  $P_1 \sim P_4$  represents path 1~4, and  $P_5$  must take a rest (represents path 0).  $P_1 \sim P_4$  will retrieve data at

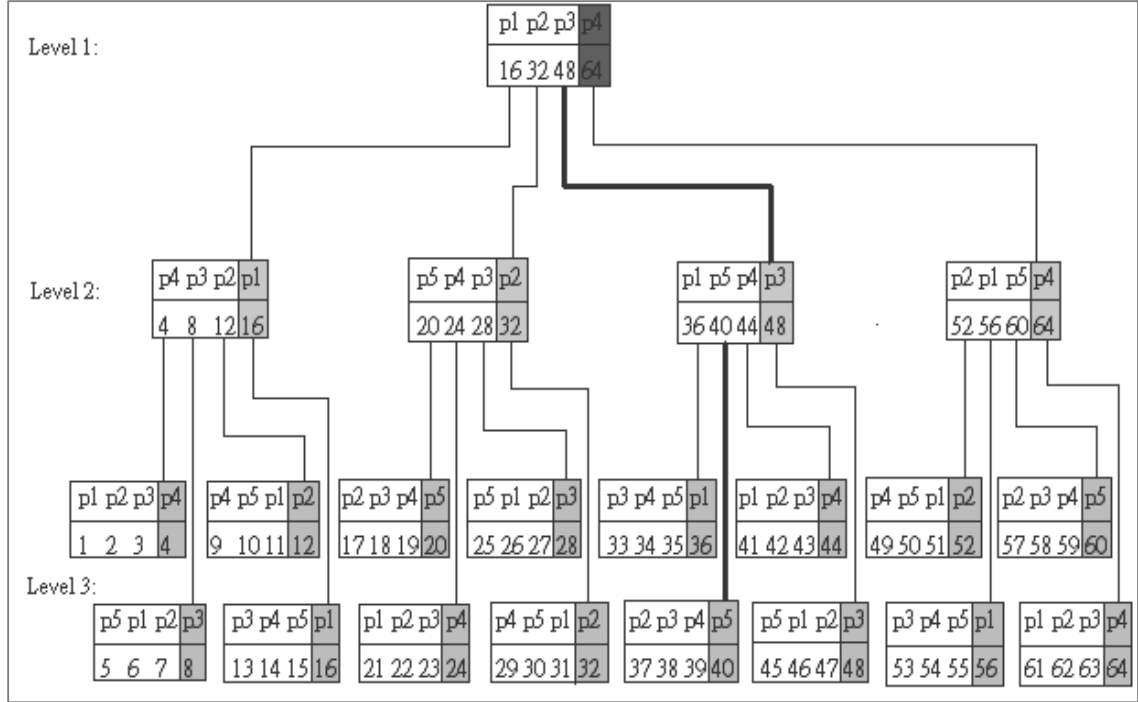




location 16, 32, 48 and 64. Finally  $P_3$  finds that the search criterion is in his range, and broadcast path = 3 to all processors.

2. At level 2, BlockSize = 4, UpperBound = 48 and UpperBound MOD  $m = 3$ .  $P_3$  represents path = 4 with location = 48. Since level number is even,

insert the data into Location=d. The detailed algorithm for all processor  $P_i$  (where  $1 \leq i \leq m$ ) is given in the following section. To implement DELETE operation, search the data for deleting and find the location to delete is Location = d. All processors must rotate data left one location from Location d+1 to Location N.



processor numbers must be decreasing when corresponding path numbers are increasing. Each processor can calculate his representing path and the location to retrieve data. In Figure 5,  $P_1$ ,  $P_5$ ,  $P_4$ ,  $P_3$  represents path 1~4, and  $P_2$  must take a rest.  $P_1$ ,  $P_5$ ,  $P_4$  and  $P_3$  will retrieve data at location 36, 40, 44 and 64.  $P_5$  finds that the search criterion is in his range, and broadcast path = 2 to all processors.

3. At level 3, BlockSize = 1, UpperBound = 40 and UpperBound MOD  $m = 0$ .  $P_5$  represents path = 4 with location = 40. Level number is odd again. In Figure 5,  $P_2$ ,  $P_3$ ,  $P_4$  and  $P_5$  represents path 1~4, and  $P_1$  must take a rest.  $P_2$ ,  $P_3$ ,  $P_4$  and  $P_5$  will retrieve data at location 37, 38, 39, and 40.  $P_3$  finds that the search criterion at location 38, and broadcast a stop signal to all processors.

## 4.2 Implement insertion and deletion

Assume communication links are existent between every two adjacent processors. Hence, data can rotate left or right among  $m$  processors. To implement INSERT operation, search and find the location to insert is Location = d. All processors must rotate data right one location from Location = N to Location = d. Then

Figure 5. The virtual tree of parallel hardware searcher.

## 4.3 The virtual tree parallel search algorithm

```

/* At beginning, host sends N = 64, m = 5, HEIGHT =
 $\lceil \log_{m-1} N \rceil = 3$  to all processors. */
BlockSize = N; /* Set up initial values. */
UpperBound = N;
DEC_m = m-1;
PATH = DEC_m;
LEVEL = 1;
/* Start to perform the operations in each level. */
while ( LEVEL <= HEIGHT )
{
    UpperBound = UpperBound - BlockSize *
        ( DEC_m - PATH );
    j = UpperBound MOD m;
    /* Find the processor  $P_j$  represents PATH = DEC_m. */
    BlockSize = BlockSize / DEC_m;
    /* New block size is used in this level. */
    if ( Level is odd )
    {
        if ( i <= j ) { PATH = i + DEC_m - j; } else { PATH
            = i - j - 1; }
    }
};

```





```

else {
    if ( i >= j ) { PATH = Dec _ m + j - i; } else
        { PATH = j - i - 1; };
};
/* Pi represents path number in PATH. */
if ( Path == 0 ) { Pi takes a rest; }
else { Location = UpperBound - BlockSize *
        ( Dec _ m - path );
};
/* Pi gets data from Location and compare with search
value. Pi sends STOP signal when criterion is found. If
Pi finds that criterion is located in his range, then Pi
broadcast his path to all processors. */
Level = Level+1;
};

```

## 5. Conclusion

A gene network expert system had been designed using a prolog machine that is incorporated with a parallel hardware searcher. Gene network databases were represented as facts, and the programs were presented as both rules and facts. All the facts and rules are treated as objects and stored in DRAM to increase access time. Java has been used to integrate the whole system. Only a small amount of qualified data needs to transfer to PC for further procession. This searcher can readily solve time consuming problem in prolog system. This expert system itself can have enough intelligence to answer gene network problems.

## 6. Reference

- [1] R. Caspi, and etc., "MetaCyc: A multiorganism database of metabolic pathways and enzymes", *Nucleic Acids Res*, Volume 34, pp. 511-516, 2006.
- [2] P. Karp, S. Paley, and P. Romero "The pathway tools software", *Bioinformatics*, Volume 18, S225-S232 2002.
- [3] T. Schrijvers, J. Wielemaker and B. Demoen, "Constraint handling rules for SWI-prolog", *Workshop on (Constraint) Logic Programming*, Ulm, February, 2005.
- [4] J. Wielemaker and A. Anjewierden. "An architecture for making object-oriented systems available from Prolog". In *Alexandre Tessier, editor, Computer Scienc, abstract 2002*. <http://lanl.arxiv.org/abs/cs.SE/0207053>.
- [5] R. BAYER and C. MCCREIGHT, "Organization and maintenance of large ordered indexes", *Acta Inf*. Volume 1, No.9, pp. 173-189, 1972.
- [6] D. Lomet. " The evolution of effective B-tree: page organization and techniques: A personal account", *ACM SIGMOD Record*, Volume 30(3), pp. 64-69, Sep. 2001.
- [7] T. Johnson and D. Sasha, "The performance of current B-tree algorithms", *ACM Transactions on Database Systems (TODS)*, Volume 18 , Issue 1, pp. 51-101, March 1993.
- [8] S. Chen, P.B. Gibbons, T.C. Mowry and G. Valentin, "Fractal prefetching B+-Trees: optimizing both cache and disk performance", *Proceedings of the ACM SIGMOD international conference on Management of data*, 2002 , Madison, Wisconsin.
- [9] S.W. Kim , H.S. Won, "Batch-construction of B+-trees", *Proceedings of the 2001 ACM symposium on Applied computing*, pp. 231-235, March 2001.
- [10] A. Guttman "R-trees: A dynamic index structure for Spatial Searching", *Proceedings of the ACM SIGMOD*, pp. 47-57, June 1984.
- [11] T. Sellis, N. Roussopoulos, and C. Faloutsos. "The R+ Tree: A dynamic index for multi-dimensional Objects", *Proceedings of the 13th VLDB Conference*, 1987.
- [12] B. Wang, H. Horinokuchi, K. Kaneko and A. Makinouchi , "Parallel R-tree search algorithm on DSVM", *Proceedings sixth International Conference on Database Systems for Advanced Applications*, pp. 237-244, April 1999.
- [13] S. Gerard and B. Chris, "Parallel text search methods", *Communications of the ACM*, Volume 31, Issue 2, pp. 202-215, February 1988.
- [14] M. Sergey, R. Sriram, Y. Beverly and G.M. Hector, "Building a distributed full-text index for the web", *ACM Transactions on Information Systems (TOIS)*, Volume 19 , Issue 3, pp.217-241, July 2001.
- [15] P. Sakti and H.K. Myoung, "Parallel processing of large node B-trees", *IEEE Transactions on Computers*, Volume 39, No.9, pp.1208-1212, September 1990.
- [16] M.A. Torres, S. Kuroyanagi and A. Iwata, "A fast parallel search method for large dictionaries", *Proceedings of International Symposium on Software Engineering for Parallel and Distributed Systems*, pp. 207 -214, 1998.





# Reconstructing protein-protein interaction networks from domain-domain interactions

Ng Ka-Lok<sup>1\*</sup>, Liu Hsueh-Chuan<sup>1</sup>, Liu Hsiang-Chuan<sup>1</sup>, Tsai J.F.<sup>2,3</sup>

<sup>1</sup>Department of Biotechnology and Bioinformatics,

<sup>2</sup>Department of Computer Science and Information Engineering,

Asia University, 500 Lioufeng Road, Wufeng Shiang, Taichung, Taiwan 413.

<sup>3</sup>Department of Computer Science, University of Illinois, Chicago, USA

\* Correspondence: klng@asia.edu.tw

## Abstract

*In this study, the domain combination pair approach is employed, introduced by Han et. al [2], to derive putative protein domain-domain interactions (DDIs) from the protein-protein interactions (PPIs) database DIP. The results of putative DDIs are computed for seven species and a combination of all seven species. For each species, a negative learning set (non-interacting protein sequences) is constructed in order to improve the accuracy of DDIs prediction. To evaluate the prediction performance, putative DDIs are compared with that of the database InterDom. Real PPI pathways are selected in order to test whether the predicted DDI results can provide potential PPIs. Furthermore, an entropy-based quantity, so called order index, is introduced to predict PPI interaction order in several pathways, and the results are rather encouraging.*

## 1. Introduction

The interaction between proteins is an important feature of protein functions. Behind PPIs there are protein domains interacting with each others to perform the necessary functions. Therefore, understanding proteins interactions at the domain level gives a global view of PIN [1].

A domain combination pair approach is employed, introduced by Han et. al [2] to rank putative DDIs pairs. This approach will allow us to predict unknown protein-protein interactions from their domains. Furthermore, cross-validation will be performed in order to test the specificity and sensitivity of this method.

In section 2, we give a description of the input PPIs data and the methods we use in this paper. Section 3 is the results for the seven species' putative DDIs, predicted PPI pathways, and the regulation order of several PPI pathways. Section 4 is the conclusions.

## 2. Materials and Methods

We employed the approaches from a recent study [2] to derive putative protein DDIs by using the PPIs database DIP [3], which recorded the PPIs data for seven

species: that is *C. elegans*, *D. melanogaster*, *E. coli*, *H. pylori*, *H. sapiens*, *M. musculus* and *S. cerevisiae*. Protein-domain annotation of DIP can be obtained from the protein domain database, Pfam [4]. Pfam is a large collection of multiple sequence alignments for each domain family and uses hidden Markov models to find domains in new proteins. Domains in Pfam-A are well defined because the corresponding multiple sequence alignments and hidden Markov models have been checked, and most of the domains have been assigned functions.

Assuming a protein contains  $n$  domains, there are  $2^n - 1$  different domain combinations according to the domain combination pair approach [2]. Then given an interacting protein pair (A,B) with  $m$  and  $n$  domains respectively, one considers that there are  $(2^m - 1) * (2^n - 1)$  possible DDIs. If a DDI is found more frequently than expected by chance, it is likely that this DDI is a true interacting domain pair. Since a protein can either has a single domain or multiple domains, combination of possible domain pairs can be derived from each of the interacting protein pair.

The results of putative DDIs are computed for seven species and a combination of all seven species. For each species, a negative learning set (non-interacting protein sequences) is constructed in order to improve the accuracy of DDI prediction. The probability that a protein pair (A,B) with  $m$  and  $n$  domains respectively could possibly interacting is estimated by the Primary Interaction Probability (PIP) value, which is determined by Eq. 7 in Reference 2.

In order to test whether the computed PIP results can provide potential PPIs links between the proteins, the complete set of proteins involved in three biological pathways' (the yeast septin, *E.coli* chemotaxis pathway, the blood coagulation pathway) are selected, their pairwise PIP values are computed and ranked, and the prediction accuracy is determined by comparing with the corresponding experimental determined network.

It is noticed that the PIP value can only give an estimation of the probability that two proteins could interact or not, it does not determine the order of the interaction. In this study, an entropy-based quantity, so called order index is introduced to predict the regulation order in several PPIs pathways, and the prediction accuracy are rather encouraging.





Given a pair of protein  $A$  and  $B$ , the order index  $OI(B|A)$  is defined by

$$OI(B|A) = \frac{MI(A,B)}{H(B)} \quad (1)$$

where  $MI(A,B)$  is the mutual Shannon entropy between proteins  $A$  and  $B$ , and  $H(B)$  is the Shannon entropy for protein  $B$ .  $MI(A,B)$  is defined by

$$MI(A,B) = H(A) + H(B) - H(A,B) \quad (2)$$

where  $H(A,B)$  is the joint entropy of proteins  $A$  and  $B$ .  $H(A,B)$  is defined by

$$H(A,B) = - \sum_{d_A \in A} \sum_{d_B \in B} p(d_A, d_B) \log_2 p(d_A, d_B) \quad (3)$$

where  $d_A$  and  $d_B$  stand for the domains belong to protein  $A$  and  $B$  respectively, and  $p(d_A, d_B)$  stands for the probability of the DDI of the domain combination pair  $d_A, d_B$ .

Given a protein pair  $A$  and  $B$ , the values of  $OI(A|B)$  and  $OI(B|A)$  are not necessary equal, this provide a mean of determine the regulation order. If the later quantity is larger then the former it is proposed that protein  $A$  is the upstream regulating protein in the pathway.

### 3. Results and Discussions

#### 3.1. Putative DDI results and InterDom

To evaluate the prediction performance, the putative DDIs results are compared, based on Jan. 16, 2006 version of DIP, with that of the database InterDom v.1.2 (June 2004). The results are depicted in Table 1.

**Table 1. The total number of PPI interactions ( $N_{PPI}$ ), total number of putative DDI ( $N_{DDI}$ ), the effective number of DDI ( $E_{DDI}$ ), the total number of matched DDI ( $E_{DDI}$ ) and the percentage of specificity ( $S_p$ ) for seven species.**

| Species                | $N_{PPI}$ | $N_{DDI}$ | $E_{DDI}$ | $M_{DDI}$ | $S_p(\%)$ |
|------------------------|-----------|-----------|-----------|-----------|-----------|
| <i>C. elegans</i>      | 4030      | 3874      | 1751      | 987       | 56.4      |
| <i>D. melanogaster</i> | 22819     | 1346      | 695       | 453       | 65.2      |
| <i>E. coli</i>         | 6966      | 4000      | 3314      | 491       | 14.8      |
| <i>H. pylori</i>       | 1420      | 894       | 276       | 245       | 88.8      |
| <i>H. sapiens</i>      | 1397      | 4000      | 1085      | 693       | 63.9      |
| <i>M. musculus</i>     | 209       | 1030      | 206       | 155       | 75.2      |
| <i>S. cerevisiae</i>   | 5952      | 4000      | 2538      | 1856      | 73.1      |

In Table 1,  $N_{PPI}$  stands for the total number of PPIs recorded by DIP. The top 4000 putative DDIs ( $N_{DDI}$ ) are selected and compared with that of the database InterDom. Some of the species'  $N_{DDI}$  value is less than 4000, such as fruit fly. The actual number of predicted DDIs could be less 4000 is mainly because of the absence of Pfam domain annotation of a PPI entry. The effective number of DDIs,  $E_{DDI}$ , denotes single domain interactions and it does not include multiple domains pair interactions, since these types of DDIs are not available in InterDom.

The predicted specificity ratio of DDI ranges from 14.8% to 88.8% for the seven species. The specificity value for *E. coli* seems to too lower in comparison to the other species, this is not a fault of our calculation. In fact,

we have performed an exactly the same calculation using the April 25, 2005 version of DIP, and it was found that the specificity value is 66.3%, this is simply because the number of *E. coli* PPIs entries recorded by DIP increased from 761 to 6966 dated on Jan. 16, 2006. A web based interface was set up, the predicted DDIs results are available at [http://210.70.80.163/kzbio2/r\\_ap.php](http://210.70.80.163/kzbio2/r_ap.php).

#### 3.2. The yeast septin complex

In the following study, several biological pathways are selected in order to test whether the computed DDIs probability can provide potential PPIs links between the components, and compared the results with the experimental determined networks and that of PreSPI. For the first example, the yeast septin complex, which composed of six subunits (CDC3, CDC10, CDC11, CDC12, GIN4, SHS1), was selected. Given the septin complex, our method correctly predicted that all the six proteins interact with each other, with a true positive (TP) value of 100%, i.e. 15/15. The same prediction accuracy is reported by InterDom. In case of taking self-interacting into account, the predicted result deteriorated, with the TP, false positive (FP) and false negative (FN) values equal to 76.2% (16/21), 23.8% (5/21) and 0% respectively. These ratios translated into a sensitivity ratio  $S_n$  of 100%, a true positive specificity ratio  $S_{p1} = TP/(TP+FP) = 76.2\%$  respectively. PreSPI returned an error message (PIP\_Value = error) for each of the 15 possible interactions. In case of adding one more component, that is the Spr28 protein, the predicted result deteriorated, with TP, FP and FN values equal to 64.3% (18/28), 35.7% (10/28) and 0%, this implies  $S_n = 100\%$  and  $S_{p1} = 18/28 = 64.3\%$ , respectively. This results outperform the PreSPI prediction and comparable to the InterDom prediction (TP, FP, FN) = ( $S_n = 16/23 = 69.6\%$ ,  $S_p = 16/21 = 76.2\%$ ).

#### 3.3. The *E. coli* chemotaxis pathway

As a second example, the *E. coli* chemotaxis pathway is selected as the study case. The chemotaxis pathway, obtained from KEGG, composed of 11 proteins (MCP (consists of trg, tap, CheD, CheM), Aer, CheA, CheB, CheR, CheW, CheY, CheZ). The predicted results are depicted in Table 2,

**Table 2. The true positive, true negative, false positive and false negative PPIs results for the *E. coli* chemotaxis pathway predicted by our calculation and PreSPI.**

|          | Our calculation | PreSPI |
|----------|-----------------|--------|
| TP       | 13              | 12     |
| TN       | 15              | 7      |
| FP       | 26              | 17     |
| FN       | 1               | 0      |
| total    | 55              | 36     |
| $S_n$    | 92.9%           | 100%   |
| $S_{p1}$ | 33.3%           | 41.4%  |
| $S_{p2}$ | 36.6%           | 29.2%  |





In general, there are  $C_{11}^{11}$ , that is 55 possible interactions among 11 proteins, 19 of those are not predicted by PreSPI, hence, these results are removed from the prediction. The sensitivity, true positive and true negative specificity values of our results are comparable to that of PreSPI. On the other hand, the InterDom database gave null results for this study.

The current prediction method is not limited to one has to know all the proteins Swissprot ID in advance. The current method can be extended to accept multiple domains, with known PfamA ID, as the input. This type of query is available at our web site, [http://210.70.80.163/kzbio2/r\\_ap.php](http://210.70.80.163/kzbio2/r_ap.php).

### 3.4. The blood coagulation pathway

For the third example, we applied the computed DDIs data to the blood coagulation pathway. Blood clotting occurs via three pathways, intrinsic, extrinsic and common pathways, in which a total of 13 proteins are involved. The blood coagulation pathway composed of 13 proteins (F1, FII, FIII, FV, FVII, FVIII, FIX, FX, FXI, FXII, FXIII, PKK, HMWK). Based on the DDIs data, the predicted results are depicted in the following table,

**Table 3. The true positive, true negative, false positive and false negative PPIs results for the coagulation pathway predicted by our calculation and PreSPI.**

|          | Our calculation | PreSPI |
|----------|-----------------|--------|
| TP       | 5               | 18     |
| TN       | 21              | 0      |
| FP       | 16              | 60     |
| FN       | 6               | 0      |
| total    | 48              | 78     |
| $S_n$    | 45%             | 100%   |
| $S_{p1}$ | 24%             | 23%    |
| $S_{p2}$ | 57%             | 0%     |

In general there are 78 possible interactions, but there are only 48 interactions can be determined in our computation. This is because not all the PPIs entries have domain annotations. Even if the domain annotation is available, such domain combination pairs may not available in our computation. In comparing our calculation with the PreSPI prediction, our prediction give a lower sensitivity ratio than PreSPI, both have a comparable true positive specificity, and we have a much better false negative specificity value. From Tables 2 and 3 results, it seems to indicate that PreSPI tend to predict more PPIs since the FN values are zeros in both Tables 2 and 3.

The predicted domain-domain interaction results are available at [http://210.70.80.163/kzbio2/r\\_ap.php](http://210.70.80.163/kzbio2/r_ap.php). Several query interfaces are implemented to facilitate data display. For example, user can input two proteins' Swissprot ID and get the probability of their interaction. In case the actual Swissprot ID is not known, user can

input domains' PfamA ID, the system will return the predicted probability of the protein interaction. To reconstruct the PPI network, user can either input a set of proteins or domains ID, the system returns a putative network of interaction.

To further characterize PPIs, the orders of PPIs in several biological pathways are predicted and the results are given in the following sub-sections.

### 3.5. Order index - *E. coli* chemotaxis pathway

As mentioned above the chemotaxis pathway composed of 11 proteins. According to KEGG annotation, the following PPIs regulation order pairs are recorded: MCP-CheA, Aer-CheA, CheA-CheB, CheA-CheY, and CheZ-CheY. Among the five regulatory relation, one of the relation has the same  $OI(A|B)$  and  $OI(B|A)$  value, therefore, one left four relations. The order index approach correctly predicted two regulatory orders out of four (i.e. CheA-CheB and CheA-CheY).

### 3.6. Order index – MAPK signaling pathway

The yeast starvation and osmolarity pathways are selected for this study. For the starvation pathway, there are six regulatory relations among the following seven proteins Sho1, Ras2, Cdc42, Ste20, Ste11, Ste7 and Kss1. The order index method correctly predicted the following three regulatory relations: Ras2-Cdc42, Ste20-Ste11 and Ste11-Ste7. Same score is obtained for Ste7-Kss1 and Kss1-Ste7. The method does not correctly predict the regulatory order for Sho1-Cdc42 and Cdc42-Ste20, hence, the accuracy of regulatory order prediction is 60% (i.e. cases have same order index values are not counted, therefore, one has 3/5).

For the osmolarity pathway, there are eight regulatory relations among the following seven proteins, Sho1, Sln1, Ste20, Ypd1, Ste11, Ssk1, Ssk2, Pbs2, and Hog1. Two of the regulatory relations, Sln1-Ypd1 and Ypd1-Ssk1, do not have DDI values available, therefore, six relations left. The order index method correctly predicts the Ste20-Ste11 and Ste11-Pbs2 relations. Same score is obtained for the Pbs2-Hog1 relation. The rest two regulatory relations are incorrectly predicted, hence, the method achieves a 50% of prediction accuracy.

### 3.7 Order index – the yeast cell cycle DNA checkpoint pathway

In this study the yeast cell cycle DNA damage checkpoint in the G2 phase is studied. There are two major problems in this study, the first one is that some of the proteins do not have their domain annotation, and the second problem is the type of DDI is not available in our computed data set. These problems could possibly affect the accuracy of our prediction.

Among the 14 regulatory relations only 13 relations are well annotated. It is because the regulatory relation among (Clb1, Cdc28) and Cks1 are not clearly defined by KEGG, so this relation is neglected.





For DNA damage checkpoint pathway, we correctly predicted the following six PPI order pairs: Mec1-Chk1, Mec1-Rad53, Rad53-Cdc5, (SCF, Met30)-Swe1, Swi5-(Clb1, Cdc28) and (HsL1, HsL7)-Swe1. The relative dependence of Cdc28 and Cak1, Gin4 and Swe1, HsL1 and Swe1 are not determined because they have the same  $OI(A|B)$  and  $OI(B|A)$  values. The rest four regulatory relations are incorrect predict, that is Cdc5-(CLB1, Cdc28), Mih1-Cdc28, Rad9-Mec1 and Swe1-(CLB1, Cdc28). Hence, the accuracy of the regulation order prediction is 60% (i.e. cases have same order index values are not counted, therefore, one has 6/10).

#### 4. Conclusion

The domain combination pair approach is employed, introduced by Han et. al [2], to derive putative protein DDIs from the protein-protein interaction database DIP. To evaluate DDIs prediction performance, our prediction results are compared with that of the database InterDom, where it can achieve a specificity value around 70% for the seven species we studied.

Three biological networks are chosen to test the prediction accuracy of our computation. The yeast septin complex, *E. coli* chemotaxis pathway and the blood coagulation pathway are reconstructed with reasonable accuracy (in comparison with InterDom and PreSPI), this indicate the merit of our calculations.

Furthermore, an entropy-based quantity, so called order index is introduced to predict the interaction order in several PPIs pathways, and the prediction accuracy are rather encouraging. This implies that the order index could be a reliable index to determine the regulation order of PPIs.

#### 5. Acknowledgement

Dr. Ka-Lok Ng's work is supported by the National Science Council of R.O.C. under the grant of NSC 95-2745-E-468-008-URD.

#### 6. References

- [1] Minghua Deng, Shipra Mehta, Fengzhu Sun, and Ting Chen, "Inferring domain-domain interactions from protein-protein interactions", *Genome Res.* 2002, 12, pp.1540-1548.
- [2] Dong-Soo Han, Hong-Soog Kim, Woo-Hyuk Jang, Sung-Doke Lee, and Jung-Keun Suh (2004). "PreSPI: a domain combination based prediction system for protein-protein interaction", *Nucleic Acids Res.* 2004, 32, pp.6312-6320.
- [3] I. Xenarios, E. Salwinski L., X.J., Duan, P. Higney, and Kim S.M. "DIP the database of interacting proteins: a research tool for studying cellular networks of protein interactions", *Nucl. Acid Res.*, 2002, 30, pp. 303-305.
- [4] A. Bateman, E. Birney, L. Cerruti, R. Durbin, L. Etwiller, S.R. Eddy, S. Griffiths-Jones, K.L. Howe, M. Marshall, E.L. Sonnhammer, "The Pfam Protein Families Database", *Nucleic Acids Res.*, 2002, 30(1), pp.276-280.
- [5] S.K. Ng, Z. Zhang, S.H. Tan, K. Lin, "InterDom: a database of putative interaction protein domains for validating predicted protein interactions and complexes", *Nucleic Acids Res.* 2003, 31(1), pp. 251-4.
- [6] S.K. Ng, Z. Zhang, S.H. Tan, "Integrative approach for computationally inferring protein domain interactions", *Bioinformatics*, 2003, 19, pp. 923-929.







# Establishment of Disease Pathway Discovery System

Hao-Chang Hsu, Yuan-Chii G. Lee, PhD\*

Graduate Institute of Biomedical Informatics, Taipei Medical University, Taiwan

\*Corresponding Author

m110094001@tmu.edu.tw, ycgl@tmu.edu.tw

## Abstract

We have incorporated genome-wide protein-protein interaction data and transcription regulators data to rebuild a "Disease Pathway Discovery System" (DPDS). This DPDS reconstructs unseen pathways using upstream regulators /downstream regulated genes information from key transcription regulators derived from protein-protein interaction network. We reckoned that if one of the upstream genes goes wrong, many in the cascade of downstream take effect and subsequently amplify the mutant signal by altering the gene expression levels or genetic mutation. These downstream genes are normally easier to be detected as their molecular copy numbers are amplified. Therefore, we aimed to reconstruct the upstream genetic pathway from 2D planar proteins (protein-protein interaction) to 3D transcriptional relationship. We have applied a familial disease Long QT syndrome (LQTS) as case study to test the efficacy of the model. Through this system, we obtained a transcriptional regulatory network from 8 previously reported LQTS related genes. After evaluation with microarray data, this network is confirmed to be heart-exclusive activity. We assume this DSDP is a feasible working model to reconstruct disease pathway.

**Key Words:** Protein-protein interactions, Protein-DNA interactions, Transcription factor, Gene expression, Pathway discovery

## 1. Introduction

If one of the upstream genes goes wrong, many in the cascade of downstream take effect and subsequently amplify the mutant signal by altering the gene expression levels or genetic mutation. Many disease-related genes have been announced to be functionally relevant to a certain disease, but as a matter of fact, these genes are most likely playing a part in a downstream of a disease. The reason for that is the gene expression signals of this kind of genes are many fold or even hundred-fold amplified. It is imperative to find a way that can be able to trace the upstream regulatory genes where the disease goes astray. Here in this study, we adopt a genome-wide functional interaction networks, transcriptional regulation networks, and high throughput gene expression profiling of paired normal and disease tissues to construct a platform from where a claimed

disease-related gene can be applied to expand its contribution to the disease by extracting its relationship with other protein or its regulatory role in transcription, so that a large-scale, coarse-grained view of cellular networks will be emerged.

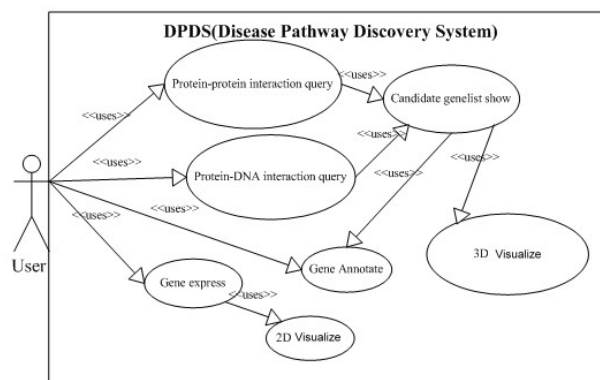


Figure 1 – Use case diagram of DPDS to describe the nature of this system from an external point of view. User may query through gene symbol. DPDS would further allow online visualizing molecular interaction networks.

Protein interactions play a critical role in systems biology of living things. Building up protein-protein interactions not only improves our knowledge towards functional genomics, but also helps predict unknown signaling pathways. High throughput approaches such as yeast two-hybrid system, mass spectrometry, and protein chips are widely applied to extract information of protein interactions in yeast, drosophila, and C. elegance. Nevertheless, the piling up information of protein interactions from PubMed is also invaluable. We will first collect all the available protein interaction databases, such as DIP [1], IntAct [2], MINT [3], BIND [4], HPRD [5], EcoCyc [6] and HIV-1 protein interactions. Among these databases, the last four have been organized, sorted and available in NCBI. Secondly, we will separate the interactions according to different source organisms. Next, we will use Entrez GeneID and Official Gene Symbol (according to the nomenclature of Human Gene Nomenclature Committee, HGNC) as the systematic name for each organism. The reason for this is to facilitate identification of orthologs (among different organisms) and paralogs (in the same organism).

Although protein-protein interactions provide important information towards network connection, yet it lacks time frame to separate the events that might not





occur at the same time. With this regards, we will add time and space scales to the network. Transcription regulation provides such real-time information as one event triggers another in cascade. To reconstitute a complete cellular network it is important to integrate it with transcription regulation data as it is widely accepted that many of the pathways in the cell are regulated both at the transcriptional (protein-DNA) and at the proteomic levels (protein-protein).

To incorporate transcription factors data into the DPDS, we use the data from commercial software, TransPath ([www.biobase.com](http://www.biobase.com)). TransPath® collects information of transcription factors' downstream regulated genes that previously published in the past 30 years. The text formatted data were parsed and then integrated into the DPDS. In addition to the time frame scale, we will add tissue-specific expression information to further depict its actual events to strengthen this network. The microarray data are adopted to fulfill the mission. Su et al., using Affymetrix microarray have made a tissue-specific pattern of transcriptomes to mapping the location for each gene [7]. We took advantage of this gene atlas as space scale. Finally, the interactive relationship among genes will be visualized using Cytoscape [8], which is a powerful bioinformatics software for viewing molecular interaction networks.

Through the combination with JWS (Java Web Start) technology, we incorporate successfully Cytoscape into our system, combining back end database with front end virtualization tool. Users can view their network results in the browser directly without further downloading any other tools. The use case diagram for the system can be seen in the Figure 1 and the flowchart of DPDS is in the Figure 2.

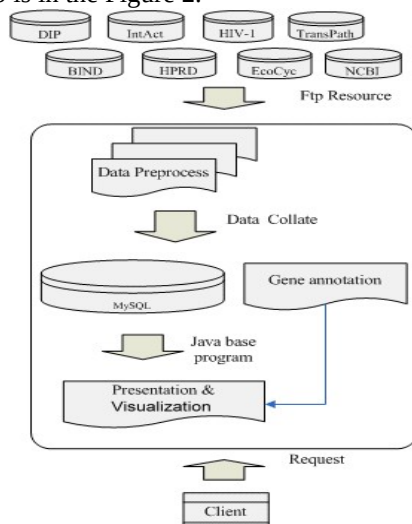


Figure 2 – Flowchart of DPDS (Disease Pathway Discovery System). DPDS is a web base query system to reconstruct disease pathway.

## 2. Results and Discussion

DPDS is aimed to reconstruct the upstream genetic pathway from 2D planar protein network (protein-protein interaction) to 3D tempo-spatial transcriptional relationship. To prove that our system is feasible and valuable, we take the Long QT syndrome (LQTS) as a case study example. LQTS is a familial disease characterized by an abnormally prolonged QT interval and stress-mediated life-threatening ventricular arrhythmias. As for subtype, there are only eight genes (LQT1 to LQT8) that found to be related to LQTS, which mostly are involved in ion channel regulation [9-14]. We started by the scarce information as only eight related genes, then built a vast protein interacting relationship, and eventually a transcriptional map was emerged. Each gene initially exists individually and finds no connection with each other. Surprisingly, by applying this protein-protein interaction database, the interaction connecting genes are emerged as common modulators and connect one and the other to form a networking web (Level one interactions from 8 LQTS genes are shown in Figure 3 and level 2 interactions in Figure 4). LQTS is thought to be a mutation in ion channels, however, through this preliminary data, it provides clue to be associated with neurological synapses, as GRIN1 is shown as one of interacting hub gene. GRIN is a critical subunit of glutamate receptor and a key protein in the synapses, connects LQT1, 2, 5, 6, 7 and 8.

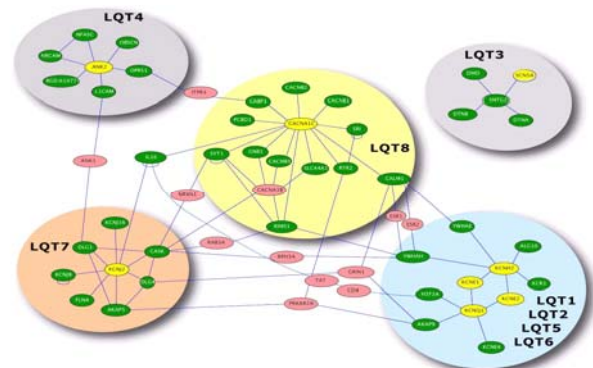


Figure 3 – Protein-protein interaction map facilitates understanding the relationship among 8 LQTS genes. The genes in pink background circle are the original LQTS associated genes, genes in green background circle are the first extended interacting neighbors from the original genes, and genes in pink background circle are the further extended interacting neighbors from genes in green circle.



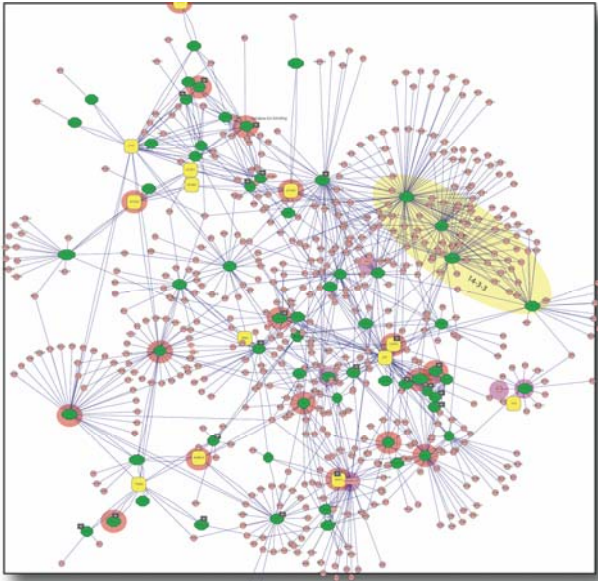


Figure 4 – Further expansion of protein-protein interaction map from 8 LQTS genes. Through DPDS, we can further inquiry associated genes related to Figure 3.

From the protein-protein network map (Figure 4), we selected 7 transcription factors as our studying focus in the sub-project, that is TBX5 (T-box 5), NKX2-5 (NK2 transcription factor related, locus 5/cardiac-specific homeobox), GATA4 (GATA binding protein 4), SMARCA4 (SWI/SNF related, matrix associated, actin dependent regulator of chromatinA4), TCF7L2 (T-cell specific, HMG-box), PAX6 (paired box gene 6), RUNX1 (runt-related transcription factor 1 (acute myeloid leukemia 1; aml1 oncogene)). Among these 7 transcription factors, NKX2-5, TBX5 and GATA4 have been reported to be involved in cardiac system, PAX6 is involved in nervous system, SMARCA4 is involved in ATP-dependent chromatin remodeling, RUNX1 is involved in hematopoiesis system. The results are promising as there are 3 out of 7 are already reported to be directly associated with cardiac system. We then extracted the cascade (upstream regulators and downstream regulated genes) information from TransPath. To demonstrate a clear expression from a complicated network, controlled vocabularies are used to classify the relationship between protein and protein and, protein and gene, such as transactivation, transrepression, phosphorylation activation, phosphorylation repression, activation, repression and DNA binding (Figure 5).

We took advantage of tissue-specific expression microarray data from 79 human tissues [7], as a means to evaluate the lines of evidence in our previously gene predicted network. Among these tissues, there include heart, atrioventricular node, cardiac myocytes. As shown in Figure 6, since NKX2-5, GATA4 and TBX5 are functioning especially in cardiac system and predicted important transcription factors in cardiac system.

We examined the expression pattern of every single gene including the transcription factors and the regulated genes in the network as shown in Figure 5. The genes are marked with blue background circle, if they are highly or exclusively expressed in heart, atrioventricular node or cardiac myocytes, as shown in Figure 7. There are 40 marked in blue circle out of 89 total gene network, indicating half of the genes presented in this predicted network are highly or exclusively expressed in heart. We conclude that the generated LQTS diseased cellular network using DPDS can be confirmed by the microarray data and will unravel the real unseen cellular world.

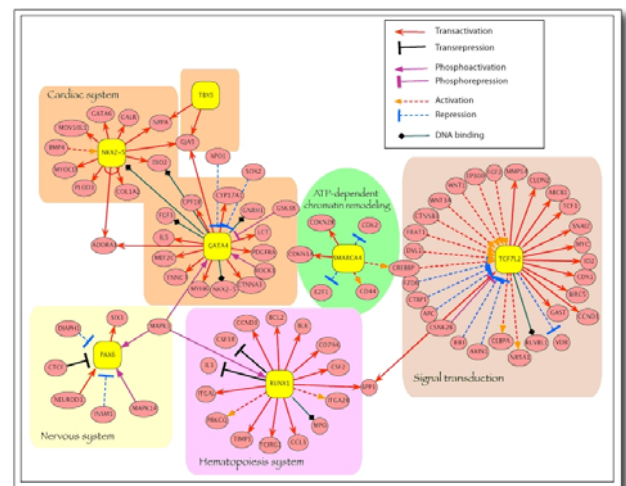


Figure 5 – Regulation of transcription factors in an interlaced web using controlled vocabularies. Seven transcription factors (in yellow square background) were selected from protein-protein interactions associated with cardiac arrhythmias as shown in Figure 4.

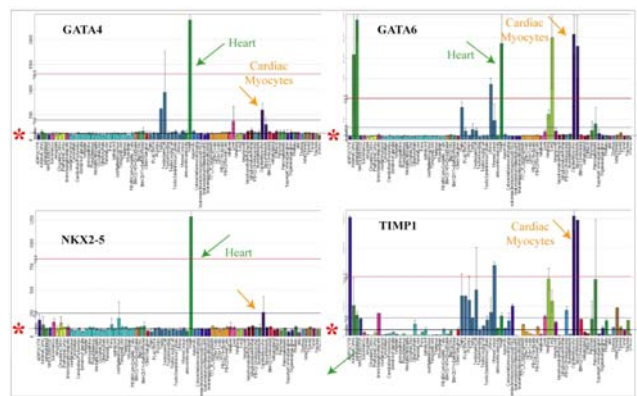


Figure 6 – Gene expression levels in 79 human tissues. The selected genes show highly expressed in heart and/or in cardiac myocytes. The asterisk indicates the median expression value calculated for each gene across all tissues.



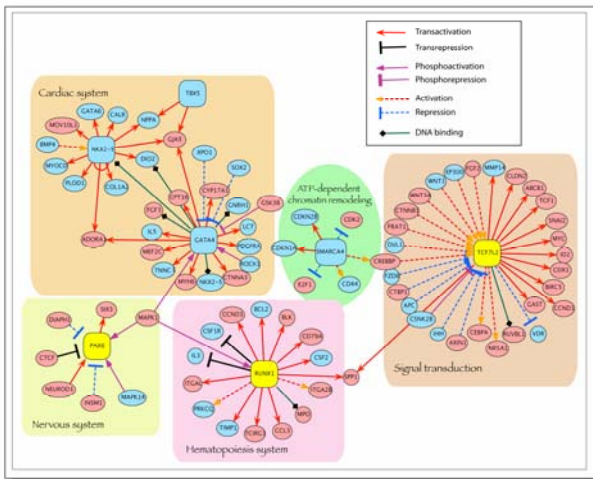


Figure 7 – Genes highly expressed in heart or heart-related tissues are marked in blue background circle in the transcription factor regulatory network. Greater ratio of genes expressed in heart when the transcription factors (shown in the center of each radial network) are highly expressed in heart.

As a means to evaluate the lines of evidence in our previously gene predicted network. Among the different 79 human tissues in Su et al. 2004 microarray dataset there include heart, atrioventricular node, cardiac myocytes. It is reasonable to infer their downstream regulated genes are expressed in higher levels in heart, cardiac myocytes or atrioventricular node. We conclude that the generated LQTS diseased cellular network using DPDS can be confirmed by the microarray data and will unravel the real unseen cellular world.

### 3. Reference

- [1] I. Xenarios, L. Salwinski, X. J. Duan, P. Higney, S. M. Kim, and D. Eisenberg, "DIP, the Database of Interacting Proteins: a research tool for studying cellular networks of protein interactions," *Nucleic Acids Res*, vol. 30, pp. 303-5, 2002.
- [2] H. Hermjakob, L. Montecchi-Palazzi, C. Lewington, S. Mudali, S. Kerrien, S. Orchard, M. Vingron, B. Roechert, P. Roepstorff, A. Valencia, H. Margalit, J. Armstrong, A. Bairoch, G. Cesareni, D. Sherman, and R. Apweiler, "IntAct: an open source molecular interaction database," *Nucleic Acids Res*, vol. 32, pp. D452-5, 2004.
- [3] A. Zanzoni, L. Montecchi-Palazzi, M. Quondam, G. Ausiello, M. Helmer-Citterich, and G. Cesareni, "MINT: a Molecular INTERaction database," *FEBS Lett*, vol. 513, pp. 135-40, 2002.
- [4] G. D. Bader, D. Betel, and C. W. Hogue, "BIND: the Biomolecular Interaction Network Database," *Nucleic Acids Res*, vol. 31, pp. 248-50, 2003.
- [5] S. Peri, J. D. Navarro, T. Z. Kristiansen, R. Amanchy, V. Surendranath, B. Muthusamy, T. K. Gandhi, K. N. Chandrika, N. Deshpande, S. Suresh, B. P. Rashmi, K. Shanker, N. Padma, V. Niranjana, H. C. Harsha, N. Talreja, B. M. Vrushabendra, M. A. Ramya, A. J. Yatish, M. Joy, H. N. Shivashankar, M. P. Kavitha, M. Menezes, D. R. Choudhury, N. Ghosh, R. Saravana, S. Chandran, S. Mohan, C. K. Jonnalagadda, C. K. Prasad, C. Kumar-Sinha, K. S. Deshpande, and A. Pandey, "Human protein reference database as a discovery resource for proteomics," *Nucleic Acids Res*, vol. 32, pp. D497-501, 2004.
- [6] I. M. Keseler, J. Collado-Vides, S. Gama-Castro, J. Ingraham, S. Paley, I. T. Paulsen, M. Peralta-Gil, and P. D. Karp, "EcoCyc: a comprehensive database resource for Escherichia coli," *Nucleic Acids Res*, vol. 33, pp. D334-7, 2005.
- [7] A. I. Su, T. Wiltshire, S. Batalov, H. Lapp, K. A. Ching, D. Block, J. Zhang, R. Soden, M. Hayakawa, G. Kreiman, M. P. Cooke, J. R. Walker, and J. B. Hogenesch, "A gene atlas of the mouse and human protein-encoding transcriptomes," *Proc Natl Acad Sci U S A*, vol. 101, pp. 6062-7, 2004.
- [8] P. Shannon, A. Markiel, O. Ozier, N. S. Baliga, J. T. Wang, D. Ramage, N. Amin, B. Schwikowski, and T. Ideker, "Cytoscape: a software environment for integrated models of biomolecular interaction networks," *Genome Res*, vol. 13, pp. 2498-504, 2003.
- [9] M. Csanady and R. Sepp, "[The long QT syndrome from the bedside to molecular genetic laboratory. The history of the first described Hungarian family]," *Orv Hetil*, vol. 146, pp. 2011-6, 2005.
- [10] H. Wedekind, T. Bajanowski, P. Friederich, G. Breithardt, T. Wulffing, C. Siebrands, B. Engeland, G. Monnig, W. Haverkamp, B. Brinkmann, and E. Schulze-Bahr, "Sudden infant death syndrome and long QT syndrome: an epidemiological and genetic study," *Int J Legal Med*, vol. 120, pp. 129-37, 2006.
- [11] H. Bundgaard, O. Havndrup, M. Christiansen, P. S. Andersen, H. K. Jensen, J. H. Svendsen, and K. P. Kjeldsen, "[Long QT syndrome: genes, mechanisms and risks; indication for genetic family screening?]," *Ugeskr Laeger*, vol. 168, pp. 2537-42, 2006.
- [12] D. M. Roden and P. C. Viswanathan, "Genetics of acquired long QT syndrome," *J Clin Invest*, vol. 115, pp. 2025-32, 2005.
- [13] L. Crotti, A. L. Lundquist, R. Insolia, M. Pedrazzini, C. Ferrandi, G. M. De Ferrari, A. Vicentini, P. Yang, D. M. Roden, A. L. George, Jr., and P. J. Schwartz, "KCNH2-K897T is a genetic modifier of latent congenital long-QT



- syndrome," *Circulation*, vol. 112, pp. 1251-8, 2005.
- [14] A. D. Paulussen, R. A. Gilissen, M. Armstrong, P. A. Doevendans, P. Verhasselt, H. J. Smeets, E. Schulze-Bahr, W. Haverkamp, G. Breithardt, N. Cohen, and J. Aerssens, "Genetic variations of KCNQ1, KCNH2, SCN5A, KCNE1, and KCNE2 in drug-induced long QT syndrome patients," *J Mol Med*, vol. 82, pp. 182-8, 2004.





# Constructing the Methionine Cycle and Its Related Disease Model with Hybrid Functional Petri Nets

Zhi En. Chen

Graduate Institute of Medical Informatics,  
Tzu-chi University  
[chen.real@gmail.com](mailto:chen.real@gmail.com)

Austin H. Chen\*

Graduate Institute of Medical Informatics,  
Tzu-chi University  
\*The Corresponding Author:  
[achen@mail.tcu.edu.tw](mailto:achen@mail.tcu.edu.tw)

## Abstract

*Methionine is a kind of essential amino acid, and it must be provided by the diet. Its derivatives play significant roles in human kinds. It has been proven that malfunctions in the methionine cycle have a relation with many severe diseases. As a result, we choose this mechanism to model for its indispensable status in human body. Petri Net is a promising tool to model metabolic pathway since we can establish a model based on biological knowledge intuitively.*

*In this paper, we construct a model of methionine cycle based on Hybrid Functional Petri Net (HFPN). Our results showed that several important effects of methionine cycle on severe diseases. Among them are excess methionine intakes result in a drastically increasing in Cystathionine that leads to disease like Down Syndrome, hepatic cancer, and liver disease. When the cycle with low CBS activity, cystathionine synthesis is stopped which causes cardiovascular disease, however, when CBS activity is too high, cystathionine increases and result in Down Syndrome.*

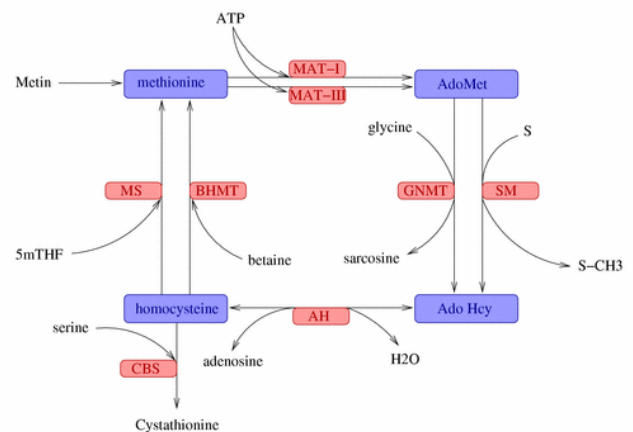
**Keywords:** metabolic pathway, methionine cycle, Petri Net, HFPN.

## 1. Introduction

Methionine is a kind of essential amino acid, meaning that it cannot be produced by the body, and must be provided by the diet. Its derivatives play significant roles in human kinds. In our body, methionine and its downstream metabolites accrue a cycle, in brief, comprising methionine, S-adenosylmethionine (AdoMet), S-adenosyl-homocysteine (AdoHcy), homocysteine, and cystathionine. The Methionine cycle, in fact, is highly complicated and supplies sulfur and other compounds required by the body for normal metabolism and growth. Normal functioning of the methionine cycle is essential for growth and development. If the methionine cycle is broken or temperately blocked, many diseases may take place. It has been proven that malfunctions in the methionine cycle have a relation with Alzheimer's disease [7], Down Syndrome [14], cardiovascular disease [17], liver disease [6], neural tube defects [4], and cancers [8, 15]. As a result, we choose this mechanism to model for its indispensable status in human body.

## The mechanism of the methionine cycle

Foods, such as fruits, meat, vegetables, nuts and legumes all contain methionine. There are several enzymes exist in the cycle, including MAT, SM, GNMT, AH, CBS, BHMT, and MS. When we take diet, methionine can first be change to AdoMet by an enzyme with isoforms, MATI and MATIII. After this reaction, SM and GNMT work to transfer methyl group and then produce AdoHcy. The enzyme called AH helps a reversible conversion of AdoHcy to Homocysteine, which has alternative fate in vivo. Homocysteine can be converted back to methionine, consequently, makes a close loop, or undergo transulfuration catalyzed by CBS to make cystathionine. In the operation of this cycle, positive and negative feedback make it more sophisticated than one considers. The first intermediate, AdoMet is served as an activator for MATIII and a repressor for MATI, however, its activity is not limited to the upstream reaction. AdoMet can also activate CBS and repress BHMT in the cycle. Furthermore, AdoHcy performs both activate and repressive properties, it can repress GNMT, SM, BHMT, however, also activate CBS.



**Figure 1.** Components of the methionine metabolism modelled by Reed et al.. The main metabolites are shown in blue boxes and the enzymes in red boxes. Abbreviations: 5mTHF=5-methyl tetrahydrofolate; AdoHcy= S-adenosyl- homocysteine; AdoMet= S-adenosyl- methionine; AH=S-adenosylhomocysteine hydrolase; BHMT = betaine:homocysteine methyltransferase; CBS=cystathionine bsynthase; GNMT=glycine N-methyltransferase; MAT= methionine adenosyl transferase; Metin=rate of methionine input;







MS=methionine synthase; S=substrates for methylation; SM=S-adenosyl-methionine-dependent methyltransferases. ([http://www.cellml.org/examples/images/reed\\_model\\_2004/reaction\\_diagram.gif](http://www.cellml.org/examples/images/reed_model_2004/reaction_diagram.gif))

## A promising theory for metabolic pathway modeling

Petri Net is a theory employed in many fields. Many extensions to the simple Petri Nets model have been developed for various modeling and simulation purposes. The major categories of Petri Net extensions are listed: (a) Hierarchical Petri Nets, which allow the previously defined net to present in a new net as an entity or process. (b) Hybrid Petri Nets, which allow the component to deal with continuous values instead of integer numbers of tokens. (c) Timed Petri Nets, which introduce the concept of deterministic time delays. (d) Stochastic Petri Nets, in which entity and process may be assigned delays which are given by a probability distribution. (e) Colored Petri Nets, which allow more complex firing rules in the processes.

However, it is until recently that Petri Net makes its huge influence on biological application, especially the metabolic pathway and other networks in biology. A novel notion of Petri net called hybrid functional Petri net (HFPN) and the enhanced version, hybrid functional Petri net with extension (HFPNe), have been developed by Masao Nagasaki and his colleague. Both methods extend and combine the notions of different kinds of Petri Nets, making it suitable for solving biological problems. As a result, in this paper, we choose Genomic Object Net, a software tool with HFPNe basis development by Nagasaki [12], to construct and simulate the complicated mechanisms of the methionine cycle.

## 2. Materials and Methods

In this section, we try to elucidate the magnificent role of Petri Net in modern simulation work. We will discuss the important roles of Petri Net, introduce the basic elements in HFPN, discuss how to construct our model, and introduce the useful tool called Genomic Object Net, or Cell Illustrator 2.0, developed by Masao Nagasaki and his colleague [2, 13].

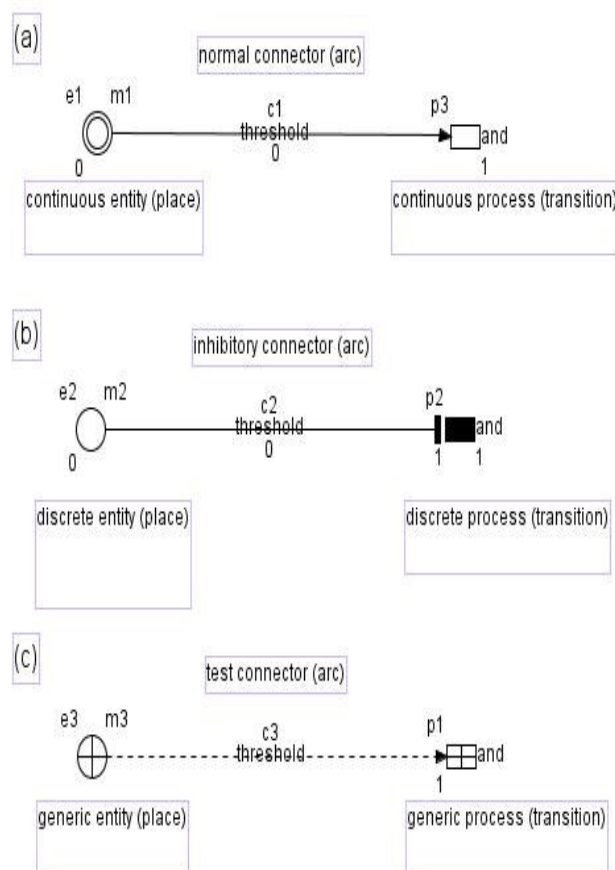
### The important role of Petri Net in biological simulation

Many theories and tools have been demonstrated for biological pathways or metabolite pathways simulation, such as E-Cell [9, 20], BioSpice [18] and so forth. Many useful simulating systems are built on the basis of ordinary differential equations (ODEs). Unfortunately, it is rather complicated to use E-Cell for modeling a simple biological pathway. Someone majors in biology won't prefer to learn complicated mathematical equations to construct their own models. It is really a chore for

biologists. Nevertheless, Petri Net holds its advantage for constructing model intuitively. In brief, Petri nets is suitable than other mathematical descriptions while simulating biological phenomena.

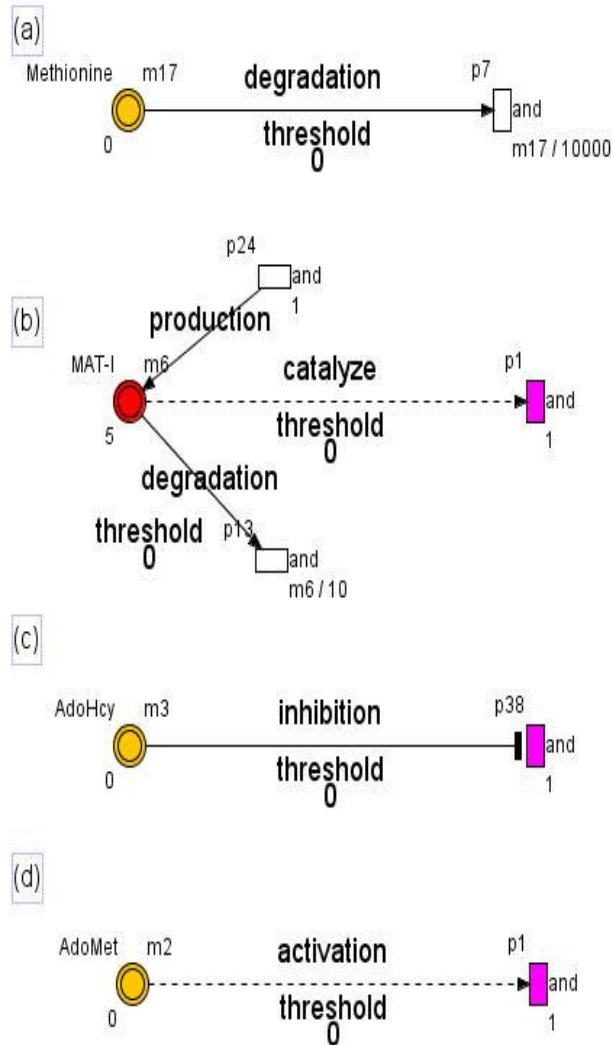
### Basic elements in Hybrid Functional Petri Net and their characteristics

HFPN contains several key components: entity (place), process (transition), connector (arc), and tokens (markings). Entity can represent protein, gene, metabolite, and any signal factor. Processes stand for reactions, binding, separation, transcription, translation and ordinary biochemical reactions. Connectors play the roles of connecting each component and making a network. Tokens, or markings, indicate the value held by each entity. Different kind of components and reactions are showed in Figure 2 and 3.



**Figure 2.** The different kinds of entities, processes, and connectors in HFPN. (a) continuous entity, normal connector, and continuous process. (b) discrete entity, inhibitory connector, and discrete process. (c) generic entity, test connector, generic process.



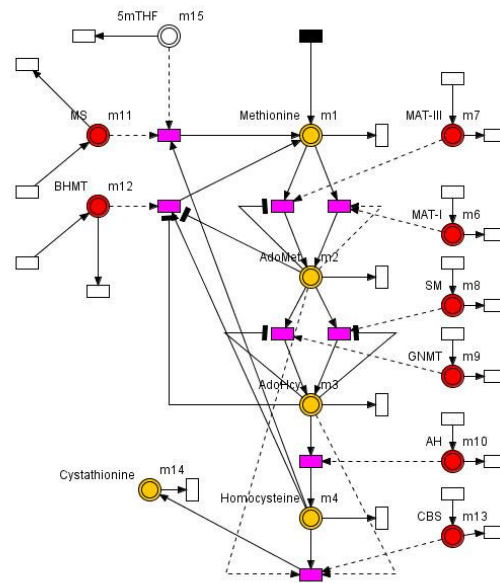


**Figure 3.** Various reactions in the model we established are expressed in (a) continuous entity, process and normal connector that represent degradation in metabolic cycle. (b) test connector that stands for enzyme catalysis and normal connector for protein production. (c) Inhibitory connectors that are used in inhibition reaction during the pathway. (d) test connectors that are used in activation.

### Constructing methionine cycle with Genomic Object Net

We construct a model of methionine cycle and apply the key parameters in the model to simulate those diseases which may be caused by abnormal methionine cycle. Entities represent enzymes and metabolites that are participating in this pathway. Continuous processes, however, represent the specific reactions occurred in this pathway. As an example, we have shown a simple structure in Figure 4. After the main structure has been established, we apply parameters provided by Reed [16] and Martinov [11], as listed in Table 1 and Table 2, to complete the whole model of the methionine cycle. As can be seen in Figure 5, we have finished a complete model of methionine cycle with all the firing rules.

### Methionine Cycle Simulation



**Figure 4.** A simple structure of methionine cycle established by HFPN.

**Table 1.** Properties of each entry in our constructed model.

| Name          | Type   | Variable   | Initial Value                              |
|---------------|--------|------------|--------------------------------------------|
| Methionine    | Double | m1         | 3.33                                       |
| AdoMet        | Double | m2         | 0                                          |
| AdoHcy        | Double | m3         | 0                                          |
| Homocysteine  | Double | m4         | 0                                          |
| Cystathionine | Double | m5         | 0                                          |
| MAT-III       | Double | m6         | 5                                          |
| MAT-I         | Double | m7         | 5                                          |
| SM            | Double | m8         | 5                                          |
| GNMT          | Double | m9         | 5                                          |
| AH            | Double | m10        | 5                                          |
| MS            | Double | m11        | 5                                          |
| BHMT          | Double | m12        | 5                                          |
| CBS           | Double | m13        | 5                                          |
| 5mTHF         | Double | m14        | 5.2                                        |
|               | Double | a1         | 100                                        |
|               | Double | a2         | 10                                         |
|               | Double | b1         | 1.7                                        |
|               | Double | b2         | 30                                         |
|               | Double | VmaxMATI   | 0.155833333                                |
|               | Double | VmaxMATIII | 6.352777778                                |
|               | Double | VmaxSM     | 1.255833333                                |
|               | Double | VmaxGNMT   | 2.944444444                                |
|               | Double | VAH        | $a1*(m3-a2*m4)/3600$                       |
|               | Double | VCBS       | $(b1*(m2+m3)-b2)*m4/3600$                  |
|               | Double | VmaxMS     | 0.138888889                                |
|               | Double | VmaxBHMT   | 0.694444444                                |
|               | Double | KmMATI     | 41                                         |
|               | Double | KmMATIII   | $2000/(1+5.7*(m2/(m2+600))*(m2/(m2+600)))$ |
|               | Double | KmSM       | $1.0*(1+(m3/4))$                           |
|               | Double | KmGNMT     | 4500                                       |
|               | Double | KmBHMT     | 12                                         |
|               | Double | KmMS       | 0.1                                        |



**Table 2.** Firing rules and properties of each process

| Name | Type       | Firing Style | Kinetic Script                                                |
|------|------------|--------------|---------------------------------------------------------------|
| p1   | continuous | and          | $V_{max}MATI/(1+(K_mMATI*(1+(m2/50))/(m1)))$                  |
| p2   | continuous | and          | $V_{max}MATIII/(1+(K_mMATIII*21.1)/(m1*m1+m1*21.1))$          |
| p3   | continuous | and          | $V_{max}SM/(1+(K_mSM/m3)+10*10*K_mSM/m3)$                     |
| p4   | continuous | and          | $(1/(1+(m3/20)))*(V_{max}GNMT/(1+(K_mGNMT/m2)*(K_mGNMT/m2)))$ |
| p5   | continuous | and          | $a1*(m3-a2*m4)/3600$                                          |
| p6   | continuous | and          | $(b1*(m2+m3)-b2*m4)/3600$                                     |
| p7   | continuous | and          | $V_{max}MS*m15*m4/(1*K_mMS+K_mMS*m15+25*m4+m15*m4)$           |
| p8   | continuous | and          | $(0.7-(0.025*(m2+m3-150)))*V_{max}BHMT*m4/(K_mBHMT+m4)$       |
| p9   | continuous | and          | $m1/10000$                                                    |
| p10  | continuous | and          | $m2/10000$                                                    |
| p11  | continuous | and          | $m3/10000$                                                    |
| p12  | continuous | and          | $m4/10000$                                                    |
| p13  | continuous | and          | $m5/10000$                                                    |
| p14  | continuous | and          | $m14/10000$                                                   |
| p15  | continuous | and          | $m6/10$                                                       |
| p16  | continuous | and          | $m7/10$                                                       |
| p17  | continuous | and          | $m8/10$                                                       |
| p18  | continuous | and          | $m9/10$                                                       |
| p19  | continuous | and          | $m10/10$                                                      |
| p20  | continuous | and          | $m11/10$                                                      |
| p21  | continuous | and          | $m12/10$                                                      |
| p22  | continuous | and          | $m13/10$                                                      |
| p23  | continuous | and          | 1                                                             |
| p24  | continuous | and          | 1                                                             |
| p25  | continuous | and          | 1                                                             |
| p26  | continuous | and          | 1                                                             |
| p27  | continuous | and          | 1                                                             |
| p28  | continuous | and          | 1                                                             |
| p29  | continuous | and          | 1                                                             |
| p30  | continuous | and          | 1                                                             |
| p31  | discrete   | and          | 3.33                                                          |

### 3. Results and Discussion

Now we can make a simulation on our methionine cycle set up by GON. We consider that the methionine in diet will be taken in gradually and we assume that the rate of methionine input (Metin) is  $3.33\mu\text{m}/\text{min}$  (about  $200\mu\text{m}/\text{h}$ ) and run our simulation for one hour (3600pt). In accordance with the model we have constructed, we can get the simulation results as: [Methionine]= $22.61\mu\text{m}$ , [AdoMet]= $102.23\mu\text{m}$ , [AdoHcy]= $10.06\mu\text{m}$ , [Homocysteine]= $0.64\mu\text{m}$ , [Cystathionine]=  $35.06\mu\text{m}$ . (Figure 6)

#### Excess methionine intakes result in serious diseases

We simulate this situation by increasing the amount of metin. Finkelstein and Martin [5] indicated that the fraction of Hcy, which is the source for cystathionine synthesis, is a function of the concentration of AdoMet. The higher intake of dietary methionine, the higher AdoMet concentration is detected. Increasing the concentration of [AdoMet] will elevate more

transulfuration fraction, and resulting in more cystathionine synthesis. We can observe a drastically increasing in [AdoMet] and [Cystathionine] in Figure 7. High level of cystathionine (cystathinuria) usually leads to some disease such as Down Syndrome [14], hepatic cancer [8], and liver disease. [6]

#### Effect of folate deficiency

After taking off the folate supply in our model, we can see that the concentration of both [Methionine] and [AdoMet] decreases slightly, at the same time, the concentration of [Homocysteine] and [Cystathionine] increases (Figure 8). This explains the phenomena described in some literatures. In our model, we have successfully proven that lack of folate will cause colorectal cancer [3, 15] and neural tube defects (NTDs) [4].

#### Effect of high and low expression of CBS activity

We modified the concentration of CBS in the model and found some interest results. Figure 9 showed that the cycle with broken CBS activity. When the cycle with low CBS activity, transulfuration function was lost and cystathionine synthesis is stopped which causes cardiovascular disease [17]. On the other hand, in Figure 10, when CBS activity is too high, cystathionine increase and result in Down Syndrome [14].

#### Methionine accumulation may cause a serious outcome

It is also interesting to understand the effect of MATI and MATIII at very low level of concentration. In our model, we assume that the initial concentration of both MATI and MATIII will be consumed within a very short time ( $\sim 1$  minute), in this situation, we can see that methionine accumulates quickly after running out of MATI and MATIII. All the subsequential intermediates syntheses are therefore completely blocked (Figure 11). This means that the whole cycle is damaged and severe results are expected.

#### Effect of betaine and methionine synthase (MS) deficiency

Betaine is a chemical that activate enzyme BHMT. In order to simulate the effect of betaine and MS deficiency, we set [BHMT] and [MS] equal to zero. The result is that [Cystathionine] arise moderately (Figure 12 and 13) and will result in diseases like higher CBS activity does.

### 4. Conclusion

Metabolic pathway is not a close and independent system. Other amino acids and intermediates may also interact with the reaction in the methionine cycle. In





this paper, we do not take them into account. A complete simulation of amino acids metabolic pathway is hard to simulate due to the complexity in this cycle. There are still many relations between amino acids and intermediates are still unknown. A partial metabolic pathway may have significant biological meanings; nevertheless, some factors that are ignored may cause an unexpected compact on the outcome. In our next step, we will take more factors into account and hopefully to explore more meaning knowledge good to human beings, especially in the cancers and diseases. KEGG [19] provides a lot of metabolic pathways, which indicate those metabolites interacting within the methionine cycle. The full-size simulation is a very challenge job and we want to improve our model to live up to this goal.

## 5. Acknowledges

We appreciate the financial supports of Tzu-Chi Research Grant TCMRC94012 and NSC 95-2221-E-320-001.

## 6. References

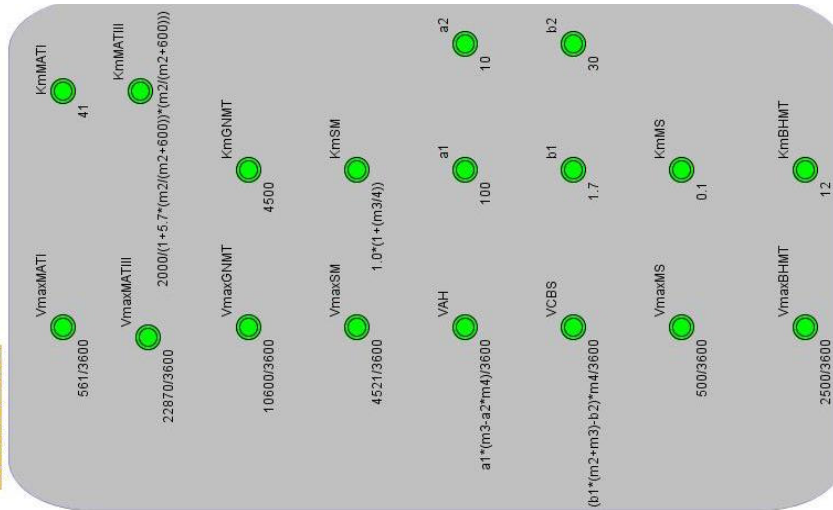
- [1] Brattström, Lars, David EL Wilcken, 2000. Homocysteine and cardiovascular disease: cause or effect? *Am J Clin Nutr.* 72, 315–23.
- [2] Doi, A., Nagasaki, M., Fujita, S., Matsuno, H., Miyano, S., 2004. Genomic Object Net:II. Modeling biopathways by hybrid functional Petri net with extension. *Applied Bioinformatics.* 2, 185– 188.
- [3] Duthie, S.J., 1999. Folic acid deficiency and cancer: mechanisms of DNA instability. *Brit. Med. Bull.* 55, 578–592.
- [4] Eskes, T.K.A.B., 1998, Neural tube defects, vitamins and homocysteine. *Eur J Pediatr.* 157, 139–141
- [5] Finkelstein, J., 1990. Methionine metabolism in mammals. *J. Nutr. Biochem.* 1, 228–237.
- [6] Finkelstein, J., 2003. Methionine metabolism in liver disease. *Am. J. Clin. Nutr.* 77, 1094–1095.
- [7] L, Miller A., 2003. The methionine-homocysteine cycle and its effects on cognitive diseases. *Altern Med Rev.* 8, 7-19.
- [8] L, Helson, Peterson R.H., Schwartz M.K., 1973. Cystathionine excess in children with hepatic cancer. *Cancer Res.* 33,1570-3.
- [9] M, Tomita, Hashimoto,K., Takahashi,K., Shimizu,T., Matsuzaki,Y., Miyoshi,F., Saito,K., Tanida,S., Yugi,K., Venter,J.C. and Hutchison, C.,1999. E-Cell: software environment for whole cell simulation. *Bioinformatics.* 15 (1), 72–84.
- [10] Matsuno, H., Tanaka, Y., Aoshima, H., Doi, A., Matsui, M., Miyano, S., 2003. Biopathways Representation and Simulation on Hybrid Functional Petri Net. *In Silico Biology.* 3(3), 389– 404.
- [11] Martinov, M., Vitvitsky, V., Mosharov, E., Banerjee, R., Ataullakhanov, F., 2000. A substrate switch: a new mode of regulation in the methionine metabolic pathway. *J. Theor. Biol.* 204, 521–532.
- [12] Nagasaki, M., Doi, A., Matsuno, H., Miyano, S., 2004. A Versatile Petri Net Based Architecture for Modeling and Simulation of Complex Biological Processes. *Genome Informatics.* 15(1), in press.
- [13] Nagasaki, M., Doi, A., Matsuno, H., Miyano, S., 2004. Genomic Object Net:I A platform for modeling and simulating biopathways. *Applied Bioinformatics.* 2, 181– 184.
- [14] Pogribna, M., Melnyk, S., Pogrybny, I., Chango, A., Yi, P., James, J.,2001. Homocysteine metabolism in children with Down Syndrome:in vitro modulation. *Am. J. Hum. Genet.* 69, 88–95.
- [15] Potter, J.D., 1999. Colorectal cancer: molecules and populations. *J. Nat. Cancer Inst.* 91, 916 – 932.
- [16] Reed, Michael C., H. Frederik Nijhout, Rachel Sparks, Cornelia M. Ulrich, 2004. A mathematical model of the methionine cycle. *Journal of Theoretical Biology.* 226, 33– 43
- [17] Refsum, H., Ueland, P.M., Nygard, O., Vollset, S.E., 1998. Homocysteine and cardiovascular disease. *Ann. Rev. Med.* 49, 31 – 62.
- [18] <https://biospice.org/index.php>
- [19] <http://www.genome.jp/kegg/>
- [20] <http://www.e-cell.org/>







# VARIABLE



# Methionine Cycle Simulation

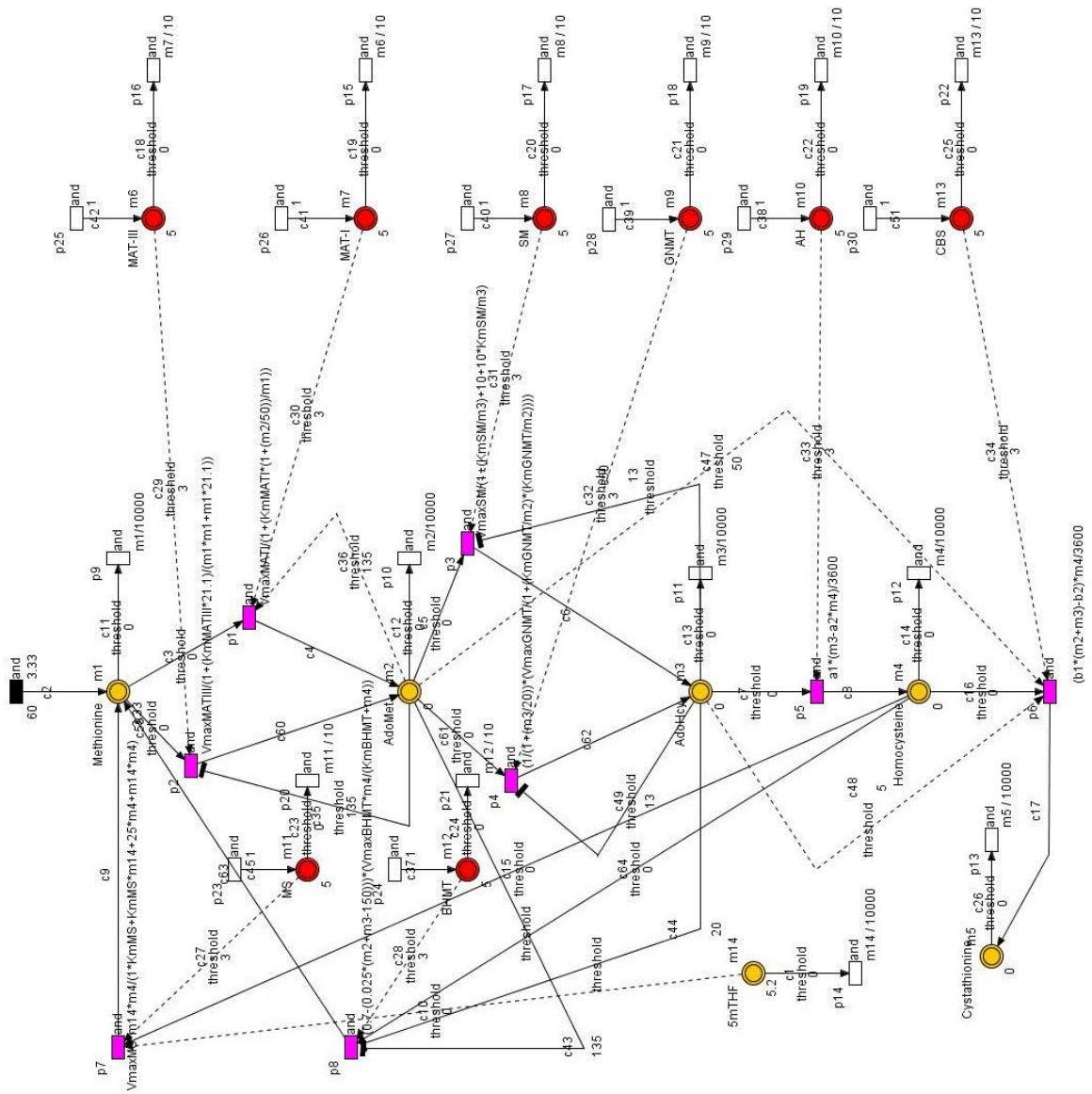
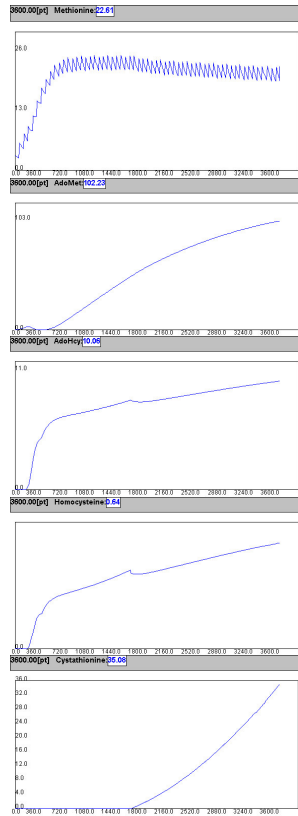
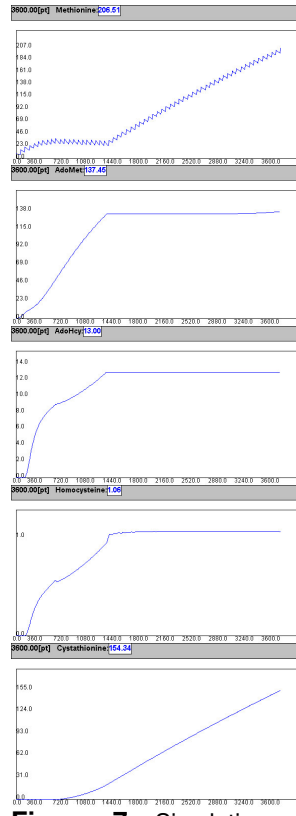


Figure 5. Simulation of methionine cycle and the variable are shown with all properties in each element.

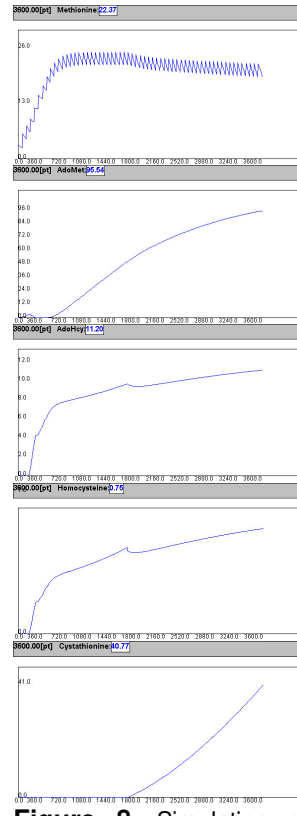




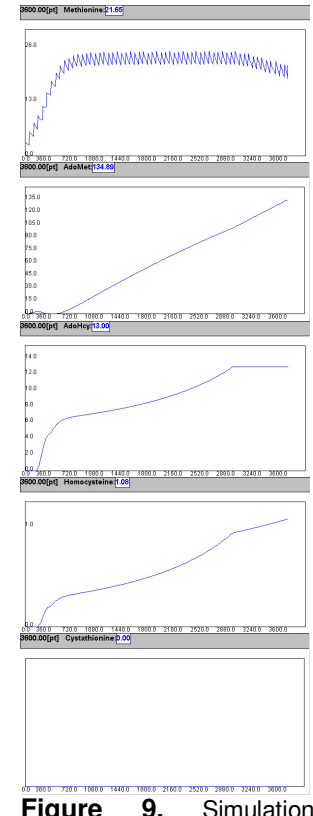
**Figure 6.** Simulation of normal methionine cycle.



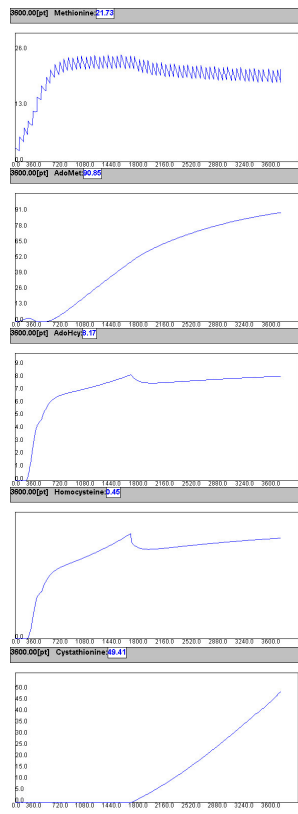
**Figure 7.** Simulation of excess methionine intake.



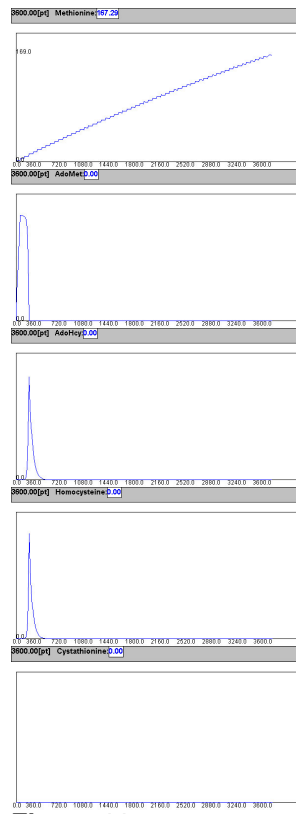
**Figure 8.** Simulation of folate deficiency.



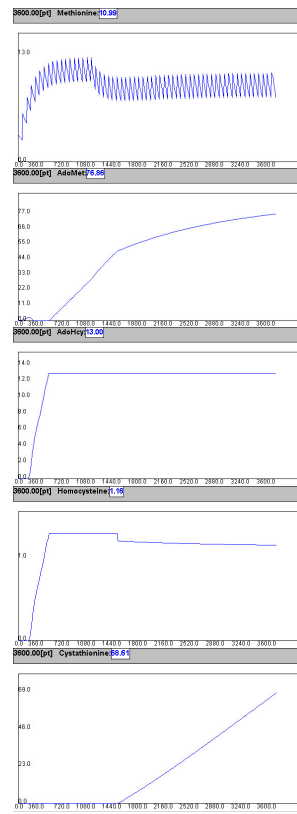
**Figure 9.** Simulation of cystathionine beta-synthase deficiency.



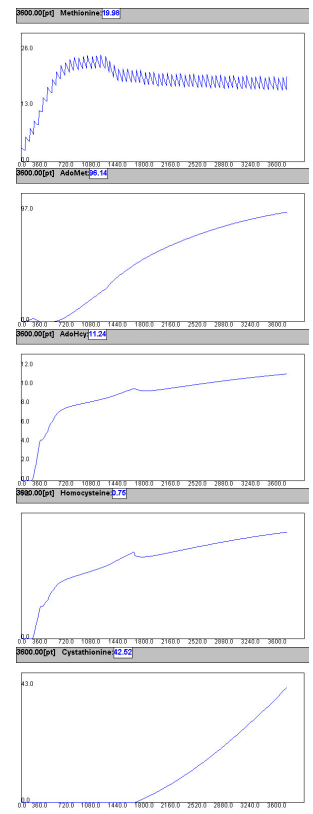
**Figure 10.** Simulation of methionine cycle in cystathionine beta-synthase deficiency.



**Figure 11.** Simulation of methionine cycle in MAT1/MATIII deficiency.



**Figure 12.** Simulation of methionine cycle in betaine deficiency.



**Figure 13.** Simulation of methionine cycle in methionine synthase deficiency..







# Computational Methods to Uncover the Protein-Protein Interactions

Fei-Hung Hung<sup>1</sup>, Hung-Wen Chiu<sup>1\*</sup>

<sup>1</sup>*Graduate Institute of Medical Informatics, Taipei Medical University, Taiwan*

<sup>\*</sup>*Correspondence E-mail: hwchiu@tmu.edu.tw*

## Abstract

*Protein-protein interaction (PPI) is an important mechanism for the cell's life. Advances on PPI information acquisitions made the prediction of PPI possible. Some high throughput experimental methods have been applied to find novel PPIs. Computational methods have been proposed for inferring PPI on the basis of known information about protein function and sequences. Once the protein-protein interaction can be predicted successfully, it will help to design new medications. In this paper, we briefly introduce some computational methods to uncover PPI and illustrate their achievements. Furthermore, we designed and demonstrate a tool to retrieve useful information about PPI.*

## 1. Introduction

Protein-protein interaction (PPI) provides the valuable information for a better understanding to cell life. Traditionally, PPI has been studied by chemical experiments [1-3]. The in-silico prediction of PPI is to know whether two proteins can interact or not by non-chemical experiment. The key point of PPI prediction is to find the rule about how two proteins interact based on known protein information. Any protein information can be used for the attempt -- the protein sequence, the function, the comparative relationship and many technical literatures for example. Depending on the work of wet lab, many useful information about PPI has been obtained and stored in databases for further research. Recently, a few high throughput approaches, e.g. yeast two-hybrid system [2,3], produces more than one half of total PPI information. A computer with higher speed process ability is essential in understanding PPI. Owing to the rapid accumulation of PPI knowledge, the rule about how two proteins interact has the chance to be extracted. Once the rule is uncovered the complex life system will be comprehended more clearly. The prediction of PPI also provides the new direction for the development of medication and treatment on molecular level. Pathogenic protein can be inhibited by its interacted protein. Moreover some researchers display PPI information using network graph to make protein-protein interrelationship more clearly. In this paper we presented some computational methods for PPI prediction and analyzed their achievements to demonstrate the feasibility of in-silico PPI prediction.

Furthermore, we presented some our developed tools to prepare the data for predicting PPI.

## 2. PPI Related Databases

### 2.1. Protein Sequence and Knowledge Databases

The Protein Information Resource (PIR, <http://pir.georgetown.edu/>) is devoted to spreading protein annotation and standardization. Now we could use PIR to classify protein families, find protein Knowledge and search literatures about protein.

SWISS-PROT (<http://ca.expasy.org/sprot/>) is a curated protein sequence database which strives to provide a high level of annotation and clearly protein sequence. Because it collects only experimental verified existences, related researchers could use it to find non-redundant information.

National Center for Biotechnology Information (NCBI, <http://www.ncbi.nih.gov/>) was established in 1988 as a national resource for molecular biology information. NCBI creates public databases, conducts research in computational biology, develops software tools for analyzing genome data, and disseminates biomedical information - all for the better understanding of molecular processes affecting human health and disease. After collection data from many different databases, researchers may discover that their data forms are not alike. At this moment, researchers can apply information in NCBI to transform them, so as to continuous.

Protein Families Database (Pfam, <http://www.sanger.ac.uk/Software/Pfam/>) is a large database covering many common protein domains and families. Investigators could look at multiple alignments of protein, view protein domain architectures and get known protein structures here. For application, researchers could use this database to check the function of one protein and apply it to find similarly functional regions. In addition, Pfam will show the visual graph of proteins the users searched.

### 2.2. PPI Databases

In addition to protein database, there are many databases for protein-protein interaction. Database of Interacting Proteins (DIP, <http://dip.doe-mbi.ucla.edu/>) database catalogs experimentally determined interactions between proteins. It combines information from a variety





of sources to create a single, consistent set of protein-protein interactions.

Biomolecular Interaction Network Database (BIND, <http://www.bind.ca/Action>) is a database collecting protein-protein interaction, cellular data and protein pathway. This web site describes PPI's characters and uses public literatures to show us.

The MIPS Mammalian Protein-Protein Interaction Database (MPPI, <http://mips.gsf.de/proj/ppi/>) is a collection of manually curated high-quality PPI data collected from the scientific literature by expert curators. They took great care to include only data from individually performed experiments since they usually provide the most reliable evidence for physical interactions.

### 2.3. Databases for PPI Prediction

These protein databases and PPI databases lead investigators to interest the prediction of PPI. Now this kinds of databases increase constantly. There is special for human PPI, besides biomolecular. Some team uses PPI to infer protein domain-domain interaction.

PreBIND (<http://prebind.bind.ca/>) is a tool of PPI prediction based on BIND and combines published papers. PreBIND contains biomolecular interaction. That is not merely protein-protein interactions. It would rather give a suitable interaction than a prediction. Returns from the PreBIND import links of related literatures to NCBI. When researchers link to its homepage, they could see the search frame. You could use protein name, NCBI accession number and PubMed ID to search. If some one wants to know a protein MSN2's interactions, he keys the protein name to this search frame and push submit down (Figure 1). The result page shows him the information of MSN2 above (Figure 2) and appears those interactions in published literatures below (Figure 3).

**Search for a protein by name and organism:**  
Enter the name of a protein (one word) and the organism it comes from.

MSN2  Any organism

For example, type "Ras1p" and select *Saccharomyces cerevisiae*.

**Figure 1. Picture of PreBIND's Search** This is the main page of PreBIND. Type your key word in the search frame, then push "Submit" bottom to search.

Protein description and name list

|                      |                                                          |
|----------------------|----------------------------------------------------------|
| protein name:        | multicopy suppressor of snf1 mutation                    |
| encoding locus name: | MSN2                                                     |
| source database:     | REFSEQ                                                   |
| accession number:    | NP_013751                                                |
| gi of protein:       | 6323680 <a href="#">SeqHound</a> (NCBI)                  |
| taxon:               | <i>Saccharomyces cerevisiae</i> ( <a href="#">4932</a> ) |

The following table lists the name(s) used to find the above gene product in the literature.

| Search | Date last searched | Number of results found | Notes |
|--------|--------------------|-------------------------|-------|
| MSN2   | Sep 02, 2002       | 45                      |       |

**Figure 2. Upper the Result of PreBIND** The frame is the basic information of this protein you key in.

| Summary of all potential interactors                                                                                                                                                                                                                                             |                                                                                                                |                          |                        |                          |                         |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------|--------------------------|------------------------|--------------------------|-------------------------|
| The list below shows all other proteins that co-occur in the literature with your query protein. The number of co-occurrence papers are listed under the column "View supporting papers". Clicking on this number will take you to a more detailed view of these co-occurrences. |                                                                                                                |                          |                        |                          |                         |
| name                                                                                                                                                                                                                                                                             | short description                                                                                              | is this interactor real? | View supporting papers | more info                | more info               |
| MSN4                                                                                                                                                                                                                                                                             | multicopy suppressor of snf1 mutation                                                                          | Probably                 | 17                     | <a href="#">SeqHound</a> | <a href="#">PreBIND</a> |
| TPK2                                                                                                                                                                                                                                                                             | Involved in nutrient control of cell growth and division                                                       | Probably                 | 2                      | <a href="#">SeqHound</a> | <a href="#">PreBIND</a> |
| SLK1                                                                                                                                                                                                                                                                             | Glucose phosphorylation                                                                                        | Probably                 | 1                      | <a href="#">SeqHound</a> | <a href="#">PreBIND</a> |
| NOT5                                                                                                                                                                                                                                                                             | member of the NOT complex, a global negative regulator of transcription                                        | Probably                 | 1                      | <a href="#">SeqHound</a> | <a href="#">PreBIND</a> |
| HSP12                                                                                                                                                                                                                                                                            | induced by heat shock, entry into stationary phase, depletion of glucose, and addition of lipids (fatty acids) | Probably                 | 2                      | <a href="#">SeqHound</a> | <a href="#">PreBIND</a> |
| DDR2                                                                                                                                                                                                                                                                             | DNA Damage Responsive                                                                                          | Probably                 | 2                      | <a href="#">SeqHound</a> | <a href="#">PreBIND</a> |
| CTT1                                                                                                                                                                                                                                                                             | cytoplasmic catalase T                                                                                         | Probably                 | 2                      | <a href="#">SeqHound</a> | <a href="#">PreBIND</a> |
| NOT3                                                                                                                                                                                                                                                                             | General negative regulator of transcription; may inhibit RNA polymerase II transcription machinery             | Probably                 | 1                      | <a href="#">SeqHound</a> | <a href="#">PreBIND</a> |
| TOR1                                                                                                                                                                                                                                                                             | Involved in cell cycle signaling and meiosis                                                                   | Unknown                  | 3                      | <a href="#">SeqHound</a> | <a href="#">PreBIND</a> |
| GLN3                                                                                                                                                                                                                                                                             | Regulates nitrogen-repressible gene products                                                                   | Unknown                  | 2                      | <a href="#">SeqHound</a> | <a href="#">PreBIND</a> |

**Figure 3. Under the Result of PreBIND** It shows all interactions of literatures on the internet.

Human Protein Interaction Database (HPID, <http://wilab.inha.ac.kr/hpid/>) is a database for human protein-protein interactions. HPID applies statistics and computations to provide researchers finding interactions. You enter their homepage and see a frame for search. Here we could type IDs of many database including NCBI to process. Figure 4 shows the procedures of search.

**www.HPID.org Human Protein Interaction Database**  
wilab.inha.ac.kr/HPID

News • Overview • Interactions •

**HPID Search**  
Human Protein ID:

**Search result**  
Ensembl Peptide ID: ENSP00000260433

**Interaction**  
Prediction Interaction: [OnlinePrediction](#)

**External Reference**

|              |                 |
|--------------|-----------------|
| EMBL         | J04127          |
| EMBL         | M18656          |
| EMBL         | M22246          |
| EMBL         | M26420          |
| EMBL         | M30796          |
| EMBL         | M30797          |
| EMBL         | M30798          |
| EMBL         | M30800          |
| EMBL         | M30801          |
| EMBL         | M30802          |
| EMBL         | M30803          |
| EMBL         | M30804          |
| EMBL         | X13569          |
| EMBL         | Y07508          |
| Ensembl_GENE | ENSG00000137869 |
| HPID         | 00144           |
| HPID         | 00488           |
| MIM          | 107910          |
| NCBI         | AAA35556        |
| NCBI         | AAA35557        |
| NCBI         | AAA35728        |
| NCBI         | AAA52132        |
| NCBI         | AAA52141        |
| NCBI         | CAA31929        |
| NCBI         | CAA68807        |
| RefSeq       | NP_000094.2     |
| RefSeq       | NP_112503.1     |
| SWISSPROT    | P11511          |

**Domain (family) information**  
[ENSP00000260433]

|          |           |
|----------|-----------|
| InterPro | IPR001128 |
| InterPro | IPR002397 |
| InterPro | IPR002401 |
| InterPro | IPR002403 |
| Pfam     | PF00067   |
| PRINTS   | PR00359   |
| PRINTS   | PR00395   |
| PRINTS   | PR00463   |
| PRINTS   | PR00465   |
| Prosite  | PS00086   |
| SCOP     | a.104.1.1 |

**HPID Online Prediction Report**  
Protein 1: ENSP00000260433

Superfamily 1: a.104.1 (Cytochrome P450)  
Superfamily's partner 1: a.104.1 [Click here to see protein\(s\) with a.104.1 structure.](#)  
Superfamily's partner 2: c.23.5 [Click here to see protein\(s\) with c.23.5 structure.](#)

Copyright 2005 Web Intelligence Lab. All rights reserved.

**Figure 4. Procedures of using HPID** (A) Researchers can type what they want in the "Human Protein ID" frame to search. (B) There is the result page of searching. Users can see its basic information and click the link "Online Prediction" to appear a new window to show the prediction of PPI. (C) The new window of PPI prediction.





Protein-Protein Interaction Database (PPID, <http://www.anc.ed.ac.uk/mscs/PPID/>) is a database covering human, mouse and rats. We can use association numbers, type of protein and their PPID number to search what you want. As the figure F, we chose the kind of input (Figure 5A) and key what you want to search (Figure 5B), then press the “Search”. In an instant, it appears the result below (Figure 5C).

Figure 5 shows the PPID database search interface. (A) The input frame has several dropdown menus: PPID (internal accession number), Accession number (Other DataBases), Name of the protein, Type of Protein (i.e. Channel), Drug (Agonist/Antagonist), and Diseases involved with proteins of PPID. (B) The search input field contains the protein ID 'P09429'. (C) The search results table shows one result found:

| PPID  | Protein Name | Type of Protein | M(kDa) | SwissProt             | UniGene               | OMIM   | Modification Date | Interaction Date |
|-------|--------------|-----------------|--------|-----------------------|-----------------------|--------|-------------------|------------------|
| A1095 | Amphotericin | Enzyme          | 25     | Human<br>Mouse<br>Rat | Human<br>Mouse<br>Rat | 163905 | 0000-00-00        | 0000-00-00       |

**Figure 5. Return of PPID** (A) Upper the input frame, researcher has several choices for input type. (B) Users can type the protein to search. (C) The result.

Database of Interacting Domain (InterDom, <http://interdom.lit.org.sg/>) is a specially database. They predict protein domain-domain interactions from PPIs. Though it is not for prediction of PPI, their approach is applied on prediction of PPI. InterDom is a database of *putative* interacting protein domains derived from multiple sources, ranging from domain fusions (Rosetta Stone), protein interactions (DIP and BIND), protein complexes, to scientific literatures. When investigators use, they should chose “Search” and it would give three selections. Researchers can use protein, domain and sequence to do their search (Figure 6A). We use protein P09429 to set an example. Here we should click the “by Protein” button, then type P09429 into the search frame (Figure 6B). When we push “Search”, we get the answer. You could see their description, domain name and domain-domain interaction (Figure 6C).

By many investigators’ striving, these related information tools have many achievements. Through communications of information, researchers can almost apply these database without misunderstanding (Table 1).

**Table 1. Data Sources of Related Protein Databases**

|          | PIR | SWISS-PROT | NCBI | Pfam | DIP | BIND |
|----------|-----|------------|------|------|-----|------|
| DIP      | *   | *          | *    |      |     |      |
| BIND     |     | *          | *    |      |     |      |
| MPPI     |     | *          | *    |      |     |      |
| PreBIND  |     |            |      |      |     | *    |
| HPID     |     |            |      |      | *   | *    |
| InterDom |     |            | *    | *    | *   | *    |

<sup>ψ</sup> It is appeared that data of PPI interflow by a wide margin. The consistence of data form means researchers are not depressed to share their research any more. This research domain carves out further.

Figure 6 shows the InterDom database search interface. (A) The main page has three search options: by Protein, by Domain, and by Sequence. (B) The search input field contains the protein ID 'P09429'. (C) The search results for High mobility group protein 1 (HMG-1) (P09429) show its function: Binds preferentially single-stranded DNA and unwinds double stranded DNA. The length is 214 AA. The protein contains one domain: HMG\_box (HMG (high mobility group) box) with a length of 69. The domain is associated with Pfam and InterDom annotations.

**Figure 6. Use InterDom** (A) At the main page, it shows three options for search. (B) Search by protein. (C) It is return for result. Under this page, it shows the information for the protein and interaction.

### 3. In-Silico Methods to Predict PPI

Because of the high throughput of protein-protein interaction, research workers will not be able to analyze these data by hand gradually. Thus researchers apply the computer and designed algorithms to analyze and then transform those data into useful information. That is the mainstream approach now.

#### 3.1. Homology

Protein-protein interaction and DNA-protein interaction is the important sources for protein research. However these data are limited. It spent much manpower and material resources at many laboratory. The final result is only specific specie. How to apply known information into other organisms? One way is homology. Now investigators know protein X and protein Y interact in organism A. In organism B, there are protein X' that is ortholog to protein X and protein Y' that is ortholog to protein Y. In that case, researchers could infer there is a interaction between protein X' and protein Y'. Such relationship as above is named the homology [4-6].

#### 3.2. Rosetta Stone

Rosetta stone is a stele. Archaeologists translate the ancient Egypt hieroglyph by it. One of protein sequence alignment approach is named it [4]. If investigators want to know whether two protein interact or not, they search similar sequence fusing by the foregoing two in other organisms. These protein sequences in other organisms are called Rosetta stone protein. In case they find similar one or more, they think the prediction exists. Now this





concept is applied for many relationships, for instance, domain-domain fusions [6].

### 3.3. Data Mining

Data mining is a method to mine useful information from a large number of data. For example, some investigators used Maximum Likelihood Estimation (MLE) to predict [7].

Text mining and graph mining are both kinds of data mining. Text mining is to use published literatures, journals and the contents of many web sites to test their key words and contexts for protein-protein interactions. Finally, researchers store the resulted relationship for get PPI or domain-domain interaction information [6].

E. Segal *et al.* point out that many cellular pathway have two characteristics. One is their gene expression are similar. Second, products of these genes often interact [8]. Besides many researchers link those identification proteins each other to turn them into a connected network graph. It supports a visual tool to appear relationship of proteins [9]. Some researcher applies graph mining to find novel protein complexes [10].

## 4. Current Achievements of PPI Prediction

Presently plenty research teams effort in PPI prediction, but all methods can not predict it perfectly. Nevertheless, the successful percentage of prediction enhances constantly. Two research teams, Uetz and Ito, used Y2H system to make high throughput of PPI even since [2,3]. The prediction of protein-protein interaction is not so far. It is known that we can get a large network including 3186 mostly novel interactions among 1705 proteins [1]. This interaction network can be applied to find interactions linking uncharacterized gene products and human disease protein about regulatory cellular pathways. It offers us a new viewpoint of protein-protein interaction. Edward M. Marcotte *et al.* brings "Rosetta Stone" approach. They apply this approach to use gene sequences for inferring PPI [4]. In the past, the protein functions depend on their structures, but investigators can recognize the whole through the observation of the part by using homology. It creates a distant path. A database was built by applying homology to extending PPIs experimentally verified in module organism to the orthologous proteins in *Homo sapiens* [5]. Besides Minghua Deng *et al.* apply data mining, e.g. MLE method, Association Method, to predict [7]. Dong-Soo Han *et al.* use more accurate computational approach, it is domain used, to do PPI prediction. In their way, the sensitivity is 77% and the specificity is 95% [11,12]. Graph mining is used by Xiao-Li Li *et al.* to build LCMA algorithm for prediction of PPI [10]. See-Kiong Ng *et al.* and Minghua Deng *et al.* infer protein domain-domain interactions from protein-protein interactions [6,7]. Their way differs from others. See-Kiong Ng's team use homology and text mining, too. They create the

InterDom database. Table 2 lists some different methods from different teams.

**Table 2. Lists of Different Methods**

| Team                                 | Methods                                                                           | Result or Verification                                                                                                  |
|--------------------------------------|-----------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------|
| Edward M. Marcotte <i>et al.</i> [4] | Rosetta Stone<br>Domain fusion<br>homology                                        | By filtering promiscuous domains, they make true interactions by 47% over the unfiltered predictions. <sup>†</sup>      |
| Xiao-Li Li <i>et al.</i> [10]        | Graph Mining<br>(LCMA algorithm)                                                  | LCMA was compared with MCODE <sup>‡</sup> .<br>For the F-measure <sup>§</sup> , LCMA is 15.99% higher score than MCODE. |
| Dong-Soo Han <i>et al.</i> [11,12]   | (domain combination<br>based protein-protein<br>interaction prediction<br>method) | Sensitivity = 77%<br>Specificity = 95%                                                                                  |
| Minghua Deng <i>et al.</i> [7]       | Data Mining<br>(MLE Method)                                                       | Compare with the result of association method and another database                                                      |
| See-Kiong Ng <i>et al.</i> [6]       | Homology<br>Rosetta Stone<br>Text Mining                                          | (Probabilistically-weighted<br>Odd Ratios)                                                                              |

<sup>†</sup> This data was about *E. coli*.

<sup>‡</sup> MCODE is a proposed created by Bader and Hogue [13].

<sup>§</sup> They defined  $Recall = |TP| / (|TP| + |FN|)$  and  $Precision = |TP| / (|TP| + |FP|)$ . F-measure takes into account of both precision and recall and is defined as  $F-measure = 2 * Precision * Recall / (Precision + Recall)$ .

## 5. Discussion

Attempts of uncovering protein-protein interactions is an accumulation of human wisdom. PPI prediction should be the greatest topic of this generation. Kinds of Protein complexes are numerous to lead to a huge degree of difficulty. Although now technology can not decode it, human continuous study and the high throughput of relative information would make it truth about eradicating cancer, stopping aging and curing inheritable disease some day in the future.

The current trend is to compare some relationship of interaction for prediction and it gets a considerable achievement. Even so, there are many job needed to be strived. Horng JT. *et al.* have supported a approach about applying computational method with data mining for protein motif. Their team wants to predict protein structure [14]. Here we try a way about applying protein motif to predict PPI without prediction of structure. Investigators know there is one or more motifs in the protein. Those motifs concern with protein functions thickly. For the reason, we find respective motifs out from known interacted proteins. Next, we may be able to look for the relationship of protein-protein interaction. Finally, we could use the relationship found to do PPI prediction. Now we have collect data of DIP, BIND, PIR and NCBI and use PHP programs we made to transfer them to our form. Now DIP and the non-redundant human PPI of BIND are completed. Total is 18,617 and store them in our MySQL database. The web-based







database (Figure 7) is building. Its functions are fragmentary. InterDom is our direction. We hope that will support another useful tool when the job is finished.

Enter what you want to search :

Example : DIP:478N → Protein\_DIP  
 SW:P05107 → SwissProt  
 PIR:IJHULM → PIR  
 GI:72068 → GI  
 NP\_000202 → RefSeq

Search result :

| Protein_DIP | SwissProt | PIR        | GI         | RefSeq    |
|-------------|-----------|------------|------------|-----------|
| DIP:5741N   | sw:P35968 | pir:I58357 | gi:9087218 | NP_002244 |

DIP's Interactions: DIP:5740N;DIP:5742N;DIP:5751N;NP\_877963;NP\_003367;NP\_005420;

BIND's Interactions: DIP:5740N;DIP:5742N;DIP:5751N;NP\_877963;NP\_003367;NP\_005420;

Figure 7. Our Database

## 6. References

- [1] Stelzl U., Worm U., Lalowski M. *et al.*, "A Human Protein-Protein Interaction Network: A Resource for Annotating the Proteome", *Cell*, 2005, Vol. 122, p. 957-968.
- [2] Uetz P., Giot L., Cagney G. *et al.*, "A comprehensive analysis of protein-protein interactions in *Saccharomyces cerevisiae*", *Nature*, 2000, Vol. 403, p. 623-627.
- [3] Ito T., Chiba T., Ozawa R. *et al.*, "A comprehensive two-hybrid analysis to explore the yeast protein interactome", *Proc. Natl. Acad. Sci.*, 2001, Vol. 98, p. 4569-4574.
- [4] Marcotte EM., Pellegrini M., Ng HL. *et al.*, "Detecting Protein Function and Protein-Protein Interactions from Genome Sequences", *SCIENCE*, 1999, Vol. 285, p. 751-753.
- [5] Cesareni G., Ceol A., Gavrila C. *et al.*, "Comparative interactomics", *FEBS Letters*, 2005, Vol. 579, p. 1828-1833.
- [6] Ng SK., Zhang Z., Tan SH. *et al.*, "InterDom: a database of putative interacting protein domains for validating predicted protein interactions and complexes", *Nucleic Acids Research*, 2003, Vol. 31(1), p. 251-254.
- [7] Deng MH., Mehta S., Sun FZ. *et al.*, "Inferring Domain-Domain Interactions From Protein-Protein Interactions", *Genome Research*, 2002, Vol. 10, p. 1540-1548.
- [8] Segal E., Wang H., Koller D. *et al.*, "Discovering molecular pathways from protein interaction and gene expression data", *Bioinformatics*, 2003, Vol. 19, p. 264-272.
- [9] Hu Z., Mellor J., Wu J. *et al.*, "VisANT: an online visualization and analysis tool for biological interaction data", *BMC Bioinformatics*, 2004, Vol. 5, p. 17-24.
- [10] Li XL., Tan SH., Ng SK. *et al.*, "Interaction Graph Mining for Protein Complexes Using Local Clique Merging", *Genome Informatics*, 2005, Vol. 16(2), p. 260-269.
- [11] Han DS., Kim HS., Seo JM. *et al.*, "A Domain Combination Based Probabilistic Framework for Protein-Protein Interaction Prediction", *Genome Informatics*, 2003, Vol. 14, p. 250-259.
- [12] Han DS., Kim HS., Jang WH. *et al.*, "PreSPI: Design and Implementation of Protein-Protein Interaction Prediction Service System", *Genome Informatics*, 2004, Vol. 15(2), p. 171-180.
- [13] Bader GD. and Hogue CWV., "An automated method for finding molecular complexes in large protein interaction networks", *BMC Bioinformatics*, 2003, Vol. 4(1):2.
- [14] Horng JT., Huang HD., Wang SH. *et al.*, "Computing Motif Correlations in Proteins", *Comput Chem*, 2003, Vol. 24, p. 2032-2043.



# Mutual Information correction for eddy current-induced Geometric artefacts In Diffusion-Tensor Tractography

<sup>1</sup>Kao-Lun Wang, <sup>1</sup>Shiou-Ping Lee, <sup>3</sup>You-Jhuan Lee, <sup>2</sup>Wen-Yih Isaac Tseng,  
<sup>1</sup>department of Medical Imaging, Far Eastern Memorial Hospital, <sup>2</sup>Center for Optoelectronic Biomedicine  
National Taiwan University, and <sup>3</sup>Institute of Medical Informatics, Taipei Medical University, Taipei, Taiwan

Correspondence: R90548010@ntu.edu.tw

## Abstract

*The paper investigated the validity of an image coregistration method, Mutual Information correction(MI) for fiber tractography. We found that mean increase was 18.56(%) in FA and 10.02(°) in first eigen vector variation angle through MI approach for White matter ROIs[1]. In addition, Plotted with the tensor ellipsoid shape, the region of gray matter is more consistent the T2 turbo spin-echo FLAIR image.*

*Tractography of Cortico-spinal was used to compare with reconstructing pathways of MI correction before and after. Utilization of MI correction of Diffusion-weighted images generate more consistency with CS fiber tracks to lateral aspect of the postcentral and contralateral gyrus. The results reveal that each tract has a unique spatial signature in term of anatomical atlas property[2].*

TR/TE = 8900/95 ms. Diffusion DTI encoding schemes were used 13 icosahedrally-oriented steps. To acquire 55 transaxial view slices covering the whole brain image. The scan time for each DTI data set was about 5 minutes.

T2 FLAIR were obtained by using the Turbo-spin-echo image sequence to finish transaxial images to include whole brain with slice thickness of 5 mm. The spatial resolution was higher than Diffusion-weighted images, the image level was more similar to Diffusion-weighted image. and TR/TE = 8500/99 ms.

We used modified SPM-based procedure that was called normalized Mutual Information to coregister DWIs to non-diffusion weighted reference images method[3-5].

Tractography algorithm used tensor deflection by an adaptive step size approach[6], tracking fiber results based on manual selecting of Cortico-spinal local seed points.

## 1. Introduction

Geometric distortion of echo-planar images produced by eddy currents induced geometric distortions that vary with the magnitude and direction of the diffusion sensitizing gradients. Such distortions derived from acquiring misalignment that different diffusion gradient strengths and orientation in turn alternate consequence.

While Diffusion-Tensor Imaging has been an important tool to reveal axonal fiber tracts in cerebral white matter noninvasively technique, single-shot diffusion gradient EPI sequence suffered from geometric distortion that will have an influence on white matter Tractography.

This aim of this work is to utilize the potential advantage of MI coregistration in the raw diffusion-weighted images and to visualize more accurate Cortico-spinal without use of complicate postprocessing in an extra hardware system.

## 2. Materials and Methods

Healthy volunteer was scanned on a 1.5T MRI system (Sonata, Siemens, Germany). A diffusion EPI sequence was used to acquire transaxial images encompassing the whole brain. Isotropic spatial resolution was obtained by setting FOV = 280 mm x 280 mm, matrix size = 128 x 128 and slice thickness = 2.2 mm. Diffusion-weighted images using the diffusion sensitivity (b-value) = 1000 s/mm<sup>2</sup> were acquired with

## 3. Results and Discussion

FA map combining first eigen vector map on 32th axial slice are shown in Figure 1. Note that the white arrow appear clarity of the rim along the brain edges in Mutual Information before and after to exhibit the high intensity difference due to image distortion.

Whereas the MI correction before and after shown an increase that was statistically significant in different white matter ROIs. The ROIs measurement (Table 1) confirmed this improvement in white matter location sampled. A mean increase of 18.56(%) in FA and 10.02 (°) in First Eigen vector variation angle was found for MI correction.

T2 FLAIR Turbo spin-echo transaxial brain image distinctly segmented into gray and white matter was shown in Figure 2(Left) comparing with reconstruction of color-coding FA map(Right) after MI correction on 32th transaxial view image.

As shown in Figure 5, the reconstruction of gray matter region FA < 0.15 is more reasonable after MI correction.

Tractography for Cortico-spinal tracks were launched from selected seed points ROIs in pons to visualize bypass to display fault tracks with red arrows in Figure 4. Compared with in Figure 5 such as anatomical atlas performed CS fiber tracts was correlated to the main pathway of fiber bundles[2].





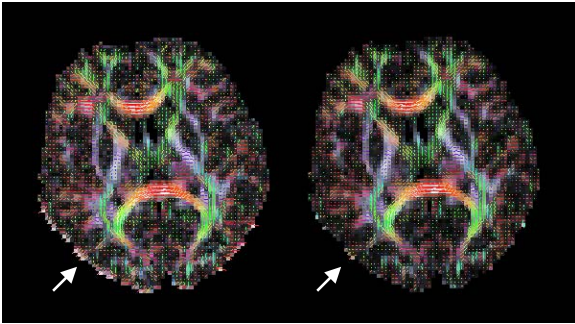


Figure 1. Reconstruction First Eigen vector superimposed FA map with MI before (Left) and after (Right).

Table 1. Illustration of difference in Mean FA and First Eigen vector as measured using MI correction in various ROIs.

| White Matter        | FA error | First Eigen vector variation |
|---------------------|----------|------------------------------|
| ROI region          | (%)      | angle (°)                    |
| Corpus Callosum     | 11.68    | 7.87                         |
| Posterior Occipital | 18.06    | 11.25                        |
| Internal Capsule    | 25.94    | 10.94                        |
| Mean                | 18.56    | 10.02                        |

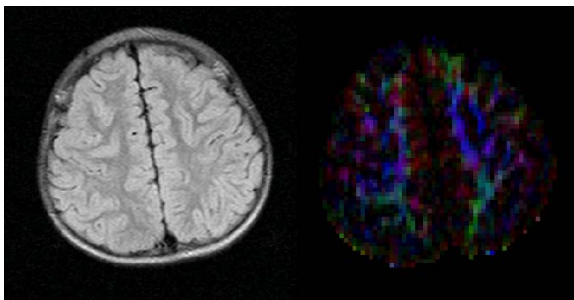


Figure 2. Plot of T2 Turbo spin-echo FLAIR and color-coding FA map on the similar to image level.

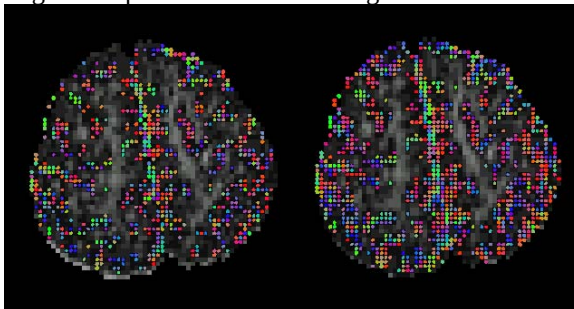


Figure 3. Comparison of Gray matter tensor ellipsoid shape superimposed FA map with MI before(Left) and after(Right).

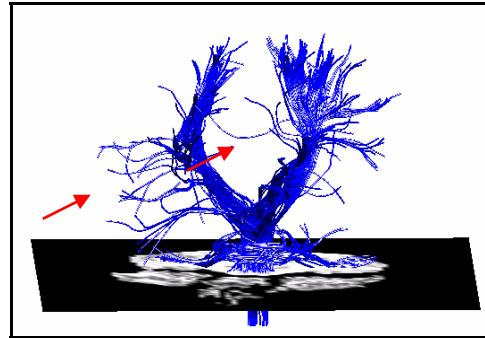


Figure 4. Tractography of Cortico-spinal tracks reconstructed from original Diffusion weighted images.

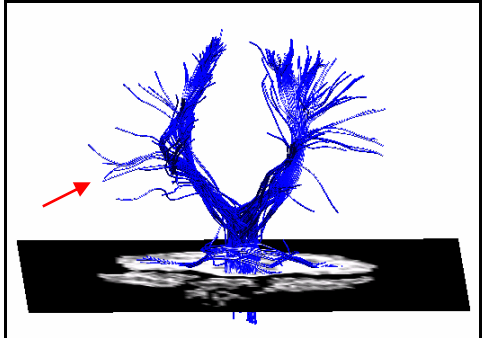


Figure 5. Tractography of Cortico-spinal tracks reconstructed through MI corrected Diffusion-weighted images.

#### 4. Conclusion

Our results indicate that using MI corrected Diffusion-weighted images for White matter Tractography provide alternative calibration approach eddy current induced phase encoding direction geometric distortion.

Despite this method raising postprocessing times, MI coregister reduction extra data acquisition pulse sequences.

#### 5. References

- [1] Yuji Shen, David J. Larkman, Serena Counsell, Ida M. Pu, David Edwards, and Joseph V. Hajnal, "Correction of High-Order Eddy Current Induced Geometric Distortion in Diffusion-Weighted Echo-Planar Images", *MRM*, 2004; 52, p1184-1189.
- [2] Mori S et al, "Diffusion-Tensor MR Imaging and Fiber Tractography: A New Method of Describing Aberrant Fiber Connections in Developmental CNS Anomalies", *RadioGraphics*, 2005; 25, p53-68.
- [3] J.-F. Mangin, C. Poupon, C. Clark, D. Le Bihan, I. Bloch, "Distortion correction and robust tensor estimation for MR diffusion imaging", *Medical Image Analysis*, 2002; 6, p191-198.
- [4] D. Kim, H.-J. Park, W.-J. Moon, E.-C. Chung, I.-Y. Kim, S. Kim, "Comparison of Registration Technologies for the Evaluation of Diffusion Tensor Imaging", *Proc ISMRM*, 2005; 13, p2340.





[5] Mark E. Bastin, "CORRECTION OF EDDY CURRENT-INDUCED ARTEFACTS IN DIFFUSION TENSOR IMAGING USING ITERATIVE CROSS-CORRELATION", MRI, 1999; 17, p1011-1024.

[6] M-C. Chou, M-L. Wu, H-W. Chung, C-Y. Wang, C-Y. Chen, "TENsor Deflection (TEND) Tractography with Sub-pixel Adaptive Step Size", Proc ISMRM, 2005; 13,p1308.



## A Virtual Bronchoscopy with Color-mapped Wall Thickness

Sang Joon Park<sup>2)3)</sup>, Jong Hyo Kim<sup>\*1)2)3)</sup>, Jin Mo Goo<sup>1)3)</sup>, Kwang Gi Kim<sup>1)3)</sup>, Sang Ho Lee<sup>2)3)</sup>

<sup>1)</sup>Department of Radiology, <sup>2)</sup>Department of Radiation Applied Life Science,

<sup>3)</sup>Institute of Radiation Medicine,

Seoul National University College of Medicine, Seoul, Korea

Correspondence: kimjh@radcom.snu.ac.kr

### Abstract

*The tracheo-bronchial tree as part of the lung is part of one of the most important organs of the human body. Even if the recent technical development improved resolution and scan velocity related to computed tomography (CT), it cannot image smaller airways and distally to obstructions, which is a limit in CT images besides real bronchoscopy. In this paper, we demonstrate a new method for wall thickness-mapped virtual bronchoscopy through measurement of airway wall thickness with a proposed algorithm. We present an enhanced virtual bronchoscopy that allows assessment of wall thickness during the virtual fly-through using below steps. Firstly, the airway trees are segmented, and then airway center lines and bifurcation points at each generation are acquired with morphological analysis. Then, we can calculate every wall thickness of each generation of the volumetric airway tree by using a proposed method based on vector analysis with perpendicular plane to airway. Finally, our own software visualizes color-mapped virtual bronchoscopy that reflects airway wall thickness, flying through the inner path of airway. Using CT scans of human cases, we present results showing that this new approach is convenient, accurate, automatic and well-visual in estimating airway wall thickness. Our proposed method can provide pre-operative planning of the intervention for lung surgery with enhanced functional virtual bronchoscopy as a complementary tool for clinical diagnosis and several diseases of airway.*

### 1. Introduction

The primary role of the lungs is gas exchange. The lungs exchange air with the external environment via the pulmonary airways. The tracheo-bronchial tree (airways) as part of the lung is part of one of the most important organs of the human body. The airways are organized as a branching network of tubes that become narrower, shorter, and more numerous as they penetrate deeper into the lung. Air enters the pharynx through the nose and is conducted to the lungs through the trachea. The trachea divides into two main branches, the left and right bronchi, which enter into the right and left lungs. The geometry of the airway tree is approximately bipodial, i.e., at each branch division (a *branchpoint* or *bifurcation*), the

parent airway branches into two child airways. The trachea, commonly defined as airway tree branch generation zero, is the root of the airway tree. As a parent airway splits into two or more child airways, a new generation of the tree is formed. The trachea is typically 1–2 cm in diameter in a normal adult. By generation three of the airway tree, the airways, called bronchioles, range in size from 1 to 5 mm in diameter. The branches from the trachea to the end of terminal bronchioles make up the first 17 generations of the airway tree. Beyond generation 17, the airways continue to branch and lead to the alveolar ducts, and ultimately terminate at the alveolar sacs, the site of gas exchange [1]–[3].

Inhaled air is distributed to the alveolus where oxygen and carbon-dioxide exchange between air and blood takes place. The tracheo-bronchial tree is complemented by a system of pulmonary venous and arterial blood vessels which transports the blood to and from the heart into the lungs.

Several pathologies can jeopardize a sufficient lung function. Among them are tumors, pulmonary embolism, collapse of the lungs (atelectasis), pneumonia, emphysema, asthma, and many more. For a proper diagnosis and treatment, the respective pathologies need to be identified and in some cases quantified. Furthermore, lung surgery requires a pre-operative planning of the intervention. Currently, this is done using a combination of computed tomography (CT) as a tool for morphological imaging of the whole lung parenchyma, and bronchoscopy as an interventional tool for inspection of the central airways and deriving tissue samples. Due to the recent technical development improving resolution and scan velocity, CT might be an promising alternative to bronchoscopy. Even because it images smaller airways and distally to obstructions, which is a limit in real bronchoscopy [4].

Quantitative analyses can be performed on the three-dimensional (3-D) airway tree to evaluate tree structure and function using a CT imaging [5], [6]. Tree geometry can be used to help physicians assess pathological conditions, such as stenoses, bronchiectasis, and tumors. For instance, an unusual reduction in the size of the airway lumen at a particular level in the airway tree may indicate a possible stenosis and an abnormal dilation of the airway represents bronchiectasis. Since an image-based analysis of the airway tree can provide precise airway size information, CT imaging can be used to



guide airway stent design [7]. In addition, unlike bronchoscopic examination of the airways, image-based analysis can provide valuable information about the condition of airways distal to a stenosis. Longitudinal analysis of the 3-D airway tree can also play a role in tracking disease. Changes in a disease process or response to therapy can be tracked over time by using the airway as a “road map” to register regions of interest. Thus, analysis of the 3-D airway structure can provide a more complete understanding of physiologic and patho-physiologic conditions.

In this contribution, we introduce our own developed the virtual bronchoscopy system to perform virtual inspections of tracheo-bronchial trees. It provides an interactive visualization of color-mapped wall thickness through measurement of airway wall thickness. The airway trees are segmented, and then airway center lines and bifurcation points at each generation are acquired with morphological analysis. Then, we can calculate every wall thickness of each generation of the volumetric airway tree by using a proposed method based on vector analysis with perpendicular plane to airway. Finally, our own software visualizes color-mapped virtual bronchoscopy that reflects airway wall thickness, flying through the inner path of airway.

## 2. Materials and Methods

Our proposed methods have been applied to 8 patients, 4 with normal and 4 with COPD (chronic obstructive pulmonary disease) cases. In each patient, 16 channel CT scan was acquired at the maximal inspiration and the dataset included the thorax from the trachea to the diaphragm since our method is focused on the airway. All our implementation and test were performed on an Intel Pentium IV PC containing 3.0 GHz and 1.0 GB of main memory. The image size of all patient datasets is 512×512.

### 2.1 Segmentation

We used 3-D confidence connected region growing (CCRG) method to extract lower and upper airways of the bronchi. This segmentation method is based on the NLM Insight Segmentation and Registration Toolkit (ITK). The algorithm is left to establish a criterion to decide whether a particular pixel should be included in the current region or not, because the most important thing is visiting neighboring pixels. The criterion used by the CCRG is based on simple statistics of the current region. First, algorithm computes the mean and standard deviation of intensity values for all the pixels currently included in the region. A user-provided factor is used to multiply the standard deviation and define a range around the mean. Neighbor pixels whose intensity values fall inside the range are accepted and included in the region. When no more neighbor pixels are found that satisfy the criterion, the algorithm is considered to have finished its first iteration. At this point, the mean and

standard deviation of the intensity levels are recomputed using all the pixels currently included in the region. This mean and standard deviation defines a new intensity range that is used to visit current region neighbors and evaluate whether their intensity falls inside the range. This iterative process is repeated until no more pixels are added or the maximum number of iterations, user defined value, is reached. As the pixel pointer walks, the output segmented image refer to corresponding pixel in the input image. As long as the number of iteration is increased, the pixel is added to the output segmentation and its neighbors become candidates for the pointer to walk.

$$I(X) \in [m - f\sigma, m + f\sigma] \quad (1)$$

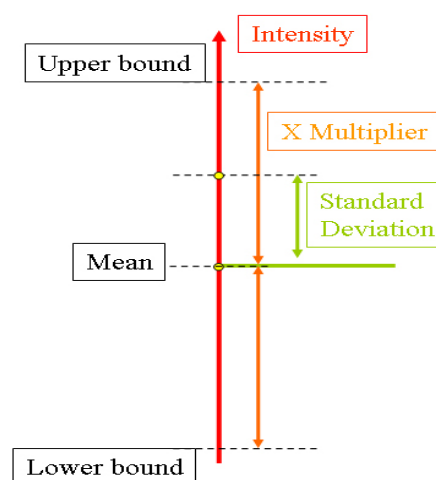


Figure 1. Schema of CCRG

where  $m$  and  $\sigma$  are the mean and standard deviation of the region intensities,  $f$  is a factor defined by the user,  $I()$  is the image and  $X$  is the position of the particular neighbor pixel being considered for inclusion in the region.

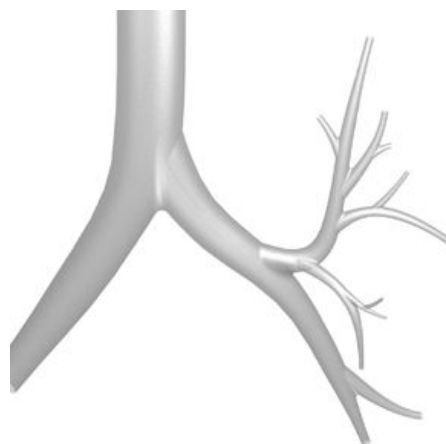


Figure 2. Segmented airway model

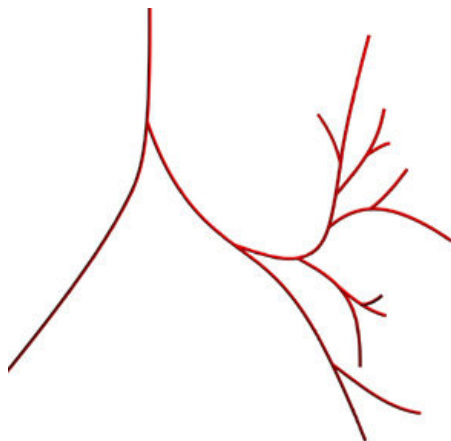


## 2.2 Center line extraction

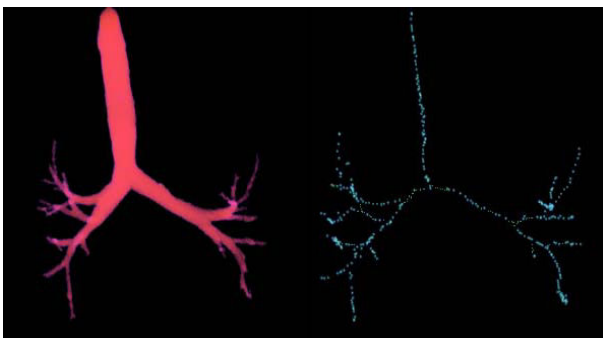
After segmentation of the airways was completed, a 3D skeletonization algorithm [8] was used to obtain the center line of the extensive airway network.

Erosion of the volume layer by layer from the outermost surface was completed using a geometry and connectivity preserving fully-parallel 3D thinning algorithm until only the center line was left. This process preserved the geometry of the original object and preserved important details such as small airway bifurcations. There were some preserving conditions that ensured the algorithm would not delete the end points of airway segments. Preservation of connectivity is very important in 3D algorithms so that large breaks do not occur, which can lead to disconnected airway segments.

This particular algorithm used a set of 38 deleting templates which were created using simple reflections and rotations around four classes of core templates [8]. Each template was applied to every object voxel within the segmented volume and if it fit, then that voxel was marked for deletion. After all voxels that satisfied a particular template were marked, they were deleted if they were not a line-end voxel. Then the next iteration began until no more voxels satisfied a particular template, and thus no more voxels were deleted.



**Figure 3. Extraction of airway center line**



**Figure 4 Segmented airway with CCRG and center line**

## 2.3 Wall thickness measurement

High Resolution X-ray computed tomography (CT) [HRCT] based imaging and quantitative measurement of peripheral bronchial airway geometry provides a valuable tool and information for assessing regional pulmonary physiology and pathophysiology.

Measurements of bronchial tree airway geometry can be used to evaluate and track the progression of disease affecting the airways, such as asthma, COPD and cystic fibrosis, and to assess the efficacy of new therapeutic approaches. Even though, the larger airways have a little effect on the airflow resistance than small airways, thickening of large airways may be a surrogate for small airway abnormalities. It is possible that the same pathophysiologic process, which results in excessive obstruction of small airways, also occurs in the larger airways. Although increased airway wall thickness in the larger airways may have little functional consequence, the ability to measure thickening using CT could prove to be a useful predictor of small airway pathology and may allow for phenotypic stratification of patients who have COPD into parenchymal - and airway- predominant categories [9].

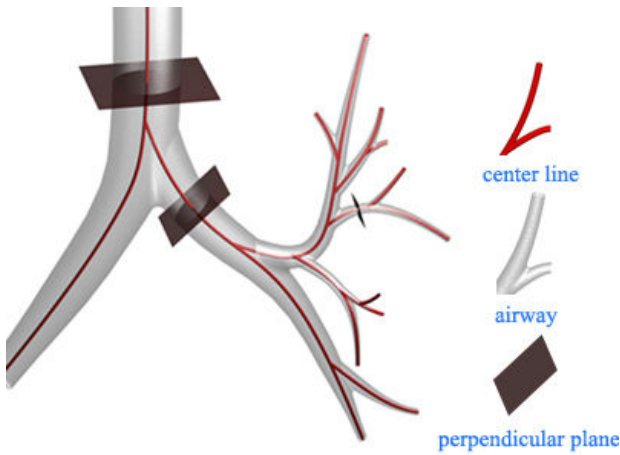
These measurements are used to identify three distinct changes in the respiratory pathways, airway size, obstructions, and airway wall thickness. These changes reflect specifically the morphology of the airways. The changes of airflow resistance can be inferred from airway size and obstructions. And we can predict or infer the pathological process from airway wall thickness.

There has been considerable interest in measuring intrathoracic airways in two or three-dimensional (3-D) sets of slices from volumetric X-ray CT. A popular airway measurement method uses the "halfmax" criteria, in which the gray level at the airway wall is estimated to be halfway between the minimum and maximum gray levels along a ray crossing the edge. However, because the scanning process introduces blurring, the half-max approach may not be applicable across all airway sizes [10]. For this reason, new sensitive and reliable methods for assessing alterations in regional lung structure and function are critically important for the investigation and treatment of pulmonary diseases.

We used a airway measurement via parameter estimation method. With pre-operated results, we applied image plane perpendicularly to airway. Then, we can calculate every wall thickness of each generation of the volumetric airway tree by using measurement via parameter estimation method. In our approach, we examine the gray-level profile of a ray crossing the airway wall and use a maximum-likelihood method to estimate the airway inner and outer radius.



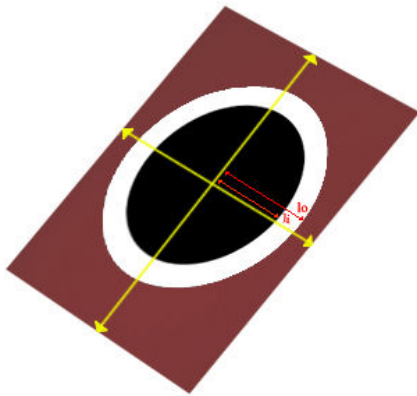




**Figure 5 Quantitative analysis**

2-D slices are re-sampled perpendicular to center line, inner border is detected, cross-sectional area, inner, and outer-radius are computed.

Fig. 6 shows a simple 2-D model of a single slice of an ideal airway cut perpendicular to the airway. The airway centroid is at the origin, and the airway has inner radius  $l_i$  and outer radius  $l_o$  ( $l_i < l_o$ ). The density of the airway wall is unity and the density of the lumen (air) is zero. The airway is surrounded by tissue of constant density (typically,  $0 \leq \alpha < 1$ ), which models the lung parenchyma.



**Figure 6. Ideal 2-D airway model with inner radius  $l_i$  and outer radius  $l_o$**

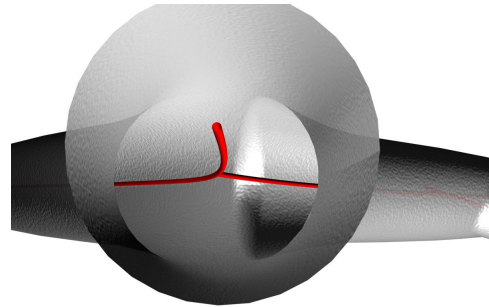
Density of airway wall (white) is unity; density of airway lumen (dark gray) is zero; tissue external to airway (light gray) has mean density  $\alpha$  (typically,  $0 \leq \alpha < 1$ ).

## 2.4 Color-mapped virtual bronchoscopy

Finally, our own software visualizes color-mapped virtual bronchoscopy that reflects airway wall thickness, flying through the inner path of airway.

For the virtual endoscopic exploration, usually we assume a viewpoint inside a bronchi. As usual, for the actual virtual bronchoscopy, we start in the trachea (Fig.

12) and traverse the tracheobronchial tree (Fig. 13 and Fig. 14) to the target region of interest using an obtained path.

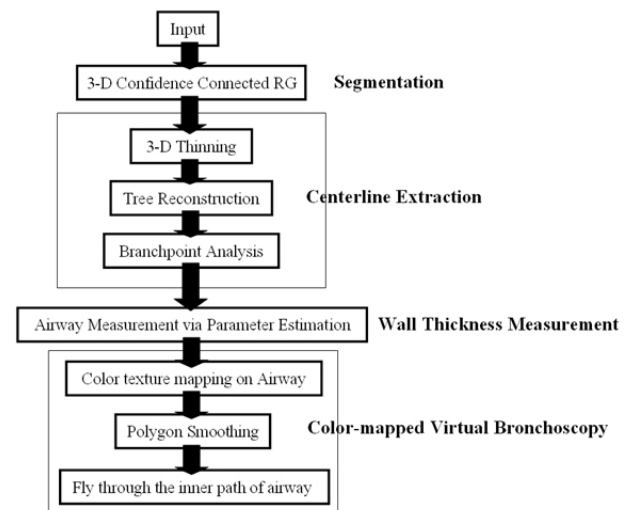


**Figure 7. Model of flying through**

## 3. Results

In general, virtual bronchoscopy is a valuable tool for the localization and measurement of stenosis for treatment planning.

We have described a new method for wall thickness-mapped virtual bronchoscopy through measurement of airway wall thickness with a proposed algorithm for 3-D X-ray CT images of the lungs. We presented an enhanced virtual bronchoscopy that allows assessment of wall thickness during the virtual fly-through using below steps. Firstly, the airway trees are segmented, and then airway center lines and bifurcation points at each generation are acquired with morphological analysis. Then, we can calculate every wall thickness of each generation of the volumetric airway tree by using a proposed method based on vector analysis with perpendicular plane to airway. Finally, our own software visualized color-mapped virtual bronchoscopy that reflects airway wall thickness, flying through the inner path of airway.







## Figure 8. Flowchart of the color-mapped virtual bronchoscopy

### 4. Discussion

Airway tree segmentation in CT images is a challenging problem because of the complex anatomy and the limitations in image quality inherent to CT image acquisition. Despite improvements in image quality accomplished by utilizing thin slices and the high spatial resolution reconstruction algorithm in CT images, if the airway long axis is not perpendicular to the imaging plane, airway segmentation is limited by partial volume effects [11]. Contrast in a CT image is determined by the differential absorption of X-rays by neighboring structures. The airway wall separates the lumen of the airway from the surrounding lung tissue and, thus, we would expect a high contrast border around the lumen. But, if the airway orientation tilts toward the image plane, the density values inside the airway lumen increase due to partial volume averaging of air and airway wall. The apparent airway lumen density increases substantially with decreasing airway diameter due to partial volume effects as well. These partial volume effects reduce the contrast between the wall and lumen, making airway wall detection difficult.

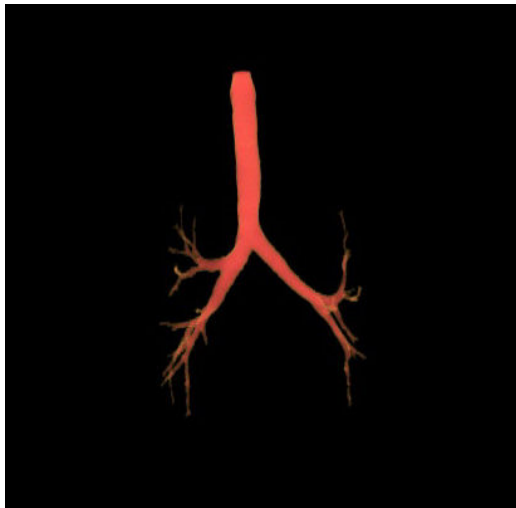
### 5. Conclusion

Using CT scans of human cases, we presented results showing that this new approach is convenient, accurate, automatic and well-visual in estimating airway wall thickness. Our proposed method can provide pre-operative planning of the intervention for lung surgery with enhanced functional virtual bronchoscopy as a complementary tool for clinical diagnosis and several airway diseases.

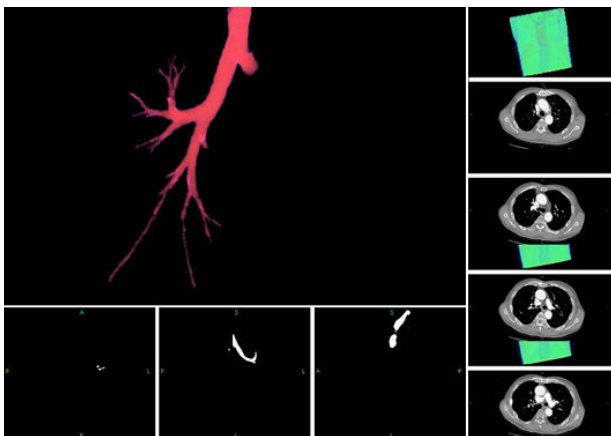
### 6. References

- [1] M. G. Levitzky, *Pulmonary Physiology*. New York: McGraw-Hill, 1999.
- [2] J. B. West, *Respiratory Physiology. The Essentials*, 4th edition. Baltimore, MD: Williams and Wilkins, 1990.
- [3] J. F. Murray, *The Normal Lung*. Philadelphia, PA: Saunders, 1986.
- [4] D. Mayer and D. Bartz and S. Ley and S. Thust and C. Heussel and H. Kauczor and W. Straßer, "Segmentation and Virtual Exploration of Tracheo-Bronchial Trees", in: *Proc. of Computer Assisted Radiology and Surgery*, 2003, pp. 35-40.
- [5] W. E. Higgins, K. Ramaswamy, R. D. Swift, G. McLennan, and E. A. Hoffman, "Virtual bronchoscopy for three-dimensional pulmonary image assessment: State of the art and feature needs," presented at the RSNA Annu. Meeting, Chicago, IL, 1998, pp. 761-778.
- [6] E. A. Hoffman and G. McLennan, "Assessment of the pulmonary structure—Function relationship and clinical outcomes measures: Quantitative volumetric CT of the lung," *Acad. Radiol.*, 1997, vol. 4, no. 11, pp. 758–776.
- [7] G. McLennan, S. Shamsolkottabi, and E. A. Hoffman, "Assessment of major airway obstruction using image analysis of digital CT information," *Proc. SPIE*, 1996, vol. 2709, pp. 197–208.
- [8] Ma C M and Sonka M, "A fully parallel 3D thinning algorithm and its applications *Comput. Vis. Image Underst.*, 1996, Vol. 64, no. 3, pp. 420–33
- [9] Nakano Y, Wong JC, de Jong PA, et al. The prediction of small airway dimensions using computed tomography. *Am J Respir Crit Care Med* 2005;171:142–146.
- [10] Reinhardt, J.M., N.D. Dsouza, and E.A. Hoffman, "Accurate measurement of intrathoracic airways", *IEEE Trans. Med. Imag.* 1997, pp. 820-827
- [11] J. M. Reinhardt, N. D. D'Souza, and E. A. Hoffman, "Accurate measurement of intra-thoracic airways," *IEEE Trans. Med. Imag.*, 1997, vol. 16, pp. 820–827, Dec.





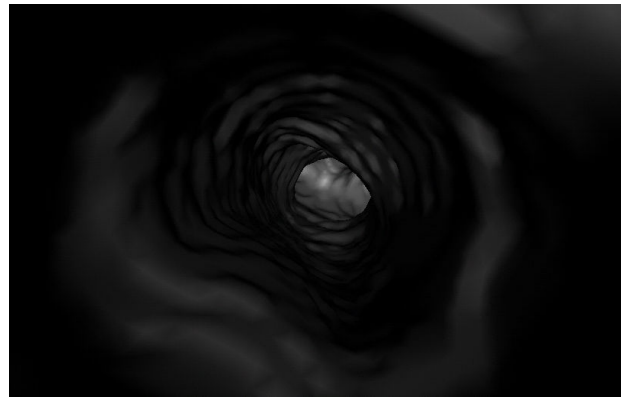
**Figure 9. Segmented airway with CCRG in normal case**



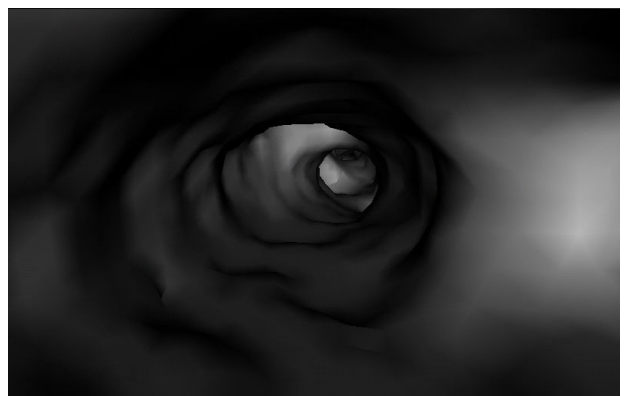
**Figure 10. Segmented airway with CCRG in COPD case**



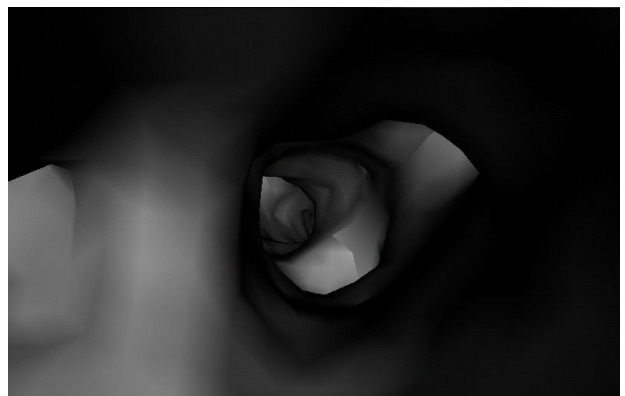
**Figure 11. Airway using geometric rendering**



**Figure 12. Virtual bronchoscopy**  
View in trachea down to main bifurcation.



**Figure 13. Virtual bronchoscopy**  
Lower airways



**Figure 14. Virtual bronchoscopy**  
Bifurcation of the bronchioles

# K-means clustering approach for kinetic pattern analysis of dynamic contrast enhancement breast MRI

Sang Ho Lee<sup>1,3</sup>, Jong Hyo Kim<sup>\*1,2,3</sup>, Kwang Gi Kim<sup>2,3</sup>, Sang Joon Park<sup>1,3</sup>, Woo Kyung Moon<sup>2,3</sup>

<sup>1</sup>Interdisciplinary Program in Radiation Applied Life Science major,

<sup>2</sup>Department of Radiology,

<sup>3</sup>Institute of Radiation Medicine,

Seoul National University College of Medicine, Korea

Correspondence: kimjh@radcom.snu.ac.kr

## Abstract

*This study introduces k-means clustering approach to automatic classification of kinetic patterns on 3-D contrast-enhanced breast MR images. Although BI-RADS criteria are widely accepted for breast MRI diagnosis, it has the limitation not to be able to consider intermediate contrast-enhancement patterns. We performed k-means cluster analysis technique that enables to effectively classify the divers enhancement and signal patterns from real world tumor cases into a reasonable number of reference pattern set that is representative of each cluster. This technique, then, detects and classifies the tumor specific patterns from the given MRI data by measuring the vector distances from the reference pattern set on each voxel. The detections are highlighted with colors reflecting likelihood of malignancy and the type of reference pattern on a 3-D volume display. This method will greatly enhance the radiologists' capability to efficiently identify and characterize the multi-dimensional patterns from tumors on DCE breast MRI.*

## 1. Introduction

Magnetic Resonance Imaging (MRI) in combination with T1 enhancement from gadolinium diethylene triamine pentaacetic acid (Gd-DTPA) [1] is a valuable imaging modality because of its high contrast sensitivity to a large number of characteristics of tissues and body fluids, including T1 relaxation times. Dynamic Contrast Enhanced MRI (DCE-MRI) is necessary to detect breast lesions, and to help differentiate malignant from benign tumors [2, 3]. Malignant breast lesions tend to enhance more rapidly than benign lesions. Observation of contrast enhancement is typically achieved using dynamic imaging techniques whereby the contrast agent is injected during acquisition of a dynamic image set.

The American College of Radiology notes that there are rare situations when traditional imaging modalities are unable to guide patient management. These include situations involving inconclusive or contradictory results on mammography or ultrasound and the search for an unknown primary tumor. Breast MRI may help to solve

these problems before definitive treatment.

The basis of DCE-MRI is the fact that tumors need many blood vessels to grow, the concentration of the contrast agent at their location will be higher than in surrounding normal tissues and they will consequently appear as brighter areas in the images. Their dynamic and morphological patterns are related to the diffusion of the contrast agent and to the shape, edges, or the internal pattern of the enhancing region [4]. However, there is a considerable overlap between dynamic and morphological patterns of benign and malignant lesions. In this paper, we will study dynamic features only.

A DCE-MRI requires the acquisition of one series of images before the injection of the contrast agent, called pre-contrast series, and of several series of images, after injection, called post-contrast series. A DCE-MRI examinations produce one hundred images or so, which have to be analyzed by ones.

American college of Radiology (ACR) Breast Imaging and Reporting Data System (BI-RADS) [5] describe the slope from pre-contrast phase to the initial enhancement phase, and the curve change in the delayed phases for kinetic curve assessment. However, BI-RADS criteria do not consider in detail the intermediate enhancement patterns in the delayed phases. Thus we performed a cluster analysis technique as a part of the role of a computer-aided diagnosis (CAD), for providing the radiologist a second opinion to take into account in difficult cases. Our clustering algorithm was a trial to explore features which are not directly human readable, reaching higher level of accuracy.

In this study, we focused on the differentiation of malignant tumor patterns extracted from DCE-MRI examinations and presented the result of k-means clustering analysis applied to the problem of classification of dynamic patterns.

## 2. Materials and Methods

### 2.1. Materials

Dynamic breast MR images from 13 patients with malignant lesions were acquired through 1.5 T Sonata (Siemens, Erlangen, Germany). First, pre-contrast T1



weighted three dimensional fast low angle shot (3D FLASH) sagittal image was obtained with fat suppression, and four consecutive post-contrast image using the same condition after the injection of 0.1 mmol/kg Gadolinium-DTPA (Magnevist, Schering, Berlin, Germany), respectively. Standard subtraction image, the image resulted after the first post-contrast image subtracts the pre-contrast image, was acquired as post-processing image.

## 2.2. Data acquisition

Signal patterns were collected on a voxel by voxel from the enhanced malignant tumor area for each patient. The total number of the signal patterns was 1734 and these signals were used as the training data for the classification of the malignant signal patterns.

## 2.3. K-means clustering

The  $k$ -means clustering is a popular method used to divide  $n$  patterns  $\{x_1, \dots, x_n\}$  in  $d$  dimensional space into  $k$  clusters. The result is a set of  $k$  centers, each of which is located at the centroid of the partitioned dataset. This algorithm can be summarized in the following steps given as Figure 1:

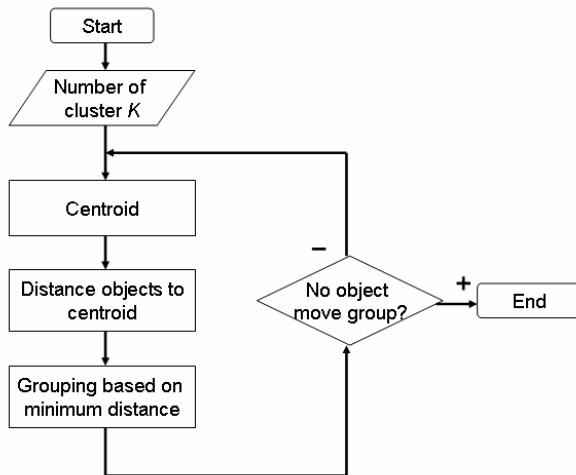


Figure 1. Steps of  $k$ -means algorithm

- (1) Choose the number of clusters  $k$  and input a dataset of  $n$  patterns  $X = \{x_1, \dots, x_n\}$ . Randomly select the initial candidates for  $k$  cluster centers matrix  $V^{(0)}$  from the dataset.
- (2) Assign each pattern to the nearest cluster using a distance measure. For each pattern  $x_i$ , compute its membership function  $m(C_j | x_i)$  in each cluster  $C_j$ . The membership function  $m(C_j | x_i)$  defines the proportion of pattern  $x_i$  that belongs to the  $j$ th cluster  $C_j$ . The  $k$ -means algorithm uses a hard membership function, that is the membership  $m(C_j | x_i) \in \{0, 1\}$ . If the pattern  $x_i$  is closest to cluster  $C_j$  (i.e., the distance between  $x_i$  and cluster center  $v_j$  is minimal), then  $m(C_j | x_i) = 1$ ; otherwise  $m(C_j | x_i) = 0$ .

- (3) Recompute the centroids (centers) of these  $k$  clusters to find new cluster centers  $v_j$ , and compute the sum of square error  $E$ .

$$v_j = \frac{\sum_{i=1}^n m(C_j | x_i) x_i}{\sum_{i=1}^n m(C_j | x_i)} \text{ for } j = 1, \dots, k. \quad (1)$$

$$E = \sum_{j=1}^k \sum_{x_i \in C_j} \|x_i - v_j\|^2 \text{ for } i=1, \dots, n; j=1, \dots, k. \quad (2)$$

- (4) Repeat step 2 and 3 until convergence. Typical convergence criteria are: no more reassignment of patterns to new clusters, the change in error function  $E$  fails below a threshold, or a predetermined number of iterations have been reached.

In  $k$ -means clustering, choosing a good set of initial cluster centers is very important to reduce the distance calculations. However, it is difficult to select a good set of initial clusters randomly. Our method used as initial clusters the centroids which are empirically acquired by training of dataset. Table 1 shows the notation used in describing the algorithm. Figure 2 shows the 10 centroids computed finally through  $k$ -means clustering analysis.

Table 1. Notation used in our algorithm

|       |                                     |
|-------|-------------------------------------|
| $k$   | The number of clusters              |
| $n$   | The number of patterns in a dataset |
| $x_i$ | The $i$ th data element (pattern)   |
| $d$   | The number of dimensions            |
| $v_j$ | The centroid of the $j$ th cluster  |
| $C_j$ | The $j$ th cluster                  |

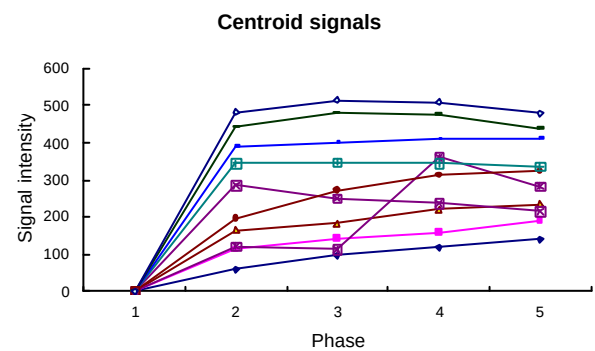
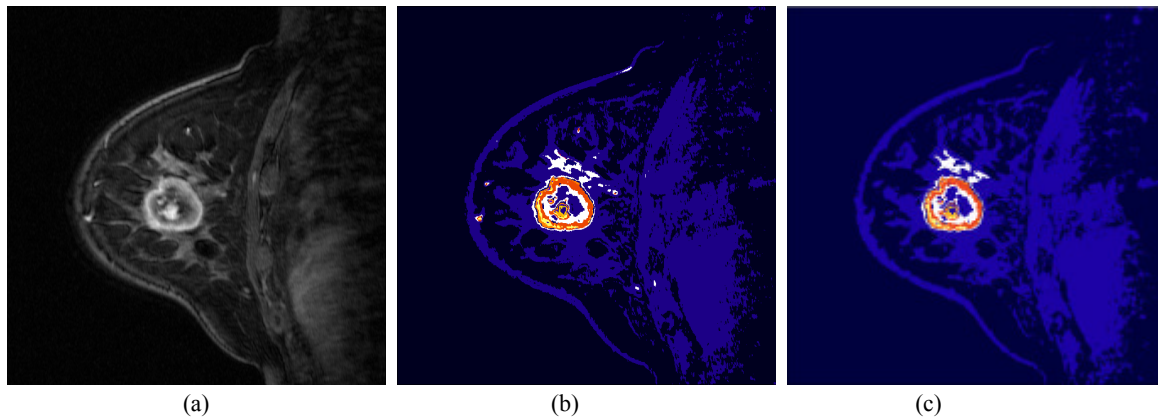


Figure 2. Ten centroid signals computed from 1734 kinetic curves in malignant tumors

## 2.4. Classified image mapping

The centroids of the signal patterns computed through  $k$ -means clustering analysis were used for classifying a random dynamic breast MR image with a malignant lesion. The criterion of classification is to



**Figure 3. An example of multi-dimensional detection on dynamic breast MR images of a patient with a single breast cancer;** (a) the DCE-MR image in the initial phase, (b) the classification result from k-means clustering, (c) the result after removing false positives

assign each pattern extracted from the MR image to the nearest reference centroid using distance measure. In this study, the number of clusters,  $k$  was 10. The voxels corresponding to the position of each signal patterns were labeled as its class name and mapped to an image.

### 2.5. Removal of false positives

The signals as the cause of the false positives resulted from our clustering algorithm can be the artery signals similar to the kinetic patterns of typical malignant tumors with rapid initial enhancement and washout delayed phases, or the signals of normal tissues resembling the benign-like kinetic patterns, that is, the persisted enhancement patterns of untypical malignant tumors. Thus, we analyzed the size and the circularity of the voxels corresponding to the false positives using ImageJ free software (National Institute of Health, USA) and adequately removed them by thresholding.

### 2.6. Development environment

In this study, the C++ programming language and Insight Segmentation and Registration Toolkit (ITK) open-source software system sponsored by the National Library of Medicine were used for image analysis, and ImageJ software for the removal of false positives.

## 3. Results

Our classification results were displayed on an image as the texture with different colors for each class of malignant signal patterns. The various classes within a malignant tumor could be visually identified. The use of the more subdivided classification than the existing BI-RADS criteria allowed radiologist the possible to be able to provide the information of more detailed cancer diagnosis. Figure 3 shows an example detected through our classification method on dynamic breast MR images of a patient with a single breast cancer.

## 4. Discussion and Conclusion

In this study, we presented the experimentation of  $k$ -means classifier based on automatic detection of tumor specific patterns, considering the intermediate kinetic patterns in the delayed phases. Such experimentation strategy can discriminate the heterogeneous patterns within malignant tumor and characterize the multi-dimensional patterns in detail, but still fail, when malignant lesions present atypical kinetic patterns similar to the signal patterns of benign tumors, as the case may be. Such observation confirms that, by itself, kinetic patterns are not able to distinguish correctly between benign and malignant lesions, and that morphological analysis of various specific tumors is to be combined.

There are several ways we plan our future work:

- (1) The short-term goal is to enrich the set of kinetic features used by the classifiers and try to improve the classification of contrast-enhancement patterns. In addition, other advanced classifiers need to be tested.
- (2) The long-term goal is to introduce morphological analysis. The combination with morphological analysis makes possible the comparison with the fully ranked classification through BI-RADS criteria.

## 5. References

- [1] A.D. Watson, S.M. Rocklage and M.J. Carvlin. *Magnetic Resonance Imaging*. St. Louis, USA, 1992.
- [2] S.H. Heywang, D. Hahn, H. Schmidt et al, "MR imaging of the breast using Gadolinium-DTPA." *J. Comp. Ass. Tomography*, 1986; 10, pp. 199-204.
- [3] W.A. Kaiser and E. Zeitler, "MR imaging of the breast: fast imaging sequences with and without Gd-DTPA." *Radiology*, 1989; 170, pp. 681-686.
- [4] C.K. Kuhl and H.H. Schild, "Dynamic image



interpretation of MRI of the breast.” J Magn Reson Imaging, 2000; 12, pp. 965-974.  
[5] ACR BI-RADS-MRI first edition American College of Radiology, 2003.

# A study on the merits of Biometrics to perform e-Health

Yasumitsu Tomioka<sup>1,2</sup>, Toshihiko Kitano<sup>1</sup>, Naoki Aida<sup>1</sup>, Kei Kakehi<sup>1</sup>,

M.A.Sadiq<sup>2</sup>, Hiroshi Juzoji<sup>2</sup>, Isao Nakajima<sup>2</sup>

<sup>1</sup>Office Tomioka Inc.,

<sup>2</sup>Tokai University School of Medicine

## Abstract

The merits of biometrics to perform e-Health are discussed in this paper. We feel that the optical sensor (fingerprint, finger-vein, face, iris, and retina ) using the videophone system is a promising biometrics system for e-Health, e.g. telemedicine. We propose the inclusion of optimized biometric methodologies and guidelines as a theme of ITU's E-Health standardization.

## 1. Objectives

The merits of biometrics to perform e-Health, e.g. telemedicine, are discussed in this paper.

We propose the inclusion of optimized biometric methodologies and guidelines as a theme of ITU's E-Health standardization.

## 2. Background

### 2-1. What is biometrics?

Biometrics is an individual identification technology utilizing the unique biological characteristics of humans, and is typically used to identify the owner of forms of identification such as cash cards, driver's licenses, and health insurance ID cards. The conventional technique is to compile information on individuals in the form of a database and then identify individuals using a PIN, password, and ID cards based on the stored information. This technique, however, is vulnerable to illegal acts such as forgery, theft, tapping, and identity theft.

The biometric method measures biological objects having the following three characteristics:

1. Universality (characteristics common to everyone)
2. Uniqueness (characteristics unique to each individual)
3. Permanence (characteristics that do not change over time)

Such individual characteristics--well-known examples include fingerprints, the retina, iris, voice, hand shape,

palm print, and face--are stored in a database for identification.

This technique simply analyzes the biological characteristics that are expressed using images and signals, and estimates the probability of pattern matching with the data in the database. As a result, the output is only a positive recognition or negative recognition, as shown in Fig. 1; note that the matching probability is not 100%.

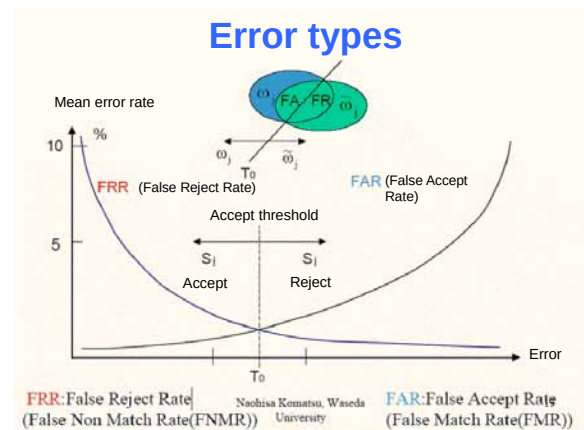


Figure 1. Error types

### 2-2. Market

Increasing global concerns about terrorist activity and identity have boosted the demand for improved security. Increased media attention has raised end users' awareness of the potential of biometric technologies in security applications. Significant market opportunities are thus emerging in application areas such as physical access control, citizen identity, network security, financial services, entitlement and healthcare. For example, InfoShop-Japan reports that the world market size for access management (keyless entry) alone was USD13 million in 2002 and will reach USD70 million in 2007(Fig. 2). As shown in Fig. 3, 50% of 2004 biometric products are fingerprint sensors. Non-contact sensors are expected to increase in popularity in the near future. If they are used in the areas of telehome care and geriatric care, the market may explode in advanced countries.



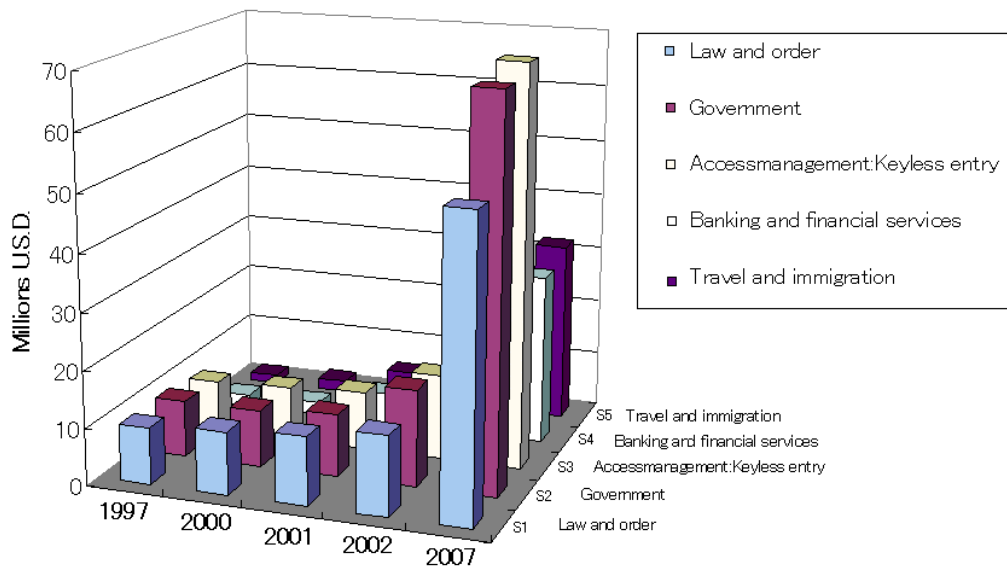


Figure 2 The World Market for Biometrics

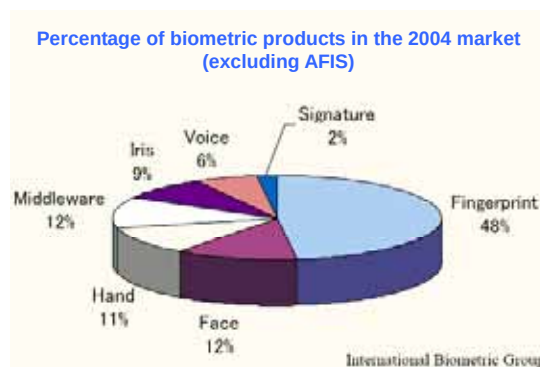


Figure 3 Percentage of biometric products

### 2-3. Necessity of biometrics in telemedicine

With the increasing importance of the comprehensive management of medical information in our aging society, individual identification is essential not only for telemedicine but for information management for individual information browsing and making insurance payments. Japan is facing the problem of an increase in the number of people who do not pay health insurance fees. As the number of clients who attempt to receive medical services in an illegal manner through identification theft is expected to grow, the importance of individual identification will be enhanced.

We have identified the following needs in telemedicine for individual identification:

1. Health insurance ID certificate (or ID card in some countries)
2. Remote consulting
3. Emergency transport
4. Health check (except for diseases)

### 5. Return of medical fees

### 6. Client chart browsing (for the client only)

The above six applications require identification accuracy as high as that needed for the online banking system.

## 3. Proposal

We feel that the following optical sensor using the videophone system is a promising biometrics system for telehomecare.

### 3-1. Face

#### A Methodology:

In telemedicine, it is possible to monitor the face of a client by using a videophone system. Since there is no compulsory or forcible feeling for the client to use the videophone system, it would be acceptable for the client to use the videophone system rather than to use the other system or applications. In the case wherein





the simple videophone system is used, it would be indispensable to develop the following technical elements. Generally in telemedicine, since the client stands so as to face the front surface of the client toward the camera of videophone, an accuracy of individual identification is relatively high by contriving the direction of lighting optimally, compared with the face identification system used in a system for administration of entering into and/or exiting from a building.

#### B Cut-off and detection of face:

It is necessary to maintain a distance between the client and the camera of videophone. It should be noted that the brightness of image is widely altered corresponding to age, gender, race and size of the client. By the general methodology, a pyramidal image was produced at first, and then a NN (Neural Network) which is common to different face size was detected from the pyramidal image.

NN methodology is sorted into two learning method, one is so-called “learning with teacher” in which the problem is gradually optimized by inputting the teacher signals (correction) and the other is so-called “learning without teacher” in which it is unnecessary to use the teacher signals. In a case where the definite solution is prepared, the “learning with teacher” is utilized, while in a case of data-clustering, the “learning without teacher” is utilized.

As a result, the dimension of the solution is reduced even any NN methodology is selected. Thus, it would be able to get a good solution by employing a relatively small amount of calculation, even though the problem involves the multi-dimensional data such as “images” and “statistics” and is not able to apply a “linear-analysis”. Accordingly, the NN methodology is widely applied to not only pattern-matching procedure and data-mining procedure but also several fields. It should be noted that, in the face certification, it is possible to resolve in high-speed by employing NN methodology.

C Face characteristics detection: In the face characteristics detection, both of “Gabor-Kernel wavelet” and “Graphic matching” methods are employed. In the “Gabor-Kernel wavelet”, as mentioned later in iris identification, since a directionality edge and a selectivity of frequency component are good, the accuracy of the edge extraction from the image is enhanced. Thus, it is possible to certainly grasp the characteristic structures on the face, such as eye, nose and mouth.

In graphic matching method, at first, 10 to 20 characteristic points in the structures of the face were previously extracted, and then these characteristic points are found out by matching.

Several different methodologies are devised and by using the opened face data-base, each of methodologies is evaluated.

#### D De-interlace:

In non de-interlacing, each of odd field and even field on an image is reduced to 1/30, while in de-interlacing to 1/60, as the scanning lines are into halves. Because of adding the odd field and even field, striped noise is produced. In order to eliminate the striped noise, it is necessary to mathematically execute the de-interlacing to the even field as well as to the odd field, after taking the striped noise by digital data. If it is not possible to execute the de-interlacing, either one of odd field and even field should be employed. In the case where either one is employed, the quality of the image would be deteriorated by 3dB in the longitudinal direction.

#### E Optical specification:

In a case where the telemedicine is executed using the face identification in future, as shown in Fig. 4, the client should stand in front of the camera of videophone, by radiated the lights from both sides of the camera, in order to erase the shadow from the face of client. By doing so, it would be possible to capture the principal constructions on the face, such as eye, nose and mouth.

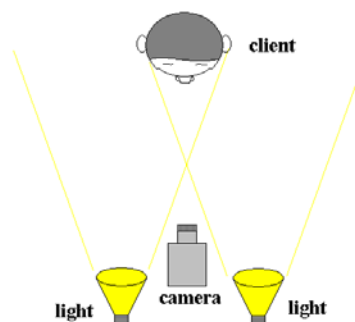


Figure 4 Face identification





### 3-2. Iris

**A Methodology:** Individuals can be identified based on the iris pattern, which is unique to each person. The iris pattern is formed within six months after birth and stabilizes within one year after birth, and does not change until death. Two devices can be used to measure the iris pattern, as follows.

1. If the videophone system is equipped with a zooming function, it can work as an iris identification device.
2. A dedicated light-reaction device can measure the speed of light reaction and examine the iris pattern simultaneously.

In the identification process, the analysis zone is

selected from the iris pattern sampled using the CCD, its profile is extracted using a Gabor filter and then coded. This code is used to select the closest iris code in the database based on the Hamming distance value. The error rate is  $2^{-244}$ , the same probability of a coin toss ending with the same result 244 times in a row.

**B Advantages:** Non-contact, easy to use, and accurate. It would be very difficult to fake an iris, even with advanced plastic surgery and eye operation techniques. In combination with the light-reaction device, it can measure the iris reaction speed (which aids in diagnosis of problems with the brain stem and oculomotor nerve) and identify the iris pattern simultaneously.

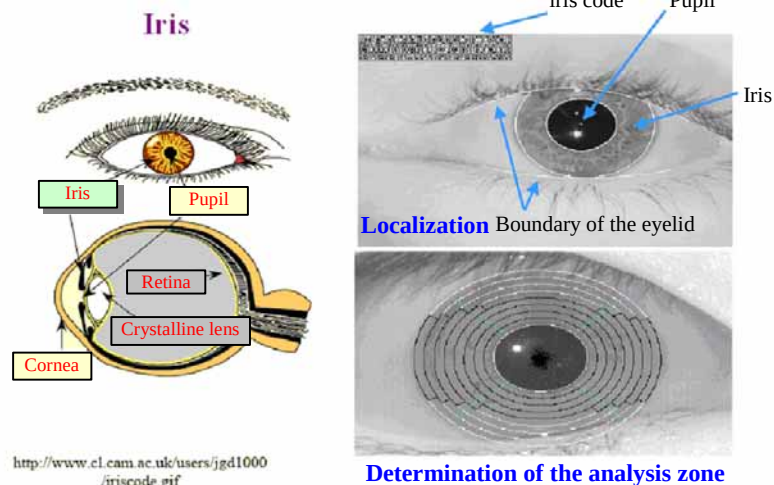


Figure 5 Analysis of the iris

### 3-3. Retina

**A Principle:** The vessels pattern of the retina can be used for identification. This vessels pattern remains unchanged unless the veins shrink due to diabetes or other illness. It is normally difficult to see the vein pattern without expanding the iris using chemicals (e.g., Mydrin). In recent years, however, a device has been developed that can see the vein pattern without the use of iris-opener chemicals. As dedicated equipment is necessary in both cases, it is not possible to conduct retina identification using a normal

videophone system alone.

**B Advantages:** Non-contact and accurate, but not easy to use. The biological origin of the retina is the same as that of the brain. In fact, it is possible to estimate the condition of the blood in the brain from that in the retina, and the brain pressure from the optic papilla on the retina. Some easy-to-use devices are expected to be developed in the future. We should focus on the retina device as a promising telehomecare device.







Figure 6 The pattern of retinal blood vessels  
(Artery and Vein)

#### 4. Prospects

##### 4-1. Reluctance to use fingerprints

The U.S. Visit Program employed by the immigration authorities in 2004, which uses fingerprints for individual identification, has been poorly evaluated by those visiting America. This is due to the fact that this FBI technique, known as AFIS (Automated Fingerprint Identification System), has been used to identify criminals in the United States. Many people may not accept the use of such a criminal identification system for totally different medical use. Most users may prefer non-contact-type biometric techniques that do not use fingerprints. Diagnosis using the videophone system provides information on the client's physical characteristics. Thus, it is natural to use such physical information for user identification.

##### 4-2. Three accuracy levels

A number of biometrics companies have developed a variety of identification techniques using fingerprints, vein patterns of the palm, face, iris, and retina characteristics. These techniques are completely unrelated to each other.

Although more than one algorithm and measurement device are available, the identification result can only be one for each individual. As in the case of mathematics, there is only one correct answer. If the identification accuracy is not sufficiently high, users will be inconvenienced. Unless an identification system is shared by many users, as many sensors and devices as the number of such systems will be necessary. This is unreasonable for users. Since the 9/11 terrorist attacks, homeland security has been focused on, so it is currently difficult to establish an international standard for individual identification. However, at least objective accuracy guidelines should be considered. We feel that the following accuracy

evaluation should be included in such guidelines.

1. Accuracy evaluation of the matching algorithm: Numeric degree of pattern matching between two input images
2. Accuracy evaluation of the matching device: Accuracy of hardware, including sensors
3. Accuracy evaluation of identity matching: Accuracy of identification of the correct person

Without these three accuracy evaluations, no identification technique can be objectively evaluated. Objective evaluation is the first step toward international standardization.

##### 4-3. Efforts for standardization

The standardization of biometrics is underway in ITU-T SG-17 Telebiometrics and ISO/IEC SC37/SC17/SC27. However, there is no forum for discussing applications to telemedicine. If E-Health is standardized by ITU-T SG-16 Q-28 in 2004-2007, discussions on the standardization of biometrics for telemedicine should be kicked off as early as possible jointly with SG-17. As mentioned above, the medical needs are diverse and the telemedicine market is growing faster than expected. We feel that biometrics for telemedicine should be included in the ITU scope. If we lose this opportunity, the importance of telemedicine standardization by ITU will fall drastically. This is due to the fact that the ICT technology will involve the elderly, who often do not go out for medical care, and to the fact that the biometric market is expected to explode. If there is no international standard, the market may be driven by a de facto standard. Although de facto standards are acceptable for applications in a free market, they may cause many problems in medical applications that should be under national control. As a result of such problems, many developing countries and the weak and elderly will be left behind. Therefore, ITU is required if standardization is to be achieved in a timely manner.

#### References

01. The World Market for Biometrics Equipment, [http://www.infoshop-japan.com/study/iz17629\\_biometrics\\_equipment\\_toc.html](http://www.infoshop-japan.com/study/iz17629_biometrics_equipment_toc.html)
02. Task Group M1.4. . Scenario Testing and Reporting Draft 3.0. Biometric Performance Testing and Reporting Part 3: INCITS M1/03-0563, INCITS 1602-D
03. Jafar M. H. Ali Aboul Ella Hassanien. An Iris Recognition System to Enhance E-security Environment Based on Wavelet Theory, AMO - Advanced Modeling and Optimization, Volume 5, Number 2, 2003.





# Development of Portable Real-time Instrument for Three-Dimensional Measurement of Scoliosis : Three-dimensional curve of spinous process

Chung pao Lin<sup>1</sup>

Chung huang Yu

Institute of Rehabilitation Science and Technology

National Yang-Ming University

<sup>1</sup>Correspondence: elemanpao@gmail.com

## Abstract

*The purpose of this study is to develop a scoliosis diagnostic instrument which is mainly designed for schools or rehabilitation settings. It provides a portable and an instant three dimensional spine curvature image unlike the traditional diagnostic method. This diagnostic instrument is non radioactive and able to measure the changes in the three dimensional spine curvatures.*

*Using simple mathematical formula to calculate algorithms of space coordinate, which improves the present traditional spine detection method, and utilize the computer to automatically estimate the severity of the scoliosis and curvature type. This system use palpation, two Webcams to capture two parallel images of "coordinate converter" in the space. Collocate digital image processing, and system of two- dimensional image is transferred to the three-dimensional space coordinate, algorithm of artificial neural networks, for calculating the coordinates of the three-dimension spinal space. Link the line by constructing the curvilinear motion of three-dimensional spinal space. The changes in the three-dimensional space curve allow the estimation of the spinal deformities, such as scoliosis, humpback, and rotation. The contours of three-dimensional space, is from the algorithm calculations and then interpret the severity*

*or type of the resulted spinal curvature.*

*Keyword: scoliosis; Cobb angle; artificial neural network ; reverse engineering*

## 1. Introduction

Scoliosis is the most common form of abnormal spinal deformity found in adolescence. This condition affects between 2% and 4% of adolescents, and between 70% and 80% of these cases have an unknown cause. Scoliosis is a deformity of the spine that involves abnormal lateral curvature and rotation of the spine in a three dimensional space and usually causes noticeable trunk distortions [1]. When one becomes scoliosis, the spine often deforms in the shape of letter "S" or "C," and it causes oppression of inner organs such as lungs and a heart resulting in heavy lumber pain or fatigue. In general, patients with curve of less than 30° are untreated. Patients with curves between 30° and 44° who are still growing may be treated with orthotic bracing. The main goal of orthotic bracing is to stop progression of deformity [2]. Patients with curves of 45° or more are likely to require surgical intervention [1]. At the moment, there are several different kinds of methods are used to measure scoliosis. Forward-bending



test is a painless examination traditionally performed on children, but this test is neither reproductive nor objective. Moreover, the inspection takes up a lot of time when applied to medical examination in schools. Radiographs are the traditional tools for assessing idiopathic scoliosis. However, there is increasing concern about the carcinogenic risks associated with them.

Radiographic methods measure only the distortion of the spine in a single plane and a second radiograph is required to extract three-dimensional information. Several research utilized motion analysis to stick marker onto the body both sides of joint. Utilized VICON motion analysis tool to build and construct whole body model, and distinguish the severity of scoliosis [3]. The researcher utilizes the motion analysis and force plate to study the severity and analyses the dependence with curved center and side [4]. But these methods must all be examined in motion analysis laboratory, which is a very inconvenient way to examine scoliosis. Imaging techniques have been used to help to identify and interpret the scoliosis deformity [5-9]. Moire' topography, the integrated shape imaging system (ISIS) and various optoelectronic and magnetic techniques have been employed to improve the three dimensional visualization of the spine. These techniques have focused primarily on the resultant back surface shape and rib hump geometry, and have not effectively described the global effect of the scoliotic spine in standing posture.

Reviews of above-mentioned discussions, the traditional methods are unable to examine the objective quantization of data clearly. Traditional radiographic increases carcinogenic risk and single one is unable to present three-dimensional coordinate of space [10]. Techniques of digital image processing are unable to describe the object surface form and lack perfect degree, and needing to algorithm complicatedly. There are some disadvantages: the high expense in fabric cost, difficult to operate, and it needs specialized expertise.

So, in this study, the detection method of scoliosis provides a simple and easy way with low cost to solve the above mentioned disadvantages of the current detection methods. This study can be measure curve of spinous and process quickly in three-dimensional space instantly.

## 2. Methods

In this research, the detection method of scoliosis can be divided into three parts: (1) Obtain the spinal space position, (2) Setting-up and analysis of the spinal curvature, and (3) Analysis and diagnosis of the clinical meaning of the curve.

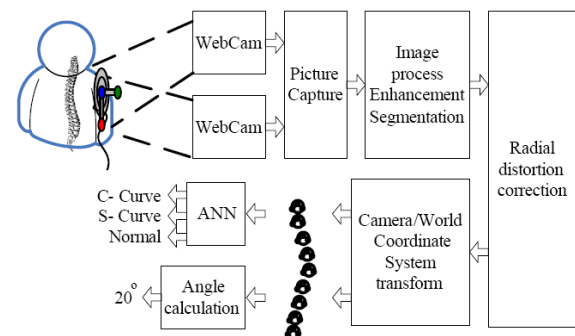


Figure 1. System flow sheet

### 2.1 Obtain the spinal space position

Utilize ' Three-dimensional space coordinate device ' made by oneself, by ways of palpation (Fig2.) Look for some characteristic positions of the spine from cervical vertebrae to lumbar vertebrae along the backbone, when find the characteristic positions, press the coordinate device to body, and the switch in the coordinate device will turn on two parallel Webcams to capture the space image of this coordinate device.

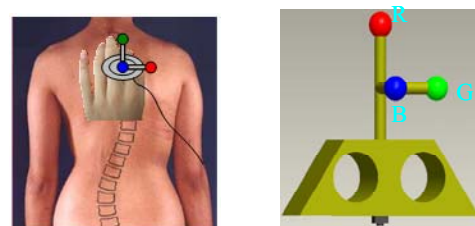


Figure 2. A, palpation; B, "coordinate converter"

Utilize two-dimensional image coordinate and algorithm of performing of the three-dimensional space coordinate, can calculate the three-dimensional coordinate in the space position of spine. It includes three parts here: three-dimensional coordinate device, digital image processing, and the space coordinate algorithm.

### 2.1.1 Three-dimensional coordinate device

While looking for the position of a point in space, we must set up a device with referencing coordinated system to represent the three axes of x, y and z in space. To set up this three-dimensional coordinate device, we use three colors which are red, blue and green to represent the three directions of axis of x, y and z in space. Using this three-dimensional coordinate device and Webcam that obtain from the market, we can capture the coordinated movement of referenced points in space.

### 2.1.2 Digital image processing

We use National Instruments LabVIEW 7.1 as the development setting, but the efficiency in calculating algorithm is poor, so we utilize Microsoft Visual Basic 6 to write the Visual Basic control of algorithm of digital image processing, and then use the ActiveX function of National Instruments LabVIEW to call out the control one to calculate the two-dimensional image coordinate of the three axis (RBG) of three-dimensional coordinate device.

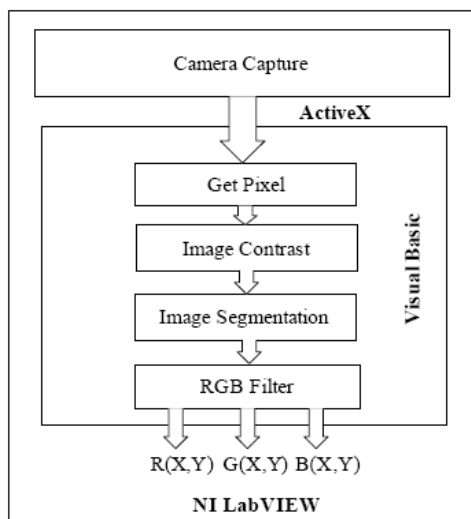


Figure 3. Image processing

However, because of the restriction of sphere form of the Camera, the image that captures from it result in twisted image (Fig4) and it will cause an error during transferring the coordinate. So, we utilize ANN (artificial neural network) to revise the error of this problem.

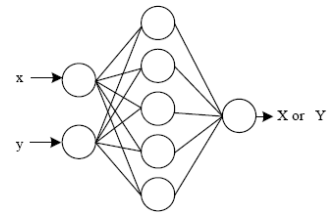
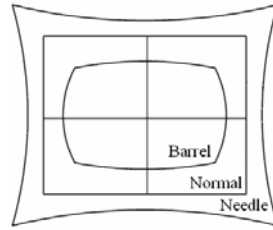


Figure 4. Image Distortion      Figure 5. ANN correct

We use Scale Conjugate Gradient Algorithm to train the ANN (fig5), and there are two input layers, five hidden layers and one output layer of EBP (Error Back Propagation). The condition is:

$$E = \frac{1}{2} \sum_k (d_k - y_k)^2 < 0.01$$

$d_k$  is the goal value and  $y_k$  is the output value of training.

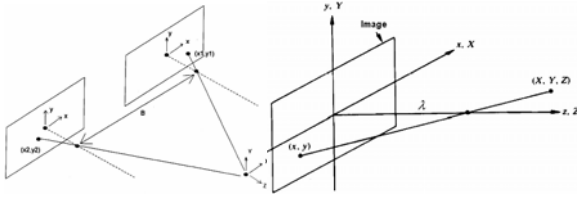
### 2.1.3 Algorithm of transferring from image coordinates to space coordinate

From Rafael C. Gonzalez, 1992, if we need to obtain object (X, Y) information (Fig6) in space. First we should get the information of the imaging depth of Z direction of the object in space. So if we can get the coordinated value of z axis, then we can calculate the coordinated value of x, y on space. The formulae as follows:

$$\begin{aligned} Z &= \lambda - \frac{\lambda B}{x_2 - x_1} \\ X &= \frac{x}{\lambda} (\lambda - Z) \\ Y &= \frac{y}{\lambda} (\lambda - Z) \end{aligned} \quad (1)$$

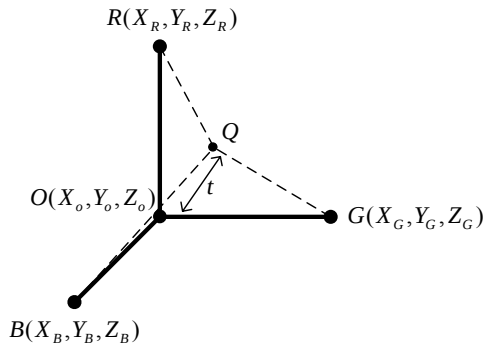
$\lambda$  : focal distance ;     $B$  : Two WebCam centre distance.





**Figure 6. Two-dimensional image is transferred to the three-dimensional space coordinate**

Utilize the formulae as mentioned above; we can calculate the coordinated value of three axes (RGB) of the coordinate device.



**Figure 7. Relation that  $R(X,Y,Z), G(X,Y,Z), B(X,Y,Z)$  and  $O(X,Y,Z)$**

However, RGB coordinate value only represents the space changes with each other. It does not locate position of switch basic coordinate in the space. But the relation in basic point was known (distance: 3cm), and all present the right angle. So, we can derive space coordinate value of point O by space vectors deductive formula, and point Q is one point that one line which point O perpendicular to the  $\Delta RGB$  level cross to the level. (Figure 7)

$$\alpha_1 = X_B - X_R; \beta_1 = Y_B - Y_R; \gamma_1 = Z_B - Z_R$$

$$\alpha_2 = X_G - X_R; \beta_2 = Y_G - Y_R; \gamma_2 = Z_G - Z_R$$

$$X_o = \frac{X_R + X_G + X_B}{3} + \begin{bmatrix} \beta_1 & \gamma_1 \\ \beta_2 & \gamma_2 \end{bmatrix} \times t$$

$$Y_o = \frac{Y_R + Y_G + Y_B}{3} + \begin{bmatrix} \gamma_1 & \alpha_1 \\ \gamma_2 & \alpha_2 \end{bmatrix} \times t$$

$$Z_o = \frac{Z_R + Z_G + Z_B}{3} + \begin{bmatrix} \alpha_1 & \beta_1 \\ \alpha_2 & \beta_2 \end{bmatrix} \times t$$

$$(X_o - X_R)^2 + (Y_o - Y_R)^2 + (Z_o - Z_R)^2 = 9$$

$$(X_o - X_G)^2 + (Y_o - Y_G)^2 + (Z_o - Z_G)^2 = 9$$

$$(X_o - X_B)^2 + (Y_o - Y_B)^2 + (Z_o - Z_B)^2 = 9 \quad (2)$$

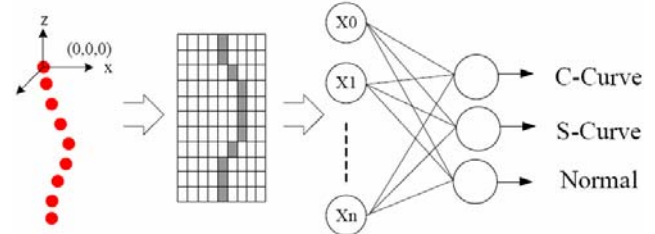
## 2.2 Setting-up and analysis of scoliosis curvature

Receive the coordinate value of X, Y and Z which changing in the space of three-dimensional coordinate device. Connect coordinate value of each point can show three-dimensional space changes of the spinal curvature.

## 2.3 Analysis and diagnosis of the spinal curvature for clinical meaning

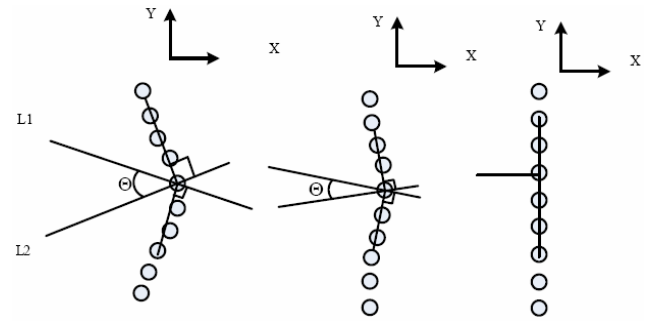
In order to analyze the clinical meaning and interpret, analyzed and interpreted each coordinate device automatically in the coordinate position in the space. It can differentiate into two categories and calculate the angle of scoliosis of the curved side.

ANN's group of learning vector quantization (LVQ) will be trained. Result will be normalized for input; Output can differentiate what type of scoliosis curvature is (C, S, Normal) (Figure 8).



**Figure 8. Learning Vector Quantization Ann**

Calculate side curved angle way of utilizing three-dimensional coordinate result, look for crooked place of vertebra most.



**Figure 9. Angle of scoliosis calculating**



$$L_1 = a_1x + b_1y + c_1 = 0$$

$$L_2 = a_2x + b_2y + c_2 = 0$$

$$\Theta = \cos^{-1} \frac{a_1a_2 + b_1b_2}{\sqrt{(a_1^2 + b_1^2)(a_2^2 + b_2^2)}} \quad (3)$$

### 3. Result

Experimental result, able to calculate the three-dimensional changes in the space coordinate of coordinate device. Simple algorithm and digital image processing and constructs the three-dimensional curvature of spine.

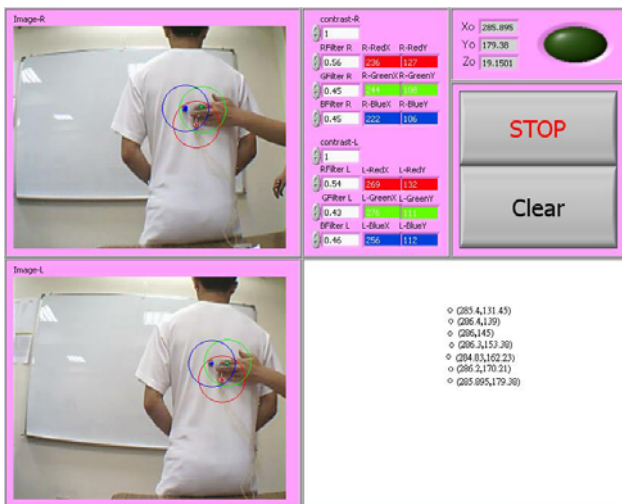


Figure 10. Experiment course

| Point | X       | Y       | Z     |
|-------|---------|---------|-------|
| 1     | 285.4   | 131.45  | 18.4  |
| 2     | 286.4   | 139     | 17.6  |
| 3     | 286     | 145     | 17.2  |
| 4     | 286.3   | 153.38  | 17.8  |
| 5     | 284.83  | 162.23  | 18.4  |
| 6     | 286.2   | 170.21  | 18.9  |
| 7     | 285.89  | 179.38  | 19.15 |
| 8     | 286.45  | 186.43  | 19.1  |
| 9     | 286.223 | 193.512 | 18.8  |
| 10    | 285.932 | 200.65  | 19.2  |
| 11    | 286.42  | 205.34  | 19.6  |
| 12    | 284.6   | 21.03   | 18.9  |

Table 1.Result

### 4. Conclusion

In the course of the experiment, factors of light of the environment and color complexity of the background are great influences of the results, thus the need of time consuming parameter adjustments of digital image processing manually (contrast ratio, thresholding value).

The calculation of the angle and traditional Cobb Angle methods are different in this research, but still can measure the severity of scoliosis, the angle measured increases as the curve become greater as well, but how to define the angle, still need further studies.

Build and construct 3D curvature, allows study of the changes on three-dimensional space of scoliosis. X-Y direction can interpret the type and angle; Y-Z direction can interpret the severity of humpback. Because we believe, scoliosis not only causes influence on X-Y direction, but also causes a great influence on Y-Z direction. Unlike the traditional method which only deal with the angle of X-Y and type, to diagnosis scoliosis.



Figure 11. coordinate converter

### 5. References

- [1] Lino Ramirez, Nelson G. Durdle, V. James Raso, and Doug L. Hill, " A Support Vector Machines Classifier to Assess the Severity of Idiopathic Scoliosis From Surface Topography" , IEEE TRANSACTIONS ON INFORMATION TECHNOLOGY IN BIOMEDICINE, VOL. 10, NO. 1, JANUARY 2006
- [2]R. B.Winter, J. E. Lonstein, J. Drogdt, and C. A. Noren, "The effectiveness of bracing in the nonoperative treatment of





- idiopathic scoliosis,” *Spine*, vol. 11, no. 8, pp. 790–791, 1986.
- [3] Kirsten L. Tulchin\*, Gerald F. Harris\*\*, Xue-Cheng Liu\*\*, John Thometz\*\*, ” NON-RADIOGRAPHIC ANALYSIS OF THE PEDIATRIC SPINE IN ADOLESCENT IDIOPATHIC SCOLIOSIS ”, 0-7803-5674-8/99, 1999 IEEE
- [4] G. Gorton, P. Masso,” Assessment of the Standing Posture of Patients with Scoliosis using Optoelectronic Measurement Techniques”, 0-7803-6469-4/00, 2000 IEEE
- [5] R. Perdrille, J. Vidal. “Morphology of scoliosis: Three-dimensional evolution.” *Orthopedics* Vol. 10, p. 909. 1987.
- [6] G.S. Bassett, M.C. Hensinger, M.D. Keiper. “Effect of Posterior Spinal Fusion on Spinal Balance in Idiopathic Scoliosis.” *JPO*. Vol. 9, pp. 672-674. 1989.
- [7] R.J. Jefferson, I. Weisz, A.R. Turner-Smith, J.D. Harris, G.R. Houghton. “Scoliosis surgery and its effect on back shape.” *JBJS* Vol. 70-B, pp. 261-266. 1988.
- [8] H. Labelle, J. Dansereau, C. Bellefleur et al. “Comparison between preoperative and postoperative three-dimensional reconstructions of idiopathic scoliosis with the Cotrel-Dubousset procedure.” *Spine* Vol. 20, pp. 2487-2492. 1995.
- [9] D.J. Pearsall, J.G. Reid, D.M. Hedden. “Comparison of three noninvasive methods for measuring scoliosis.” *Physical Therapy* Vol. 72, pp. 648-656. 1992.
- [10] P.O. Ajemba, N.G. Durdle, D.L. Hill, and V.J. Raso, "RE-POSITIONING EFFECTS OF A FULL TORSO IMAGING SYSTEM FOR THE ASSESSMENT OF SCOLIOSIS", CCECE 2004- CCGEI 2004, Niagara Falls, May 2004 IEEE.
- [11] Rafael C. Gonzalez and Richard E. Woods, ”Digital Image Processing”, 1992.



# Diagnoses and Surgical Planning of Lung Cancer in Color-Enhanced 3D Visualization System

Horng-Shyang Liao<sup>1,2</sup> Po-Ying Li<sup>1</sup> Cheng-Wei Ku<sup>1</sup> Kuen-Long Tsai<sup>1</sup> Chia-Yang Sun<sup>1</sup>

<sup>1</sup>National Center for High-Performance Computing, Taiwan, ROC

<sup>2</sup>Department of Computer Science, National Chiao Tung University, Taiwan, ROC

E-mail: {hsliao|bylee|heresy|xjapan666|morrisun}@nchc.org.tw

## Abstract

*A medical images visualization system for lung cancer detection, diagnoses, and surgical planning is proposed in this paper. Our system provides the feature of 3D color-enhanced volume rendering with 2D images pre-segmentation. This system also allows the traditional 2D and 3D navigation and measurement, which let users easily visualize the data. This visualization system can be divided into two parts: pre-processing and post-processing. In the pre-processing part, we develop a semi-automatic program with a friendly graphical user interface (GUI) to segment lung region apart from the original thorax computed tomography (CT) images. In the post-processing part, we develop an interactive volume visualization program that supports pre-integrated high quality rendering, multi-volumes rendering, and transfer function classification. A 3D OpenGL-based graphical user interface for transfer function design is given in the volume visualization program. Based on the transfer function modification, different features inside the volume data will be enhanced in the visual results. With the above features, our proposed system can provide a very useful and decent visual effect for the region of lung in the thorax CT images. A trained medical staff can use this software system to locate the tumor insides the lung, or perform a surgical planning in a 3D scene. In our opinion, this system is also feasible for surgical training.*

**Keyword:** volume visualization, transfer function, medical imaging, helical CT, image segmentation.

## 1. Introduction

From the USA report of cancer statistics [23], lung cancer is the worldwide top one fatal cancer. Based on the observation and clinical study cases, the pulmonary doctors feel the needs of assistant tools to retrieve more information from the CT images during the diagnoses. In this project, there are three major problems that we want to solve. The first problem is the delayed detection of tumors. Large tumors inside the lungs are obvious and visible in CT images, but at the early stage of lung cancer when the tumors are small, this is not quite the

same. Other kinds of visual enhancement based on statistics is needed in early stage diagnoses. The second problem is the difficulty to compare the size of the same tumor in the different sets of CT images taken at different time. Doctors need good measuring tools to provide an insight about the variant size of tumors between different time. And the last problem is that conventional surgical planning is done based on the stack of 2D CT images. It would be extremely helpful if we can use volume visualization to provides a 3D display that is close to realism.

An intuitive method to solve the above problems and to achieve acceptable diagnostic response time is volume rendering by graphics hardware-based texture mapping [2, 3, 9]. In order to reduce sampling artifacts, we also implement the pre-integrated method [8] by OpenGL [19] with GLSL [13]. However, if we simply use volume rendering to show a whole set of CT images, there are too many visual informations and will clutter the display. We need to highlight the regions of interests, including tumors and some bones, organs, vessels, or tissues that are useful for diagnoses or surgical planning. To cope with this issue, we adopt the transfer function concept to design a GUI. This GUI provide an easy way to modify the color table which maps different density values into different RGBA colors and classifies meaningful regions by enhanced visual effects. Unfortunately, for the detection of lung cancer, the density values of tumors and skin are very close. It is hard to distinguish them by applying transfer function during the classification step of volume rendering. In order to solve this problem, we develop a semi-automatic pre-processing method to segment the lung regions apart from the thorax CT images. We modify our volume rendering system to render original thorax images and processed lung images with individual transfer functions at the same time. Finally, an interactive, high-quality, and color-enhanced volume rendering system is presented and its transfer function design, measurement, and 3D viewing can achieve our three goals to help doctors to observe lung cancers.

## 2. Related Work

In this section, we will discuss the fundamental knowledge of volume rendering that is used in our system and some previous works about lung visualization in 3D environment.





## 2.1. Volume Rendering

Volume rendering is a useful tool for scientific or medical visualization. It can be categorized into two different kinds of methods: direct volume rendering and indirect volume rendering. Direct volume rendering will render volume data directly, such as ray casting [15] or voxel projection [5]. Indirect volume rendering will do the geometry reconstruction from a chosen iso-value inside the volume data, such as marching cube [17, 24].

Following the trend of graphics hardware in recent years, texture-based volume rendering is an efficient direct volume rendering method [2, 3, 6, 7, 9]. When the pre-integrated method is proposed, interactive texture-based volume rendering also can render high quality result [8].

Features detection is another important problem in volume rendering. It is usually to design a color mapping transfer function to highlight features by visual effects. It is easy to map density values into RGBA colors but the difficult part is to design a good color table to highlight the desired features. This non-trivial problem is listed in Johnson's "Top Scientific Visualization Research Problems" [12]. Pfister et al. hold a race in 2001 and classify all proposed methods into four categories [21]:

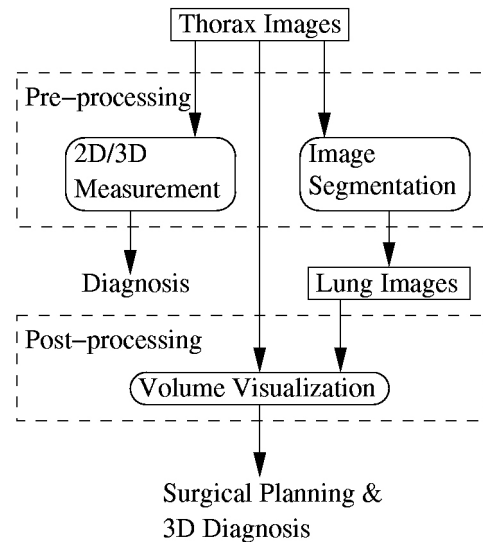
1. trial and error, with minimum computer aid;
2. data-centric, with no underlying assumed model;
3. data-centric, using an underlying data model; and
4. image-centric, using organized sampling.

It is hard to find general rules to design good transfer function automatically. A semi-automatic method with friendly graphical user interface is the most common method.

## 2.2 Lung Cancer Visualization

Based on users requests and technical evolution of medical imaging and volume visualization, some integrated libraries, such as VTK [18] and ITK [11] have been developed. Based on these libraries, many research groups start to create specific medical assistant tools. Some of them focus on diagnoses and surgical planning of lung cancer [4, 10, 25]. The input data are CT images of DICOM format. These tools follow the very similar processing procedure. The first step is to segment the region of interests for diagnoses. The second step is to create 3D visualization display for surgical planning. All of them use direct volume rendering and Zhang et al. also provide an indirect volume rendering. The differences between these methods are the details of processing algorithms which they applied to their own system.

## 3. System Overview



**Figure 1. System overview.**

**Figure 1** is the processing overview of our proposed visualization system. Input data are a stack of thorax CT images of DICOM format. Our pre-processing program has a set of measuring tools and a semi-automatic lung region segmentation tool. A parametric segmentation algorithm is applied. After some interaction with user, a stack of lung images will be generated. The detail description is presented in **Section 4**. Another part in pre-processing program is measuring tools. Most of them are 2D measuring tools, such as the measuring the length or the angle, except 3D seeds growing for computing the size of the volume object. However, only 2D images are displayed at this stage. Our post-processing program is a volume visualization subsystem. Multiple volumes with individual transfer functions can be rendered together. An interactive navigation and transfer function adjustment are also available. Therefore, users can do simple surgical planning and diagnoses in 3D environment. **Section 5** will explain the details.

## 4. Image Segmentation

In 1999, an automatic image processing method for pulmonary nodules' detection is proposed by Armato et al. [1]. Our image segmentation functionality is based on their lung segmentation part but ignore their nodules detection part because our nodules detection problem is left to our volume visualization part. Unfortunately, two major problems arise when we apply their method to our input images. The cause of these two problems is that the images from patients of lung cancer don't have perfect lung region. The first problem comes from their threshold selection method. They use a cumulative gray-level profile along the slope from a single 2D CT image's upper-left corner to center. The spatial information is difficult to carry in patient's CT images, especially when the nodules are on the line. The other problem is the



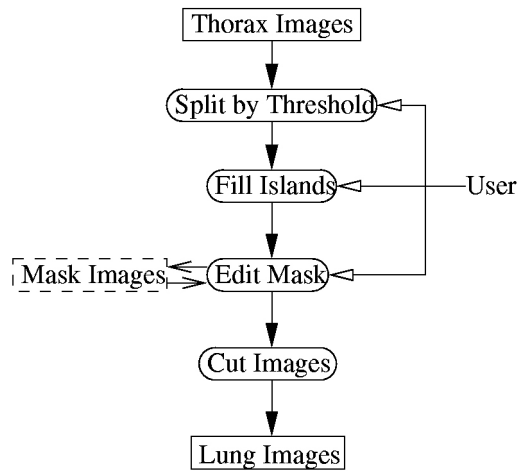


Figure 2. Segmentation procedure.

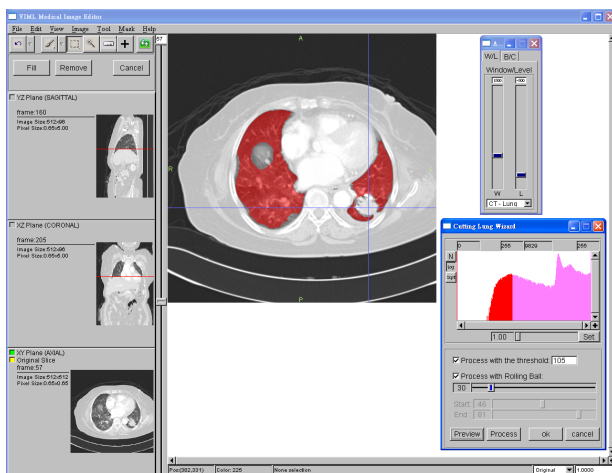
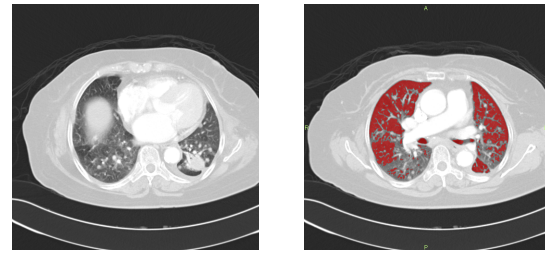


Figure 3. Segmentation GUI.

ambiguity of lung region when the nodules are very close to the boundary of lung or to the heart.

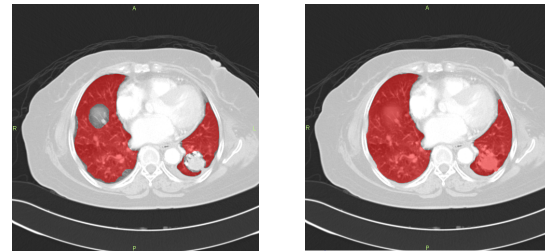
In order to prevent our system from the above problems, we develop our image segmentation sub-system by modifying partial algorithms. The working procedure presents in Figure 2 and the working environment looks like Figure 3. The input images are a stack of CT thorax images. A single slice of input image looks like Figure 4(a). Our first step will split lung region from thorax by threshold. We choose Otsu's algorithm [20] to replace the original cumulative profile method. Otsu's algorithm will choose threshold in the position of the balancing distribution of image's histogram and we use the chosen threshold to separate image. This splitting step has three small steps.

1. Thorax is split from original CT image based on the threshold determined by Otsu's algorithm.
2. The biggest contour of thorax is identified.
3. The region covered by the previous contour is processed by the Otsu's algorithm again to split lung region from thorax region by threshold.



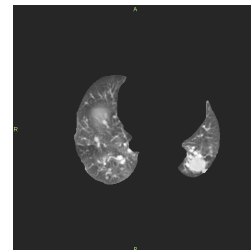
(a) Original image.

(b) Split by threshold.



(c) Fill islands.

(d) Edit mask.



(e) Lung image.

Figure 4. Segmentation example.

After the splitting by threshold, we can get an image like Figure 4(b). The second step is to fill islands. Island means a small fractal area inside a larger area. In this step, we also use the rolling ball algorithm as Armato et al. do. The processing result of this step is depicted in Figure 4(c). The third step is the mask editing. We provide a set of editing tools, including drawing, erasing, selecting, filling, and magic wand selecting. User can modify mask images of lung region arbitrarily, save mask images, or load previous saved mask images. One of the modified example is in Figure 4(d). The final step is to cut the original images by mask images and then generates a stack of lung images. A single slice of lung images is in Figure 4(e).

## 5. Volume Visualization

Figure 5 presents the procedure of our volume visualization sub-system. Besides the rendering pipeline, this procedure also shows the interactive functionalities. The input data are the thorax images and lung images and the output data are rendering results. User can observe and navigate the rendering results, and modify

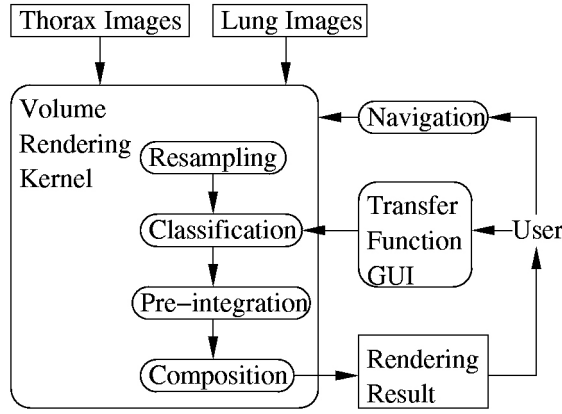


Figure 5. Volume visualization procedure.

the transfer functions through a graphical user interface. The volume rendering kernel is the most important part because the system takes the advantage of the state-of-the-art graphics hardware, we have to consider the constraints from graphics hardware while bring graphics hardware into full utilization. We will describe our volume rendering pipeline in **Subsection 5.1**, and more details about classification in **Subsection 5.2**. The last part of this section will present the measuring method in our 3D scene.

### 5.1. Volume Rendering Pipeline

The first step in our volume rendering kernel is resampling because we choose view align sampling planes instead of axis align planes. We use 3D texture that is supported by the current graphics hardware to store the volume data. This allows the graphics hardware to perform tri-linear interpolation when our system re-samples the data.

Classification is the second step and it is in charge of color mapping from density values to RGBA colors. This is a useful tool to highlight the desired range of density values. It can be a mathematical equation or just a color table. In our implementation, we provide a graphical interface for user to define the transfer function and create a color table from it. Color table can be treated as texture in graphics hardware. However, we don't use this color table for mapping because we want to improve rendering quality by pre-integrated method [8].

Our third step is pre-integration which is implemented using GLSL. Engel et al. propose this algorithm in programmable graphics hardware to prevent sampling artefacts caused by insufficient Nyquist frequency [8]. This algorithm will consider the integration between adjacent slices. In other words, slab is used to replace slice. After this step, a 2D color table will be generated. Axes represent front density values and back density values separately, so each value in this table is an integrated result from front to back. We use this pre-integrated table to do the color mapping.

The final step in our volume rendering pipeline is composition. We follow the volume rendering

composition equations that are derived by Levoy [15]. We can do the alpha blending recursively from back to front by applying the following equations:

$$c_{out} = a_{source} \times c_{source} + (1 - a_{source}) \times c_{in}$$

$$a_{out} = a_{source} + (1 - a_{source}) \times a_{in}$$

where  $a_{in}$  and  $c_{in}$  are previous integrated alpha and RGB colors.  $a_{source}$  and  $c_{source}$  are alpha and RGB colors from incoming voxel.  $a_{out}$  and  $c_{out}$  are current integrated alpha and RGB colors. But, when we use the pre-integrated method, the equations have to be revised as the following equations:

$$c_{out} = c_{preint} + (1 - a_{preint}) \times c_{in}$$

$$a_{out} = a_{preint} + (1 - a_{preint}) \times a_{in}$$

where  $a_{in}$ ,  $c_{in}$ ,  $a_{out}$ ,  $c_{out}$  are the same as previous definitions.  $a_{preint}$  and  $c_{preint}$  are the alpha and RGB colors from pre-integrated computing.

In our implementation of volume rendering pipeline, we also do the following two designs.

1. We also sort all view align planes from all volumes because there are two volumes as our input data.
2. We don't calculate gradient and shading to get better visual effects because we choose the better performance with less memory cost. We will discuss this issue in **Section 7**.

### 5.2. Classification

Classification is a color mapping procedure to enhance visual effects. The mapping function, also known as transfer function, is a set of mathematical equations or a color table. We use piecewise linear interpolation concept to develop our graphical user interface for color table generation.

Figure 6 is the proposed graphical user interface. Its controlling panel can be simply divided into three major parts:

- **Up-left block.** This block displays the histogram of density logarithm values and distribution of RGBA colors from the selected volume data set. The histogram chart is a useful guide for user to find features. The distribution of RGBA colors use four curves to present the results separately. This block shows the color distribution in real time.
- **Bottom-left block.** This is an interactive block. There is a density values' mapping bar in the bottom and can add or delete control points. We can set arbitrary RGBA colors for each control point. Afterward, the system will use linear interpolation to generate RGBA colors for other density values in between. The four bars in the middle are used to specify RGBA values separately for the selected control point. In upper-left corner, volume's name is displayed. Under the name, there is a box containing the color of current selected control

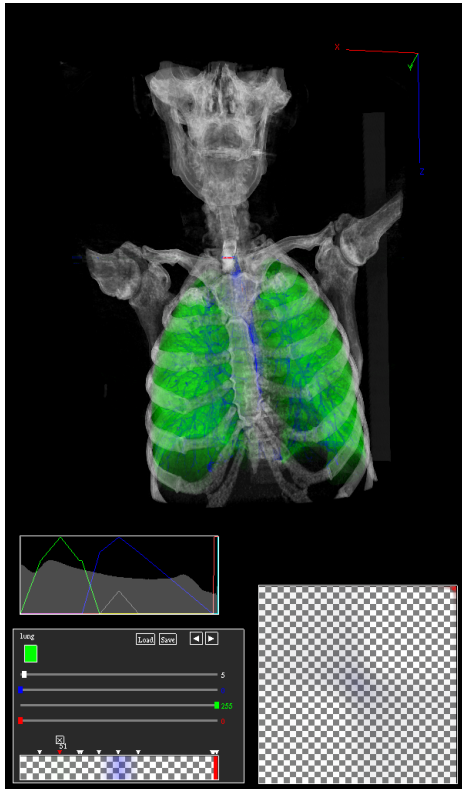






point. In upper-right corner, there are four buttons. From left to right, their functions are "load transfer function", "save transfer function", "jump to previous volume", and "jump to next volume".

- **Right block.** This block contains the pre-integrated table. It is a reference information without any interactive mechanism.



**Figure 6. Graphical user interface of transfer function.**

**Figure 7** is an example of the mapping between the designed transfer function and rendering result. From top to bottom, there are histogram of thorax volume data, volume rendering result, and histogram of lung volume data. In this case, we redistribute the density values from 0 to 255. Bone is identified in the volume data of thorax. Lungs, airways, tumors are identified in the volume data of lung. We also show the benefits of having individual transfer function for every volume data set. More features can be identified in the regions that users want to look at.

### 5.3. Measurement

For 3D measurement, we develop a 3D seed growing algorithm to calculate volumetric size. User controls a 3D probe to locate the seed and chooses a threshold of tolerance. After the seed growing procedure, we multiply the marked voxels by the size of single voxel, and will get the size of marked region. We highlight the



**Figure 7. A transfer function example.**

marked region on the screen for user to verify the result.

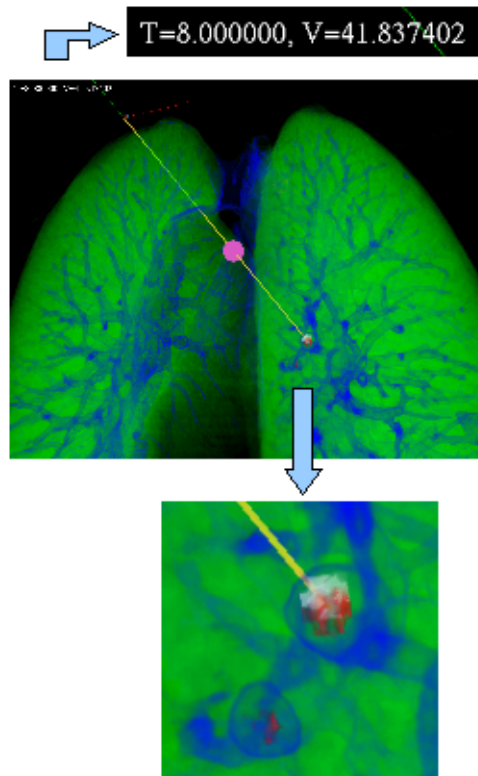
**Figure 8** is an example. A probe located in assistant coordinate axes is pointed out to a tumor. A purple ball that can slide along the probe gives a concept about the threshold of tolerance. The selected tumor is marked as white. Values of threshold and volumetric size display in the upper-left corner.

## 6. Results

Our system is based on OpenGL [19], GLSL [13], FLTK [22], and 3D VR Engine [16]. It is developed using Microsoft Visual C++ 6.0 in win32 platform. Our testing platform is a PC with dual AMD Athlon MP 2800+ CPU, 3GB Register ECC DDR266 memory, and NVIDIA GeForce 6800 GT graphics card with 256MB graphics memory.

We have 50 patients' data sets and there are more than 2 data sets for each patient. We randomly test more than 10 data sets and the system works well. We select two data sets to present our results. The first case has some



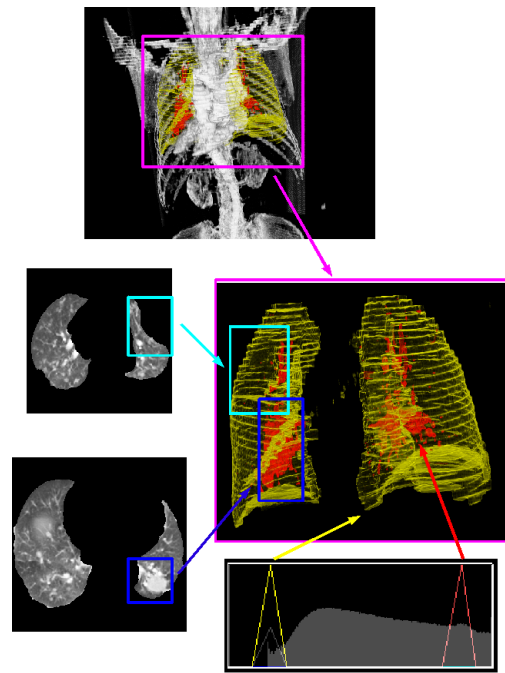


**Figure 8. Statistics of volumetric size.**

small tumors and the second case has small tumors and big tumors. In the first case, lung region can be detected automatically. User only needs little editing to adjust the mask. In the second case, the automatic detection result is less than desired because the tumor is too large and right next to the heart or the boundary of thorax. **Figure 4** is the example of the second case. User needs some editing to fix the boundary of the lung. Nevertheless, our system still gives a convenient way to segment lung region from the thorax CT images.

We put the previous segmented result of first test case into our visualization sub-system. **Figure 7** is the rendering result.  $512 \times 512 \times 512$  is the original thorax volume's resolution and  $512 \times 512 \times 256$  is lung volume's resolution. The size of single voxel is  $0.66015625 \times 0.66015625 \times 1.0 \text{ mm}^3$ . The 3D visualized volume data is redistributed into 8-bits unsigned char images. Rendering performance of this case is about 3.8 frame per second (FPS) in  $1024 \times 768$  screen resolution and 1.5 FPS in  $1280 \times 1024$  screen resolution.

The second test case has lower volume resolution but more complex diagnoses. Its original thorax volume's resolution is  $512 \times 512 \times 96$  and segmented lung volume's resolution is  $512 \times 512 \times 35$ . Voxel size is  $0.65234375 \times 0.65234375 \times 5.0 \text{ mm}^3$ . The rendering and partial diagnoses results are presented in **Figure 9**. Because the resolution of volume is low, the rendered image is not smooth. And due to the contrast agents in these images, vessels and bones have high density



**Figure 9. A diagnoses example in our 3D visualization environment.**

values. We use bone frame to locate the lung position and then turn it off on the rendering frame by adjusting its transfer function. We color the boundary of lung as yellow and partial range of blood as red. A dent region and a cancer inside the lung are observed. Its 3D visualized data is also redistributed into 8-bits unsigned char images. Rendering performance of the second case is about 4.5 FPS in  $1024 \times 768$  screen resolution and 2.5 FPS in  $1280 \times 1024$  screen resolution.

## 7. Discussions

In this section, we want to discuss our proposed features detection method in volume rendering step. The proposed method can identify spatial features easily and decrease the memory cost in graphics hardware. However, we decide to sacrifice some visual effects to gain the rendering performance. We will discuss the details in the following sub-sections and some concerns especially based on our study cases.

### 7.1. Advantages

There are two advantages to use our method:

1. **Specific spatial features.** Traditional method for classification usually specify features by the values of density, first derivative, or second derivative. However, some features from our study cases are hard to be distinguished by the traditional method, which means that the values of density, first derivative, and second derivative between different features are very similar. For example, patient's skin and tumors





have similar density values in CT images. Tumors are close to sphere and skin is surrounding the whole body. Therefore, first derivative and second derivative are also similar. We need to find other ways to separate similar densities between different spatial regions. 3D seed growing is an intuitive method to identify spatial features. However, it is hard to give the proper growing seeds for the purpose to detect unknown features. We need a tool to find out all discontinuous and doubtful tumors inside lung region. Therefore, we use our proposed method to locate the region of interests in spatial domain by segmentation and find detailed features by individual transfer function for each segmented volume. It is quite useful for doctor to detect features in specific regions without interfering signals.

2. **Low memory cost.** The memory of graphics card is relatively smaller but faster than main memory for rendering. However, texture-based volume rendering is consuming a lot of graphics memory. The trade-off for reducing memory usage is to double the size of the original volume data.

## 7.2. Disadvantages

There are two disadvantages to use our method:

1. **Without shading.** We don't calculate first derivative from densities of volume, so we can't do the shading effect. Shading effect is useful to provide the depth concept. For the issue of depth concept, we think it's good enough to use attenuation in volume rendering. Shading is good for clear surface rendering. However, we usually need to look inside the surface when we choose direct volume rendering. If the depth information is very critical, our system can render left-eye image and right-eye image simultaneously in a stereo viewing virtual reality environment. It will provide a pretty decent visual depth cue for users.
2. **Unclear rendering boundary.** Levoy's "isovalue contour surface" opacity function [15] can be controlled easily by the triangular classification widget proposed by Kniss et al. [14]. They can identify very clear iso-surface inside volume data. However, if we choose to calculate first derivative on the fly, it is a time consuming calculation. If we choose to calculate first derivative in advance, it needs large memory to store the computing results. For the issues from memory cost and rendering performance, we sacrifice the benefit from first derivative. It is an important future work to draw clear boundary with low memory cost.

## 8. Conclusion

In order to provide an alternative method to detect tumors, to measure tumors' size, and to do the 3D surgical planning, we propose a lung cancer visualization system with 2D pre-process and 3D post-process. Our 2D pre-processing sub-system is capable of the semi-automatic lung region segmentation and fundamental 2D measurement, and our 3D post-processing sub-system can display results using volume visualization with lively color effects and 3D measurement. Our lung cancer detection method is not intended to totally replace the current diagnoses, but it is another assistant tool to improve the diagnoses by identifying the features. Less interference caused by other tissues or organs also provides clearer rendered image. For surgical planning and training, an interactive 3D volume visualization can be very helpful.

In the near future, we have five issues to resolve in order to improve our system:

1. Clear boundary without high memory cost is needed.
2. Rendering performance of large volume has to be improved.
3. VR tracking devices are useful assistant equipments.
4. More coordinate information from DICOM format has to be provided.
5. Clipping plane of any arbitrary orientation is classical but useful tool to help surgical planning. It is valuable to be implemented.

Many improvements are needed for medical visualization. Such as the automatic and intelligent algorithms for segmentation or classification are very critical to enhance the usability of medical visualization system. Rendering quality, speed, and correctness to help diagnoses are always the major concern in 3D volume rendering, especially for large and high resolution image data, which could be realized in the future.

## Acknowledgements

We thank Chang Gung Memorial Hospital for data supporting. We also thank Tan-Chi Ho, Shou-Cheun Ting, Ming-Jing Li, Charlie H. Chang, Dr. Ching-Yao Lin, and San-Liang Chu for partial implementation and discussions.

## References

- [1] S. Armato, M. Giger, C. Moran, J. Blackburn, K. Doi, and H. MacMahon, "Computerized Detection of Pulmonary Nodules on CT Scans", In *Radiographics*, 1999; volume 19, pp. 1303–1311.
- [2] B. Cabral, N. Cam, and J. Foran, "Accelerated Volume Rendering and Tomographic Reconstruction Using Texture





- Mapping Hardware”, In *Proceedings of the 1994 Symposium on Volume Visualization*, 1994; pp. 91–98.
- [3] T. J. Cullip and U. Neumann, “Accelerating Volume Reconstruction with 3D Texture Hardware”, *Technical report*, 1994.
- [4] A. Delegacz, S.-C. B. Lo, M. T. Freedman, and S. K. Mun, “Feature Extraction, Analysis, and 3D Visualization of Local Lung Regions in Volumetric CT Images”, In *Medical Imaging 2001: Visualization, Image-Guided Procedures, and Display*, 2001; volume 4319, pp. 76–81.
- [5] R. A. Drebin, L. Carpenter, and P. Hanrahan, “Volume Rendering”, In *Proceedings of Computer Graphics*, 1988; volume 22, pp. 65–74.
- [6] K. Engel, M. Hadwiger, J. M. Kniss, A. E. Lefohn, C. Rezk-Salama, and D. Weiskopf, “Real-Time Volume Graphics”, *SIGGRAPH'04 course notes*, 2004; volume 28.
- [7] K. Engel, M. Hadwiger, J. M. Kniss, and C. Rezk-Salama, “High-Quality Volume Graphics on Consumer PC Hardware”, *SIGGRAPH'02 course notes*, 2002; volume 42.
- [8] K. Engel, M. Kraus, and T. Ertl, “High-Quality Preintegrated Volume Rendering Using Hardware-Accelerated Pixel Shading”, In *Proceedings of the ACM SIGGRAPH/EUROGRAPHICS Workshop on Graphics Hardware*, 2001; pp. 9–16.
- [9] A. V. Gelder and K. Kim, “Direct Volume Rendering with Shading via Three-Dimensional Textures”, In *Proceedings of the 1996 Symposium on Volume Visualization*, 1996; pp. 23–31.
- [10] B. M. Hemminger, P. L. Molina, T. M. Egan, F. C. Detterbeck, K. E. Muller, C. S. Coffey, and Joseph K. T. Lee, “Assessment of Real-Time 3D Visualization for Cardiothoracic Diagnostic Evaluation and Surgery Planning”, *Journal of Digital Imaging*, 2005; volume 18, issue 2, pp. 145–153.
- [11] Insight Software Consortium. Insight Segmentation and Registration Toolkit (ITK), 1999–2003, <http://www.itk.org/>.
- [12] C. R. Johnson, “Top Scientific Visualization Research Problems”, *IEEE Computer Graphics and Applications*, 2004; volume 24, issue 4, pp. 13–17.
- [13] J. Kessenich, D. Baldwin, and R. Rost, “The OpenGL Shading Language”, <http://www.opengl.org/documentation/oglsl.html>.
- [14] J. Kniss, G. Kindlmann, and C. Hansen, “Multi-Dimensional Transfer Functions for Interactive Volume Rendering”, *IEEE Transactions on Visualization and Computer Graphics*, 2002; volume 8, issue 3, pp. 270–285.
- [15] M. Levoy, “Display of Surfaces from Volume Data”, *IEEE Computer Graphics and Applications*, 1988; volume 8, issue 5.
- [16] H.-S. Liao, P.-Y. Li, C. H. Chang, and S.-L. Chu, “3D VR Engine”, In *Proceedings of 7th International Conference on High Performance Computing and Grid in Asia Pacific Region (HPC Asia 2004)*, 2004.
- [17] W. E. Lorensen and H. E. Cline, “Marching Cubes: A High Resolution 3D Surface Construction Algorithm”, In *Proceedings of Computer Graphics*, 1987; volume 21, pp. 163–169.
- [18] K. Martin, W. Schroeder, and B. Lorensen, “The Visualization Toolkit (VTK)”, 1993–2002; <http://public.kitware.com/VTK/>.
- [19] OpenGL Architecture Review Board, D. Shreiner, M. Woo, J. Neider, T. Davis, and D. Shreiner, *OpenGL Programming Guide: The Official Guide to Learning OpenGL, Version 1.4*, Addison-Wesley, 4th edition, 2003.
- [20] N. Otsu, “A Threshold Selection Method from Gray-Level Histograms”, *IEEE Transaction on Systems, Man, and Cybernetics*, 1979; volume 9, issue 1.
- [21] H. Pfister, B. Lorensen, C. Bajaj, G. Kindlmann, W. Schroeder, L. S. Avila, K. Martin, R. Machiraju, and J. Lee, “The Transfer Function Bake-Off”, *IEEE Computer Graphics and Applications*, 2001; volume 21, issue 3, pp. 16–22.
- [22] B. Spitzak et al., “Fast Light Toolkit (FLTK)”, 1998–2005; <http://www.fltk.org/>.
- [23] P. A. Wingo, T. Tong, and S. Bolden, “Cancer statistics 1995”, *CA: A Cancer Journal for Clinicians*, 1995; volume 45, pp. 8–30.
- [24] G. Wyvill, C. McPheeters, and B. Wyvill, “Data Structure for Soft Objects”, *The Visual Computer*, 1986; volume 2, issue 4, pp. 227–234.
- [25] H. Zhang, X. Sun, and H. Duan, “3D segmentation and visualization of lung volume using CT”, In *Medical Imaging 2005: Visualization, Image-Guided Procedures, and Display*, 2005; volume 5744, pp. 538–546.







# The use of geographic information systems(GIS) for rapid assessment of health facilities following a disaster: the case of the tsunami disaster in the province of Aceh

Anis Fuad\*, Hari Kusnanto\*, Adi Utarini\*, JP van Dijk\*\*, JW Groothoff \*\*

\*Public Health Department, Faculty of Medicine, Gadjah Mada University, Jogjakarta, Indonesia

\*\*Department of Social Medicine, University Medical Center Groningen, University of Groningen, The Netherlands

Correspondence: fuadanis@yahoo.fr

## Abstract

**Background:** The tsunami disaster in the province of Aceh in December 2004 caused more than 200.000 casualties, 500.000 people became homeless and a lot of infrastructures was destroyed including health facilities. A health facilities assessment was needed to provide detailed information with regards to rehabilitation and reconstruction of health services provision after the disaster. This article aims to describe the process of the rapid assessment, analyse the field results and discussing the use of free GIS software during the whole process.

**Methods:** A cross sectional survey in 326 out of 354 health facilities (92,1%) in 20 of the 21 districts in the province of Aceh was conducted two months after the disaster. Data concerning health staff, medical equipment, functional services as well as geographic coordinates of each health facility were collected. Spatial data were obtained from the Indonesian Remote Sensing and Geographic Information Systems (RSGIS Forum) mailing list, United Nation-Health Information Center (UN-HIC) and other NGOs. Epimap was used to examine the disaster impact in terms of the condition of the health infrastructure, staffs and functions. Information of each health facility were published online using open-source MapServer software.

**Results:** About 20% of the health facilities were destroyed with different intensity. The most affected areas were the Banda Aceh municipality, the district of Aceh Besar, Aceh Jaya and Aceh Barat. Spatial analyses showed that the damaged area influenced the damaged severity of health facilities. A web-based GIS application was developed to provide this information accessible by the public.

**Conclusion:** It is possible to integrate GIS method for data collection, analysis, and online visualization of health facilities assessment. The use of free software and collaboration with multiple organization enhanced the implementation. Strengthening local capacity with low cost GIS method is important for future disaster preparedness especially for developing countries. b-based GIS application to make it accessible by public.

## 1. Introduction

Geographic information systems (GIS) is a system composed by hardware and software to collect, process, analyze data (tabular and spatial) to produce geographic based information in order to improve decision making. In the public health domain, GIS methods were mostly used for outbreak investigation and epidemiologic studies. It has been used to define physical access to malaria treatment in Kenya[1], measurement tuberculosis catchments area in the South Africa[2], describing the epidemiological pattern of hepatitis C in Bangladesh[3] and many others[4]. The use of GIS is popular in the developed and developing countries as well, especially with the support of advanced information technologies

With the increasing use of GIS in several public health functions[5], however its utilization for humanitarian emergency is limited. Kaiser *et al* proposed several GIS potentials following the disaster i.e: rapid health assessment, estimation of mortality, resource allocation, planning and logistics as well as epidemiologic control[6]. On the other hand, Noji proposed the use GIS method also useful for estimating the disaster-affected population[7]

The tsunami disaster in December 2004 has caused more than 200.000 casualties, more than 500.000 people became homeless and a great deal of the infrastructure was destroyed including health facilities. Public health responses following the disaster can be categorized as emergency relief, rehabilitation and reconstruction of the health system. Information technology can facilitate an effective public health response through several methods. Arnold *et al*, especially reported the importance of information-sharing to enable rapid information collection, processing and distribution[8]. Mathew emphasized the use of information technology architecture (wireless network), hardware (personal digital assistant) and application systems (software) to facilitate this function in out of hospital condition[9].

However, until recently, lack of empirical evidence exists in describing the use of GIS for health facilities assessment following a disaster. Only fourteen papers could be found focusing on the simultaneously occurrence of the terms 'geographical information systems' and 'natural disasters' and only one of them is







related with the needs assessment. Existing publications of post disaster needs assessment mostly concerned with human needs to feed the emergency action, such as the work of Waring *et al*[10]. During the tropical storm Allison in Texas in 2001, Waring *et al* utilized GIS for indicating the survey cluster and analyzed the impact of the storm based on the cluster.

The importance of this research is justified because of while most of the published articles were about disaster-impact on human; our report is about its impact on health facilities. During the rehabilitative and reconstructive phase, public health authorities are often having difficulties whether to rebuild the existing or to develop new health facilities in the post disaster area. Detailed information of the disaster impact on health facilities including health facilities damaged level, population catchments area as well as its staffs and equipment, is hypothetically will help in this situation. Objective of this article is to report (1) the process of rapid assessment of health facilities following the tsunami disaster in Aceh, (2) the result of the field assessment analysis and (3) the lesson learnt of free GIS software utilisation on data collection, analysis and dissemination of the result.

## 2. Material and methods

A cross sectional study was conducted in Aceh province two months after the tsunami. Eight teams (each consisted of 1 public health surveyor, 1 civil engineer and 1 GIS surveyor) visited 20 out of 21 districts in two weeks period. The survey aimed to collect data and information to assess health facilities functioning, number, types and presence of human resources, types of medical equipments and supplies, physical infrastructure, photograph as well as geographic coordinate. This data is important to measure to what extent the impact of the disaster on existing health infrastructures. Health facilities which have been visited were public hospitals, district health office, primary health centers, district pharmacy warehouse and health laboratories. Each teams was equipped with paper based questionnaires, laptop, digital camera and a global positioning system receiver. GPS Garmin 72 S was used to capture geographic coordinate with 15 m accuracy in latitude/longitude. During the survey, each team member performed different task. The public health surveyor was responsible to interview the respondents (mostly head of the health facilities), the civil engineer surveyor measured physical infrastructure condition and GIS surveyor captured the coordinate.

Electronic data entry was facilitated by a web based electronic form. Application was developed with PHP (Personal Hypertext Processor) v 4.3.10 programming software combined with mysql database v 4.1.9 (both are free softwares). As in other developing countries, obtaining spatial data is very hard and even if available the cost is expensive. After the tsunami, several versions

of Aceh spatial data existed, however, the most accurate spatial data were obtained from Indonesian Remote Sensing and GIS (RSGIS Forum) mailing list. This version contains the most complete and up-to-date administrative boundaries, main road and river as well as topographic elevation. Shapefile of damaged area was obtained from United Nation Humanitarian Information Center (UN-HIC). Epimap, a GIS module in a public domain epidemiological software of Epi Info 3.3.2, was used to create layers of health facilities types and to superimpose with other layers.

In spite of being analyzed descriptively, assessment result was visualized online. The aim was to make the result available immediately and accessible anytime, anywhere to the public and especially to the stakeholders who are involved in Aceh rehabilitation activities. We asked for permission from the Ministry of Health to install a web-based GIS into one of its servers. Previously, there was only one web server and one mail server at the MOH Data Center. We selected one server to be installed with RedHat 6.2 operating systems before installing UMN MapServer v 4.4.1, an open source web-based GIS application. Soon after installation, the webgis was accessible at <http://map.depkes.go.id>.

## 3. Results

Aceh province is located at the top of Sumatra island which is inhabited by approximately 4.3 million population who are distributed 55,390 km<sup>2</sup> area composed by 21 districts, 228 subdistricts and 5,928 villages. The gap of population density is very high, so that with average density of 77.7 people/km<sup>2</sup>, the most populated district is Banda Aceh municipality (density of 2,916/km<sup>2</sup>), while the lowest is Gayo Lues (with the density of 13/km<sup>2</sup>). People seems to live in the northern and eastern part of Aceh and capital province not elsewhere. Before the tsunami, median life expectancy of Aceh people in general is about similar to other province in Indonesia. While the self morbidity was not differently significant to other area, however, proportion of the population more than 5 kilometres to health facilities were higher than average (38% vs 27.75%) than other part of Indonesia[14].

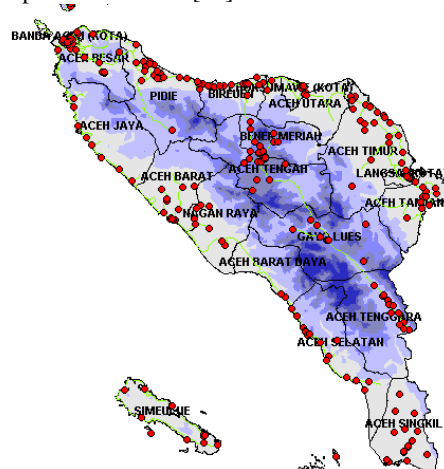


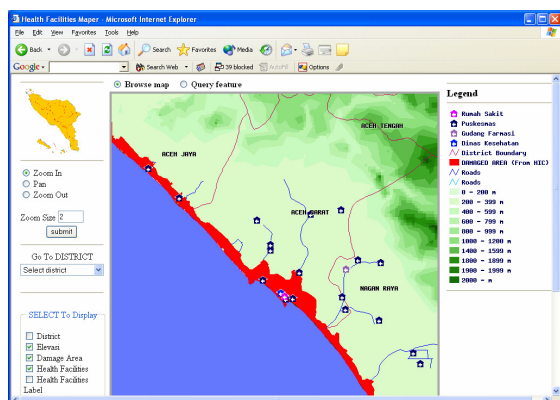
Figure 1. Health facilities distribution in Aceh Province





According to the APHO database there were 259 public health centers in Aceh province. This survey then collected a provincial-wide data from nearly all of health facilities in Aceh based on the APHO database. We have visited 301 health facilities including 233 public health centers, 31 public and private hospitals, 20 district health offices, 16 district pharmacy warehouses and 1 provincial level tuberculosis treatment center. Most of them are primary health centers since which are dispersed usually in each subdistricts to provide primary health care program for 8,000-45,000 of population. Public health center ratio to population is tend to smaller in urban areas than in rural.

Damaged intensity of physical infrastructure was classified within two categories; I (mild to moderated damaged) and II (heavy damaged). We found 26 (8,6%) health facilities were within category I and 24(7,9%) in category II. Most of the health facility with category III was located in Banda Aceh municipality, Aceh Besar district, Aceh Jaya district, Aceh Barat district and Simeuleu district.



**Figure 2.** Web based application of post tsunami health facility assessment (with zooming in Aceh Barat Daya district, superimposed with damaged area and elevation)

#### 4. Discussion

During the disaster, information about health facility is one of the most wanted information by many parties and NGOs. Therefore, providing this information is crucial during emergency and rehabilitative phase. We have explained the utility of geographic information systems to support health facilities assessment following the disaster. Geographic information systems was involved in this needs assessment from the collection of health facilities geographic location, more detailed analysis and data sharing.

However our method was not sterile from the limitation, for example the double burden of data entry from the manual and electronic. Even if each team has been prepared with the computer, in reality it was very hard to complete the electronic data entry within the field. In this concern, WHO has initiated Service Availability Mapping[15] project which intent to collect

health facility information supported by a handheld gadget equipped with GPS. This PDA could serve as an electronic questionnaire and capture geographic location coordinate as well. This method is promised to perform more accurate and faster data collection than the conventional one. GIS team of WHO also has similar interest with health facilities mapping. activities in collecting health facilities database. Healthmapper software, (free, available under request to WHO), has been used as the main platform to store health facility database which then could be superimposed with other health indicators.

Actually, Indonesia conducts regularly, several national-wide surveys and census which in some part collects health facilities data including its coordinates (see, Indonesian Family Life Survey, Demographic Health Survey, National Social Economic Census). However, until present there is no initiative to integrate all of the health facility data. The only web-GIS application of those data was only the thematic map named as National Spatial Information Systems[16]. With this regards, our method to store online health facility database is relevant.

Another issue is the limited capacity of health sector in conducting health facility survey. Therefore emphasis on training is a must do agenda. Our personal experience showed that district health office staff were mostly not well informed about the use of GPS and the software to analyse the information. We underlined previous research which recommended the strengthening the capacity of district and provincial level on the use of GIS. In Indonesia, the use of GIS has been hampered with the limited human resources, unavailability of affordable GIS software, limited inter sectoral coordination and information sharing. Sipe and Dale recommended the importance of using ArcView (a GIS commercial software) and special training for public health staff[16]. However to implement a web-based GIS software with ArcView platform is still too expensive for some provincial health office

At the central level, MOH should also be encouraged to improve their capacity for managing a web-based GIS of health facilities in collaboration with the Central Bureau of Statistics. Before that, development standard of identifier for each health facilities should be put in place. For a big countries as Indonesia, until present no health facilities standard identifier exist for approximately 1200 hospitals, 7000 public health centers and others.

One of the important key success of this implementation is a good collaboration with multiple organizations. Even if, this survey could not be carried out immediately after the disaster, because of some administrative procedures and preparation of the teams, however good coordination with provincial level and district level, informal groups (mailing list) as well as the role of World Health Organization and the UN-HIC was very important. It emphasized the Malilay experience that coordination is the most important factor for the post disaster public health program[17].





In general, GIS methodology could provide more detailed analysis of in a post disaster situation beyond administrative boundaries. Our experience shows that the use of free GIS software is sufficient to present better insight on the tsunami impact to health facilities. This experience showed that limited budget is not the obstacle to implement GIS application in the public health domain, especially in the post disaster situation.

## 5. Conclusion

Health facilities assessment following the disaster can be enhanced with the GIS methodology. The use of free GIS software and collaboration with multiple organizations are the key success of the implementation. Capacity building of public health offices concerning the use of free GIS application remains important issue for the developing countries.

## 10. References

- [1]Noor, AM, Zurovac, D, Hay, SI, Ochola, SA, Snow RW. Defining equity in physical access to clinical services using geographical information systems as part of malaria planning and monitoring in Kenya. *Tropical Medicine and International Health* 2003, 8(10): 917–926
- [2]Tanser F and Wilkinson D. Spatial implications of the tuberculosis DOTS strategy in rural South Africa: a novel application of geographical information system and global positioning system technologies. *Tropical Medicine and International Health*. 1999, 4(10) 634–638
- [3]Mujeeb SA, Shahab S, Hyder AA. Geographical display of health information: study of hepatitis C infection in Karachi, Pakistan. *Public Health*. 2000, 114, 413–415
- [4]Ricketts TC. Geographic information systems and public health. *Annu. Rev. Public Health* 2003. 24:1–6
- [5]Richards TB, Croner CM, Novick, LF. Geographic Information Systems (GIS) for State and Local Public Health Practitioners. *J Public Health Management Practice*, 1999, 5(2), 73–76
- [6]Kaiser, R; Spiegel, P; Hendersen, A; Gerber, M. The application of geographic information systems and global positioning systems in humanitarian emergencies: lessons learned, programme implications and future research. *Disasters*; 2003 Vol. 27, p127-140
- [7]Noji, Ek. Estimating population size in emergencies. *Bull World Health Organ* 2005; 83(3)
- [8]Arnold, JL, Levine, BN, Mnmatha R, Lee, F. Information-sharing in out-of-hospital disaster response: the future role of information technology. *Prehosp Disast Med* 2004;19(2):2001-2007
- [9]Mathew, D. Information Technology and Public Health Management of Disasters—A Model for South Asian Countries. *Prehosp Disast Med* 2005;20(1):54–60
- [10]Waring S, Zakos-Feliberti A, Wood R, Stone M, Padgett P, Arafat R. The utility of geographic information systems (GIS) in rapid epidemiological assessments following weather-related disasters: methodological issues based on the Tropical Storm Allison Experience. *Int J Hyg Environ Health*. 2005;208(1-2):109-16.
- [11]Available at <http://inisiatif.rsgisforum.net/index.html>
- [12]Available at <http://www.cdc.gov/epi/>
- [13]Boulos, MN, Honda, K. Web GIS in practice IV: publishing your health maps and connecting to remote WMS sources using the Open Source UMN MapServer and DM Solutions MapLab. *International Journal of Health Geographics* 2006, 5:6
- [14]National Social Economic Census, 2003
- [15]Anonim. Service Availability Mapping. Available at <http://www.who.int/healthinfo/sam/en/index.html>
- [16]Anonim. Sistem Informasi Spasial Nasional (English: National Spatial Information System) available at <http://202.155.86.41/kpk/TableProv.asp>
- [16]Sipe, NG, Dale, P. Challenges in using geographic information systems (GIS) to understand and control malaria in Indonesia. *Malaria Journal* 2003, 2:36
- [17]Malilay, JPublic health assessments in disaster settings: recommendations for a multidisciplinary approach. *Prehospital Disaster Med*. 2000 Oct-Dec;15(4):167-72.



## Dynamic PubMed: Self-updating software for PubMed bibliometric research

*Kumara Mendis, Rasika Amarasiri*  
*University of Sydney, Monash University*  
*kmendis@med.usyd.edu.au*

### Abstract

Bibliometric analyses are increasingly being used to monitor research quality assessment and for synthesizing research evidence. The basic unit in bibliometric analysis is the journal article that is indexed in electronic bibliographic knowledge bases such as PubMed. Our aim is to demonstrate a methodology to auto-update PubMed bibliometric analysis that measures the quantity of publications and displays in a tabular and graphical format in a web interface.

PubMed e-Utilities along with a normal PubMed query was used to interrogate the database. A web-server based MySQL stored the queries and results. An online PC acted as a scheduler client also has several Java applications that are executed at scheduled time intervals to update the system. The Java applications query the MySQL database located at the web server to retrieve the queries that need to be executed. It then formulates the automated queries to be sent to the e-utilities interface with the three second time intervals. The resulting data are stored back in the MySQL database. Once all the queries have been executed, the Java application executes the scripts in the web server to update the graphs with the new information.

The system automatically updates many types of queries both yearly and monthly intervals. We have automated country specific total PubMed publications, discipline specific updates such as 'general practice' or 'health informatics' or a particular fast developing field such as 'biotechnology and genetics'. Comparison between countries can also be depicted as line charts.

The incorrect or incomplete use of the author, institution and country affiliation remains major problems for the accuracy. Errors on updating the PubMed queries were dependant on the time of the update. This was maximal at peak-times when the server load was highest.

### 1. Introduction

Economically developed nations are increasing their investments on health and medical research (HMR) at a rapid rate. Australia [1] and United Kingdom [2] are at the forefront of the Organization of Economic Co-operation and Development (OECD) countries which have almost doubled the HMR expenditure every few years. This huge increase in HMR funding makes tracking, assessment of funding imperative. Research quality assessment (RQA) processes has been under much discussion in the U.K. and Australia. Australia's

RQA process is largely based on the UK model [3]. Traditional peer-review methods for RQA have come under scrutiny. Metric-based methods which include research income, publications, citations and the number of research students are increasingly being proposed for consideration as an alternative to the traditional peer-review processes [4] because they cost less and have been found to predict the same results as the peer-review methods.

The widening research-practice gap is a major concern for the practice of evidence based health care [5]. Traditional medical textbooks become out of date even before they are published. Investing on journals and electronic sources that synthesize evidence from the critically appraised top journals and is updated at least one a year is recommended for clinicians [6].

For synthesizing research evidence or for monitoring research funding using metric methods, the basic unit is the journal article that is published in electronic bibliographic knowledge bases such as PubMed [7], PsycInfo [8] or ISI Web of Science [9]. Publications are not the only, but certainly very important elements in the medical knowledge exchange process.

PubMed is the web interface of Medline in public domain that is commonly used by clinicians, academics and the public. Medline is the largest database that is accessed when using PubMed and currently contains more than 15 million bibliographic citations. PubMed uses the term 'citation' to refer to individual publications. To avoid confusion we will use the word 'publication' exclusively from this point. PubMed indexes more than 4,800 biomedical journals published more than 70 other countries. PubMed is essentially dynamic with an addition of between 1500-3500 completed references are added each day [10].

Bibliometrics is defined as 'the use of statistical methods in the analysis of a body of literature to reveal the historical development of subject fields and patterns of authorship, publication, and use. It was formerly called statistical bibliography' [11]. Bibliometric studies form a relatively minor group in PubMed publications with a total of approximately 1700 articles up to April 2006. It has however increased from 158 in 2000 to 217 in 2005. The types of bibliometric analysis are wide and cover many domains. It ranges from assessing what different countries get for their research spending [12], analysis of a specific domain like Parasitology [13] or to





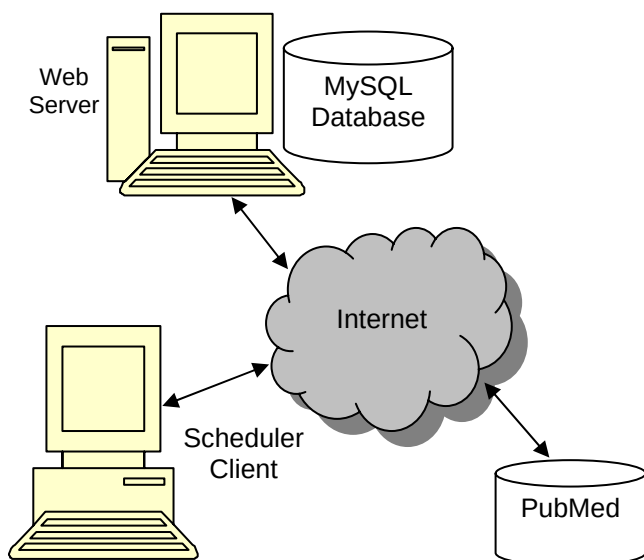


a single medical journal [14]. How PubMed citations can be used to track HMR funding in Australia was another publication using bibliometric methods to track expenditure [15]. In the debate over whether the academics are losing control over clinical research, the much needed empirical data is worked out using metric analyses [16]

However the continuous updating of bibliographic databases results in bibliometric studies that will be out of date no sooner the number of publications is retrieved to complete an article. Our aim is to demonstrate a methodology to auto-update PubMed bibliometric analysis that measures the quantity of publications in tabular and a graphical web-based interface.

## 2. Methods

### Architecture of the Dynamic PubMed Graphing System



**Figure 1 - Architecture of the Dynamic PubMed Graphing System**

The basic architecture of the Dynamic PubMed Graphing System is illustrated in Figure 1. The system consists of three main components:

1. The PubMed online database.
2. A web server that is used to deliver the dynamic graphs over the web.
3. A scheduler client that updates the graphs at scheduled time intervals.

PubMed publication consists of many so-called 'tags' that are abbreviated names for the different fields. The tags are structured in a way that is similar to a structured abstract of a journal article but they are more comprehensive. The fields that were used in this analysis

are: author affiliation (AD), publication type (PT), medical subject heading (MeSH), title (TI) and all fields [ALL].

The AD tag is the field that includes institutional affiliation and address of the first author. The information in this field was used to obtain the publications from Australia. We searched for the word 'Australia' and also for all state and territory names in the AD tag. The AD tag was also used to count the publications originating from the universities. The PT tag was used to pick out a specific publication type - 'clinical trials'.

The MeSH tags are the main key words of the Medline database. Every year the MeSH words are revised - some are added and others deleted. Currently Medline has about 22,997 MeSH words. The specificity of PubMed can be increased when MeSH words are used to query Medline

The PubMed online database is the source for all the main information presented in the graphs generated by the Dynamic PubMed Graphing System. The e-utils service [17] provided by PubMed was used to query the database. The EGQuery option of the e-utils suite returns the count of records found in each and every database in the NCBI. A query sent to the EGQuery interface formulated as a HTTP post request results in an XML output with the following format:

```
<?xml version="1.0"?>
.....
<Result>
  <Term>stem cells</Term>
  <eGQueryResult>
    <ResultItem>
      <DbName>pubmed</DbName>
      <MenuName>PubMed</MenuName>
      <Count>151496</Count>
      <Status>0k</Status>
    </ResultItem>
    .....
  </eGQueryResult>
</Result>
```

An instance of <ResultItem> element is generated for each database entry. The <Count> element carries the number of citations in that database and the Status entity gives a feedback as to what resulted from the query of that database.

The queries that need to be executed are stored in a MySQL database located in the web server. This database also stores the queried results and other details for generating the graphs. Scripts in the web server are used to generate the dynamic graphs.

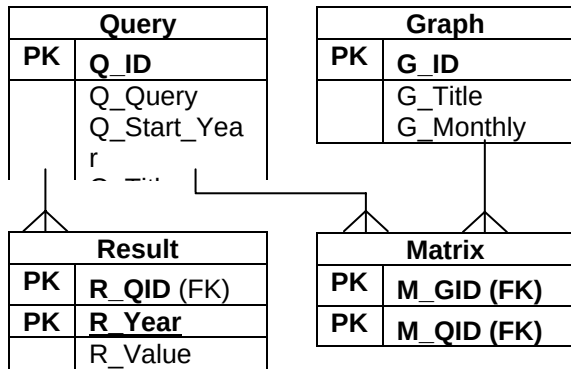
An online PC is used as the scheduler client. This machine has several java applications that are executed at scheduled time intervals to update the system. The Java applications query the MySQL database located at







the web server to retrieve the queries that need to be executed. It then formulates the automated queries to be sent to the e-utils interface with the correct time intervals. The resulting data are stored back in the MySQL database. Once all the queries have been



**Figure 2. Database structure for the Dynamic PubMed Graphing System.**

executed, the Java application executes the scripts in the web server to update the graphs with the new information.

The system is designed to use the results from the same query in multiple graphs. To achieve this, the database was designed to have separate tables for the queries and the graphs. A linking table was used to connect the queries to the graphs. This architecture eliminates the requirement of updating the same query repeatedly from PubMed.



**Figure 3. Query Editor of the Dynamic PubMed Graphing System**

A screenshot of a typical dynamically update graph is illustrated in Figure 4. The generation of the graphs is done using a PHP script in the web server. In order to reduce the load on the server on generating the graphs for each and every request for a graph, the graphs and the relevant pages are generated only once after the database has been updated. Static images and HTML pages are created by a PHP script in the server that is called by the Java application that updates the database.

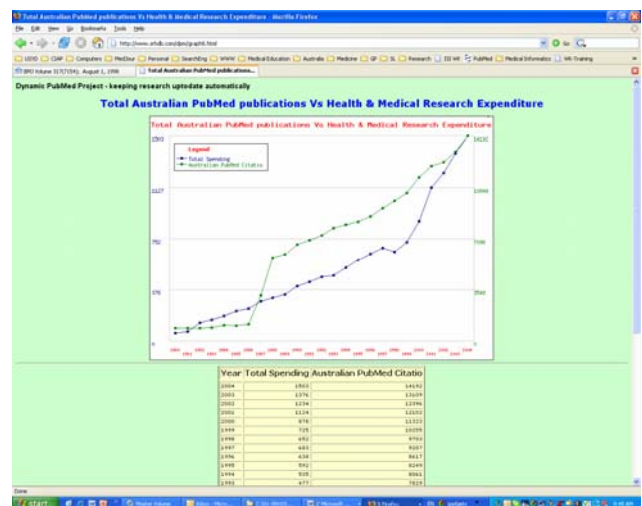
The structure of the database is illustrated in Figure 2. The query table stores the details of the queries that will be used to update the graphs. The query is stored in the format it is presented to the PubMed e-utils interface except for the date range. The starting year and a title for the query are also stored in the table.

The graph table carries information about the graph. This includes the title of the graph and an indication whether it is a monthly graph or a yearly graph.

The matrix table is used to add multiple queries to a graph. The Dynamic PubMed Graphing System can display results from up to 5 queries in a single graph.

The results from the queries are stored in the result table. A similar table is used to store the results from the monthly queries.

A web based administration interface was also developed to add and modify the queries and the graphs that would be used in the system. A screenshot of adding a new query to the system is illustrated in Figure 3.



**Figure 4 A typical dynamically generated chart with part of the table**

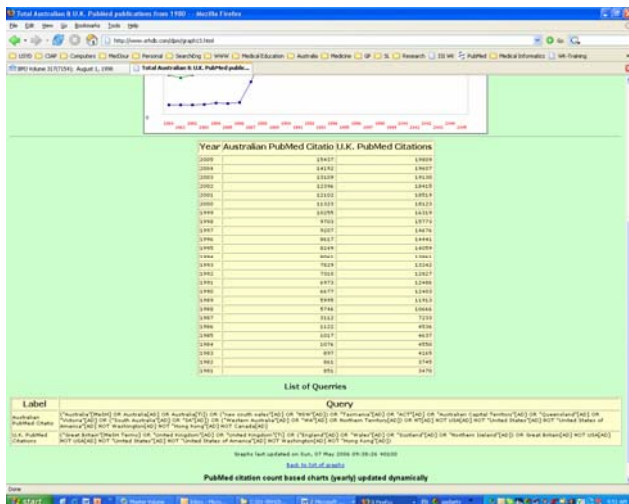
### 3. Results

Typical Dynamic PubMed web pages include: (Fig 4)

- The query or more than one query that was used to produce the results.
- The results in Tabular and graphical form.
- The date and time of last update.

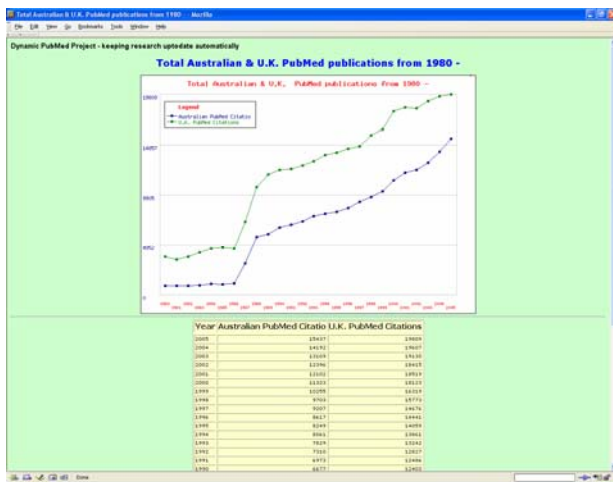
Dynamic PubMed can also include data that do not get updated regularly - for e.g. the annual HMR non-governmental expenditure of Australia alongside dynamically updating data. These can be manually updated into the MySQL tables periodically.





**Figure 5** A typical dynamically generated table, related query and time of last update (with part of the chart)

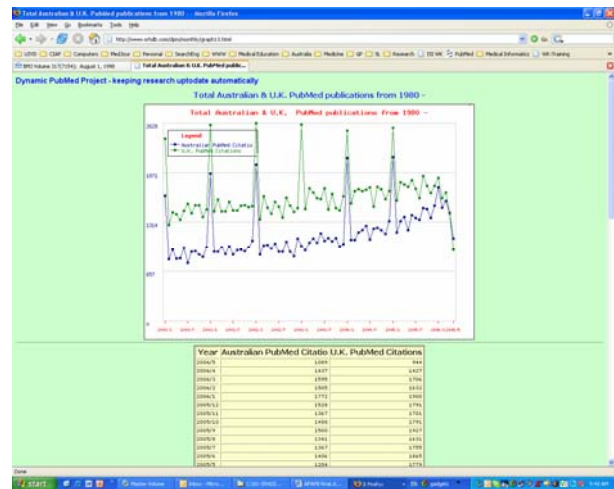
We have also designed to have two types of charts - yearly and monthly. The annual update (for e.g. between two countries publications) gives a long-term trends (Fig 6) and the monthly update show the current trends (Fig. 7).



**Figure 6** Total annual Australian and U.K. PubMed publications from 1960 - 2006

#### 4. Discussion

Dynamic-PubMed software can automatically run a PubMed query at a given time and frequency and write the publication counts to a relational database. This data is displayed as a line charts on a web page. In addition the results are also shown in tabular form. The results of a query can be displayed in two modes – a long-term summary over more than a decade and a more detailed monthly summary to view the current variation in trends.



**Figure 7** Total monthly Australian and U.K. PubMed publications from 2000 - 2006

#### Problems encountered during implementation

During the implementation of the system several problems were encountered. These included problems associated with access to the Internet and problems in the PubMed server. Some of these problems and the suggested remedies are as follows:

##### 1) Accessing the e-utilities interface

Since the web service provider did not allow accessing the Internet from the web server, the system had to be broken down into two parts. One to actually display the graphs on the web server and the other to dynamically query the PubMed database via the e-utilities interface.

The solution adopted was to develop the part that updated the graphs as a separate Java application. The Java application would access the e-utilities interface and update the MySQL database in the web server. The Java application was installed on a PC and scheduled to run daily at a pre-determined time interval that was recommended by PubMed.

The Java application was developed in such a way that the queries to the e-utilities interface were done at intervals of three seconds as recommended by PubMed. This was to avoid the application from flooding the e-utilities interface with queries.

##### 2) HTTP errors from e-utilities

While accessing the e-utilities interface it was noted that at times the result from the EGQuery interface did not work. A detailed analysis showed that the interface was returning a HTTP 500 error. The code corresponds to an internal error in the system where the software responsible for generating the dynamic results has generated an error.

As there was no valid count that could be generated from a page containing an HTTP 500 error, the Java application was modified to detect these errors and retry the query until a valid result was returned.





Recently several of the graphs started showing incomplete results and a further investigation revealed that the EGQuery interface was now generating some HTTP 502 errors. These errors occur when the server gets overloaded. The error is reported from the proxy or a load balancer that may be sitting in front of the web server.

Similar to the HTTP 500 error, the Java application was modified to detect this error and retry the query after a short time interval.

### 3) Reducing the number of queries

It was noted that citations that were older than about five years were rarely getting updated. Therefore, it was unnecessary to update these values daily. Some of the graphs had queries dating back to 1961 and these queries seldom returned a modified value. In order to reduce the number of queries sent to e-utilities it was decided to update the values of queries related to citations older than year 2000 only once a month.

A new field was added to the query table to record the last full update. If the last full update was done more than a month ago, the query will be executed for the full date range and from year 2000 onwards otherwise.

The above modification resulted in a significant reduction in the number of queries sent to PubMed and an increase in the performance.

### 4) Port blockings

Due to the security setups in the academic network it was not possible to implement the system initially. The web access had to be done through a proxy server and access to the MySQL database on the web server was not possible at all. These had to be resolved with the assistance from the network administrators.

### PubMed and Bibliometrics

PubMed was used as it is one of the largest biomedical bibliographic databases in the public domain and facilitates access to the 'raw' data of indexed biomedical publications. PubMed improved access to the database recently by enhancing the functionality of e-utilities with code examples to the ENTEZ engine [18]. This makes it easier when designing software.

Bibliometric assessment of research performance is based on one central assumption - 'scientists who have to something worthwhile to contribute publish their findings in the international journals'. One of the largest biomedical knowledgebase in public domain is PubMed and the high visibility of Australian [19] publications in international biomedical literature make this combination suitable for Australia. Furthermore bibliometric analyses performed at the macro-level (e.g., a whole country) yield at best general assessments of fields as a whole [20], for instance, how good a country's performance is in physics, chemistry, psychology or immunology, without a reliable breakdown to the individual research groups or programs.

Querying bibliographic databases is a reliable way of retrieving information. More than three quarters of the studies included in systematic reviews are identified through electronic bibliographic databases that are searched using Boolean logic [21]. However, the limitations imposed by the absence of important journals and papers from a single database analysis such as PubMed may be of concern.

Most health-related bibliometric analyses use the ISI databases and particularly Impact Factor related metrics. However, the impact factor has its own critics [22]. Prevention Research Centre is one of the Centers for Disease Control (CDC) largest funded programs [23]. The PRC's influence factor (the number of PRC peer-reviewed publications that considered a journal highly influential) was only weakly correlated with the ISI impact factor. This suggested that the ISI metric and the PRC judgment are not closely aligned.

It has also been pointed out that counting publications (without citations) would not give a good indication of the quality or the impact. It has been difficult anyhow to ascertain the real impact on clinical practice by even tracking highly promising articles. One bibliometric study concluded that [24] 'one in four promising technologies resulted in a published randomized control trial and fewer than one in ten entered routine clinical use within 20 years of the index basic science publication'.

Evaluation of scientific research is crucial with the increase amounts of funding that goes into HMR research. Peer review is typically a predominantly qualitative assessment of research performance and can be influenced by personal prejudices and preferences. Bibliometric indicators discussed here represent the quantitative measurements. But quantitative elements may be also present in peer review, e.g., number of publications in high prestige scientific journals. Conversely, citations given to research work can be seen as judgments, "votes" of colleague-scientists in favour of the work cited.

Bibliometric analyses with PubMed data is dependant on the accuracy provided in the relevant tags. The accuracy of indexing for authorship and country is dependant on the information included on the [AD] tag. This includes the first authors information. The similar state names (Victoria in Australia and Canada) and abbreviations (WA – Western Australia, Washington) can give erroneous counts even at country level. Similarly the [AU] tag has the author name information. The inconsistency of the name submitted to journal article may adversely affect the individual publication counts.

There have been other bibliographic software utilities like BIRS [25] that assists the end users to get better





understanding of their search domain, formulate and expand their search queries, and visualize the bibliographic research results. The Dynamic-PubMed will be just one step the beginning of automating bibliometric analysis.

Dynamic PubMed may be used as a tool that automatically keeps track of the number of PubMed research publications to ascertain both long and short term trends. It will be a good tool to track the results of short term HMR funding at country, institution or specific specialty or domain level.

## References

- [1] The Hon Tony Abbott MHR - Media Releases. \$10 million boost for health research- 21/03/2006.  
<http://www.health.gov.au/internet/ministers/publishing.nsf/Content/health-mediarele-yr2006-ta-abb033.htm?OpenDocument&yr=2006&mt=3> (Accessed 14-05-2006)
- [2] Improving the nation's health, increasing the nation's wealth: A new deal for medical research HM Treasury.  
[http://www.hmtreasury.gov.uk/newsroom\\_and\\_speeches/press/2005/press\\_100\\_05.cfm](http://www.hmtreasury.gov.uk/newsroom_and_speeches/press/2005/press_100_05.cfm)  
(Accessed 14-05-2006)
- [3] Shewan LG, Coats AJ. The Research Quality Framework and its implications for health and medical research: time to take stock? *Med J Aust.* 2006 May 1;184(9):463-6
- [4] Hobbs FDR, Stewart PM. How should we rate research? *BMJ* 2006;332:983-984
- [5] Haynes RB, Hayward RS, Lomas J. Bridges between health care research evidence and clinical practice. *J Am Med Inform Assoc.* 1995 Nov-Dec;2 (6):342-50.
- [6] Sharon E. Straus, W. Scott Richardson, Paul Glasziou, R. Brian Haynes. *Evidence Based Medicine* 3rd Edition. Churchill Livingstone. 2005
- [7] PubMed. <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi>  
(Accessed on 15-05-2006)
- [8] PsycINFO  
<http://www.apa.org/psycinfo/> (Accessed on 15-5-2006)
- [9] Institute of Scientific Information <http://www.isinet.com/>  
(Accessed on 15-5-2006)
- [10] MEDLINE – Fact sheet  
<http://www.nlm.nih.gov/pubs/factsheets/medline.html>  
(Accessed on 15-05-2006)
- [11] National Library of Medicine. Medical Subject Headings.  
[http://www.nlm.nih.gov/cgi/mesh/2006/MB\\_cgi?mode=&index=15026&field=all&HM=&II=&PA=&form=&input=](http://www.nlm.nih.gov/cgi/mesh/2006/MB_cgi?mode=&index=15026&field=all&HM=&II=&PA=&form=&input=)  
(Accessed on 15-05-2006)
- [12] King D. Scientific impact of nations. *Nature.* VOL 430 15 JULY 2004
- [13] Falagas MF, Papastamatakis PA, Bliziotis IA. A bibliometric analysis of research productivity in Parasitology by different world regions during a 9-year period (1995–2003). *BMC Infectious Diseases* 2006, 6:56
- [14] Mendis K, Solagarchchi I, Weerabaddana C. Three decades of the Ceylon Medical Journal— analysis using MEDLINE (PubMed). *Ceylon Medical Journal.* Vol. 50. No. 1, March 2005
- [15] Mendis K, McLean R. PubMed citations as an Index of Outcome for Research: Monitoring the Response to Australian Health and Medical Research Expenditure. (Accepted for publication in the *Medical Journal of Australia*)
- [16] Patsopoulos NA, Ioannidis JPA, Analatos AA. Origin and funding of the most frequently cited papers in medicine: database analysis. *BMJ* 2006;332:1061-1064
- [17] Sayers E, Wheeler D. Building Customized Data Pipelines Using the Entrez Programming Utilities (eUtils).  
<http://www.ncbi.nlm.nih.gov/books/bv.fcgi?rid=courswork.chapter.eutils> (Accessed 14-05-2006)
- [18] ENTEZ The life sciences search engine.  
<http://www.ncbi.nlm.nih.gov/gquery/gquery.fcgi>  
(Accessed on 15-05-2006)
- [19] Butler L. Explaining Australia's increased share of ISI publications-the effects of a funding formula based on publication counts. *Research Policy* 2003; 32: 143-155
- [20] van Raan AFJ. The use of bibliometric analysis in research performance assessment and monitoring of interdisciplinary scientific developments. *Technikfolgenabschätzung -Theorie und Praxis Jahrgang - März 2003, S. 20-29.*  
<http://www.itas.fzk.de/tatup/031/raan03a.htm>  
(Accessed on 15-05-2006)
- [21] Sampson M, Barrowman NJ, Moher D, Clifford TJ, Platt RW, Morrison A, Klassen TP, Zhang L. Can electronic search engines optimize screening of search results in systematic reviews: an empirical study. *BMC Med Res Methodol.* 2006 Feb 24;6:7.
- [22] Lundberg G. The "omnipotent" Science Citation Index Impact Factor *Med J Aust* 2003 178: 253-254
- [23] Franks AL, Simoes EJ, Singh R, Sajor Gray B. Assessing prevention research impact a bibliometric analysis. *Am J Prev Med.* 2006 Mar;30(3):211-6.
- [24] Contopoulos-Ioannidis DG, Ntzani E, Ioannidis JP. Translation of highly promising basic science research into clinical applications, *Am J Med* 2003; 114: 477–484
- [25] Ding Y, Chowdhury CG, Foo S, Qian W. Bibliometric Information Retrieval System (BIRS): A Web Search Interface Utilizing Bibliometric. *Journal of the american society for information science.* 51(13):1190–1204, 2000





# Visualizing the Health Resources of Taiwan: the application of Web-GIS

Po-Huang Chiang

Chi-Pang Wen

Yaw-Tang Shih

Correspondence: [chiangp@nhri.org.tw](mailto:chiangp@nhri.org.tw)

## Abstract

To understand the health resources allocation and to discover the health disparity issues in Taiwan, we developed a web-based GIS application system in National Health Research Institutes (NHRI) which provides flexible way to investigate the health resources locally. We use the Area Resource File (ARF) of the township level of Taiwan from the government website and official reports. The Area Resource file contains the data from 369 townships or 25 counties since 1996. It includes demographic, medical services and health outcomes information. We create shape file for health resource areas and to disseminate the underlying data and maps over Internet by using ArcIMS from ESRI.

The results of interactive ARF system provide significant information discovery. It not only allows researchers to study the health resource allocation locally, but also allows them to examine the health and environmental disparity issues through the system.

### Keyword :

geographic information system、Internet、Web-GIS、Area Resource File、public health informatics

## Introduction

To fast improve our understanding of health factors and mechanisms that affect our health is a fundamental necessity for planning appropriate health policy and preventive interventions. The elucidation of health issues is beset by well design health research with adequate health data. The research design tries to combines health demographic, services and outcome information integration by using Geographic Information System and creates effective information presentation system to deliver the powerful policy solution available for building sustainable, flexible infrastructures to power the real-time health policy research task. Research teams can discover and share information assets as well as seamlessly present those information assets throughout the health community. .

The Health Data Team at Center for Health Policy Research (CHPR) and Development, National Health

Research Institutes (NHRI) conducts research in the field of health policy by providing innovative evidence-based information and systems and is dedicated to the generation of new knowledge about the nature of health information problems, the development of new information resources and health policy analysis technique and tools to support evidence-based health research [1], and the evaluation of various innovations in overcoming health information problems.

This study tries to combine health Area Resource File (ARF) of Taiwan which was collected from the government website and official reports and create new healthcare information presentation system for policy analysis. Geographic Information System (GIS) was chosen since it provides a powerful, logical, and intuitive means to store, manipulate, and retrieve data [2-3]. It provides the ability to visualize in a map form, only those features or objects that meet specific selection criteria. Through web GIS technology, you can visually identify features and a geographic representation in an instant. Interest in Web GIS application in health fields has grown considerably over the past few years. Recent advances in the application of web GIS technology have improved, and will continue to revolutionize the spatial analysis of diseases, environmental contamination, and social/demographic information. Many articles focus on primary care [4-7], monitoring [8], surveillance[9-10], eliminate health disparity [11], surveying [12] and health resources searching [13]. The growing availability of health, demographic, and environmental databases containing local, regional, national, and international information are propelling major advances in the use of GIS and computer mapping.

## Materials and Methods

The county and township specific Area Resource File (ARF) from the project of “International Collaborative Network for Health Policy Research” has been collected available public and secondary data sets for all researchers at the CHPR since 2003. The mission is to integrate different data sources and construct a comprehensive data base for research use. Types of data sets collected by the data team were



including government statistics report, vital statistics, national insurance claims data and survey or Census information. (Figure 1), (Table 1).

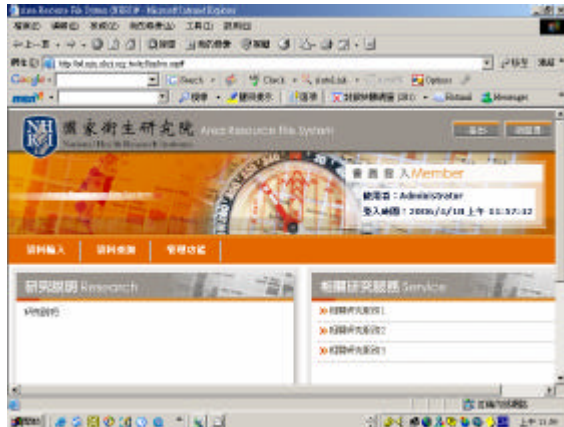


Figure 1. ARF system

Table 1: Contents of ARF

|                                                        |                                                     |
|--------------------------------------------------------|-----------------------------------------------------|
| <b>Summary reports</b>                                 | Statistics Abstract/Statistics Book for 25 counties |
|                                                        | Social Indicators, R.O.C                            |
|                                                        | DOH statistics reports and others                   |
| <b>Population statistics</b>                           |                                                     |
| <b>Mortality data</b>                                  |                                                     |
| <b>National Health Insurance claims data</b>           |                                                     |
| <b>2002 KAP (Knowledge, Attitude, Practice) survey</b> |                                                     |
| <b>2005 National Health Interview Survey</b>           |                                                     |
| <b>Others</b>                                          |                                                     |

We used the ESRI ArcIMS 9 for our development. Its architecture and functionality have been engineered specifically to publish maps, data, and metadata on the Web. The software is designed so that it is easy to create maps, develop Web pages that communicate with the maps, and administer a Web mapping site [14]. ArcIMS is a client-server distributed system. On the client side, these custom viewers, HTML Viewer and Java Viewer allow visitors accessing Web site to view high-quality, interactive maps that feature tools including seamless pan, dynamic zoom, MapTips, and keyboard shortcuts. The Server side includes ArcIMS Spatial Server, ArcIMS Application Server, ArcIMS Application Server Connector and ArcIMS Manager, provide the capability to publish a map onto Internet. The ArcIMS Application Server runs as a background process and handles the load distribution of incoming

requests. It also catalogs which services are running on which ArcIMS Spatial Servers. Using this information, the Application Server dispatches an incoming request to the appropriate Spatial Server. Application Server Connector is used to connect between Web Server and ArcIMS Application Server which provides different kinds of format, including Servlet Connector, Cold Fusion Connector, Java Connector and ActiveX Connector. ArcIMS Manager consists Author, Design and Administrator which can provide the ability to create and manage your application. ArcXML is the protocol for communicating with the ArcIMS Spatial Server. An ArcIMS Spatial Server is the backbone of ArcIMS and provides the functional capabilities for accessing and bundling maps and data into the appropriate format before sending the data back to a client. The ArcIMS software is also designed to be distributed across a network and to be scalable as the demand for maps increases. ArcIMS has a multi-tier architecture consisting of presentation, business logic, and data tiers. In addition, ArcIMS has a set of applications for managing a Web mapping site. Overview of the ArcIMS architecture sees Figure 2.

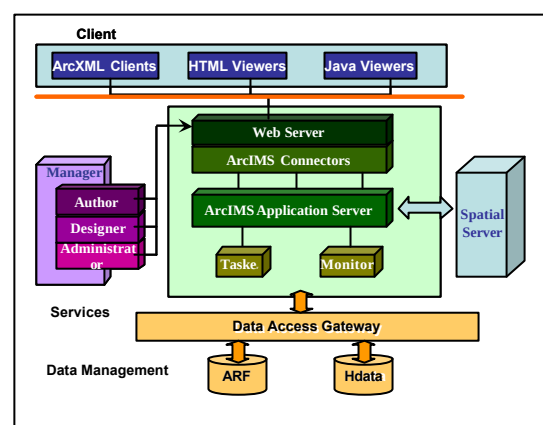


Figure 2. Architecture of ArcIMS

## Results

The Health Resources Internet Mapping system is initiated by select the County and Township level boundary. The data layer has Demographic, Health/Medical Services and Health Outcome for area information presentation layers (Figure 3). Detail selection can be chosen through select toolbar (Figure 4). The GIS map layer control the active and visible layer to perform an identify and other analytical operation (Figure 5). It also provides facilities and transportation (point) information (ex. Hospital, Public Services, and Station etc.) (Figure 6a, 6b). Several customized tools buttons are located in the menu bare above the interactive map.



Figure 3



Figure 4.



Figure 5.



Figure 6a



Figure 6b

ArcXML is the communicating bridge between client and server side. Data and map can be visualize through the interactive system. In this study, we use HTML for easy access for user across the Internet. (Figure 7a, 7b, 7c)

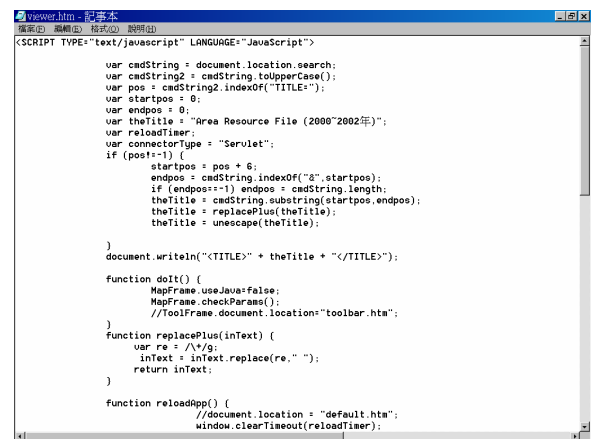


Figure 7a Front page viewer script

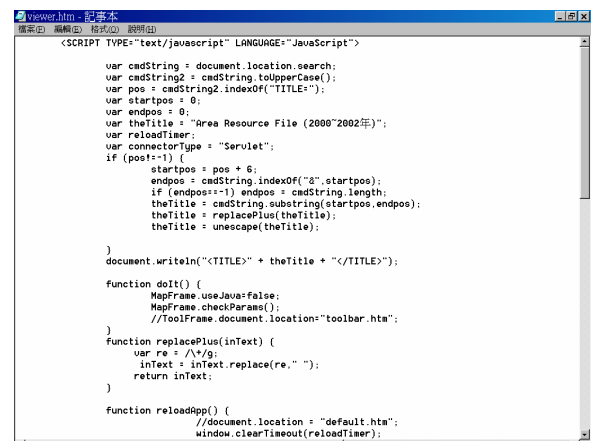


Figure 7b HTML script



```

Functions for sending XML requests and XML responses
=====
// global variables
// change these in aimsCustom.js to send XML response to custom function.
// use numbers >= 1000
var selectXMLMode = 6;
var identifyXMLMode = 7;
var queryXMLMode = 8;
var findXMLMode = 14;
var hyperlinkXMLMode = 15;

// encoding for XML header
var charEncoding = "UTF-8";
var localeEncoding = 'encoding=' + charEncoding + ' ' + ' ';

// charset for dynamic HTML pages - static pages must be changed manually
var charset = "BIG5";

var formColor = "#000000";

// common dynamic variables
var XMLMode = 1;
var okToSend = true;

var xHalf = xDistance/2;
var yHalf = yDistance/2;

// ending position to start parse scan of XML string
var xmlEndPos = 0;

```

Figure 7c Java script

## Discussion and Conclusions

The Web GIS provides a suite of tools for examining the health outcome related to location and distribution of service providers with respect to health and demographic characteristics and other GIS reference layers (e.g., administrative and political boundaries). The data visualization available through this project enables the public, healthcare providers, and policy makers to better understand the distribution of services, and improves the ability of decision-makers to understand the implications of economic, social and demographic factors on the distribution of services.

The system was not able to demonstrate the analysis capability in health disparity issues yet due to the sources for data entry was unavailability and lack of township level data. Some Inconsistency of elements among counties and years was also found in different data sources. The continuous effort for better data quality of ARF was needed. Improved understanding the health issues will come from an abundant and complete health database, then better health care and safety medical service can be reached.

The Health Data Team at Center for Health Policy Research and Development, National Health Research Institutes (NHRI) conducts research in the field of health policy by providing innovative evidence-based information and systems and is dedicated to the generation of new knowledge about the nature of health information problems, the development of new information resources and health policy analysis technique and tools to support evidence-based health research, and the evaluation of various innovations in overcoming health information problems. Other research topic like community resource accessibility can also be conducted if the township level data are available. Each sub-unit represent by its population

weighted centroid (the centre of population in the area rather than the geometric centroid) and the travel time taken to each community resource (for example, a hospital) along the road network can be calculated using the network functionality in ArcInfo GIS. There are so many potential applications for us to explore and benefit the community.

With this integrate new design of healthy community at their disposal, researchers will be able to search for existing health care information, identify needed population, develop new economic welfare policy research questions and seamlessly integrate information assets throughout the policy research society. Through the establishment of information alternation platform, the health ARF, National Health Insurance Database, along with other health survey and government census data, provides a unique opportunity to examine health issues and the health industry as a whole. Health researchers and managers can focus not only on health care utilization and expenditure but also on social and economical issues. Medial practitioners can identify associations between health events and drugs or devices.

The elucidation of health issues is beset by well design health research with adequate and accurate health data. Our goal of national health is hindered by lack of secure and effective information sharing systems and the clearness of the premise and provisions. The area of health informatics is being revolutionized by new technological advancement in data exchange and internet development. Improved understanding of health issues will be a direct result from an abundant and complete health database. Well-designed health research along with adequate and accurate health data will assist improving our health care delivery system and the overall health of Taiwan.

## Acknowledgement

This project is administered by the Centre for Health Policy Research and Development, NHRI. Part of the research reported in this paper was supported by a Department Of Health grant of National Health Information Infrastructure (NHII) to NHRI. We would also like to thank Dr. Der-Ming Liou, National Yang-Ming University, Ms. Shu-Fang Shih, Ms. Shu-I Chiu, Miss Hsiao-Lei Chen, Miss Hsiao-Hui Chen and Mr. Ta-Chien Chan, Center for Health Policy Research and Development, NHRI, for their support to this study.

## References

1. Waters E, Doyle J. "Evidence-based public health: Cochrane update." *J Public Health Med.* 2003; 25:72–5.
2. Croner CM. "Public health, GIS, and the Internet. *Annu Rev Public Health.*" 2003; 24:57–82.
3. Richards TB, Croner CM, Rushton G, Brown CK, Fowler L. "Geographic information systems and public health: mapping the future. *Public Health.*" Rep. 1999; 114:359–73.
4. Luo W, Wang F, Douglass C., "Temporal Changes of Access to Primary Health Care in Illinois (1990–2000) and Policy Implications", *Journal of Medical Systems* 2004; 28:287-99.
5. Margarethe Theseira, "Using Internet GIS technology for sharing health and health related data for West Midlands Region", *Health & Place* 2002; 8:37-46.
6. David C, Goodman DC, Mick SS, Bott D, Stukel T, Chang C, Marth N, Poage J, Carretta HJ. "Primary Care Service Areas: A New Tool for The Evaluation of Primary Care Services." *Health Services Research.* 2003; 38:287–309.
7. Higgs G, Richards W. "The use of geographical information systems in examining variations in sociodemographic profiles of dental practice catchments: a case study of a Swansea practice." *Prim Dent Care.* 2002; 9:63–9.
8. Dimitrios Ptochos, Dimitrios Panopoulos, Kostas Metaxiotis, Dimitrios Askounis, John Psarras. "Using internet GIS technology for early warning, response and controlling the quality of the public health sector." *International Journal of Electronic Healthcare (IJEH)*, Vol. 1, No. 1, 2004.
9. Maged N Kamel Boulos. "Towards evidence-based, GIS-driven national spatial health information infrastructure and surveillance services in the United Kingdom" *Int J Health Geogr.* 2004; 3: 1.
10. Boulos MN. "Towards evidence-based, GIS-driven national spatial health information infrastructure and surveillance services in the United Kingdom." *Int J Health Geogr.* 2004;3:1
11. Christopher L. Fulcher, Catherine E. Kaukinen. "Visualizing the Infrastructure of US Healthcare Using Internet GIS: A Community Health Informatics Approach for Reducing Health Disparities." *MEDINFO 2004.* Amsterdam: IOS Press.
12. Smith DP, Gould MI, Higgs G. "(Re) surveying the uses of Geographical Information Systems in Health Authorities." 1991–2001. *Area.* 2003; 35:74–83.
13. 網際網路醫療資源地理資訊查詢系統. Available from URL: <http://203.65.100.159/dohgis/> (Accessed: August 19, 2006)
14. An ESRI white paper, May 2005. ArcIMS 9 Architecture and functionality. Available from URL: [http://www.esri.com/library/whitepapers/pdfs/arc\\_ims9-architecture.pdf](http://www.esri.com/library/whitepapers/pdfs/arc_ims9-architecture.pdf) (Accessed: August 19,2006)
15. 張郁生, "ESRI ArcGIS 地理資訊系統簡介", 國土資訊系統通訊:49 Available from URL: <http://ngis.moi.gov.tw/TheFiles/journal/49/008.htm> . (Accessed: August 19,2006)
16. 蔡博文, "GIS 作業平台功能簡介", 國土資訊系統網站. Available from URL: [http://ngis.moi.gov.tw/trainfrom/10202\\_1.aspx?class\\_seq=6](http://ngis.moi.gov.tw/trainfrom/10202_1.aspx?class_seq=6) . (Accessed: August 19,2006)



## A forecast modelling and decision support system for vaccine demand Decision Making

**Chen k. Feng(author)**

Fu Jen Catholic University, Taiwan

[leo0803@im.fju.edu.tw](mailto:leo0803@im.fju.edu.tw)

**Ruey k. Chiu (corresponding author)**

Fu Jen Catholic University, Taiwan

[im1004@mails.fju.edu.tw](mailto:im1004@mails.fju.edu.tw)

### Abstract

*The objective of this research aims to propose a computer-based forecast and decision support system to forecast and allocate the annual vaccine demand for specific vaccines of Taiwan so that the governmental immunization authority can take this result to make a better decision for budgeting and purchasing for these annual vaccine requirement.*

*The research firstly uses the models of ARIMA and Neural Network, respectively to forecast the relative variables for these specific vaccines. The results generated by these two models are then further use to compute the estimation of annual workload and the number of vaccine demand, so that the more suitable and reasonable demand is selected for decision making and vaccine allocation.*

### 1. Introduction

Nowadays, medical computer science was a field of study involved in the application of computer science based algorithms, methodologies, and design processes into medical-based research and clinical practice. Medical informatics is not just the application of computers into the medical field; mostly it involves the integration of computers into medicine. The collection of structured and unstructured data had to be stored in databases and should be available to health professionals for the purpose of best patient care and support better decision, without the loss of confidentiality or security of information. Information technologies (IT) and computer-based decision support system may be a strategic tool to improve the quality of decision and patient safety, and so forth.

Many countries have already devoted to develop forecast models of vaccine and put in many resources including human and capital. A medical research team concluding Longini etc.(2004)built a simulation model to simulate the condition of flu virus, and this model can suggest executive how many antiviral agents you must prepare for your population. In Europe many countries adopt this research result and prepare their strategic resources against H5N1. Another team concluding Viboud etc.(2003) use a method of analogues to predict the incidence of FLU spread. This method is to use a

time series as input and compute the similarity with past data to determine the incidence of current period. These medical research teams can predict trend of epidemic basic on computer science objective information. It can make the quality of decision better and reduce the mistakes of insufficient of data.

Until now, the immunization affair of Taiwan's Center for Disease Control (CDC) hasn't yet adopted a computer-based model and method to determine its annual demand of national immunization vaccines. Instead, they take an experience rule by referring to the previous supplied volume and current population growing rate to estimate the demand for its variety of vaccines. In this research, we firstly build a decision model to forecast the demand for DPT and OPV vaccines by applying the forecast models built on ARIMA and Neural Network, respectively. Next, we apply these values and the calculation formula of annual vaccine demand released by World Health Organization (WHO) to calculate the annual demands for each vaccine to provide a better information for the decision making of government's immunization authority of Taiwan to conduct its annual vaccine purchasing. By the aids of these decision models to estimate the volume of annual vaccine demand and to conduct the vaccine allocation, we expect the annual cost of expenditures for purchasing vaccines can be under well controlled and the waste of vaccines can be reduced as well.

### 2. Literature Reviews and Research Adoption

#### 2.1. Impact Factors of Vaccine Demand

Regarding to the research of vaccine demand for the national immunization program of Taiwan's government, in the past only Chen (2003) brought up an optimal purchasing forecast model for the national vaccine demand. Chen (2003) indicated that the impact factors of determining vaccine demand could be referred from Alexander (2002) who used the factors to estimate vaccine usage in the Republic of Ukraine. These factors are vaccine-preventable disease trends, public commitment to infectious disease control and prevention, national immunization schedule and mandated coverage, target vaccine-inoculation population size, vaccine wastage rates, and so on. Chen







established his significant impact factors used for conducting the forecast by using a group appraisal and decision method. He not only considered the factors mentioned in Alexander's research but also added the net migration of population and actual vaccine usage over the past four or more years to help to forecast the annual vaccine purchase volume. Chiu and Kuo (2006) further adopted these decision factors and added birth cohort, vaccine wastage rate, vaccination complete rate, and so on to conduct the forecast for the annual demand of Diphtheria Tetanus Pertussis (DTP), Measles Vaccine (MV) and Oral Poliovirus (OPV). The variable used are further classified into vaccine demand relative variables and population growth relative variables. The results of forecast generated were better than Chen(2003).

To choose and decide the sufficient and necessary decision factors to build a forecasting model is always a prior issue of carrying on modeling a decision. In this research, we increase two parts of construct to support our forecast. The one is the construct of population including the birth of infant, the migration of population and the total population. These demographic variables can affect the demand of vaccine so this research would take these data to assist analysis. The other construct is psychology including the level of education, conjugal condition and the number of contract hospital. Zimmerman(2003) indicate that the relationship between level of education and whether inoculated is positive. Nexoe, Kragstun & Sogaard(1999) found that the relationship between whether live with family and whether inoculated is positive. As to number of contract hospital, if the distance to hospital is shorter, the willing to be inoculated is stronger. Base on above reasons, this research try to add these variables to improve the preciseness of forecast

## 2.2 Method of Estimating Vaccine Demand

World Health Organization (WHO, 2002) announced the estimation model of total vaccine demand which could be computed as Formula 1.

$$N = \frac{B \times I \times D \times (1 + R)}{1 - W} - S \quad (1)$$

where,

N = the total amount of vaccine demand in doses for next year.

B = the number of annual birth rate for next year.

I = the expected immunization coverage rate in percentage for next year.

R = the reserve stock in percentage for next year.

W = the wastage rate in percentage for current year.

S = the number of vaccine doses in stock for current year.

As we see from Formula 1, if there is no stock on hand in the first year's estimation, then the value of S is 0. Because the vaccine we discuss here is the immunization vaccine under current statutory policy, we must put the S into consideration. To determine vaccine

demand, we must also consider the expected immunization volume for a period. Furthermore, we must also consider vaccine wastage rate from the past to supplement the vaccine volume which is caused by damage. Lastly it needs to add reservation stock into consideration in case an abnormally unpredictable requirement such a sudden infectious disease outbreak.

Instead of using a fixed  $I \times B \times D \times (1 - C)$  to calculate the total number of inoculation, in this research we can also use a forecast variable J to substitute  $I \times B \times D \times (1 - C)$  in Formula 1 so that we may obtain a new formula which is shown in Formula 2. Hence, the variable J refers to the demand of total number of inoculation for the next year. The value of Variable J is obtained by forecast mainly on the basis of the forecast values of B and I. To deserve to mentioned, C is be added to accept formula into our immunization conditions. Because the life in Taiwan become better, more people are willing to pay money for inoculation by them self in hospital. So replace rate of vaccine at one's won expense (C) is more and more high and must be considered as a factor in demand formula.

$$J = I \times B \times D \times (1 - C) \quad (2)$$

Then, we can use formula 2 to compute the total amount of vaccine demand in doses for next year (N). J can be estimated by forecasting directly or derived from  $I \times B \times D \times (1 - C)$ . We can choose a better estimation and take it into formula 3. Formula 3 can computer the demand of vaccine according to the reserve stock in percentage for next year (R), the wastage rate in percentage for current year (W) and the number of vaccine doses in stock for current year (S).

$$N = \frac{J \times (1 + R)}{1 - W} - S \quad (3)$$

J = the amount of inoculation in doses in DHC

C = replace rate of vaccine at one's won expense.

## 2.3. Forecasting Modeling

There are two modeling methods which are applied for modeling the forecast model of this research. One is the method of time series. The other one is neural networks. The two methods are very often used by the researchers and practitioners to conduct the forecast which are closely related to the time changes of the historical data in the past. Because our intentions to the forecasts of the annual new baby birth, the expected immunization coverage rate, and so on are all related to the time changes, therefore we decide choose these two methods. However, we also desire to select a better result from results generated by these two methods to compute the final volume of annual vaccine demand.

### 2.3.1. Auto-Regressive Integrated Moving Average

Time series model have been develop into many different types. In this research, we adopt the Auto-Regressive Integrated Moving Average (ARIMA) model which was introduced by Box and Jenkins in 1976 and 1994 [Box and Jenkins, 1976; Box, et al., 1994]. As we





understand, ARIMA is also the most commonly used so far among many types of time series models.

The ARIMA is generally referred as auto-regressive integrated moving average model of (p, d, q) which is shorted as ARIMA (p, d, q) for short. The word of "Integrated" refers to if the raw series of data is not stationary, it can use the differencing method to make the data series stationary and then use the stationary data series to build the ARIMA (p, d, q) model.

In this research, we use multiple-variant ARIMA to build the forecast model for vaccines of DPT and OPV. For adding time series variable X to assist subscribing time series Y, Box-Jenkins in 1977 issued transfer function. ARIMA with transfer function can not only deal with the data set of predict variable and input variables but also subscribe their relationship. It can make ARIMA adapt more condition with flexibility to forecast. In the practice of inoculation, the number of inoculation is almost a stationary series that suitable to fit an ARIMA model for predicting, so this research would use ARIMA with its transfer function to build a forecast model to improve the quality of decision for vaccine decision making.

As we can see from formula 4, ARIMA original only subscribes the change of  $y_i$  and fits it as ARMA(p,q). ARIMA with transfer function(ARIMAT) can subscribe the change of  $y_i$  and  $x_{ki}$ , and series of  $x_{ki}$  can explain the part of variation of  $y_i$ . So we can see that after  $y_i$  was differentiated and subtracted by  $u$ , the value can be distribution by  $x_{ki}$  and residual of  $y_i$ .

$$(1-B)^d y_i - u = \sum_{k=1}^{K} \frac{\omega_{kl}(B)B^b}{\delta_{kr}(B)} (1-B)^d x_{ki} + \frac{\theta_q(B)}{\phi_p(B)} e_i \quad (3)$$

$\omega_{kl}$  = the parameter of limited lag

$\delta_{kr}$  = the parameter of limitless lag

$b$  = the level of lag

### 2.3.1. Neural Network

A generic model of neural network mode consists of three types of layers, which are input, hidden, and output layers which are illustrated as Figure 1 (Negnevsky, 2002). Each layer may consist of many artificial neurons or viewed as processing elements. Each processing element is designed to emulate the function analogous to a biological neuron in human brain. The number of neurons in each layer can be as many as tens, hundreds, even thousands depending upon the problem. Again, each neuron can have several different inputs and may send its output to many other neurons. As it is shown in Figure 1, a neuron may be connected by weighted links passing signals from one neuron to another. The output signal is transmitted through the neuron's outgoing connection. The outgoing connection splits into a number of branches that transmit the same signal. The outgoing branches terminate at the incoming connections of other neurons in the network. Hence, neurons are wired up in a 3-

dimensional pattern. The features of each layer are briefly described as follows:

(1) Input Layer: The initial inputs are presented to this layers. The number of processing elements depends on the problem.

(2) Hidden Layers: It shows the interaction among processing elements. There is no standard rule to decide the number of processing elements. In order to find out the best number of elements used in the neural network model, we need to verify each network is generated by repeatedly testing. The neural network can have not only more than one hidden layer but also none of hidden layer.

(3) Output Layer: It shows the output variables which are decided when the network is in convergent state through a number of simulation run. The number of processing element of output also depends on the problem.

Commercial neural networks incorporate three and sometimes four layers, including one or two hidden layers. Each layer can contain from ten to thousand neurons. Experimental neural networks may have five or even six layers, including three or four hidden layers, and utilise millions of neurons.

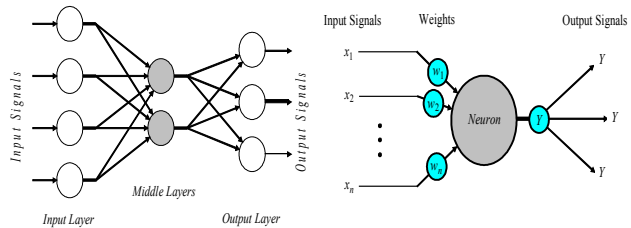


Figure 1. A Generic Model of Back Propagation Neural Network

Neural network model also has many different forms (Fish et al., 1995). Among these forms, researchers and practitioners (Lippmann, 1987; Fish et al., 1995) agreed that nowadays the back propagation neural network (BPNN) is the most representative and universal model in many different models of Neural Networks. BPNN is the part of supervised learning network which is commonly used to build decision model for diagnosing and forecasting. Lippmann (1987) presented a proof that the forecast results by using BPNN are better than using traditional statistics analysis by applying it in several forecasts of real cases.

Neural networks have been used in wide areas such as investment analysis, signature analysis, process control, monitoring, marketing, and various kinds of trend analysis. In this research, we intend to use the Back Propagation neural network model predict the related variable values of forecasting the vaccine demand is a type of trend analysis as time changes of the historical data in the past.





### 3. Research Framework

In this section, we will discuss the framework of this research from the aspects of research method, the decisive variables selected for forecasting, and forecast model and procedure of the research so that the readers can comprehensively understand the process of this research.

#### 3.1 Research Method and Process

##### 3.1.1. Research Method

This research takes forecast models to predict the determining related factors with vaccine demand. In addition, it uses ARIMAT predicting the “total number of inoculation” for next year and uses BPNN predicting the “coverage rate of inoculation”.

As we discuss in the prior sections, the Multi-variant ARIMA model of time series is adopted to build the forecast model to forecast the total doses of demand for certain vaccine in the next year by inputting one variable “the number of baby-birth”.

The workload of immunization in current year is used as a referential value to decide the amount of vaccine demand they purchase. So, this research uses several variables to increase the accuracy of the forecast of workload of immunization in the next year. It obtains great results in many single variable statistic methods (Shieh, 2001).

In order to obtain a better result of forecast, we attempt to combine the advantage of ARIMA and BPNN. BPNN can also accept multiple factors of influencing vaccine demand as the input variables. Lippmann (1987) has shown that the great forecast results by using BPNN with multivariable inputs. And ARIMA can handle single number list great. So we can build a integrate forecast model and base on the property of variables we use different mode to predict to get a better accuracy.

##### 3.1.2. Research Process

We establish single/multiple-dose vaccine demand forecast and optimal allocation model through 4 stages of process, the detail research architecture and procedure are illustrated in Figure 2.

The first stage we use ARIMA, ARIMAT and BPNN models to forecast all related variables about vaccine demand. In ARIMA, we use models to forecast “the numbers of baby-birth” and “replace rate of vaccine”. The result of “the numbers of baby-birth” can be used to build a transfer function to assist ARIMAT to predict “the number of inoculation”. In BPNN, this research intends to use “conjugal condition” and “the number of migration” to be input variables. Then we train BPNN to predict “coverage rate of inoculation”. These predictions will be use to computer the vaccine demand.

The second stage then takes the value generated from the first stage and take into formula 1 and 2. We can obtain two estimations of the amount of inoculation in doses in district health centre(DHC) (J). These two estimations will be compare to actual value and take one

that has a smaller error to vaccine demand formula. After that, We can compute the vaccine demand in next year (N).

The third stage then takes the workload of inoculation of DHC in the past year to get the vaccine demand of DHC ( $Z_i$ ), and each DHC can use his  $Z_i$  to compute the volume of his own vaccine demand next year.

The fourth stage will take vaccine demand forecast result from third stage to compute the vaccine amount of allocation for each DHC and the cost for each centre's allocation and the total cost for the entire nation for this vaccine demand under some specified constraints which is regulated by the rules defined by CDC's immunization management operating authority. The constraints and rules are shown in Figure 2.

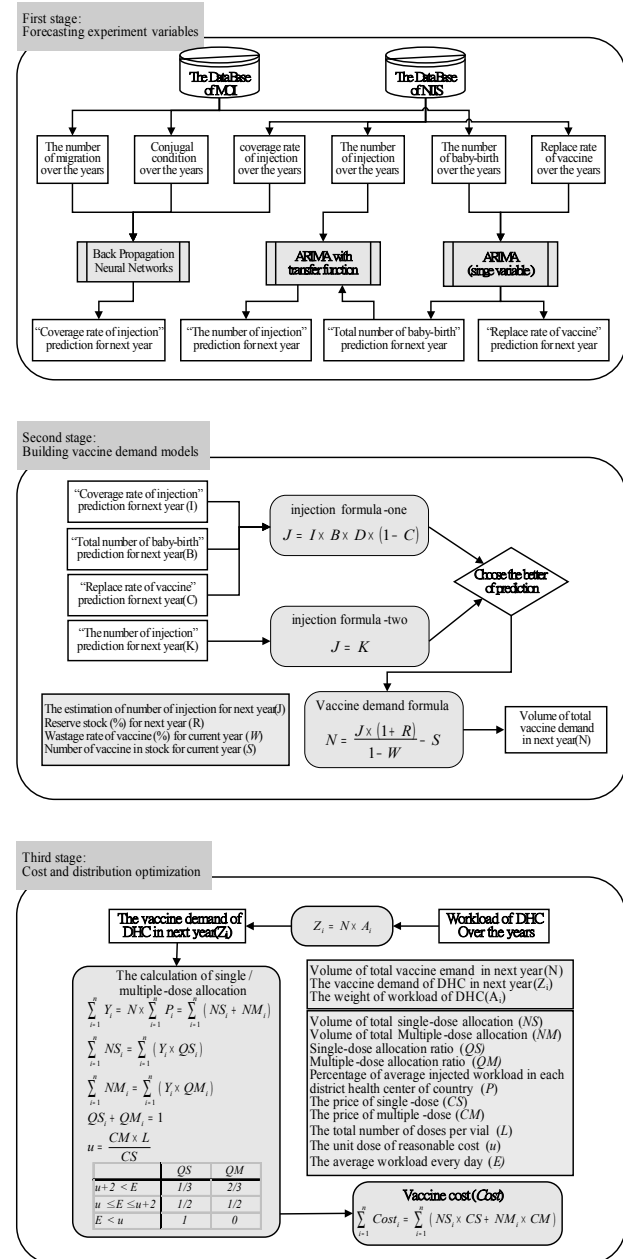


Figure 2. Forecast model and procedure for this research



### 3.2 Decision Variables

This research initially obtains many considerable impact variables which may impact the decision of vaccine demand through initially investigating from various reports and literatures (Alexander, 2002; Chen, 2003; WHO, 2002). This research classified these input variables for forecasting model into major categories which include the people relative factors and the psychology relative factors. These factors are shown and described in Table 1. The predict variable is “the number of inoculation”, these input variables must be filtered before they can be used as the true factors for forecasting. In addition, there are simulation variables used to adjust the number of vaccine purchased base on the vaccine demand of local unit. Each local unit can set these simulate variables including reserved stock, wastage rate and the number of vaccine in stock. It can let local health centers to obtain proper demand based on their individual demands and situations.

We only take the data from the Taipei county, which is the largest one county and also has most population in Taipei. Taipei County not only it is the most populated county in Taiwan but also its geographical environment may be thought as a window of Taiwan which includes the densely populated city areas, thinly populated country areas, and far-reaching mountain areas. It contains twenty-seven district health centres which are dispersedly located around the entire county. We firstly build the model to forecast the demand for each district health centers, and then sum them up to estimate the demand for the entire county. We think if the model can correctly estimate the need of Taipei county, the results shall be able to be further applied for the entire country. Consequently we choose the most populated Taipei County to be the research target to be the experimental county. The data sources relative to the population growth relative factors described in Table 1 used to forecast vaccine demand are extracted from the entire 29 local district health centres in the data base of National Immunization Information System (NIIS) of Taipei County. The relative data of population factors for the past years are solicited from the Ministry of the Interior (MOI) of Taiwan’s government. The time period of data sampling is from 1993 to 2004. We use the data from 1992 to 2003 as the training data to the forecasting model and the data of 2004 as the proofing data.

At first, we investigated the trend of the workload of immunization operation for each vaccine from every local district health centers in Taipei County over the time period from 1993 to 2002. We attempt to find if they have a trend that can be followed so that the forecast can properly be applied. Interestingly, we discovered a very similar trend for different vaccines throughout the twelve months of different years. Among these investigations, we found three vaccines having a stable presentation of workload trend in which are Measles-Mumps-Rubella (MMR). It should belong to

the forecast of historic data trend for the vaccine demand. Hence, in this research we firstly select MMR to be the experimental vaccine to conduct the demand forecast. In our mind, if we can fine a proper forecast model to accurately forecast the demand of total number of inoculation of MMR, then we can apply it to forecast other two kinds of vaccines.

According to our understanding, the MMR normally inoculates only 1 dose for a new-born baby during his/her one year old period. Whereas in each year there exists some makeup MMR inoculation for other ages of children (usually in his/her 1<sup>st</sup> grade of elementary school) who missed the normal inoculation during his/her must-be immunization period and for the other special cases like women at childbearing age. Because these cases of makeup inoculation are few and data collection are not precisely, therefore although the government do not take into account of these inoculators for the annual vaccine immunization coverage rate but we still put them into the consideration of the input sources of annual total number of inoculation estimate the workload of immunization. This is the reason why we set the forecast target on “total number of inoculation” but not “the number of already-inoculated-people” that may be taken from the vaccine immunization coverage database is that especially consider non-new-born baby or other public immunization demand to be more close to the condition of real demand.

|                                      | Decisive Variables                            | Explanations and Data Sources                                                                                                                                                                                                                         |
|--------------------------------------|-----------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Vaccine demand relative variables    | Total number of inoculation                   | It also means workload of vaccine inoculation. The value is obtained from NIIS database and is also verified against the data obtained the workload of vaccine inoculation each month over the years in Taipei county.                                |
|                                      | Immunization coverage rate                    | The value is assumed to be constant. According to the average immunization coverage rate released from the CDC immunization administration office of Taipei county in the last three years, the average coverage rate of DTP vaccine is about 95.62%. |
|                                      | Vaccine wastage rate                          | The value is thought to be constant. Because there is no related records in NIIS database, so we adopt 25% as the general wastage rate of DTP according to WHO’s reports.                                                                             |
|                                      | The number of vaccine in stock                | The value will be the surplus stock in the end of every year, but NIIS is just installed in operation in 2003, so it has no actual record of vaccine in stock before the years of 2002.                                                               |
|                                      | Vaccine reserve stock                         | The value is thought to be constant. Because the policy of Taiwan stipulates it must maintain the stock for being used for three-month period so the percentage of vaccine reserve stock is set to be 25%.                                            |
|                                      | The number of doses per fully immunized child | The value is thought to be constant. A new-born baby must inoculate DTP total four doses which come to 2, 4, 6, 18 months to inoculate one dose per each after being born.                                                                            |
|                                      | The price of single/multiple dose             | The value is thought to be constant. Because the price of vaccine is different every year, so the value will be obtained through CDC of Taiwan to obtain the purchase price of DTP in 2002.                                                           |
| Population growth relative variables | The number of baby-birth                      | The value may be obtained form ministry of interior                                                                                                                                                                                                   |
|                                      | Conjugal condition                            | The value may be obtained form ministry of interior                                                                                                                                                                                                   |
|                                      | The number of migration                       | The value may be obtained form ministry of interior                                                                                                                                                                                                   |

Table 1. The Decisive Variables and Explanations





#### 4. Experimental Result Analysis and Estimation

Through different model predict different variables, we can obtain the result showed table 2. We use ARIMA model to predict “the number of baby-birth” and get a prediction 31202. it has a error rate 4.52%. this is a small error rate for predicting. Moreover, we use “the number of baby-birth” building a transfer function to add the accuracy of ARIMAT of “the amount of inoculation” and take 31202 as input variable to model. The result of predicting of “the amount of inoculation” is 30568 and has error rate 9.42%.

We also use ARIMA model to predict “replace rate”. The result of predicting is 90.56% and has error rate 4.45%. Besides, the variable “coverage rate” use BPNN model to forecast. The BPNN model use two input variables to forecast including “conjugal condition” and “the number of migration”. The result of prediction is 93.47% and has error rate 2.85%. These result can be used to compute the vaccine demand of Taipei county.

| variable      | prediction | error rate | model  |
|---------------|------------|------------|--------|
| baby-birth    | 31202      | 4.52%      | ARIMA  |
| inoculation   | 30568      | 9.42%      | ARIMAT |
| replace rate  | 90.56%     | 4.45%      | ARIMA  |
| coverage rate | 93.47%     | 2.85%      | BPNN   |

Table 2. the result of forecasting of related variables

Using above prediction we can estimate the number of workload of DHC in next year. By formula-one, we can compute the estimation 26,411 and has error rate 5.46% with actual workload in 2004.

$$J = 93.47\% \times 31202 \times 1 \times 90.56\% = 26411$$

In formula-two, we can use the prediction of “the number of inoculation” as our estimation, so the estimation is 30568 and has error rate 9.42%. we can take small error rate, 26411, into our demand formula.

|            | formula-one | formula-two |
|------------|-------------|-------------|
| estimation | 26411       | 30568       |
| error rate | 5.46%       | 9.42%       |

Table 3. the comparison of estimation of formula

$$N = \frac{26411 \times 1.8333}{1 - 0.4} - 21547 = 59152$$

Finally, we compute the vaccine demand in Taipei county in 2004 and the number is 59,152.

When we get the number of vaccine demand, we can distribute the quantity to each DHC according to their workload over the years. Moreover, following single/multi dose vaccine principle as Figure 2, we can get the distribution of single/multi dose listed table 4. The table includes the cost of single/multi dose vaccine.

Cost of single dose vaccine is 3,736,128; Cost of multi dose vaccine is 2,145,459; Total cost is 5,881,587.

|                     | price | amount | cost      | total cost |
|---------------------|-------|--------|-----------|------------|
| single-dose vaccine | 116   | 32,208 | 3,736,128 | 5,881,587  |
| multi-dose vaccine  | 79.6  | 26,953 | 2,145,459 |            |

Table 4. The Result of vaccine distribution

#### 5 Conclusion and Suggestion

In chapter 4 the Experimental Result shows that the error rate of “baby-birth”, “replace rate” and “coverage rate” are under 5%, those are good predictions and have a great accuracy. A great prediction is important for vaccine demand model because accurate prediction can derive effective estimation.

After building the forecasting model to predict the relative variables with vaccine demand, we begin to estimate the annual workload of DHC. The result of research shows the combination of forecast models has a better performance than single forecast model. So formula-one combines three predictions and has a smaller error rate 5.46%. Using the estimation of formula-one can get a vaccine demand 26,411. the number can be used to advice by decision maker.

Finally, this research presented a “single/multi dose vaccine principles” for CDC to decide the distribution of single/multi dose vaccine, which can reduce the wastage of vaccine and optimizing the distribution. Using it, we can compute the cost of vaccine and the total cost of MMR vaccine in Taipei county in 2004 is 5,881,587. This result can be an advice number for CDC.

#### Reference

- [1]Alexander, T. (2002), The National Immunization Program of Ukraine: An Assessment of Performance, Financing and Resource Allocation Options, Ukraine Country Assessment Report, <http://www.synergyaids.com/resources.asp?id=4094>.
- [2]Box, G.E. and Jenkins, G.M. (1976), Time Series Analysis: Forecasting and Control, *Time Series and Digital Processing*, Second Edition, Holden Day Inc., San Francisco, CA.
- [3]Box, G.E., Jenkins, G.M. and Reinsel, G.C. (1994), *Time Series Analysis: Forecasting and Control*, 3rd Edition. Prentice Hall, Inc.







- [4]Chen, C.H., Dolan, Matt D., and Lin, B. (2004), Improve processes on healthcare: current issues and future trends, *Int. J. of Electronic Healthcare*, Vol. 1 No. 2, pp. 149-164.
- [5]Chen, S.Y. (2003), *The Study of Building the Best-Fitting Forecast Model for National Immunization Vaccine Procurement*, Master's Thesis of Institute of Information Management of Fu-Jen Catholic University (unpublished), Taipei, Taiwan.
- [6]Chiu, R.K. and Kuo, C.W. "A decision support model of demand forecast for national immunization vaccines", *Int. J. Electronic Healthcare*, Vol. 2, No. 1, 2006, pp.92-104.
- [7]Fish, K.E., James, H.B. and Milam A. (1995), Artificial Neural Networks: A New Methodology for Industrial Market Segmentation, *Industrial Marketing Management*, Vol.24, pp.431-438.
- [8]PATH Annual Report. (2003),
- [9][http://www.path.org/files/RMER\\_annual\\_report\\_2003.pdf](http://www.path.org/files/RMER_annual_report_2003.pdf).
- [10]Longini, I. M., Elizabeth, M. H., Nizam, A. and Yang, Y.(2004), Containing Pandemic Influenza with Antiviral Agents, *American Journal of Epidemiology*, Vol 159, No 7.
- [11]Lippmann, R.P. (1987), An introduction to computing with neural nets, *IEEE Acoustics, Speech and Signal Processing Magazine*, Vol. 3, No. 4, pp. 4-22, 1987
- [12]Negnevisky, M. (2004), Artificial Intelligence: A Guide to Intelligent Systems, Addison Wesley Publisher Co., An Imprint of Pearson Education, 2<sup>nd</sup> Edition.
- [13] Xexoe, J., kragstrup , J. S Sogard, J.(1999) Decision on FLU vaccination among the elderly: A questionnaire study based on the health belief model and the multidimensional locus of control theory. *Scand j prim Health care*, 17, 105-110.
- [14]Shieh, B.C. (2001), *Introduction and Application of Data Mining*, Information Consultant Corporation Publishing Co., Taipei, Taiwan.
- [15] Viboud, C., Boëlle, P. Y., Carrat, F., Valleron, A., and Flahault A.(2003), Prediction of the Spread of Influenza Epidemics by the Method of Analogues, *American Journal of Epidemiology*, Vol 158, No 10.
- [16]WHO (2002), *Guidelines for Estimating Costs of Introducing New Vaccines into the National Immunization System*, <http://www.who.org/>.
- [17]Zimmerman, R. K., Santibanez, T. A., Janosky, J. E., Fine, M. J., Raymond, M., Wilson, S.A. et al.(2003) What affects FLU vaccination rates among older patients? An analysis from inner-city, suburban, rural, and veterans affairs practices. *The American Journal of Medicine*, 114, 31-38.



# Rasch Analysis Assists BNHI with Revealing Most Misfitting Strings of Hospital Medical Service Information

Tsair-Wei Chien      Ou Lydia Liu      Wen-Chung Wang

Chi-Mei Medical Center      University of California, Berkeley      National Chung Cheng University

The corresponding author: smile@mail.chimei.org.tw

## Abstract

*The Bureau of National Health Insurance (BNHI) in Taiwan provides weekly hospital medical service information to the public. The authors downloaded datasheets from the BNHI website including growth rates of 569 hospitals comparing the counterpart-weeks of Year 2005 and 2004. Rasch analysis (1960) was conducted to estimate hospital growth rate abilities and to calibrate item difficulties using the computer program Winsteps. Results showed that the averaged growth rate of outpatient service among hospitals in Taiwan displayed a negative growth in 2005, either north sub-division bureau or the region hospital showed the most negative growth rate in its group respectively. The most unusual misfitting profile of any hospital generated by Winsteps was demonstrated to show up distorted growth on the specified item (week) and clearly expose the unexpected response in various week of outpatient service for putting more efforts on further detection.*

## 1. Introduction

The Bureau of National Health Insurance (BNHI) in Taiwan weekly publishes hospital medical service information to the public since January of 2005. Rasch analysis of item response theory, which permits traits to be measured by comparisons within tests or within items, may assist BNHI with not only ranking levels of hospitals' growth rate but also revealing most misfitting strings of its service distortion under global budgeting system of health care.

How Rasch[1] analysis can easily and accurately assist BNHI with identifying the most misfitting strings of hospital medical service to the public and hospital managers for warning or tracing overgrowth will be demonstrated in this study.

## 2. Methods/Analysis

### 2.1. Subjects

The study sample was recruited from 569 hospitals' outpatient (outpatient department, OPD) growth rates resulted from the counterpart-weeks of year 2005 and 2004 on website datasheets of BNHI. Effective data were included in the study if they met the following criteria: (1) no missing datum on each compared week; (2) at least 20 effective growth rates calculated in a year.

### 2.2. Procedure

Data were downloaded and collected in MS-Excel to form a 2x2 column-row datasheet representing 52 weeks and 477 hospitals respectively.

Using the concept of questionnaire survey with Likert [9] type multi-scoring scale, we divided each growth rate into 6 categories by the five thresholds of -3.5%, 7.0% and 0%, such as 1 representing less than -7%, 6 being more than 7% on each week's outpatient growth rate.

### 2.3. Measures

The 6-point items constitute the 52-week (indicator) inventory, which was developed to assess hospital outpatient growth activities.

### 2.4. Data Analysis

The analysis consisted of two parts. First, the unidimensional assumption of the evaluation inventory with appropriate categories was examined using Rasch analysis. The WINSTEPS computer program[2] was used to perform the Rasch analysis under the rating scale model.[3,4] Second, Rasch analysis with Winsteps computer program was applied to estimate hospital growth rate abilities, to calibrate item difficulties (i.e. weekly growth rate response), and to identify the most misfitting strings ( $p < .05$ ) for each hospital in order to investigate the most growing hospital and the most abnormal weeks in that hospital.

To examine unidimensionality of the evaluation inventory scale, the infit and outfit statistics were used to indicate whether the data fit the model's expectation. The infit mean square (MNSQ) is sensitive to unexpected behavior affecting responses to items near the hospital's ability measure; the outfit MNSQ is sensitive to unexpected behavior by persons on items far from the hospital's ability level.[4,5] MNSQ can be transformed to a t statistic, termed the standardized Z value (ZSTD), which follows approximately the t or standard normal distribution when the items fit the model's expectation.

In this study items with both infit and outfit MNSQ beyond the range of 0.5 and 1.5 were considered poor fitting. When items fit the model's expectation, the residuals (observed scores minus expected scores) should be randomly distributed. A factor analysis was conducted to verify whether any dominant component existed among the residuals. The unidimensionality



assumption held if no dominant component was found.[6]

### 3. Results

A total of 477 hospitals participated in this study. The characteristics of the study sample are shown in Table 1.

Table 1. Characteristics of the Hospitals (n=477)

| Types    | Med.Cntr | Region     | District    | Others   | Total |
|----------|----------|------------|-------------|----------|-------|
| Taipei   | 7(6.60%) | 15(14.15%) | 82(77.36%)  | 2(1.89%) | 106   |
| Northern | 1(1.64%) | 12(19.67%) | 47(77.05%)  | 1(1.64%) | 61    |
| Central  | 3(3.16%) | 11(11.58%) | 80(84.21%)  | 1(1.05%) | 95    |
| Southern | 3(3.95%) | 13(17.11%) | 54(71.05%)  | 6(7.89%) | 76    |
| Kao-Pin  | 3(2.44%) | 12(9.76%)  | 106(86.18%) | 2(1.63%) | 123   |
| Eastern  | 1(6.25%) | 5(31.25%)  | 9(56.25%)   | 1(6.25%) | 16    |
| Total    | 18       | 68         | 378         | 13       | 477   |

#### 3.1. Unidimensionality of Simplified Scale

When the Rasch model was applied to examine the 52 polytomous items, 9 misfit items (weeks) had outfit MNSQ statistics beyond the range of 0.5 and 1.5, which were statistically significant at the 0.05 level. These 9 items were excluded from the rest of the analyses. A factor analysis on the residuals of Rasch-transformed scores showed no dominant factors. The first and second factors accounted for only 10.4% and 6.7% of the residual variance, respectively. These results indicated that there was a good model-data fit and that the assumption of unidimensionality held for these 43 polytomous items when assessed at this one year comparison of growth rates. We further examined the psychometric properties of the 43-item scale. The person measures (ranging from  $-2.60$  to  $1.62$ ) had a mean of  $-0.35$  and a SD of  $0.63$ . The person reliability was  $0.88$  (which can be similarly interpreted as Cronbach's  $\alpha$ ), indicating that these items yielded reasonably good estimates for the hospitals.

The Figure 1 shows the map of persons and items along the continuum. The item difficulties were spread out across the hospitals, indicating that these items could well differentiate these patients.

The highest and lowest abilities of hospitals in growth rates shown in Figure 1 were the total of upper 7 and bottom 10 representing BNHI branches and hospital levels at first and second byte respectively.

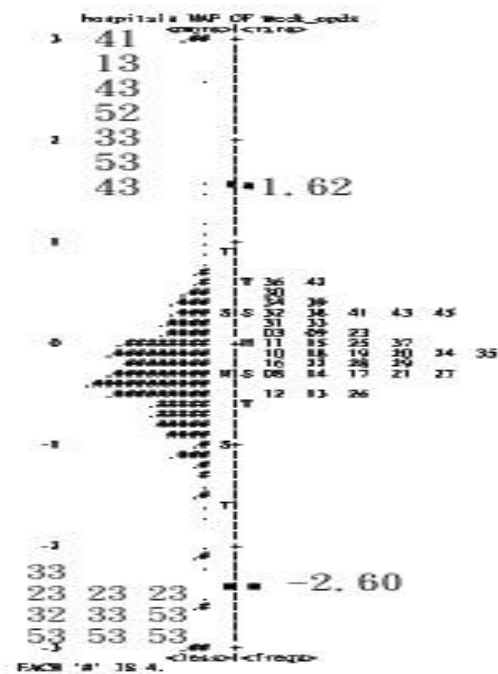


Figure 1. Item and person map under the Rasch model

#### 3.2. Ranks of Hospitals' OPD Growth Rates

The ranks of BNHI's 6 region branches were highest 48.5(Eastern) to lowest 44.2(Northern) after measures in logit( $M=0$ ;  $SD=1$ ) transformed into T-score ( $M=50$ ;  $SD=10$ ) to show up box plots from SPSS statistical software.

A box-plot of figure 3 shows that hospitals in Eastern region branch of BNHI are asymmetry. Some hospitals were shown as outliers or extremes with symbol o or \* respectively in figure 2 and figure 3. Inter-quartile range of box-plot is appropriately set at 1.348 standard deviation occupied half of samples.[7] The probability of the outliers and extremes out from  $\pm 1.5$  and  $\pm 2$  boxes in a box-plot is around .01 and .0005 respectively.[7] and stands for the growth rates much milder or severer needed a further investigation. Those hospitals out of boxes but inside outliers and extremes in box-plot are top or bottom 25% samples distributed in spread, symmetry or tail length referring to the shape of symbol  $\perp$ .

Figure 4 shows the ranks of hospital levels are other-type (48.3), district (46.9), region (46.25) and medical center (45.45). There are two medical centers located beyond extremes and one out of outliers, which should be further inspected or investigated for reasons.



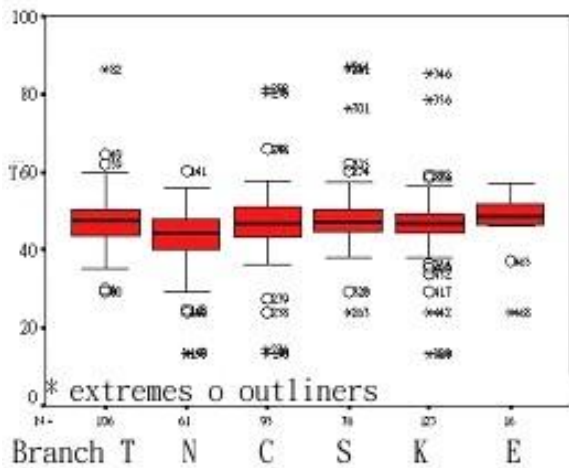


Figure 2. Growth rates between BNHI region branches.

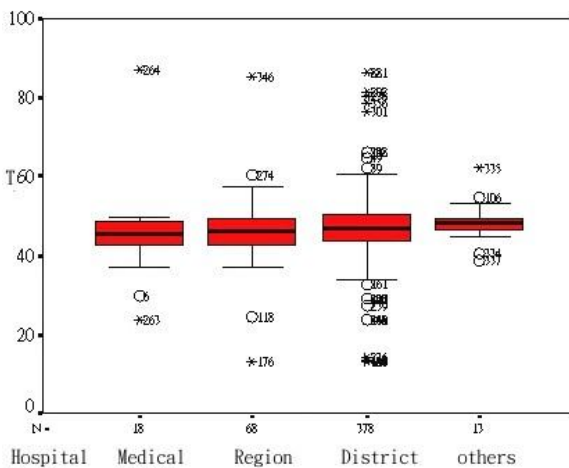


Figure 3. Growth rates between hospital levels.

### 3.3. Most Misfitting Strings of a Hospital

Besides knowing the abnormal BNHI branches or hospital levels in growth rates, most misfitting strings generated by Rasch analysis were provided with helping managers detect the abnormal weeks(items), which were statistically significant at the 0.05 level, in growth rates incurred in a hospital.

We showed a district hospital in Central region branch of BNHI in Figure 4. As for the demo hospital's ability at 1.01 and outfit MNSQ 4.3 bigger than the golden standard of 1.5[8], 3 weeks in Figure 4 shown up were most misfitting strings ( $p < .05$ ) which indicated that the matched ability response at logit -1 reflecting the abnormal response in growth rates needed to be further inspected and understood.

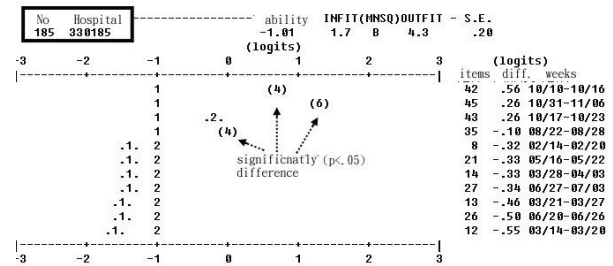


Figure 4. most misfitting strings shown in a hospital

### 3.4. Abnormal Phenomena on Internet

In order to allow all hospitals to easily monitor their abnormal items incurred to growth rates, we need to develop computer programs to be of practical usage.

An abnormal growth week on intra-hospital or inter-hospital was shown in Figure 5 and Figure 6 respectively.

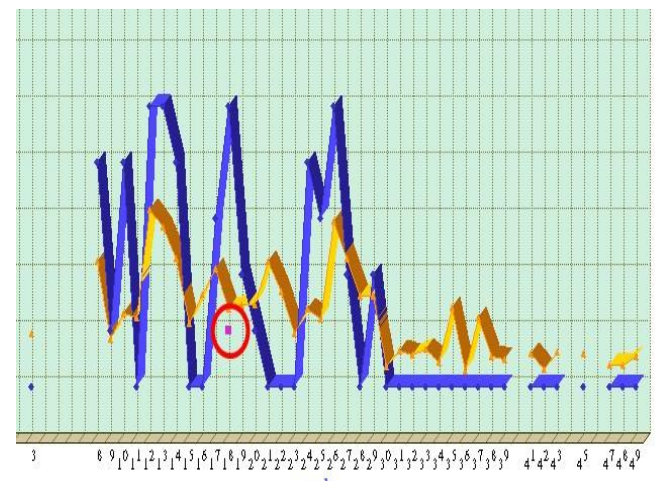


Figure 5. a circle shown abnormal on intra-hospital

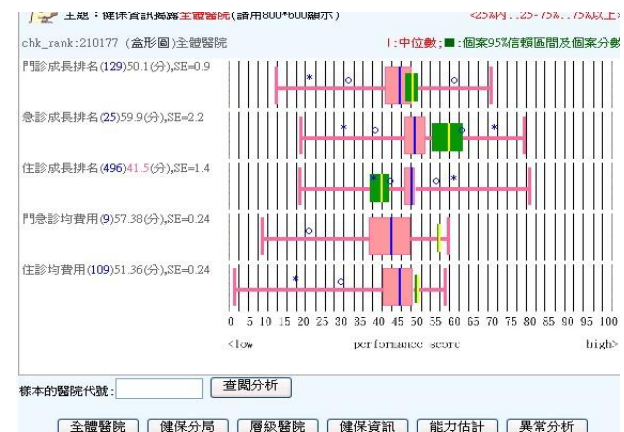


Figure 6. A box-plot shown abnormal on inter-hospitals

### 4. Discussion and Conclusion

Evaluation indices can provide a measurement tool with an enhanced range and sensitivity for assessing



comprehensive hospital abilities in growth rates, which is especially useful when applied to hospitals with abnormal phenomena under global budgeting. Using Rasch analysis [1], this study found that such a measurement tool can be utilized to detect and delete noise growth weeks in order to make the measurement to be a unidimensional construct.

Because the items of the scale have been shown to measure comprehensive growth rates as a single, unidimensional construct for hospitals in a year, it is worth displaying the results on internet for mutual comparisons.

The Rasch-transformed scores can be viewed as interval-level measures [5,6], whereas the observed 6 categories raw scores are ordinal data. Since most statistical techniques assume that the data are at least on an interval scale, the Rasch-transformed scores of the growth evaluation inventory scale are suitable for comparing abilities between or within groups.[10] In addition, comparisons on internet should be widely adopted and developed through line-chart plots for easily monitoring the abnormal growth rates under global budgeting.

To use the new technique proposed in this article, the users can first obtain the knowledge of both psychometrics and computer science and then look up the ZSCORE [6] beyond  $\pm 2$  in a bit to detect the abnormal items incurred in a hospital.

There have been insufficient empirical investigations of the OPD growth rates between or within hospitals in such overall perspectives as this study offering an example not just of how research could be applied to inpatient or emergency departments for more detailed information of the phenomena of healthcare, but also of a regression model could be developed for the abilities of hospitals' OPD growth rates by factors of some predictive attributes in a future.

## 5. References

- [1] G. Rasch, Probabilistic Models for Some Intelligent and Attainment Tests, Copenhagen, Denmark: Institute of Educational Research, 1960.
- [2] J.M. Linacre, *WINSTEPS* [computer program]. Chicago, IL: <http://www.winsteps.com>; 2006.
- [3] D. Andrich., "A rating formulation for ordered response categories", *Psychometrika*, 43, 1978, pp. 561-573.
- [4] B.D. Wright, G.N. Masters, *Rating Scale Analysis*, Chicago, Ill: MESA Press, 1982.
- [5] B.D. Wright, M. Mok, "Rasch models overview", *J Appl Meas*, 1:1, 2000, pp.83-106.
- [6] J.M. Linacre, "Detecting multidimensionality: which residual data-type works best?", *J Outcome Meas*, 2:3, 1998, pp.266-283.
- [7] Tukey, J. W., "Box-and-Whisker Plots." §2C in *Exploratory Data Analysis*. Reading, MA: Addison-Wesley, 1977, pp. 39-43. Retrieved at [http://exploringdata.cqu.edu.au/box\\_norm.htm](http://exploringdata.cqu.edu.au/box_norm.htm)
- [8] T.G. Bond, C.M. Fox, C.M., *Applying the rasch model: fundamental measurement in the human sciences*, Lawrence Erlbaum Assoc Inc, 2001.
- [9] R. Likert, "A Technique for the Measurement of Attitudes", *Archives of Psychology*, 140, 1932.
- [10] S. E. Embretson, Item response theory models and spurious interaction effects in factorial ANOVA designs. *Applied Psychological Measurement*, 20:3, 1996, pp.201-212.





## The Design and Development of a Physician-Oriented Information Infrastructure for the Enhancement of e-Hospital Facilities.

Hsu-Hua Ho<sup>1</sup>, Tzu-Yi Yu<sup>2</sup>, Hung-Pin Hou<sup>3</sup>, Wen-Yung Wu<sup>4</sup>  
*Director of Internal Medicine Department, Puli Christian Hospital<sup>1</sup>*  
*Assistant Professor, National Chi Nan University<sup>2</sup>*  
*Director of Administrative Department, Puli Christian Hospital<sup>3</sup>*  
*Director of Radiological Department, Puli Christian Hospital<sup>4</sup>*  
hua@mail.pch.org.tw<sup>1</sup>, tyyu@ncnu.edu.tw<sup>2</sup>, joy@mail.pch.org.tw<sup>3</sup>

### Abstract

*Recent advancements in modern medical diagnoses have necessitated a huge increase in the frequency of the use of equipment such as CT and ultrasound machines. Correspondingly, the storage and dissemination of these medical images have become an important issue to medical professionals. Management of these images has traditionally been slow and cumbersome. With the prevalence of the personal computer, however, along with increased network bandwidth, it is now possible to handle this information electronically as well as wirelessly. The Picture Archiving and Communication System (PACS) is at the forefront of this revolution. Yet, commercially available PACS software is generally very expensive and hard to afford for hospitals with limited financial resources. A dilemma among many hospitals is deciding how to acquire and implement the proper PACS system without unduly affecting the budget. In this paper, a full-function, efficient, and economical PACS system is presented as a viable, non-compromising option for many small and medium-sized hospitals. This system, designed and developed mainly by the physicians and technicians of the Puli Christian Hospital (PCH), with the cooperation from academics, allows for customization to fit the needs of the individual hospital. This system can be used as the foundation of the health information infrastructure and to enhance e-Hospital service.*

### 1. Introduction

There are always unexpected catastrophes or severe infectious diseases which might occur from time to time. Unforgettable experiences such as the 911 attack in New York, the SARS epidemic in Taiwan during 2003, and the hurricane Katrina in New Orleans Louisiana and Mississippi in 2005 have been very prominent news items. These incidents not only created unimaginable chaos, but it also impacted against the local Medical and Medicare systems. From reading the review reports of these tragedies, one of the important conclusions that can be reached is that if the documentation/information is not stored electronically, it can be destroyed by these unexpected events (such as by explosion or flood) and cannot be recovered or can be recovered only with the

expenditure of much time and effort. This conclusion actually agrees with one of the major goals of the “National Health Information Infrastructure” (NHII) project presented by the US Department of Health and Human Services [1]. It hopes to reach the “Paperless Healthcare” by the year of 2010. Similar to this project, the National Health Service (NHS) of the Great Britain has the “National Program for IT” project [2], and the Department of Health of Canada also has the “Electronic Health Record (EHR)” program which tries to promote the digitalization of patient records [3] to create a paperless environment. The goal of the EHR is not only to save and store patients’ information more securely, but also to transfer electronic more easily when patients are referred to another hospital. These electronic records will give the doctors a better picture of their patients’ medical history, reduce the waste associated with unnecessarily repeated examinations. Furthermore, with electronic health records, hospitals can easily provide useful statistical reports to the Health Department.

In light of these global tendencies, the Department of Health (DoH), Executive Yuan, R.O.C. (Taiwan) is sponsoring the “National Health Information Infrastructure” (NHII) project which focus is on the integration and extension of some original carry on projects [4, 5, 6, 7]. The main visions of this NHII project is to “establish a national-wide e-Health environment and to build-up new information sharing model for Healthcare and Medicare services”. The DoH is also trying to encourage medical facilities to utilize emerging internet and computer technology to develop and enhance in-house computer infrastructures. This is the basis of the motivation for the Puli Christian Hospital to design the “Picture Archiving and Communication System” (PACS). PACS was first announced in 1980s, and its intent was to achieve a paperless/filmless environment for the hospitals and clinics. This goal actually matches the objective of the NHII project. As a matter of fact, the PACS not only digitalizes the patients’ personal information and their medical histories, but it also stores and retrieves medical images obtained from any medical examination electronically. Even though controversy has followed the use and developments of the PACS [8, 9, 10, 12, 13], the Taichung Veterans General Hospital first presented their PACS system in 1999 [11]. The PCH was inspired



and motivated by this successful experience, but it came to the conclusion that most commercial PACS are not only prohibitively expensive, but they lack the customization feature required for each hospital. Furthermore, one of the important missions for PCH is to provide healthcare services for the surrounding mountainous region and remote areas where medical resources are severely lacking. Only by developing an in-house PACS can costs be greatly reduced while still providing exemplary health care for the patients and the community around the Nantou, Puli County – a region that is far removed from the nearest metropolitan city.

## 2. System Architecture

As is the case with many other regional hospitals, Puli Christian Hospital lacks the financial resources to purchase expensive PACS facilities. In order to reduce costs and use the current available resource, and also because of the availability of the personal computer, this PCH-PACS system is developed based on the PC-Windows platform. One immediate and obvious advantage to this is that there is virtually no training expenses involved since the Windows-based operating system is very popular in this country. Basically, this system can be divided into two parts: database design and image processing (including acquisition and retrieval). This system evolved in several distinct stages. Originally it was built under *Visual Basic* and *Access* databases. Later, the *VBA (Visual Basic Applications)* embedded with the *Access* component was used as the front end program, and the *SQL* server was used as the back end database. After the *Microsoft* Company came out with their new Operating System and software, this system incorporated the advantage of the continuous update feature, giving it better stability under the *Microsoft XP .NET* framework. Figure 1 shows this in-house self-made software system.



Figure 1 The self-made PACS system used in PCH.

## 2.1. Medical Database Design

In order to synchronize the PACS system with the original Healthcare Information System (HIS) and to reduce the development costs, the minimum requirement for the database is that the PACS system database schema should be designed to fit for implementation in a relational database. In addition, the medical database management systems and applications developed to manipulate data stored in databases should provide reporting functions easily and communicating with other packages effortlessly. The *Access* and *SQL* database were chosen because of these flexibilities. *Access* is embedded with *VBA* as the front end program, and the *SQL* server is used as the back end database server. The *SQL* server is a robust database; it can handle huge dataset even if they are located on different hard drives. They are *Microsoft* software and can embed other *Microsoft* components. Their statistical functions are useful and can be adopted for medical statistical reports. The data integrity features allow them to maintain consistent medical data constraint when new records are added to the database. The data security feature is capable of protecting sensitive information with varying levels of access privileges. The concurrency control allows different physicians to access, manipulate and update this database simultaneous at the same time in different locations. Automatic periodically backup is performed to protect this medical database. Either from the viewing workstation or from the server, the strong backup system protects the electronic health records and the medical images well. After years of operations under real-life situations, it has been determined that a robust database is the foundation of a successful PACS system. This medical database not only may be used to store the information securely, but it can also function as a communication channel to provide fast and accurate data retrieval and display.

## 2.2. Medical Images Processing

Most of the available PACS systems in medical facilities today are used to acquire, retrieve, and store the medical images. This information is generally in the form of still images (not the dynamical image such as the Audio Video Interactive (avi) file). They are obtained from various sources, such as from the Computed Tomography (CT), Magnetic Resonance Imaging (MRI), Breast Ultrasound, Echocardiography and Cardiac Doppler Ultrasound and from Transcranial Color-Coded Sonography, ... etc. Therefore, the major concerns are how to obtain high quality images with affordable devices and how to configure the system to retrieve images instantaneously from the database and then display them fluently throughout the intranet. Thanks fully the notable progress in hardware development, spurred on by market competition, has resulted in graphics and acquisition cards that are not only more affordable, but has much better quality than



ever before. This in-house PACS system adopts a *PCI* (Peripheral Component Interconnect)-based acquisition card to capture most of the images. Most images are stored as still images, i.e., the single image format in the medical database. However, a Cardiac Ultrasound is much more useful if it can display how the blood flow moves between the atrium and ventricle of the heart. Hence, the dynamical image (also known as animation pictures) file format is needed. Based on different requirements, this system can store images in still or animated formats. In other words, when the images are being acquired, the system can store them into BMP, JEP or DICOM (Digital Imaging and Communication in Medicine) files format (still image file format) or AVI, WMV or MOTION\_JPEG (the animated file format).

There are instances when physicians need to manipulate the acquired images. Editing functions such as the zooming in/out, brightness and contrast adjustments, local region enlargement for detail texture, distance or area measurement ... etc., are all provided through a user-friendly interface. Another useful function is that all the images acquired or retrieved can be linked together into a slideshow for demonstration purposes. For those medical instruments (such as endoscope and ultrasound) that produce non-DICOM images, a special multimedia video card is used to digitize and capture the medical video image into dynamical or statlcal file format and then store them into the PACS image database. The related functionalities and the features will be discussed in the next section. Figure 2 shows the system architecture of the PACS used in PCH.

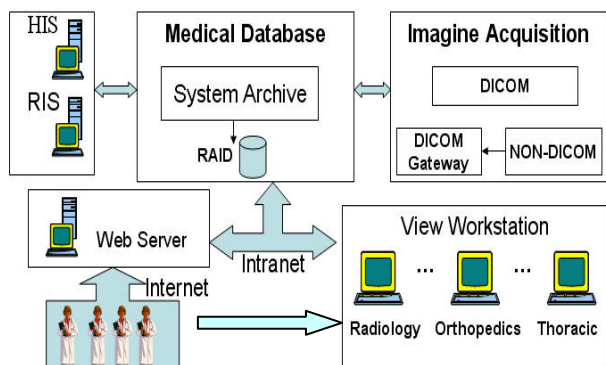


Figure 2 The PACS system Architecture in PCH.

### 3. System Features

As previously mentioned, this in-house software was primarily developed by the attending physicians of the *PCH*. The graphics user interface is very friendly and follows the conventional window software architecture. As Figure 1 shows, the GUI is displayed with Chinese characters, making it easy to manipulate even for the non-technical users. When a physician desires to access the system to obtain information on a patient, all the examination results completed the same day will be displayed, even if the patient had visited more than one

doctor in different departments. Data deserving of urgent or immediate attention will automatically be displayed in red. If any image-related tests were performed, those images can also be shown on a high resolution computer monitor. A complicated examination might be composed of more than one image. This system allows the physician to view pictures in either single image or multiple (4 or 16) image mode. If the same examination was performed previously, this system can also display these two sets of images together, allowing the doctor to track the patient's prognosis and to follow the progress of the disease to determine if the patient is getting better or worse. Figure 3 demonstrates this functionality.



Figure 3 Multi-image display mode of the in-house PACS system.

Preparing the diagnostic reports with medical images is very tedious work. This new system allows the physicians to embed the medical images in the report by selecting the image and then clicking the "Report" option. The selected image(s) can be placed either vertically or horizontally. Text can easily be made to flow around the images. Since most patients' symptoms can be classified, and some of them are either similar or the same, the physicians can use the system to pre-define these symptoms, along with their suggested method of treatment and/or comments. In specifying the parameters of the reports, these pre-defined items can be included to reduce the amount of typing. Furthermore, in order to take advantage of the electronic referral procedure, the physician (after inputting administrative permission and security clearance) can use this system to easily attach the medical images along with the diagnostic reports, and then utilize *Microsoft* email system (such as *Outlook Express* or *Outlook*) to disseminate the information to designated personal in other hospitals. Even though sending the information through the email system could cause the security concern, yet the encryption approach, request return receipt and the digital signature can be used to meet the authentication, integrity and nonrepudiation security purpose. These features are demonstrated in Figure 4 and Figure 5 respectively.







Figure 4. Easy report generation feature.



Figure 5. The software integrates email system for e-referral procedure.

When the PACS was first announced, one of its goals was to create a paperless working environment. However, after several years of development, realistic expectations have been re-focused towards increasing the rate of correct diagnosis in order to diagnose and treat the patients as quickly as possible. The way this system is developing is in this direction. It possesses very useful and efficient manipulative functions when examining the desired images. For example, the physician can use the mouse to select the interested area, and then enlarge this region to view more detailed information. By using the scrolling function of the mouse wheel, various properties of the image, such as the brightness, contrast or quality, can be easily adjusted. In many cases, the distance or the size of particular image, such as the tumor size, means a lot to the medical treatment. This system also provides the flexibility to allow the doctor to measure the distance and area multiple times over the image directly. This distance information is also preserved when exported to the diagnostic report. In addition, if further image manipulation function is needed, this system can evoke other image manipulation software such as *PhotoShop* to perform the editing. The images can also be copied and

pasted into other software packages to make illustration material for training purposes. These functions are shown in Figure 6.

Even though this system was originally designed for the attending physician, resident physicians or any clinicians who are outside of their area of specialty can easily use this system. However, there is always the desire to increase the practicability and convenience for the PACS. As medical science becomes more diversified and specialized, even an experienced physician might have difficulty in making correct judgments when examining results from other specialties. To overcome this obstacle for physicians or other technicians using this software, when the user selects an image, the appropriate image database with related documentation will be displayed. The user can then compare the images in order to obtain necessary information relating to the image and/or to confirm the symptoms of a particular disease. This practical function is demonstrated in Figure 7 and Figure 8.

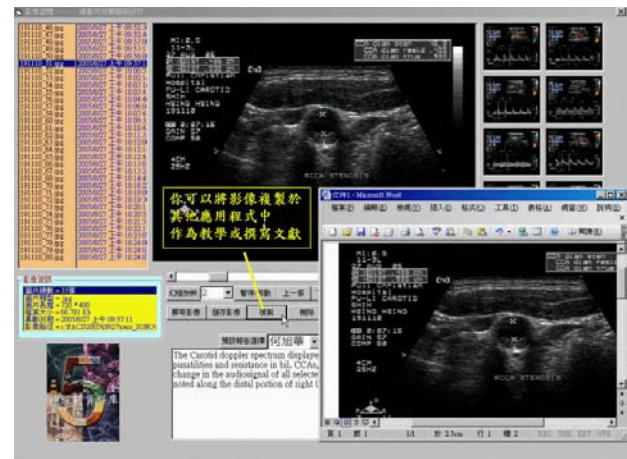


Figure 6. The system has strong image manipulation capability.

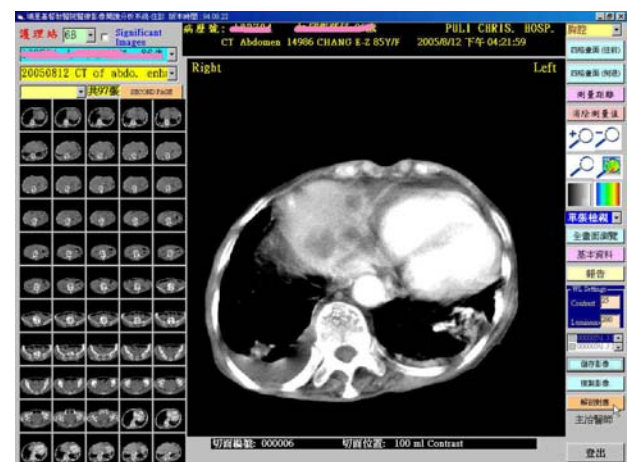


Figure 7. The system displays the examined images which are then compared with the appropriate medical images.



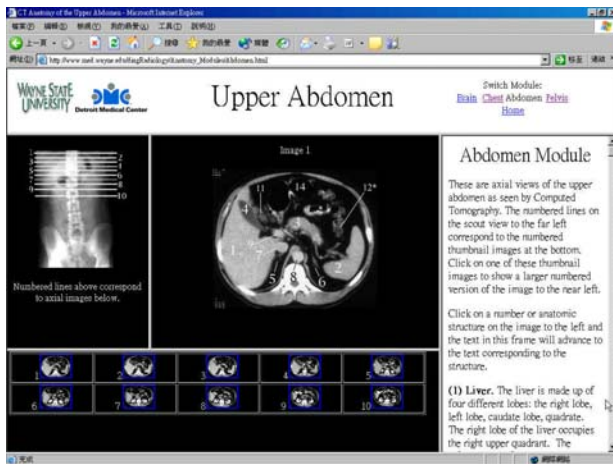


Figure 8. The system can bring in standard images with detail documentation for comparison so that the physicians can have better confirmation and understanding of patients' conditions.

After several years of evolution, the internet activities have become an important part of people's regular daily life. According to the recent statistical report, the time spent engaged in internet browsing activities for people in Taiwan ranks among the top five worldwide. The fact that the internet has been deeply embraced indicates that the web browser can be a very convenient tool for collecting the information needed. With this in mind, this PACS system allows the world wide web-based viewer to access the database server through the Local Area Network (LAN). This web-based viewing system allows physicians to check examination results and retrieve image archived through the hospital internet into the PACS servers. It is a design based on the DICOM 3.0 standard format over TCP/IP communications protocol. With tight security control (such as firewall and password protection and access privilege settings), the system provides a convenient way for the attending physician or any clinician to verify a patient's condition at any time. The system also automatically tracks all activities when a physician uses this web-based viewing software, and saves the histories into the system log. The system administrator can monitor all accesses through this history log to prevent any illegal web actions. These functions are demonstrated in Figure 9 and Figure 10.

One of the on going projects is to integrate 3D medical database browsing into this system. Two major graphics libraries are used to develop this 3D browsing feature –the *vtk* [14, 15] and *itk* [16]. The *vtk* (Visualization ToolKit) and *itk* (National Library of Medicine Insight Segmentation and Registration Toolkit) are open-source, freely available software systems for computer graphics, image processing, and visualization. It supports hundreds of algorithms in visualization and image processing fields. These two libraries were designed primarily for medical research. They also support different computer language bindings, and are

available on different computer platforms such as *Linux* or *Windows*. The user can operate the arrow keys to move the sliding plane for the desired cutting plane or to rotate the mouse to visualize the entire medical object in three dimensions freely.



Figure 9. The login interface for the web-based viewing system.

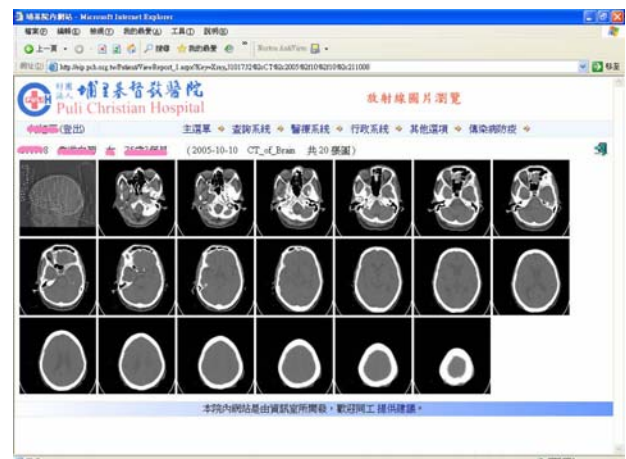


Figure 10. Web-based viewing interface allows online browsing of patient examination results.

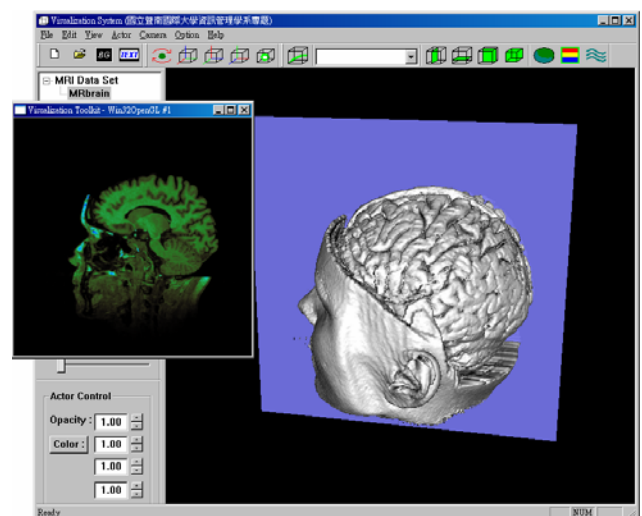


Figure 11. The interactive three dimensional browsing feature for medical data set.







#### 4. System Performance

Puli Christian Hospital has in place a high speed internet and intranet communication system that has been in operation for a number of years. Recently, the 10/100/1000 Base-T Ethernet was installed. With this high speed communication capability, image retrieval and display processes can be carried out very smoothly at each viewing terminal. The average response time is 0.2 seconds for retrieval of a CT image and 0.08 seconds for retrieval of an ultrasound or endoscopic image. There are currently over 40 terminals that can access the medical database at the same time without notable lag. Since this PACS is an in-house developed software, the cost of the entire system is proportional to the hardware (mainly, the viewing terminals). For each single unit, the hardware is equipped with a PC (*Cerelon* CPU, 512 MB RAM) and one 19-inch LCD monitor with 1280 x 1024 resolution, with a cost of approximately \$900 USD. The *HP ProLiant*, which is being used as the image server, is equipped with a 200 Gigabyte hard drive storage capacity. For data backup purposes, the *RAID* (Redundant Array of Independent Disks) system with four 300 Gigabyte hard drives is used.

As it was mentioned previously, the image quality is of particular importance for the PACS system. For those images that are available in digital format (such as the Computed Tomography), there is no need to do any conversion. However, for those non-digital sources, there is a need to capture those images with a special graphical device. An “ASUS Extreme AX 700 PRO” acquisition card (also known as Frame Grabber) is used to obtain high quality images. The price for this card is around \$250 USD each, and this device is currently installed for the following five procedures: the Cerebrovascular Ultrasound, Cardiac Ultrasound, Abdomen Ultrasound, Kidney Ultrasound, and the Upper and Lower alimentary endoscope. Because these five different departments have different hardware equipments, the organic interface necessary for these machines to capture images is very important. A user-friendly interface was designed for these non-DICOM apparatus to capture these images and store them in the medical database. When images from these sources are acquired, the system can convert them to JPG, BMP, DICOM or any animated formats (AVI or WMV). Table 1 shows the image quality and the hardware specification used, and Table 2 shows the number of patients and number of images acquired and processed in one month (March 01 to March 31, 2006).

Most commercial PACS systems are sold based on either the number of desired licenses or the “KeyPro” purchase which locks the view terminals. Hence, increasing the number of view station not only increases the hardware cost but also the software expense, resulting in a heavy financial load to the medical

facilities. For government-supported clinics or metropolitan hospitals, the financial pressure may not be a major concern. However, because regional medical facilities handles not only local residents’ medical problems but also has the responsibility to provide healthcare services for the mountainous area or remote communities, the in-house self-made PACS means a lot since it can be freely duplicated without incurring excessive hospital expenses.

Table 1. Image quality and the hardware specification.

| Imaging Modality    | Hardware Spec. | Method of Image Acquisition | Spatial Resolution (Pixels per image) | Brightness (bits per pixel) |
|---------------------|----------------|-----------------------------|---------------------------------------|-----------------------------|
| Computed Tomography | SIEMENS        | DICOM                       | 512x512                               | 12                          |
| Ultrasound          | HP 8500        | Frame Grabbing              | 720x480                               | 8                           |
| Ultrasound          | HP 4500        | Frame Grabbing              | 720x480                               | 8                           |
| Ultrasound          | TOSHIBA        | Frame Grabbing              | 720x480                               | 8                           |
| Endoscope           | Olympus CV-200 | Frame Grabbing              | 720x480                               | 8                           |
| Endoscope           | Olympus CV-260 | Frame Grabbing              | 720x480                               | 8                           |

Table 2. Images acquired during March, 2006

| Imaging Modality | Number of Patients | Number of Images | Amount of storage used |
|------------------|--------------------|------------------|------------------------|
| CT               | 221                | 7004             | 3820 Mbyte             |
| Ultrasound       | 460                | 2886             | 193 Mbyte              |
| Endoscope        | 120                | 1080             | 74 Mbyte               |
| Total            | 801                | 10970            | 4087 Mbyte             |

#### 5. Conclusion

When PACS was first announced in the 80s, many people, especially the radiologists, expressed high interest and expected that it could help the hospital achieve a filmless environment. Even though there were controversies during the development period, PACS has become an indispensable tool in hospital operations. However, the main obstacles preventing wide scale implementation of PACS among all the hospitals and clinics are its cost and lack of flexibility. Most PACSs were developed by vendors and software companies, and can generally be afforded only by metropolitan hospitals. In addition, because they were developed by software companies who were anticipating that they would be used by a full-service hospital, these PACSs may not completely fit the needs of diverse or smaller hospitals. For regional hospitals with limited budget, it has been even more difficult to promote the system in serving the needs of their patients. According to statistics reports from the Department of Health, Executive Yuan, R.O.C., only around 7% of hospitals in Taiwan utilize the PACS.





To overcome this restriction in Puli Christian Hospital, a customized PACS system has been developed by their in-hospital physicians and technicians from the Information Department. It is well known that in order to develop productive software, a good and thoughtful system analysis needs to be performed prior to building a PACS that will fit the needs of the hospital. In order to achieve this, the on-site physicians were deemed to be the best system analysts. This full-function PACS was initiated three years ago, and has achieved a relatively mature state in providing hospital-wide service. Because this system was developed and implemented by the PCH physicians, its maintenance and upgrade capabilities are much stronger than any commercially-developed PACS software. With regular feedbacks from the users (the clinicians, nurses, and technologist), modifications can be made quickly and easily. For the Puli Christian Hospital, this PACS not only saved considerable expense compared with purchasing commercial software, it has also provided the following contributions and benefits to PCH:

- Instead of manually handling medical results, this PACS provides systematic access to the patient's information on a hospital-wide basis. This has sped up the administrative process and medical treatments considerably.
- Medical information, examination results, and treatment histories have been stored electronically, thereby greatly reducing required physical storage space while increasing the capability of data preservation.
- Integration of the original Healthcare Information System (HIS) and the Radiology Information System (RIS) has increased hospital administrative efficiency; management abilities have improved enormously while reducing operation costs.
- This easily manipulated system has helped physicians/clinicians to make more accurate diagnosis, resulting in enhanced medical service. The ability to show medical results through the PACS has also earned the trust of patients.
- Simplification of the referral process for patients to transfer in or out of the hospital. Sharing the medical resources has also facilitated network treatments.
- The majority of PACS systems in Taiwan are commercially-developed products. In stark contrast, in developing this customized system, considerable valuable experience has been gained through designing the system analysis, software architecture design, debug process, system maintenance and developing updates, ... etc. It is strongly believed these experiences can be used to

showcase and promote further hospital software development in Taiwan.

## References

- [1] U.S. Project of the National Health Information Infrastructure. <http://aspe.hhs.gov/sp/NHII/>
- [2] Homepage of National Health Service of the Great Britain. <http://www.nhs.uk/>
- [3] Homepage of Health of Canada. [http://www.hc-sc.gc.ca/hcs-sss/ehealth-esante/ehr-dse/biblio\\_can\\_e.html](http://www.hc-sc.gc.ca/hcs-sss/ehealth-esante/ehr-dse/biblio_can_e.html)
- [4] Homepage of Department of Health, Executive Yuan, Taiwan, ROC. <http://www.doh.gov.tw/cht/index.aspx>
- [5] Technical Report, National Science Council (NSC), Taiwan, ROC. <http://www.nsc.gov.tw/pub/yearbook/yearbook93/main/4/4-2.htm>
- [6] Fu-Zhong Wang, "The Study of National Health Information Infrastructure in Taiwan.", Technical Report, Bureau of National Health Insurance, Taiwan, R.O.C., Dec, 2005.
- [7] Cheng-Hua Lee, "Health Care and Medical Insurance: The National Health Information Infrastructure.", Technical Report, Bureau of National Health Insurance, Taiwan, R.O.C., Sep 20, 2003.
- [8] Steckel RJ, "The Current Application of PACS to Radiology Practice.", *Radiology* 1994; vol. 190(3), pp 50A-52A.
- [9] Friedenbergr RM, "PACS in the Clinical Setting: Revisited.", *Radiology* 1994; Vol 190(3): 53A-54A.
- [10] Davide Caramella, "Is PACS Research and Development Still Necessary?" International Congress Series 1281, 2005, pp. 11-14.
- [11] Tain-Chen Wu, San-Kan Lee, Chen-Hsing Peng, Chia-Hsien Wen and Shu-Kun Huang, "An Economical, Personal Computer-based Picture Archiving and Communication System", *RadioGraphics*, 1999, Vol. 19(2), pp 523-530.
- [12] I-Chiu Chang, Hsin-Ginn Hwang, David C. Yen, J.W. Lian, "Critical Factors for Adopting PACS in Taiwan: Views of Radiology Department Directors", *Journal of Decision Support System*.
- [13] Mervyn D. Cohen, MB, ChB, MD, Lori L. Rumreich, MBA, Kimberley M. Garriot, S. Gregory Jennings, MD "Planning for PACS: A Comprehensive Guide to Non-technical Considerations.", *Journal of the*





American College of Radiology, April 2005, Vol. 2 No. 4, pp 327-337.

[14] William J. Schroeder, Kenneth M. Martin, and William E. Lorensen. "The Design and Implementation of an Object-Oriented Toolkit for 3D Graphics and Visualization". In *IEEE Visualization*, pp. 93-100, 1996.

[15] *vtk* homepage, <http://www.vtk.org/>

[16] *itk* homepage, <http://www.itk.org/>





## Economic Evaluation of Japanese e-Health Systems: A Cost-Benefit Analysis

Masatsugu Tsuji

Graduate School of Applied Informatics Kyoto University of Education Faculty of Urban Science  
University of Hyogo  
tsuji@ai.u-hyogo.ac.jp

Fumio Taoka

Masaaki Teshima

Meijo University

### Abstract

*The e-health system is a type of telecare that transmits health-related data of its users such as blood pressure, ECG, and blood oxygen to a remote medical institution via a telecommunications network. 100 Japanese local governments are already implementing such systems, utilizing a total of more than 12,000 devices (as of March 2005). This paper analyzes the economic benefits of the e-health system in the following four regions in Japan: Kamaishi City, Iwate Prefecture; Nishiaizu Town and Katsurao Village, Fukushima Prefecture; and Sangawa Town, Kagawa Prefecture. The Contingent Valuation Method (CVM), which has been recently widely adopted in the fields of Environmental Economics and Health Economics to measure the benefits of services not traded in the market, is used. We conducted a survey of 348 users in Kamaishi City regarding their Willingness to pay (WTP) to use the service. Using their WTP, we estimate the demand function of the system, which is assumed to be a logistic curve. The average WTP calculated is 4,519 yen (approximately US\$37). We then compare the benefits with the costs of the system, which consist of equipment, salaries of doctors and nurses, and other operations. The ratio of benefits over costs (B/C ratio) is about 1.07, which implies that the system of Kamaishi City yields more benefits than costs. This result is rather surprising when compared with the systems of other regions. The paper also analyzes how the benefits expressed in terms of WTP are attributed in exact monetary terms to factors such as (a) reduced anxiety in day-to-day life, (b) stabilization of illness, (c) enhancement of health consciousness, and (d) decrease in medical expenditures. The costs borne by each of the involved parties are identified; namely, the respective amounts paid by individual users and public funds such as tax and medical insurance. This paper suggests reimbursing the e-health system through medical insurance.*

### 1. Introduction

The e-health system transmits health-related data of its users, such as blood pressure, ECG, and blood oxygen, to a remote medical institution via a telecommunications network. To date, eighty Japanese local governments are currently implementing such systems, utilizing a total of more than 12,000 devices – the largest number of any country. Several household electric appliance manufacturers such as Panasonic, NEC, Fujitsu, Sanyo, and Hitachi, produce and sell remote

monitors at prices of US\$2,000 to US\$4,000 per set (see [1] for details). This system is equipped with a simple device which, when used continuously, records an elderly person's condition or a patient's illness in graphs, which are then used for diagnosis and consultation. Reports sent by the medical institution also help users to enhance their daily health consciousness and maintain good health. These positive effects have been identified through field surveys.

The e-health system in Japan has already passed the experimental stage, and is entering the diffusion stage. In order for the system to be diffused further, it is necessary to prove its cost-effectiveness by comparing its benefits and costs. The latter consist of equipment such as servers and peripheral devices, salaries and wages of doctors and nurses, and maintenance fees such as telephone charges and other miscellaneous operating costs. On the other hand, to indicate its concrete benefits in monetary terms is analytically difficult, since the benefits mainly come from users' subjective satisfaction which is difficult to measure. Without a definite confirmation of the e-health system's cost-effectiveness, the system's future sustainability cannot be guaranteed.

This paper aims to carry out a Cost-Benefit Analysis of the e-health system in the following regions: Kamaishi City, Iwate Prefecture; Nishiaizu Town and Katsurao Village, Fukushima Prefecture; and Sangawa Town (currently part of Sanuki City), Kagawa Prefecture. Benefits are expressed in terms of WTP (willingness to pay), based on the CVM (Contingent Valuation Method), and costs are calculated as the sum of equipment, salaries of doctors and nurses, and other operations. The benefits and costs are compared with a B/C ratio. We also discuss policy measures to increase this ratio by means of reimbursement through medical insurance, as well as direct subsidies from tax revenue. For other economic evaluations of telemedicine, refer to [2], [3] and [4.]

### 2. e-Health systems

First we briefly introduce e-health systems and describe an example. Kamaishi City, Iwate Prefecture, which started its e-health system in 1993, was chosen for the survey data, as it has one of the longest records of implementation. In addition, the city's system is operated by a private medical corporation named *Rakuzankai* at a monthly charge of 2,500 yen (approximately US\$22.73) per family of up to four persons, and a users' association that conducts many events to promote usage. Since the system uses the CATV network, transmission charges are zero. The peripheral device used by Kamaishi City, called "Urara," is manufactured by *Nasa* Corporation. The device is





equipped with memory, an electric sphygmomanometer, an electro-cardiograph, an electric signboard, and a button for answering questions. At the hospital, nurses check transmitted data and report to the doctor as well as to the device's users whenever they notice abnormal symptoms. The hospital sends monthly reports to all users with comments, which are used for their health management. There are 211 devices, and the total number of users is 409.

Nishi-aizu and Sangawa have almost identical systems based on "Urara," since they imitated Kamaishi's system.

Katsurao Village has a different system, using NEC's "Sukoyaka-mate" device. Katsurao Village distributes terminals to all families who have person(s) older than sixty-five years old, and thus it has the largest number of users. In Table 1, we compare the systems of four regions. We will see that these four regions yield different results according to the differences in their systems.

**Table 1. Comparison of systems**

|            | No. of users | No. of devices | Terminal | Network      | Charges   |
|------------|--------------|----------------|----------|--------------|-----------|
| Kamaishi   | 348          | 211            | Urara    | CATV         | 2,500 yen |
| Nishi-aizu | 518          | 400            | Urara    | CATV         | none      |
| Katsurao   | 926          | 325            | NEC      | Tele. (ISDN) | none      |
| Sangawa    | 384          | 225            | Urara    | CATV         | none      |

### 3. Methods

#### 3.1. Contingent valuation method (CVM)

In order to measure the benefits of services not traded on the market, the following methods are often utilized: (a) travel cost method; (b) replacement costs method; (c) hedonic approach; and (d) CVM. In what follows, we use CVM, which has been recently widely adopted in the fields of Health Economics and Environmental Economics. In CVM, the benefits to users are measured in terms of WTP, which is the monetary amount users are willing to pay for receiving a service. By asking the WTP of each user, we can then construct the surrogate demand function for the e-health system. Although CVM and WTP have strong theoretical bases, CVM tends to introduce bias, because it asks for concrete valuation and choice under fictitious circumstances. Care should be taken to clarify what kind of bias is introduced, and to remove this bias (for further discussion on biases, see [5]).

#### 3.2. Questionnaire

For the surveys, we use questionnaires. Let us take the example of Kamaishi City. We interviewed 348 users of the e-health system questions pertaining to the following: (a) WTP; (b) effectiveness; (c) frequency of usage; and (d) user characteristics such as age, gender, income, education, and health condition. The portion of the questionnaire related to WTP is as follows. We begin by asking whether they would be willing to pay monthly charges of 5,500 yen (US\$45) to use the system. If their answer is "yes," we then ask whether they would be willing to pay 7,500 yen (US\$62.5). If they reply "yes" again to 7,500 yen, their WTP is 7,500 yen. If "no," then we lower the amount to 6,500 yen (US\$54.17). If they

reply "yes" to 6,500 yen, then that is their WTP. If again their answer is "no," we lower the amount further to 5,500 yen. We repeat this process until their WTP is determined. Out of a total of 348 replies, 291 had no missing value. The distribution of WTP from the survey is as follows: 10,000 yen (16 users), 8,000 yen (1), 7,500 yen (12), 6,500 yen (11), 5,500 yen (62), 4,500 yen (8), 3,500 yen (69), and 2,500 yen (112). The distribution of replies is shown in Table 2. Those of the other regions are shown in Tables 3, 4, and 5.

#### 3.3. Estimation of demand function and WTP

Based on the above WTP of each user, we estimate the demand function of the system; more precisely, we estimate the probability of acceptance to amounts questioned and the number of users who will agree to pay. The functional form of demand to be estimated is assumed to be logistic, namely,

$$\text{Probability of acceptance} = 1 - 1/(1 + \exp(-\alpha - \beta \log WTP)).$$

The probability of acceptance is the ratio of the number of users who reply that they are willing to use the device at the amount of charges provided in the questions. The estimated coefficients  $\alpha$  and  $\beta$  are summarized in Table 6, 7, 8, and 9 for four regions.

The estimated demand function for Kamaishi's e-health system is shown in Fig. 1. The average WTP is calculated as the area under this demand function, which results in being 4,519 yen (approximately US\$37) per user per month. The demand functions of other regions have the same form.

In sum, WTP of Kamaishi, Nishiaizu, Katsurao, and Sangawa are 4,519 yen, 3,177 yen, 1,640 yen, and 2,955 yen, respectively.





**Table 2. Distributions of replies: Kamaishi**

|              |       |       |       |       |       |       |       |        |
|--------------|-------|-------|-------|-------|-------|-------|-------|--------|
| WTP (yen)    | 2,500 | 3,500 | 4,500 | 5,500 | 6,500 | 7,500 | 8,000 | 10,000 |
| No. of users | 112   | 69    | 8     | 62    | 11    | 12    | 1     | 16     |

**Table 3. Distributions of replies: Nishi-aizu**

|              |       |       |       |       |       |       |       |        |
|--------------|-------|-------|-------|-------|-------|-------|-------|--------|
| WTP          | 0     | 500   | 750   | 800   | 1,000 | 1,500 | 2,000 |        |
| No. of users | 141   | 4     | 2     | 1     | 54    | 7     | 3     |        |
| WTP          | 2,500 | 3,500 | 4,500 | 5,500 | 6,500 | 7,500 | 8,000 | 10,000 |
| No. of users | 67    | 38    | 1     | 69    | 6     | 13    | 9     | 5      |

**Table 4. Distributions of replies: Katsurao**

|              |       |       |       |       |       |       |       |        |
|--------------|-------|-------|-------|-------|-------|-------|-------|--------|
| WTP (yen)    | 0     | 500   | 1,000 | 1,500 | 2,000 | 2,500 | 3,000 |        |
| No. of users | 85    | 95    | 113   | 25    | 50    | 3     | 14    |        |
| WTP (yen)    | 3,500 | 4,000 | 4,500 | 5,000 | 5,500 | 6,500 | 8,000 | 10,000 |
| No. of users | 2     | 1     | 1     | 9     | 2     | 1     | 1     | 7      |

**Table 5. Distributions of replies: Sangawa**

|              |       |       |       |       |       |       |        |
|--------------|-------|-------|-------|-------|-------|-------|--------|
| WTP          | 0     | 500   | 1,000 | 1,500 | 2,000 | 2,500 | 3,000  |
| No. of users | 93    | 3     | 14    | 38    | 3     | 23    | 5      |
| WTP          | 3,500 | 4,000 | 4,500 | 5,000 | 5,500 | 6,500 | 10,000 |
| No. of users | 6     | 0     | 17    | 1     | 2     | 2     | 2      |

**Table 6. Results of estimation of demand function: Kamaishi**

|          | Estimated | Standard error | <i>t</i> -value | <i>p</i> -value |
|----------|-----------|----------------|-----------------|-----------------|
| $\alpha$ | 27.441134 | 2.0739667      | 13.231          | 0.0000          |
| $\beta$  | 3.3033572 | 0.2474318      | 13.351          | 0.0000          |

**Table 7. Results of estimation of demand function: Nishi-aizu**

|          | Estimated | Standard error | <i>t</i> -value | <i>p</i> -value |
|----------|-----------|----------------|-----------------|-----------------|
| $\alpha$ | 29.866890 | 1.9972607      | 14.954          | 0.0000          |
| $\beta$  | 3.7547249 | 0.2382865      | 15.757          | 0.0000          |

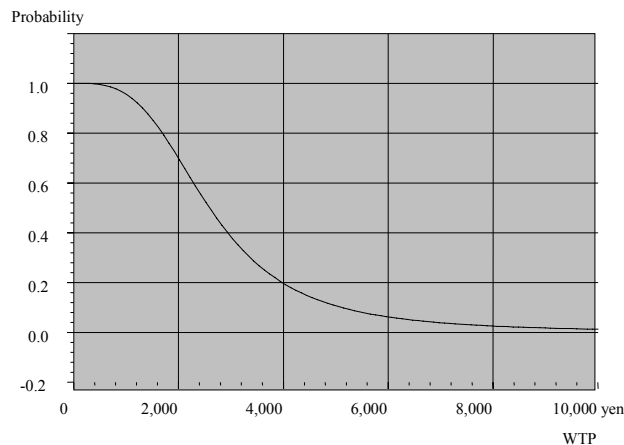


**Table 8. Results of estimation of demand function: Katsurao**

|          | Estimated | Standard error | t-value | p-value |
|----------|-----------|----------------|---------|---------|
| $\alpha$ | 15.800151 | 1.8418040      | 11.775  | 0.0000  |
| $\beta$  | 2.2491025 | 0.1794782      | 12.498  | 0.0000  |

**Table 9. Results of estimation of demand function: Sangawa**

|          | Estimated | Standard error | t-value | p-value |
|----------|-----------|----------------|---------|---------|
| $\alpha$ | 25.470392 | 2.5995525      | 9.798   | 0.0000  |
| $\beta$  | 3.2403337 | 0.3281845      | 9.874   | 0.0000  |



**Figure 1. Estimation of demand function: Kamaishi**

## 4. Cost-Benefit Analysis

### 4.1. Total Benefits

In our Cost-Benefit Analysis, total benefits and costs are compared over a period of several years. In this paper, the time span is set at six years because all device parts are held as inventory for six years. WTP obtained above is per user per month, and the total yearly value is obtained by multiplying by the number of users and by 12 months. At the time of the survey (October 2000), 348 users were registered. One-year benefits total approximately 18,871,344 yen (US\$157,261). In order to obtain six years' worth of benefits, we calculate the present value of six years' benefits with a 4% discount rate, and we assume that the number of registered

users remains constant for six years. This yields a six-year benefit total of 95,782,869 yen (US\$798,182). In the same way, total benefits of the other three are also calculated, which are shown in Table 10.

### 4.2. Total Costs

The total costs of the system include costs for equipment, salaries and wages of doctors and nurses, and other miscellaneous costs which include maintenance. The cost of host computers at the hospital is about 1,700,000 yen (US\$14,166) which includes instalment fees of 250,000 yen (US\$2,083) per device. The total costs for 200 devices are about 37,600,000 yen (US\$313,333). Development of software prior to instalment costs 4,000,000 yen (US\$33,333), which is also considered part of the initial costs. We



assume that 10% of the value of equipment can be considered its current value after six years of operation; therefore 90% of the value of equipment is included in the costs, which amounts to 38,970,000 yen (US\$324,750). This is paid for with a one-time payment at the beginning.

Regarding salaries and wages, a part-time doctor is paid about 1,728,000 yen (US\$14,400) per year, a full-time nurse is paid about 5,040,000 yen (US\$42,000), and one part-time worker receives about 1,800,000 yen (US\$15,000). The doctor and part-time worker spend half of their working hours on this system. Thus, total yearly salaries are 8,568,000 yen (US\$71,400). Other maintenance costs, such as those for printer toner and postage, are about 1,851,600 yen (US\$15,430). As mentioned earlier, there are no telecommunications charges. Thus, total annual operational costs are about 10,419,600 yen (US\$86,830). According to this calculation of six years' operational costs, as well as those of equipment (with the 4% discount rate), the costs of the e-health system in Kamaishi City total 95,782,869 yen (US\$789,190). The total costs of the other regions are summarized in Table 11.

### 4.3. B/C Ratio

From the above calculation, benefits total 95,782,869 yen (US\$798,182), whereas costs total 95,782,869 yen (US\$789,190) over the period of six years. Therefore, the B/C ratio is 1.07; that is, benefits exceed costs. This is a rather surprising result, because all other local governments where field research was conducted show ratios of less than 1. The B/C ratio being larger than 1 does not necessarily indicate that the medical corporation, *Rakuzankai*, earns a positive profit. Since its revenues consist of user charges of 2,500 yen, this amount is smaller than that of WTP; more precisely, 211 peripheral devices are being leased, and 2,500 yen is being charged per family. The monthly revenue from rental devices thus totals 527,500 yen (US\$4,396), with annual revenue of 6,330,000 yen (US\$52,750). On the other hand, annual operational costs as

obtained in the previous subsection are approximately 10,419,600 yen (US\$86,830). Thus, revenues are far less than operational costs. Though *Rakuzankai* is losing money through this operation, losses are covered by other hospital income. The B/C ratios of the other regions are shown in Table 11.

### 4.4. B/C Ratio for Local Governments

If the e-health system of Kamaishi City were considered as a private business, its profits would be negative; however, as a public project, it accrues more benefits to society than costs. For the other three local governments, the costs are larger than benefits. Why, one might ask, do these other regions implement such projects? This is a common question facing all local governments in Japan. For our answer, let us consider the matter from the viewpoint of the local governments themselves. Most e-health systems in Japan are supported by the central government, through subsidies to local governments that match the cost of equipment.

For all local governments other than Kamaishi City, the B/C ratio is smaller than 1. This difference is due to the costs of equipment; Katsurao Village purchased rather sophisticated and expensive devices capable of transmitting motion pictures. Nishiaizu and Sangawa received subsidies from different Ministries at different times, and since these devices cannot be utilized as one single system, it was necessary to install various kinds of equipment for each system. Kamaishi City purchased the lowest-cost devices; this is the primary reason that its B/C ratio is the highest. Due to central government subsidies, local governments must only cover operating costs such as salaries and wages, and maintenance costs. The B/C ratios recalculated in this manner are labelled as (B/C)\* in Table 11. All (B/C)\* values are larger than 1. This is the answer to the above question – for local governments with access to central-government subsidies, e-health projects accrue more benefits than costs.

**Table 10. Comparison of Benefits**

Unit: yen

|           | No. of users | WTP               | Benefit (6 years)              | Fee   |
|-----------|--------------|-------------------|--------------------------------|-------|
| Kamaishi  | 405          | 4,519 (US\$41.08) | 102.9mil. (US\$953.5 thousand) | 2,500 |
| Katsurao  | 926          | 1,640 (\$14.91)   | 99.4 (\$904.6)                 | none  |
| Nishiaizu | 518          | 3,177 (\$28.88)   | 107.7 (\$979.1)                | none  |
| Sangawa   | 484          | 2,955 (\$26.86)   | 106.5 (\$968.2)                | none  |



**Table 11. Comparison of costs and B/C ratio**

|                 | Kamaishi                         | Nishiaizu                           | Katsurao                            | Sangawa                             |
|-----------------|----------------------------------|-------------------------------------|-------------------------------------|-------------------------------------|
| No. of devices  | 211                              | 400                                 | 325                                 | 225                                 |
| Equipment       | 39.9 mil. yen<br>(\$36.3 thou.)  | 136.7 mil. yen<br>(\$1,243thou.)    | 111.4 mil. yen<br>(\$1,012.7thou.)  | 133.5 mil. yen<br>(\$1,213.6 thou.) |
| Salaries        | 8.6 mil. yen<br>(\$78.2 thou.)   | 3.7 mil. yen<br>(\$33.6 thou.)      | 3.36 mil. yen<br>(\$30.5 thou.)     | 4.5 mil. yen<br>(\$40.9 thou.)      |
| Others          | 1.9 mil. yen<br>(\$17.3 thou.)   | 1.9 mil. yen<br>(\$17.3 thou.)      | 10.4 mil. yen<br>(\$94.5 thou.)     | 3.0 mil. yen<br>(\$27.3 thou.)      |
| Costs (6 years) | 95.5 mil. yen<br>(\$868.2 thou.) | 184.5 mil. yen<br>(\$1,677.3 thou.) | 184.2 mil. yen<br>(\$1,674.5 thou.) | 174.3 mil. yen<br>(\$1,584.5 thou.) |
| B/C             | 1.07                             | 0.58                                | 0.54                                | 0.61                                |
| (B/C)*          | 1.87                             | 2.31                                | 1.42                                | 2.60                                |

the four cases are summarized in Table 16.

## 5. Cost-sharing among parties

### 5.1. Decomposition of WTP into effects

The previous result shows that WTP in Kamaishi is 4,519 yen. In this section, we decompose this WTP into the various discrete effects of the e-health system. We then calculate who bears the costs of the e-health system examined in the previous section, and to what extent.

In the survey of users, we asked whether the system led to four positive effects: (a) reduced anxiety in day-to-day life, (b) stabilization of illness, (c) enhancement of health consciousness, and (d) decrease in medical expenses. We regressed the users' stated WTP based on their replies to these questions. Thus, we estimated the following equation:

$$W = ax_1 + bx_2 + cx_3 + dx_4 + e,$$

where  $a$ ,  $b$ ,  $c$ , and  $d$  are coefficients to be estimated,  $x_1$ ,  $x_2$ ,  $x_3$ , and  $x_4$  are dummy variables that each take a value of 1 if a reply is "yes" and 0 if it is "no," and  $e$  is an error term. The results are shown in Table 12, 13, 14, and 15 for four cases.

Based on the above estimation, the extent to which each effect had monetary value was calculated; that is, how the WTP is decomposed into these four components. The p-value of "decrease in medical expenses" was not significant, so this component was ignored. The results of

### 5.2. Exact Amount of Reimbursement

The effects of enhancement of health consciousness and reduced anxiety in day-to-day life offer benefits to the users themselves, since these users' personal lives are enriched. Stabilization of illness has an additional benefit to society, as it results in decreased medical expenditures, conservation of medical resources, etc. That is, this effect carries a positive externality. This argument leads us to suggest who should bear the cost of the e-health system. If an effect benefits individual users, then they are willing to pay that amount. On the other hand, if it benefits all of society, then society can collectively bear that fraction of the cost. According to the above figures, in Kamaishi City individual users bear 2,763 yen in cost, while society pays 349 yen as reimbursement. In Nishiaizu, reimbursement is 439 yen. Surprisingly, the amount of 2,763 yen obtained here is very close to the amount of 2,500 yen charged by Kamaishi City (for detailed discussion on reimbursement, see [6] and [7]).



**Table 12. Result of estimation: Kamaishi**

|                                     | Co-efficient | Stand. Error | <i>t</i> -value | <i>p</i> -value |
|-------------------------------------|--------------|--------------|-----------------|-----------------|
| Stabilization of illness            | 979.0        | 298.6        | 3.3             | < 0.001         |
| Enhancement of health consciousness | 2612.9       | 224.1        | 11.7            | < 0.001         |
| Reduced anxiety in day-to-day life  | 1535.7       | 264.3        | 5.8             | < 0.001         |
| Decrease in medical expenses        | 767.9        | 701.8        | 1.1             | <0.28           |

**Table 13. Result of estimation: Nishi-aizu**

|                                     | Co-efficient | Stand. Error | <i>t</i> -value | <i>p</i> -value |
|-------------------------------------|--------------|--------------|-----------------|-----------------|
| Stabilization of illness            | 753.1        | 299.3        | 2.51            | < 0.012         |
| Enhancement of health consciousness | 1340.5       | 340.4        | 3.93            | < 0.000         |
| Reduced anxiety in day-to-day life  | 909.6        | 366.7        | 3.48            | < 0.014         |
| Decrease in medical expenses        | 493.3        | 348.5        | 1.42            | <0.158          |

**Table 14. Result of estimation: Katsurao**

|                                     | Co-efficient | Stand. Error | <i>t</i> -value | <i>p</i> -value |
|-------------------------------------|--------------|--------------|-----------------|-----------------|
| Stabilization of illness            | 329.5        | 463.5        | 0.71            | < 0.478         |
| Enhancement of health consciousness | 446.0        | 256.4        | 1.74            | < 0.083         |
| Reduced anxiety in day-to-day life  | 1075.8       | 268.7        | 4.00            | < 0.000         |
| Decrease in medical expenses        | 887.2        | 1217.8       | 0.73            | <0.487          |

**Table 15. Result of estimation: Sangawa**

|                                     | Co-efficient | Stand. Error | <i>t</i> -value | <i>p</i> -value |
|-------------------------------------|--------------|--------------|-----------------|-----------------|
| Stabilization of illness            | 402.0        | 409.9        | 0.98            | < 0.329         |
| Enhancement of health consciousness | 302.6        | 380.5        | 0.80            | < 0.428         |
| Reduced anxiety in day-to-day life  | 2190.2       | 321.4        | 6.81            | < 0.000         |
| Decrease in medical expenses        | -130.0       | 522.4        | -0.24           | <0.804          |

**Table 16. Decomposition of WTP into effects**

| Effect                              | Value (yen)     |                 |                 |                 |
|-------------------------------------|-----------------|-----------------|-----------------|-----------------|
|                                     | Kamaishi        | Nishi-aizu      | Katsurao        | Sangawa         |
| Stabilization of illness            | 349             | 439             | not significant | not significant |
| Enhancement of health consciousness | 1,834           | 1,075           | 179             | not significant |
| Reduced anxiety in day-to-day life  | 929             | 680             | 475             | 774             |
| Decrease in medical expenses        | not significant | not significant | not significant | not significant |







## 6. Conclusion

So far, we have conducted surveys of four local governments. Except for Kamaishi City, their B/C ratios are approximately 0.5; that is, benefits cover only half of the costs. In addition, regarding the frequency of usage of the device, Kamaishi City has a much higher ratio than the other local governments. Thus, Kamaishi City demonstrates very distinctive characteristics. This is due to their efforts to promote usage, such as a users' association that organizes events to enhance health consciousness, and to the participation in the system of medical doctors, which increases the users' reliance on the system.

It is clear from our previous studies that the e-health system is useful for consultation and for maintaining the good health of the elderly and patients suffering from chronic diseases who are in stable condition, but that e-health is not effective for curing disease. E-health has psychological benefits, such as providing a sense of relief to its users by the knowledge of being monitored by a medical institution 24 hours a day. This makes it difficult to estimate its benefits in concrete terms. We are only able to provide estimates of these amounts in this paper.

To date, Japanese local governments implementing this system have not charged any fee, except for Kamaishi City, instead receiving all their money in subsidies from the central government. Kamaishi City itself receives subsidies from the Ministry of Welfare through collaboration in two of its projects – the Special Project for Promoting Regional Health, and the Pilot Project for Promoting Tele-Medicine. Katsurao Village receives subsidies from the Ministries of Welfare Agriculture, Nishiaizu from the Ministry of Welfare, the Land Planning Agency, and the Ministry of Agriculture, and Sangawa from the Ministries of Welfare and Agriculture. Due to the current severe fiscal situation in Japan, however, local governments can no longer rely on such subsidies. For sustainability of the e-health system from a financial point of view, a new framework is required. Reimbursement using medical insurance is one possibility. Analyses such as the Cost-Benefit Analysis carried out in this paper provides a theoretical foundation for such reimbursement to be realized.

In this paper we selected four regions, applied CVM to each e-health system, and made a comparison of the results. This, however, is not the most appropriate method for such an estimation. It is necessary to pool the data of the four regions, and then proceed with the analysis. This is our task for the future.

## 10. References

- [1] M. Tsuji, "The Telehomecare/e-health System in Japan," *Business Briefing: Global HealthCare*, pp. 72-6, London, Jan. 2002
- [2] S. Miyahara, M. Tsuji, C. Iizuka, and F. Taoka "The Willingness of Japanese Citizens to pay for e-Health Systems," *Journal on Information Technology in Healthcare*, Vol. 4, No. 2, April 2006
- [3] S. Miyahara, M. Tsuji, C. Iizuka, T. Hasegawa, F. Taoka, and M. Teshima, "An economic evaluation of Japanese telemedicine, focusing on teleradiology and telepathology," forthcoming from *Journal of Telemedicine and Telecare*.
- [4] S. Miyahara, M. Tsuji, C. Iizuka, T. Hasegawa, and F. Taoka, "On the Evaluation of Economic Benefits of Japanese Telemedicine and Factors for Its Promotion," forthcoming from *Journal of Telemedicine and eHealth*.
- [5] M. Tsuji, W. Suzuki and F. aoka "The application of CVM for assessing the tele-health system: an analysis of the discrepancy between WTP and WTA based on survey data," in *Assets, Beliefs, and Equilibria in Economic Dynamics: Essays in Honor of Mordecai Kurz*, ed. by C. D. Aliprantis, K. Arrow, P. Hammond, F. Kubler, H.-M. Wu, N. Yannelis, pp. 493-506, Springer, September.
- [6] M. Tsuji, W. Suzuki, F. Taoka, K. Kamata, H. Ohsaka, and T. Teshima, "An Empirical Analysis of the Assessment of the Tele-Home-Care System and Burden of Costs: Based on a Survey on Kamaishi City" (in Japanese), *Japan Journal of Medical Informatics*, Vol. 22, No. 1, pp. 76-89, April 2002.
- [7] M. Tsuji, W. Suzuki, and F. Taoka,, "An Empirical Analysis of a Telehealth System in term of Cost-sharing," *Journal of Telemedicine and Telecare*, Vol. 9, Suppl. 1, pp. 41-43, 2003





# Health Information on Cellphones over Wireless Application Protocol (WAP): Issues, Obstacles, and Strategies for the Philippines

Alvin B. Marcelo, MD, UP College of Medicine-Medical Informatics Unit, 547 Herran St, Ermita, Manila, PHILIPPINES 1000 [alvinbmarcelo@yahoo.com](mailto:alvinbmarcelo@yahoo.com)

## Abstract

Medical records are the backbone of the health care service sector. Automating medical records has been one of the challenges in the IT industry and has been shown to bring numerous benefits for patients, health providers, and facilities. Globally however, stakeholders in developing countries have been slow in adopting information and communications technology into their health careflow because of the significant investment in hardware and software. The Philippines is the texting capital of the world with more than 25 million people with access to a cellphone. This study demonstrates a prototype electronic medical record that is accessible with WAP-enabled cellphones (wireless application protocol). Technical, economic and clinical considerations attendant with such portable wireless medical records are discussed as well as strategic directions for future development in the Philippines.

## 1. Introduction

Medical records are the backbone of any health facility. It is a document that contains all the transactions that are related to the care of an individual patient. Its uses range from patient care, to research, to teaching, and even to medicolegal cases.

In the past fifty years, as the field of computer science matured, there have been attempts at shifting the medical records system from paper-based to electronic format. Called electronic medical records (EMR), these new systems were proffered to be better than paper-based records because of the enhancements that come with automation such as speed, reliability, legibility, and efficiency. Together with this shift to EMR however, was an increase in mobility of patients and the re-orientation of many health systems to managed care. This created a demand to make their medical records portable across different health facilities. Health data portability is such a desired state

that in the United States that a special law was passed to ensure this feature.<sup>1</sup> Economic benefits of portability have been cited by experts in the range of billions of dollars.<sup>2</sup>

This study proposes providing portability to electronic medical records using cellphones as the access devices. The popularity of cellphones in the Philippine setting offers a unique opportunity to enhance its use beyond simply commercial messaging.

## 2. Objectives

Develop a prototype application for accessing an electronic medical record system using cellphones and the Internet

Determine factors which influence usability of cellphones for accessing electronic medical records over the Internet

Enumerate opportunities for future work in the area of portable medical records in the Philippine setting

## 3. Method

The author for this project adopted an existing electronic medical record system. The Community Health Information Tracking System (CHITS at [www.chits.info](http://www.chits.info)) is a product of the University of the Philippines Medical Informatics Unit through its lead developer, Dr. Herman Tolentino.<sup>3</sup> CHITS is a web-based application that can be used on any networked computers running Linux. It requires a server running Apache, MySQL, and PHP. The system has been functional in two health centers in Pasay City (Lagrosa and Malibay



Health centers in Pasay) and will be installed citywide in Marikina.

Although considered a successful project (a health technology project still used beyond one year based on Heeks criteria), CHITS has several concerns. First, it requires substantial investments in hardware. In a typical setup, CHITS will require four terminals: two at the reception area, one in the doctor's office, and another at the treatment room. These workstations should all be connected in a local area network with one designated as server. The system also necessitated a significant amount of training to be given to the end users who are mostly midwives and nurses who have not used computers at all in their whole life. Lastly, technical support and maintenance personnel must be readily available to fix problems as they occur. The value of CHITS decreased for every day that the application was down.

These concerns led the investigator to look for other devices that may be used to access CHITS. One such device is the cellphone. The cellphone is as ubiquitous to Philippine society as any other communications device. It is affordable (some selling as low as Php 1500) and is a networked device right out of the box. Technical support is readily available. And health workers without IT skills would invariably have some exposure to a cellphone than to a desktop computer and are less averse to the technology.

The key maneuver is to enable the cellphone to communicate with a web server that runs CHITS. The HTML (hypertext markup language) and WML (wireless markup language) hybrid adapted Web server (or HAWHAW at <http://www.hawhaw.de>) is a toolkit to create such universal mobile applications. Using HAWHAW's application programming interface (API), some elements of CHITS was ported into a WAP interface and delivered through a standard micro display viewed using a Siemens SL55 desktop emulator (see figure). A general packet radio services (GPRS) connection from myGlobe

(Globe telecommunications) was used in the actual prototype running on a Nokia 6230.

A typical CHITS application has the following modules:

Master patient index registration – adding a patient into the system  
Consult – starting and ending a consult for a patient already registered in the system  
National Tuberculosis Program (based on the Directly Observed Therapy Short Course of the Department of Health)  
Maternal Care module  
Expanded Programme on Immunization (EPI) module

There are other modules but for this proof of concept the NTP drug intake module was chosen.

For these data elements, the following functionalities were developed:

- Registering a patient into the master patient index
- Searching for a patient and his/her medical record number in the registry
- Starting and ending a consult
- Choosing a patient from the list of active consults
- Submitting an NTP drug intake report

For purposes of this study, only these functionalities were chosen as they represented the major components of electronic medical record systems. As simple as these components seem to be, they already brought out several issues and concerns related to the use of EMR in cellphones (see screenshot series).

#### 4. Observations

Observations were made during development and during implementation (alpha testing). The following parameters were noted during development: ease of development and display issues. During implementation the following





parameters were recorded: ease of use, data integrity, storage, security, ethics, quality, training, portability, reliability of connectivity, and cost.

## • Observations during development

### *Ease of development*

The HAWHAW application programming interface is fully documented and easy to use. There are two methods for developing applications with HAWHAW: using PHP or XML (eXtensible Markup Language [www.w3c.org](http://www.w3c.org)). The author opted to use XML because it is easier for novices to understand and because the language is supported by international standards governing bodies like the W3C consortium ([www.w3c.org](http://www.w3c.org)). It should be noted that some PHP still had to be used for processing the data that has been collected. Figure below shows how HAWHAW was placed within the CHITS application framework.



A slightly difficult part in the programming is the integration of logic into the application in anticipation of human-prone errors. An example of this is when a registered patient is re-registered inadvertently by the user. Logic must exist in the application to warn the end user that a duplicate entry of the same name will be created if he or she proceeds with the registration. As in any health application, end users have the privilege to override the logic if needed.

### *Display*

The Siemens SL55 test interface is limited to 4096 colors and 101x80 pixels (Nokia 6230 is 128x128 pixels). This is similar to most cellphones on the market although Nokia and other smartphones have larger displays. Due to this limitation in display, efficiency of information has to be considered when delivering content to the cellphone. Delivering standard web pages created for desktops will be impractical for cellphones. For one, the display will require significant amount of scrolling vertically and horizontally to view the full page. For another, the large sizes of web pages designed for desktops will consume substantial bandwidth and time if delivered in raw format to cellphones. It was for this reason that HAWHAW developer put a limit on the size of files to be displayed. The most logical option was to perform an abstraction of relevant data elements in CHITS and create the instruction set in XML that will eventually be delivered to the cellphone to display.



## Observations during implementation

Internal alpha testing revealed several issues as follows:

### *Display*

With an effective display area of 101x80 pixels, viewing the medical record becomes the most prominent issue when it comes to delivering health information to cellphones.





The small display size severely limits the amount of information that can be delivered to the end user<sup>4</sup>. In effect, completeness of data could not be guaranteed and the end user encounters great difficulty putting the piece-meal information into one coherent picture. To mitigate this limitation, an option is to deliver content in installments (several consecutive pages). Many other space saving techniques were employed like the use of small fonts, text wrapping, and removal of external field labels and placing them inside the fields as default values instead<sup>5</sup>. This last technique can save up to 33 to 50% of the display area but carries a higher risk for errors in data entry. (See figure below).

External labels (top); Labels are placed inside input boxes (bottom)



### ***Data integrity (synchronization with the master patient index)***

A fundamental component of any medical record system is the master patient index (MPI - the number that uniquely identifies every patient within the health facility). This number has to be recorded in every transaction between

facility and patient. Using a web-based database-driven system, synchronization with the MPI is inherent because there is only one central master patient index for all end users that they can access over the Internet. This synchronization is not easily achievable with stand-alone applications that do not share a central database.

Keeping the unique patient identifier in the cellphone in sync with the central master patient index then is easier with a web-based solution since in practical terms, there is only one database that is being accessed by all. Transactions made for and on behalf of the patient from any cellphone device are then recorded consistently with the same numerical identifier making later consolidation of data for any particular patient seamless and error-free. These unique identifiers are usually hidden from end users, as they are only machine-readable. End users are presented with the more familiar identifiers like name and age.

### ***Storage***

Using the web-based system, no data is actually kept in the cellphone. All data are stored in the central database and none is retained in the cellphone. The only data that may be saved is the web address of the application. No patient identifying information is saved. This is consistent with ethical use of health information.

### ***Security***

Security of the web-based application is made possible with usernames and passwords.

### ***Ethics***

Only authorized personnel with the proper usernames and passwords are able to proceed further and access the other features of the medical record. An audit trail is available





where records retrieved by end users are documented. Uses of health information outside of the application however cannot be documented and this is a weakness of portable records.

### **Quality**

Validation of forms may be embedded within the application to ensure quality of data entry; i.e., if the data entered is of the wrong format (e.g., entering letters when the data required are numerical), the application can alert the end user of the error and the correct response can be given.

### **Training**

Some amount of training is necessary in setting up GPRS services and in navigating web pages on cellphones. Another study is needed to determine the ease of use of the WAP interface to workers at various levels of the health system.

### **Portability**

Access to the most current version of the record is made possible if a centralized database is made available over the Internet. The prototype application was accessed in the Philippine General Hospital, in Quezon City, in Jose Rizal Provincial Hospital in Roxas City, Capiz, and theoretically, anywhere Globe GPRS services are available.

### **Reliability of connection**

There were instances when an error message "Network is not available" was seen on screen during testing. These errors were temporary in nature and the application would once again be available at a later time. One possible reason is cellular service network congestion if there are

too many concurrent users.

### **Costs**

Globe GPRS charges twenty-five centavos per kilobyte (one kilobyte of message has about 1000 characters). This would translate to around five centavos per NTP drug intake report. For a total of 180 days (the six month duration of TB drug intake) this would result to nine pesos charges incurred to report the drug intake of every patient using the cellphone.

### **Discussion: Opportunities for Future Research**

Several issues with regard delivery of medical information into cellphones were noted. These issues must be discussed as they have significant impact on the development strategy and on marketing of the end product.

### **Display**

Viewing medical records in cellphones are not only impractical but could also be dangerous. With small screens, it is difficult to get a complete picture of the patient's data and makes browsing tedious and ineffective. Efficient navigation of medical records through small interfaces will require health workers to undergo intensive training to search and to retrieve relevant information.

Cellphones are not being able to display all relevant data at the same time because of the limitation of the screen size. This becomes an issue to the physician wishing to see relationship of laboratory data with, perhaps, physical examination or clinical history. Such relationships are difficult to deliver on a small screen<sup>6</sup>. As such, the effect of small screen size





is the inability to deliver comprehensive information enough to make a clinical decision. The cellphone may not be an effective tool for viewing complex data. However, for smaller discrete data (e.g., latest vital signs or the latest serum potassium levels, or a warning of a possible drug interaction), the cellphone may be an adequate display tool.

In this study, the investigator opted to use the cellphone as an input device rather as a display device. This limited the usability of the cellphone for comprehensive medical record purposes but defined a safer and more effective way of employing this technology for health. In the future, cellphones may be best suited for collecting health data from the field and entering them directly into a central database. Other interface paradigms, like that of the personal digital assistant, a slightly larger interface common to smartphones, may also be used and tested <sup>7</sup>.

### ***Stand-alone application versus network-based***

In a multi-user and multi-protocol environment (such as a rural health unit in the Philippines), a central database system where there is consistent access to a master patient index is crucial for data integrity and optimal research. It is therefore important for portable wireless applications (such as the CHITS over cellphones) to query central databases. Stand-alone portable applications therefore have limited use in the aforementioned environments.

Although the prototype used in this study was made using wireless application protocol (network-based), a similar application with the same functionalities may be made using a programming language like Java. The output of such an activity will be a standalone program that can work on the cellphone even without a network connection. This may seem

advantageous since a reliable connection is not required and allows the end users to create and edit medical records even if they are offline. However, there are significant issues in consistency of data when several users try to edit the same record at separate times. Synchronization takes great effort and may become burdensome to the end users. Data integrity comes into question and may not be resolved unless sophisticated protocols and priority schedules are established.

On the other hand, a network-based solution where all users access a central database ensures that there is always only one consistent copy of any record on the system at any point in time regardless of the device used. Duplication is limited and efforts at consolidation are minimized. Maintenance of data and records is simplified.

### ***Connectivity and costs***

By using cellphones and the widely available wireless application protocol, this technology leverages a communication network that exists around the country. This opens up possibilities for the deployment of cellphone devices as tools for discrete data entry.

Eliminating the tedious paper trail and using a central database model, the model proposes shows a more timely, effective, and efficient way of collecting National Tuberculosis Program data directly from the field. Such a model may be employed with other similar programs of the Department of Health like malaria, HIV/AIDS, and dengue to name a few. It does not require sophisticated equipment like desktops or laptops connected to network devices (hubs and switches). All it requires is a WAP-enabled phone that may be purchased from any cellphone store around the country.

### ***Conclusion***





The ubiquity of cellphones in the Philippines makes them attractive for accessing electronic medical records. Although electronic medical records are in their early stages of development in the country, possibilities for portability, if built into these EMRs, can open up a large market among the health sector. Cellphones are well suited for discrete data entry especially in the field. The wide coverage of the telecommunications network companies makes this possible.

Devices with small displays however, are not optimally built to provide a comprehensive look into the whole electronic medical record. It is safer to limit their use for data entry and for field data encoding.





<sup>1</sup> Health Insurance Portability and Accountability Act (Kennedy-Kassebaum).

<http://www.hipaadvisory.com/regs/HIPAAprimer.htm>

<sup>2</sup><http://informationweek.com/story/showArticle.jhtml?articleID=161601166>

<sup>3</sup>Linking Primary Care Information Systems and Public Health Information Networks: Lessons from the Philippines. Herman Tolentino, M.D., Alvin Marcelo, M.D., Portia Marcelo, M.D., M.P.H., Inocencio Maramba, M.D., M.Sc. [Stud Health Technol Inform](#). 2005;116:955-60. ([http://iospress.metapress.com/\(o1czznmbz15u2uqym03dfju4\)/app/home/contribution.asp?referrer=parent&backto=issue,169,182;journal,3,13;linkingpublicationresults,1:300379,1](http://iospress.metapress.com/(o1czznmbz15u2uqym03dfju4)/app/home/contribution.asp?referrer=parent&backto=issue,169,182;journal,3,13;linkingpublicationresults,1:300379,1))

<sup>4</sup>Improving Web interaction on small displays. Matt Jones, Gary Marsden , Norliza Mohd-Nasir , Kevin Boone , George Buchanan. *Computer Networks* 31 (1999) 1129–1137

<sup>5</sup>An evaluation of integrated zooming and scrolling on small screens Steve Jonesa, Matt Jonesa, Gary Marsdenb, Dynal Patel b, Andy Cockburn. *Int. J. Human-Computer Studies* 63 (2005) 271–303

<sup>6</sup>Effective search results summary size and device screen size: Is there a relationship? Simon Sweeney, Fabio Crestani. *Information Processing and Management* Volume 42, Issue 4 , July 2006, Pages 1056-1074

<sup>7</sup>Improving web search on small screen devices. Matt Jonesa, George Buchananb, Harold Thimblebyc,\* . *Interacting with Computers* 15 (2003) 479–495.





# Designing a ubiquitous computer environment for medical education

Hyung Hoi Kim<sup>1,2</sup>, Tung Tran<sup>1</sup>, Hwa Sun Kim<sup>1</sup>, and Hune Cho<sup>1</sup>

<sup>1</sup>Dept. of Medical Informatics, Kyungpook National Univ., Republic of Korea

<sup>2</sup> Dept. of Laboratory Medicine and Biomedical informatics, Pusan National Univ., Republic of Korea  
 hhkim@pusan.ac.kr; ttadm20031@yahoo.com; polala@knu.ac.kr; hunecho@knu.ac.kr

## Abstract

*Ubiquitous access to computers and networks has led to major gains in the integration of technologies into the teaching process. The laptop requirement was initiated to produce a new generation of teachers who are well prepared to use the new tools for learning in their teaching practice. It was designed to immerse pre-service teachers in a technology-rich learning environment that provides ubiquitous anytime/anywhere access to technology tools, Internet-based resources, and an online communication environment to support and build communities of practices among all players.*

## 1. Introduction

Various learning systems have been developed in the past decade, and the majority of these systems are implemented with a client-server or are based on a centralized server. The client-server and centralized server approaches are metaphors for student-teacher and repository-centric, which reflect real-world learning scenarios in which teachers act as the content producers, while students act as the content consumers.

The ubiquitous learning environment provides mainframe computing, which is pervasive and interoperable. Ubiquitous computing is the universal interconnection of virtually any wireless/wired device nationwide, and beyond, a condition gradually being realized by ongoing advances in the Internet, smart appliances, and related technologies. Ubiquitous communication tools can be anything from mobile devices to tiny communication-capable sensors connected to high-speed networks, enabling information to be accessed and sent anytime, anywhere, in other words, ubiquitously. Today, with the development of wireless connectivity, mobile devices such as laptops, personal digital assistants (PDAs), and phones are now able to communicate from almost anywhere, inexpensively and at very high speeds, giving users great freedom in choosing when, where, and how to be connected.

The evolution of education and training at a distance can be characterized as a move from d-Learning (distance learning) to e-Learning (electronic learning) to m-Learning (mobile learning). These three stages of

development correspond to the influence of the Industrial Revolution of the 18th to 19th centuries, the Electronics Revolution of the 1980s, and the Wireless Revolution of the last years of the 20th Century on society[1, 2].

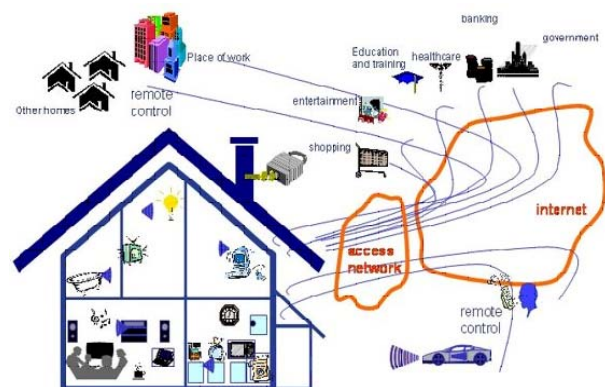


Figure 1. The ubiquitous computing model.

Currently, as more teachers and students use laptops or handheld computers on a daily basis, it is essential that policy makers and practitioners alike understand the effect that ubiquitous computing has on the classroom environment.

This paper presents the architecture of a prototype ubiquitous application for medical education. This paper is organized as follows. In the first sections, we present the general design principles used to specify the system's architecture, and the second section, we developed XML schemas that can be used to create and validate general XML documents for encoding essential data elements required for medical educations. The third section describes the process of generating XML schemas for educational templates. The next section reviews some of the issues encountered when adapting previous approaches to include additional template constraints. Finally, the fifth section includes conclusions and discussion for further work.

## 2. Materials and Methods

### 2.1 Design of system architecture

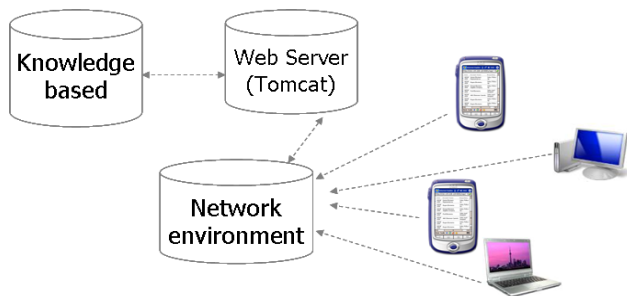
Ubiquitous computing and pervasive computing are emerging disciplines linking elements from distributed systems, mobile computing, embedded systems, human





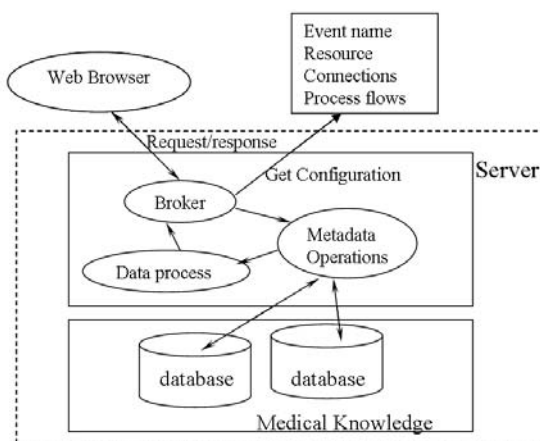


computer interaction, computer vision, and many other fields. Their vision is grounded in the belief that processors are becoming so small and inexpensive that they will eventually be embedded in almost everything. More specifically, the system will interface with the devices shown in Figure 2. In Figure 2, the Web server is a computer, which has Tomcat version 5.5.17 installed as a Web server application. Therefore, we can develop servlets that allow students or users to access information and carry on discussions with other users. We developed web modules that can obtain information from a knowledge base as the user requires. Then, it formats the retrieved information into the following XML schema [3]. In addition, the Extensible Markup Language (XML) data are transformed to HyperText Markup Language (HTML) using Extensible Stylesheet Language (XSL) to facilitate viewing on PDAs, notebooks, and desktop computers [4].



**Figure 2. Ubiquitous system architecture.**

Figure 3 is expressing the system architecture for medicine education based knowledge. In this architecture, client use web browser to send a request to server. And the event sender requires a reply. For example, user wants to know heart information. In the server side, Broker receive event from client request and lead Metadata operations module to query data from database.



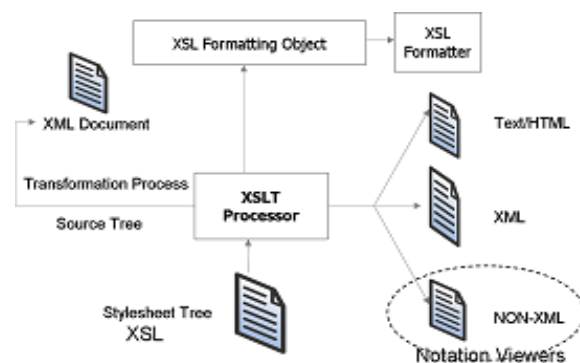
**Figure 3. Education system architecture.**

The database contains medical knowledge. The database receive query and return information according commands. The metadata operations module receives

data from database, and sends them to data process module. The data process module receives data and generates data to XML document following XML Schema. XML Schemas describe the structure and meaning of an XML document, in much the same way that a class describes an object in a program. To work with an XML document in a program we would like to map its components directly to a set of objects that reflect the document's meaning according to its schema. We can achieve this by compiling the schema into a set of derived content classes that can be validated. Data binding thus allows XML-enabled programs to be written at the same conceptual level as the documents they manipulate, rather than at the more primitive level of parser events or parse trees. Eventually, data process transform XML document to HTML using XSL and send it as pure HTML to browser.

## 2.2 XML schema for medical education

The W3C designate XML as the standard for the next –generation web document in 1998. XML is essentially a language describing another language, that is, a meta-language. XML separates the structure of the document into three parts: the part defining the structure of the document, the form of the expressing document, and the composition of the document. Thus it is important to composing the frame of the document, and Data Type Definition (DTD) solves this. However, because of the weak point that the validation of the data provided through the limited data type is impossible, the XML schema was emerged as new standard. Furthermore, XSL, XSLT and others through which transforms and transmits the mere XML documents suited to the environment of the user are considered as the related standards. Here the XSLT processor shares XPath which applies XSL to XML and with the result creates trees and can reads the desired contents using query [5].



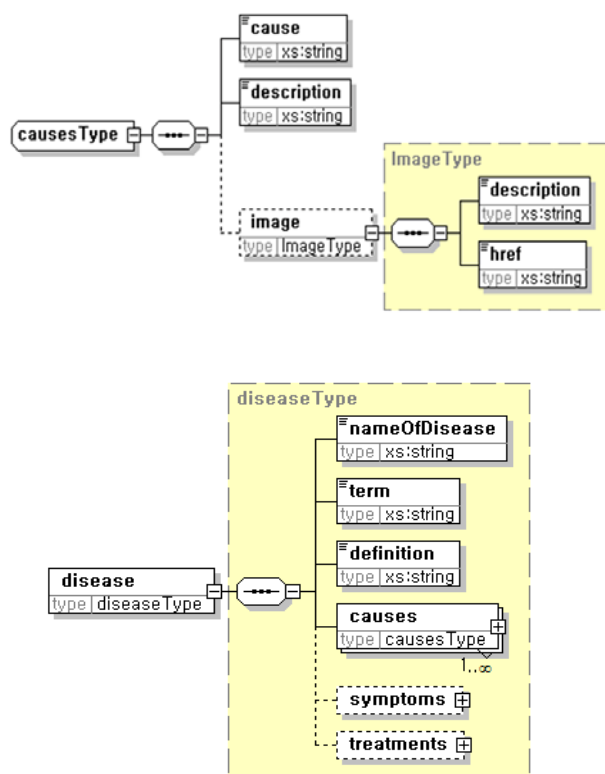
**Figure 4. XML Transformation process.**

Xpath provides the structural addressing mechanism for XSLT and the elements of the Xpointer-based XML documents is expressing the XML standard techniques



mentioned above and the object-oriented database storage which easily holds the meta-data of the XML document form was added. This indicates that the stored XML is shown through the web interface employing XSL, and the tasks are easily carried out through Xquery.

We can access knowledge bases for querying and obtaining information, anytime and anywhere. A personalized repository organizes the information that interests us, keeps track of how we access that information, and contains annotations about its relevance, quality, and comprehensibility. We access our repositories in our own ways. When we want more information, we can access repositories that trusted friends and associates are willing to share. Currently, the knowledge base stores information on more than 100 diseases, such as heart disease, kidney disease, and lung disease [6, 7].



**Figure 5. XML schema for medical education**

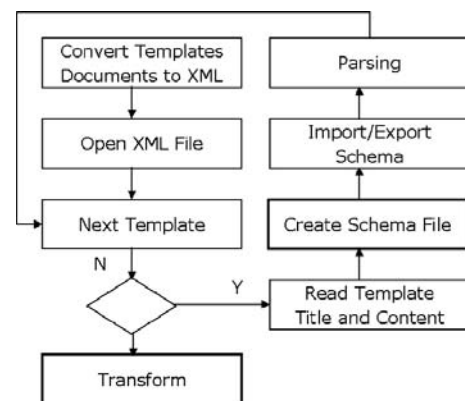
We used XML Spy to display the details of the schema for the content tree, as illustrated in Figure 5, which is a useful way of visualizing schema fragments. The schema <causesType>, <disease> and its children elements serve as the meta-model for the schema of the templates.

### 2.3 Data Extraction Module

In the case of Oracle 9i, there needs to be a process that maps XML elements and attributes them to the table

field. Since the XML document creates tables according to its complexity, even a single XML needs table mapping. A tree-type document object model (DOM) is used for extracting data from the relational database that supports open database connectivity (ODBC) [8].

Application programs are provided so that common words can be defined using internal DOM application programming interface (API) when storing data. Where a need occurs for a standard method of interaction with the data, DOM allows access to the content regardless of the structure, providing an object model for modeling an XML document. The discussion in the last section describes the basic XML schema for medical education template generation. The process consists of Conversion and Transformation. Conversion of template tables is an intermediate XML representation that attempts to recover some of the structural information inherent in the templates (Figure 6).



**Figure 6. XML schema generation process.**

This step is necessary because templates are presently defined with text and tables in a Microsoft Word document. Transformation of XML representations to XML schemas is accomplished with a JAVA program. This can also be accomplished with XSLT transformations.

### 3. Results

The knowledge access subsystem supports the natural ways people access information. In particular, it supports personalized, collaborative, and shared knowledge, to find the information that people need. It observes and adapts to its users, to meet their needs better. The subsystem integrates a variety components to gather and store data, to monitor user access patterns, and to answer queries and interpret data.

**Data representation:** The subsystem stores information encountered by its users using an extensible data model that links arbitrary objects via arbitrarily named arcs. There are no restrictions on object types or names. Users and the system alike can aggregate useful information regardless of its form (text, speech, images, video). The





arcs, which are also objects, represent relational (database-type) information, as well as associative (hypertext-like) links.

**Data acquisition:** The subsystem gathers as much information as possible about the topic of interest to a user. It does so through the raw acquisition of data objects, by analyzing the acquired information, by observing people's use of it, by encouraging direct human input, and by tuning access to the user.

**Automated observers:** Subsystems watch the queries that users make, the results they dwell on, the files they edit, the XML they send and receive, the documents they read, and the information they save. The system exploits observations of query behavior by converting query results into an XML document according to the XML schema, and using XSL to transform it into HTML to represent information (Figure 7, 8). The subsystem generates the results as an XML document after aggregating information from the database. In addition, it also creates an XSL style sheet in order to describe how the XML document should be displayed.

```
<?xml version="1.0" encoding="UTF-8"?>
<xsl:stylesheet version="1.0" xmlns:xsl="http://www.w3.org/1999/XSL/Transform" xmlns:fo="http://www.w3.org/1999/XSL/Transform" >
  <xsl:template match="diseaseCollection">
    <center>
      <h1><xsl:value-of select="disease/nameOfDisease"/></h1>
      <p><b>Medical Terminology:</b><xsl:value-of select="disease/term"/></p>
      <p><b>What is a</b><xsl:value-of select="disease/nameOfDisease"/><b>?</b></p>
      <blockquote><xsl:value-of select="disease/definition"/></blockquote>
      <br>
      <p><b>What causes a heart attack?</b></p>
      <blockquote>
        <ul>
          <xsl:for-each select="disease/causes">
            <li><p><b><xsl:value-of select="cause"/></b></p>
              <p><blockquote><xsl:value-of select="description"/></p>
              <img/><xsl:attribute name="src"><xsl:value-of select="image/href"/></xsl:attribute></img>
            </li>
          </xsl:for-each>
        </ul>
      </blockquote>
      <br>
      <p><b>What are the symptoms?</b></p>
      <blockquote>
        <ul>
          <xsl:for-each select="disease/symptoms/symptom">
            <li><p><b><xsl:value-of select="symptom"/></b></p>
              <xsl:for-each>

```

Figure 7. XSL for medical XML document.

In the existing medical education sites on the web, there are education data in their own way. Most of the data are stored on the certain tables of the DBMS systems. They are providing the lectures to the users through transforming the data into the form of HTML by using the server-side scripts of ASP, JSP, and PHP and sending them to the browser. In the XML system, XML schema is located on the outside, not in the lecture sites in which XML and XSL reside. In other words, the definition for the lecture data is specified by another standard enterprise. DOM loads XML and XSL located in the education sites, and before loading XML, this XML data tests the validity according to XML schema specified by the standard enterprise, and XSL is applied to XML data by the DOM's XSLT processor, with which result HTML is created and transmitted to the client browser by the web server, so that there happens no blockade in the data sharing.

the higher blood levels of adrenaline released from the adrenal glands during the morning discussed, may contribute to rupture of cholesterol plaques. Approximately 50% of patients with warning symptoms such as exertional angina or rest angina prior to their heart attacks.

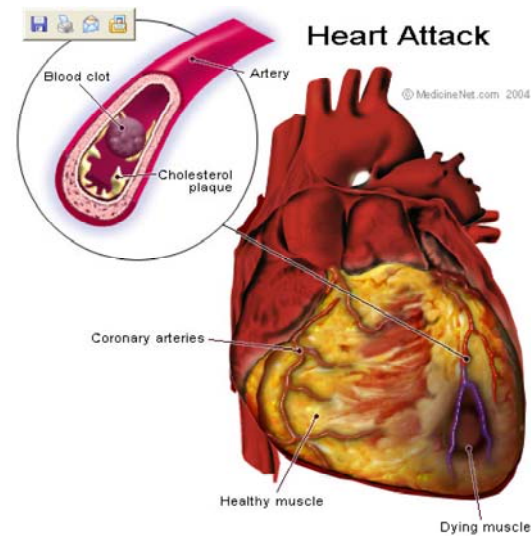


Figure 8. XSL describes how the XML document should be displayed.

#### 4. Conclusion and Future works

Ubiquitous computing leads to ubiquitous learning and technology provide abundant opportunities for sharing information, constructing knowledge, and stimulating personal growth. It is pervasive and prevalent in many interactions among students, faculty, administration, and staff. Web-based learning is currently showing its technological potential as an environment for learning. In recent years, medical schools have responded in a variety of ways to the challenges of modernizing traditional disciplinary-based courses [9-12]. The trend has been move to various integrated curriculum models, including problem-based learning (PBL). The availability of computer-based systems to support curriculum management and the delivery of learning materials (e-learning), may provide on way of supporting students undertaking PBL without increasing existing teaching loads.

In the future, we will extend the system so that it provides more clinical data, and users who are connected to the server can discuss it in real time.

The system proposed and implemented in this study bears the following characteristics:

- Ability to understand and solve problems in healthcare for students can be expected to increase.
- Students participation to the discussion in the cyber classes can be expected to increase compared with the traditional methods





- If two research teams are successful in developing and implementing course material by using PBL approach, they can apply similar approaches to other courses offered by university for healthcare.
- Knowledge and information sharing can be achieved for the joint research team.

## Acknowledgements

This work was supported by the Brain Korea 21 program from the Ministry of Education and Pusan National University Research Grant, 2005.

## References

- [1] B. Van Grimbergen, G. Thienpont, G. J. De Moor, and H. Buysse, "Test-based E-Learning: a new path in training medical students," *Stud Health Technol Inform*, vol. 93, pp. 33-8, 2002.
- [2] J. M. Renard, S. Leroy, H. Camus, M. Picavet, and R. Beuscart, "Personalising e-learning modules: targeting Rasmussen levels using XML," *Stud Health Technol Inform*, vol. 95, pp. 732-7, 2003.
- [3] P. Spencer, *Professional XML Design and Implementation*, Wrox, Chicago, IL, 1999.
- [4] Extensible Style-sheet Language, available at: <http://www.w3.org/Style/XSL/>, Dec. 2004.
- [5] Developer's Guide – Chapter 4 XML Query Structure, available at : <http://umlsks.nlm.nih.gov/>, June 12, 2006.
- [6] Information technology, Knowledge base, available at : <http://kb.iu.edu/>, July 17, 2006.
- [7] eMedicine Clinical Knowledge Base, available at: <http://www.emedicine.com/>, May 23, 2006.
- [8] R. Ramakrishnan and J. Gehrke, *Database Management Systems*, McGraw-Hill International edition, New York, NY, 2002.
- [9] K. Schilling, J. Wiecha, D. Polineni, and S. Khalil, "An interactive web-based curriculum on evidence-based medicine: design and effectiveness," *Fam Med*, vol. 38, pp. 126-32, 2006.
- [10] J. G. Ruiz, M. J. Mintzer, and R. M. Leipzig, "The impact of E-learning in medical education," *Acad Med*, vol. 81, pp. 207-12, 2006.
- [11] P. Moule, "E-learning for healthcare students: developing the communities of practice framework," *J Adv Nurs*, vol. 54, pp. 370-80, 2006.
- [12] I. Heeteby, M. Hatcher, and H. Tabriziani, "Web based Health Education, E-learning, for weight management," *J Med Syst*, vol. 29, pp. 611-7, 2005.





# Virtual Site Visit Abroad for EMR/PACS system by Broad Banded Teleconference system

Naoki Nakashima<sup>a</sup>, Shuji Shimizu<sup>b</sup>, Koji Okamura<sup>c</sup>, Ho-Seong Han<sup>g</sup>, Michio Kimura<sup>h</sup>

<sup>a</sup> Department of Medical Informatics, Kyushu University Faculty of Medicine, Fukuoka, Japan

<sup>b</sup> Department of Endoscopic Diagnostics and Therapeutics, Kyushu University Faculty of Medicine, Fukuoka, Japan

<sup>c</sup> Computing and Communications Center, Kyushu University, Fukuoka, Japan

<sup>g</sup> Department of Surgery, Seoul National University Bundang Hospital, Seoul, Korea

<sup>h</sup> Department of Medical Informatics, Hamamatsu University Hospital, Hamamatsu, Japan

Correspondence: [nnaoki@info.med.kyushu-u.ac.jp](mailto:nnaoki@info.med.kyushu-u.ac.jp)

## Abstract

*In order to promote exchange of knowledge and standardization of the medical technique, social medical systems, medical ethics, new medical items, medical science, and medical education in Asia-Pacific area, we have established a medical network with high-quality moving image over broad-banded Internet lines since February 2003. Broadcasting real-time teleconference and live demonstration with medical-quality videos using Digital Video Transport System (DVTS) were so useful to learn advanced medical techniques and equipments including electronic medical record (EMR) / picture archiving and communication system (PACS) beyond borders. We have done 58 events until March 2006. In this project, we conducted a virtual site visit at Seoul National University Bundang Hospital from Yokohama venue of 25<sup>th</sup> Joint Conference of Medical Informatics to see the advanced EMR/PACS by the network in November 2005. A remote lecture about Korean medical insurance system was transmitted from Seoul. The session was given by Japanese language. Over than 260 participants, including medical doctors, medical engineers, hospital office workers, vender company staffs, and students, in Yokohama venue enjoyed and learned the system remotely for 2 hours. The result of questionnaire survey suggested high satisfaction on the session.*

*We concluded that the medical informatics is one of suitable fields for the broad-banded telemedicine network as the content.*

## 1. Introduction

Level of medical care is varied among the countries because of a disparity in economic power and differentials in policies, religions and customs. The medical information does not distribute easily beyond the borders, and medical staffs usually are not aware of the existence of these differentials among countries. We may resolve part of the problem by teleconference system with high quality moving image (1).

A new network communication system, digital video (DV) over Internet protocol (IP), can transmit a real-time high quality moving image by low cost. Recently, we can use international submarine optic fiber cable networks in the Asia-Pacific area for broad-banded transmission. We have thought that we could visit at a lot of hospitals abroad virtually, and inspect variety of medical techniques and equipments easily by this network system. We have mainly transmitted the contents of surgical field so far (2). In this time, we conducted "virtual site visit at Korea for medical informatics system" in the biggest annual meeting for medical informatics in Japan. Here we report the record of the Asia-Pacific medical network project we have established, and the virtual site visit session including network configuration and the effect of the achievement.

## 2. Material and Methods

### 2.1 Organization

The Hyeonhae/ Genkai project was established to use the Korea-Japan Cable Network (KJCN) for development of informatics research and friendship between Korea and Japan in 2002. As a subproject of the Hyunhae/ Genkai project in medical field, we joined it to use KJCN for medical purpose in February 2003 (2, 3).

### 2.2 Network







We connected the venue in Pacifico Yokohama to Otemachi AP of NTT communications Co. in Tokyo by Gigastream (<http://www.gigast.com/en/index.html>). We used Japan Gigabit Network (JGN2: <http://www.jgn.nict.go.jp/index.html>) from Otemachi AP to Nagahama AP of KDDI (<http://www.kddi.com/english/index.html>), and connected to Asia-Pacific information infrastructure (APII) test bed through KJCN. We used Korea Advanced Research Network (KOREN, <http://www.koren21.net/eng/index.php>) in Korea (1Gbps) up to SNU Bundang hospital.

### 2.3. Terminal systems organization for moving image transmission

We used Digital Video Transfer System (DVTS) over Internet protocol for mutual communication. DVTS is open sourced free ware and we can download it from Web site (<http://www.dvts.jp/>). We also used AREL Spotlight (<http://www.arelcom.com/products.html>) for image sharing on display of EMR /PACS in the Yokohama venue and SNU Bundang hospital.

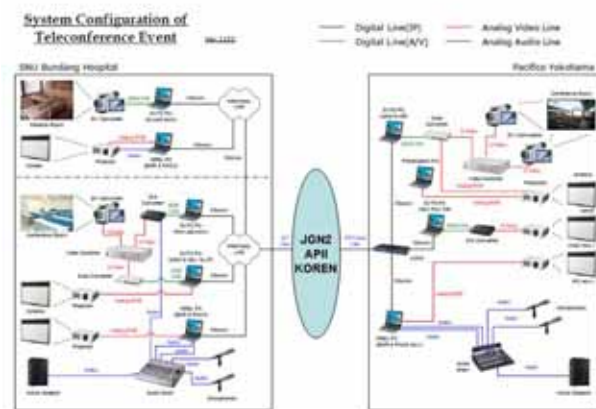


Figure 1 Network configuration of the event.

### 2.4. Security system

We usually use a security system by VPN for high speed transmission, when we transmit patient's live image in the case of project to protect patient's privacy (4). However, we asked a student in SNU to be a faked patient for demonstration in this event, thus, we did not use security system.

### 2.5. EMR and PACS system

EMR system in SNU Bundang hospital was developed by ezCaretech Co. (<http://www.ezcaretech.com/>), and PACS was purchased from Agfa Co. (<http://www.agfa.com/en/co/index.jsp>). High resolution displays attached to PACS was from WIDE Co. (<http://www.wideusa.com/>).

### 2.6. Content of remote lecture

We used power point (Microsoft) on Ariel Spotlight for presentation in the remote lecture.

### 2.7. Questionnaire survey for attendants in the venue

We conducted a questionnaire survey to evaluate the event by a paper after the event. We could recover 26 from the attendants.

## 3. Results

### 3.1. Record of the network project

We have conducted 58 times of telemedical event in Asia-Pacific area since February 2003 until March 2006.

We conducted a cross-strait teleconference session in the APAMI2003 in Daegu, Korea (2). That was 5<sup>th</sup> event in the history of our project. This time was 50<sup>th</sup> event in it (figure 2).

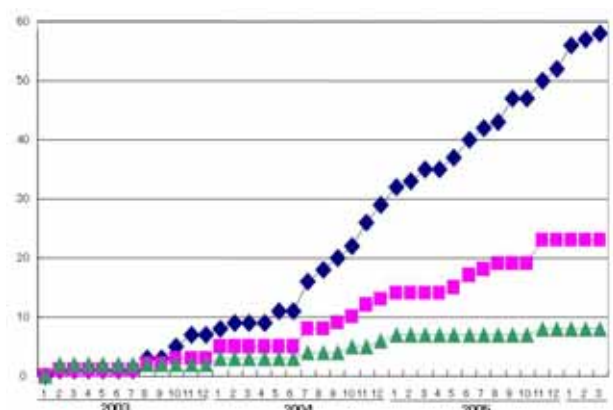


Figure 2 The events record in our project for three years (2/2003-3/2006). Diamonds show event number, squares show number of connected institutes, and triangles show number of connected countries.

### 3.2. The virtual site visit at the SNU Bundang hospital from the Yokohama venue by the network

We conducted the session for 2 hours from 4pm on 25<sup>th</sup> November, 2005, which was sponsored by the standardization committee in the Japanese association of Medical Informatics (JAMI) after several rehearsals. Every greeting and presentation was given with Japanese translation by subtitle on the screen or a translator.



Figure 3a View of the Yokohama venue of the event in 25<sup>th</sup> Joint Conference of Medical Informatics.





Over than 260 attendants had an experience as users for EMR/PACS in SNU Bundang hospital by the network.

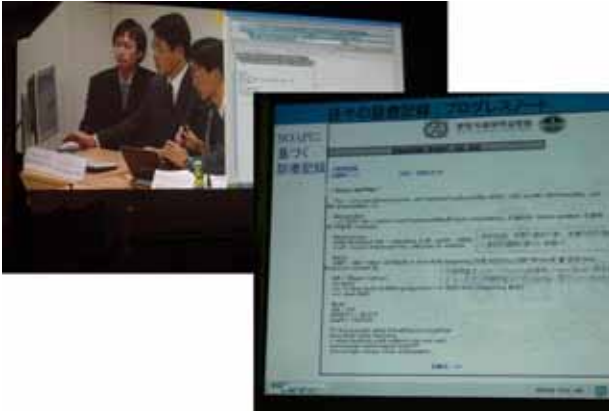


Figure 3b Explanation with operation on the progress note in EMR by medical doctors in SNU Bundang hospital. Usually, they use dual monitor in EMR system in Bundang hospital.

We had dual screens in the front of the venue. The left screen showed the DVTS image for Korean side medical staffs. The right screen presented the Ariel Spotlight image for EMR/PACS system. We had the third small screen for explanation of the system or translation on the front-left side.

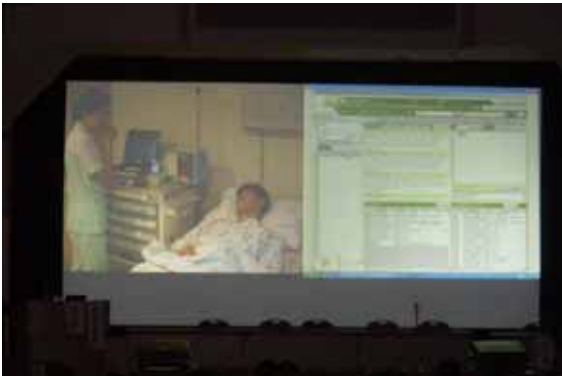


Figure 3c Live Demonstration of nursing care system in EMR in an inpatient room on a faked patient (a SNU student).

After the explanation of the operation of EMR in the conference room in Bundang hospital, they switched the image from an inpatient room to present a nursing care system on a faked patient.



Figure 3d Presentation of PACS on the network. They use 5M monochrome high resolution display in PACS in Bundang hospital.

PACS demonstration was done on CT and MRI still images. They operated 3D reconstruction and analysis with the images in Bundang hospital.

### 3.3. Remote lecture

Prof. Sang-Yo Nam (Institute of Health & Welfare, Yuhan College) gave a remote lecture for the Yokohama venue about Korean medical insurance system including reimbursement for PACS installation in Japanese language. He answered a remote question from the venue.

### 3.4. Questionnaire survey

Although the recovery ratio of the questionnaire paper was low (10%), the results of the survey was meaningful. As shown in figure 4a, we had 260 attendants from a variety of the field in the Yokohama venue. Unfortunately, there was another big session for nurses in other room in the same time, nurse attendants were few in the event.

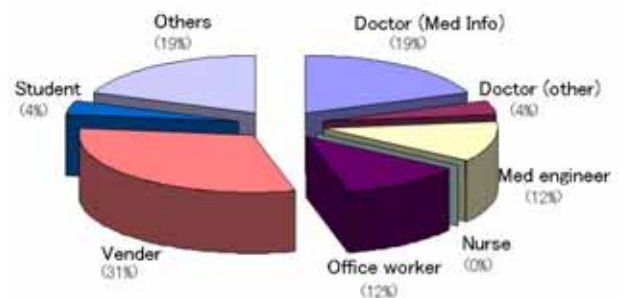


Figure 4a Occupation of Participants (n=26).

Major part of the attendants felt interest in the Internet teleconference after the session as shown in figure 4b.



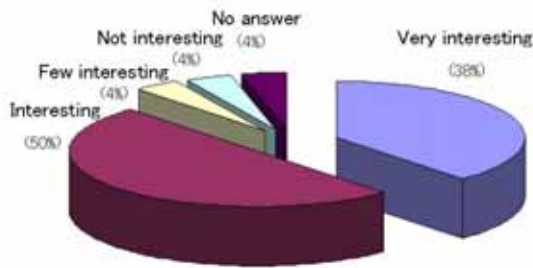


Figure 4b Interest for Internet Conference (n=26)

We could get a high evaluation from the attendants as shown in figure 4c.

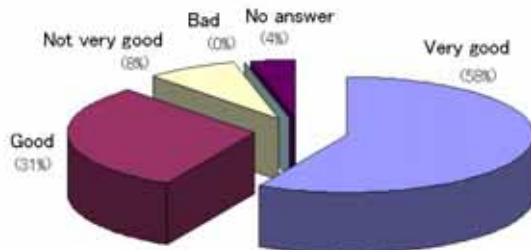


Figure 4c Evaluation of the event (n=26)

The Japanese attendants in the venue might feel strong interest in the medical system in Korea including medical informatics and medical insurance system as shown in figure 4d.

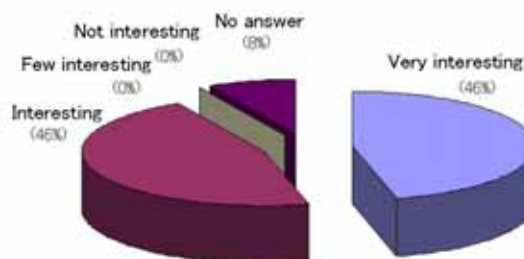


Figure 4d Interest for Korean Med Info (n=26)

## 4. Discussion

The knowledge of medical science is uniform and can easily spread via one-way transmission by scientific journals. In contrast, medical information about new techniques, equipments, medicines, policy, ethics, so on are often unique in local regions and vary in areas and areas. The present system will help to remove a barrier of medical communications and help to standardize them using high quality moving images. We can communicate mutually and exchange medical information over the national border with minimum stresses.

Our main contents in the project have been surgical subjects including endoscopic surgery and microsurgery so far. In the event this time, we showed we could extend the project contents to medical informatics and social medical system fields.

In this event, we used Japanese language in all presentation, because the major parts of the attendants were not familiar with English. The language problems should be cared in international conference. Additionally, we always think about voice quality in the teleconference, especially for strong echo and howling to proper communication. DVTS has no compression process. Thus the time delay is enough small (about 0.3 sec between Yokohama and Seoul) to make a smooth conversation, although we need to have at least 30 Mbps bandwidth by full frame DVTS transmission.

Along with a recent spread of the service of international optic fiber cable network in Asia-Pacific area, we have connected 8 countries (Japan-Korea; 2/2003, -Unite States; 1/2004, -Australia; 7/2004, China; 10/2004, -Taiwan; 12/2004, -Thailand; 1/2005, and -Singapore; 11/2005) by broad-banded network for 3 years in the project. We believe the network promotes the standardization of medical information in this area by both online and offline site visit at abroad institution.

We have established high-quality video transmission system over Internet protocol among Asia-Pacific countries, which is easy to perform, reliable and economical. In this time, we could use medical informatics and social medical system as contents in the network infrastructure. This will be a promising and very helpful tool of network for the regional and worldwide standardization of medical information system, medical techniques and social medical system.

## 5. Footnotes

We express our sincere appreciation to Prof. Heung-Sik Kang, the president of SNU Bundang hospital and Prof. Chul-Hee Lee, the CEO of ezCaretech Co., Prof. Sang-Yo Nam and Prof. Takayuki Nishiyama, Yuhan College, and the committee members of standardization in JAMI for great cooperation for the event. We also deeply thanks to Network NOC members of JGN2, APII, and KOREN. We also thank to Mr. Jun-Hun Lee and other staffs in Bundang hospital. We also appreciate the engineering staffs from QIC for continuous support for the project. The ezCaretech Co., WIDE Co., NTT communication Co., and National Institute of





Information and Communications Technology (NICT) cosponsored to the event and helped us. This project has been partly supported by “Development and Operation of the Next Generation Internet Technologies”, one of the Core University programs conducted by the Japan Society for the Promotion of Science.

## References

- [1] T. Takashi, “The present and future of telemedicine in Japan”, *Med Info* 2001; 61, 131-137.
- [2] S. Shimizu, J.S. Hahm, Y.W. Kim, Y.J. Park, J.H. Lee, S.K. Youm, K. Okamura, N. Nakashima, H.S. Choi, C.H. Kang, M. Kawamoto, M. Tanaka, “Telemedicine with digital video transport system over the Korea-Japan cable network” *Journal of Korean Medical Informatics* 2003; 9, S322-S326.
- [3] N. Nakashima, K. Okamura, J.S. Hahm, Y.W. Kim, H. Mizushima, H. Tatsumi, B.I. Moon, H.S. Han, Y.J. Park, J.H. Lee, S.K. Youm, C.H. Kang, S. Shimizu, “Telemedicine with digital video transport system in Asia-Pacific area” *Proceeding of Advanced Information Networking and Applications (AINA) 2005* 2005; 253-257.
- [4] N. Nakashima, S. Shimizu, K. Okamura, J.S. Hahm, Y.W. Kim, B.I. Moon, H.S. Han, J.H. Lee, “Rapid growth of broad banded telemedicine network in Asia-Pacific area” *Proceeding of the 7<sup>th</sup> China-Japan-Korea Medical Informatics Conference* 2005; 35-37.







# A Collaborative Workspace for Telemedicine System

Augustine Tsai, Hung-Wen Tung  
 Institute for Information Industry, Taiwan  
 Email: {atsai, hwtung}@iii.org.tw

## Abstract

*This paper is based on the previous research on collaborative session and presents the design of a collaborative workspace in a telemedicine system. In this collaborative workspace, medical images, reports, device data are being transferred using the standard for medical image exchange. We use an open instant messaging protocol to conduct the remote consulting activity. With this collaborative workspace, multiple physicians located at different sites can examine one patient at the same time. With image annotation and structured report being transferred, physicians can exchange their opinions and findings with each other. Specialized or experienced physician can provide clinical guidelines to the bed-site physician. Device data provides real time information about the patient. With this system, diagnostic health care service can be provided remotely to areas lacking sufficient medical resource.*

## 1. Introduction

With the growth of Internet and the development of hardware and software technology, applying IT to medical care has been one of the important issues in current days. Many medical devices have been integrated with computer systems to analysis the data. Standards for medical information are being established to facilitate the exchange of information. When facing the problem of unbalanced medical resources between rural or less-developed areas or cities and developed ones, how to provide essential health care service to those rural and less-developed areas drives the idea of using a telemedicine system to deliver the services from remote site.

To this end, we utilize the idea of collaborative workspace [1] when designing a remote consulting system. In this system, multiple health care providers located in different sites could work together to give diagnostics to the patients in rural or less-developed areas. The underlying communication between users of this system is based on an instant messaging protocol. The integration of medical resources, including medical records, images, mechanic devices, and guidelines, makes the multiple health care providers can examine the same patient information concurrently. The rest of the paper is divided as follows: Section 2 introduces the instant messaging adopted by our system. Section 3 presents the standards for medical images and reports that our system followed. Section 4 presents the design

of the collaborative workspace. In Section 5, we develop the conclusions of our work.

## 2. Instant Messaging

The use of instant messaging (IM) has become popular in these years. There are several popular consumer IM services, like AIM, ICQ, MSN and Yahoo, exist today. The Internet Engineering Task Force (IETF) has formalized the core XML streaming protocol as an approved instant messaging and presence technology under the name of Extensible Messaging and Presence Protocol (XMPP), and the XMPP specification have been published as RFC 3920 and RFC 3921 [2] [3].

In the security aspect, unlike some legacy protocols, like POP3 and SMTP, which do not provide a way to choose nodal relationships, i.e. permission to send or receive to or from a node. The presence notion in the XMPP protocol does not allow nodes who have not established a mutual agreement to send or receive messages and/or streams of data. Figure 1 shows a simple message in XMPP format. There are previous works utilizing IM in telemedicine system [7], mostly in an information transfer aspect. In this paper, we utilize IM to put some add-on features in our collaborative workspace. For organizations wishing to ensure that all data are encrypted, they can configure the collaborative session to ensure encryption of message transfers.

```
<message to="JamesBond@server1/JETI" id="JETI_0" type="chat">
  <thread>22b0f010b4a9782fc</thread>
  <body>hello</body>
  <x xmlns="jabber:x:event">
    <composing/>
  </x>
  <html xmlns="http://www.w3.org/1999/xhtml">
    <body>
      <span style="font-family: Arial; background-color:
    </body>
  </html>
</message>
```

Figure 1 An example of XMPP message.

## 3. Standards for Medical Images and Reports

The Digital Imaging and Communications in Medicine (DICOM) was created by the national Electrical Manufactures Association (NMEA) [4], and is now commonly used for the transfer and storage of medical images. For transmission and storage of clinical documents, the structured report (SR) standard provides capability to link text and other data to particular images or waveforms and to store the coordinates of finding. A







SR document not only describes the specific features contained in images or waveforms but can also refer to any number of images or waveforms. Therefore, SR bridges the traditional gap between imaging system and information system [8].

The medical images and documents used in our system are compatible with the DICOM image and SR document standards, therefore the system can fit into any hospital that conforms to these standards. We also use the DICOM standard to facilitate the storage and retrieval of medical image and structured report at remote sites. In the security aspect, the data transferred can just be encrypted using standard mechanisms and utilities.

#### 4. Collaborative Workspace

In our previous works [1], we have discussed the concept of session management in a collaborative workspace. In this paper, we present the design of our system to realize the concept of session management.

Figure 2 illustrates the collaborative workspace of our telemedicine system. On the left side demonstrates the bed-site physician who is requiring consultant when he/she examines the x-ray image of the patient. So he/she login to the system and invite a remote physician who is specialized in that particular area to join the collaborative workspace the system provided. The content of the remote consultant invitation includes the patient identification, the medical image identification, etc. When remote physician accepts the invitation, he/she can retrieve the related images and records from the DICOM server using the identifications in the invitation and then enter the collaboration workspace. Therefore in collaborative workspace, both physicians see the same user interface that includes the medical record, medical images, and data from the mechanic devices connected to the patient. Both physicians then can make annotations to the medical images, and update the medical record according to the information they acquired. In this way, the patient could receive diagnosis from a specialist that is located on the remote site. The following discusses the detail of this design.

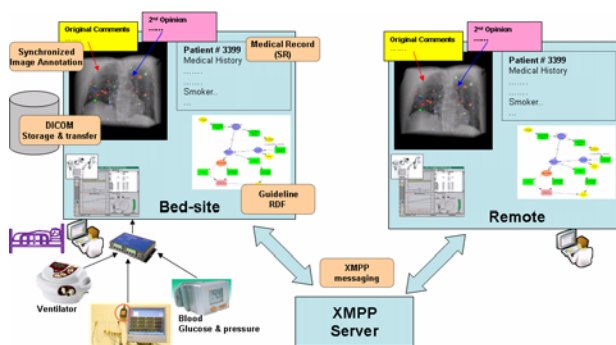


Figure 2 Collaborative workspace.

#### 4.1. Session Protocols in Collaborative Workspace

Figure 3 illustrates the instant message processing in the collaborative workspace. All communication underlying the workspace is based on the XMPP protocol. The XMPP protocol used includes the instant messaging for remote consultant invitation, image annotation, findings, clinical guidelines, and device data. We design a message processor plug-in to the instant messeng client. Each message transferred to the collaborative workspace is parsed by the message processor and categorized according to the message tags.

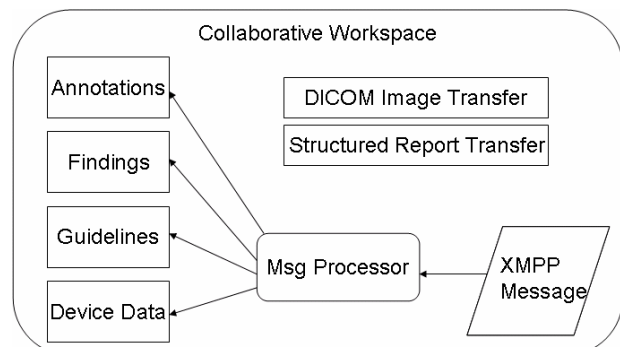


Figure 3 Instant message processing in collaborative workspace.

We use the synchronized annotation function to describe how we design the session protocols in collaborative workspace. In initial phase, bed-site physician retrieve the DICOM image from the server in the hospital using the retrieval command defined in the DICOM standard. And then he/she invites the remote physician to join his/her collaborative workspace. When the remote physician accepts the invitation and joins the collaborative workspace, the patient's medical image and document are retrieved from the DICOM server by the patient identification provided in the invitation message. Figure 4 demonstrates the remote consulting user interface when the physician joins the collaborative workspace.

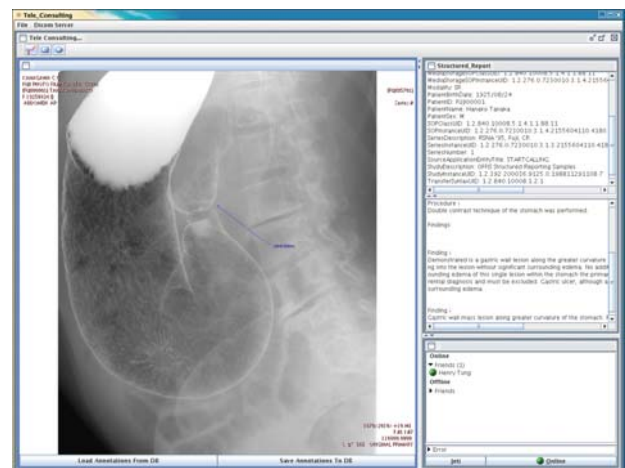


Figure 4 Remote Consulting User Interface



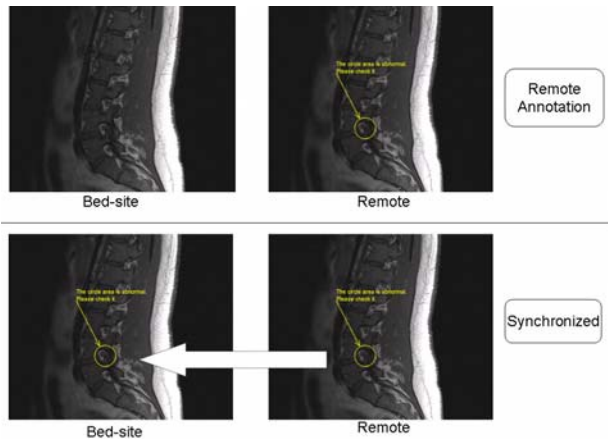


Figure 5 Synchronizing Image Annotation.

As figure 5 shows, when the remote physician finds that something's wrong within the DICOM image, he/she can use the tool provided by the workspace to annotate on the image. When remote physician submits the annotation, the workspace would present it to the screen of bed-site physician.

To transfer such an annotation to another physician, the data representation of the annotation is formatted using XML and is inserted to the body of XMPP message. As shown in figure 6, when the system parses the annotation tag in the body of the XMPP message, it will automatically process this message as an annotation instead of normal chat message. And the receiving site will trigger the drawing utility to re-produce the annotation on the screen to synchronize with remote-site, which is the result in the lower part of figure 5. The receiving site can also make a counter-comment of the image as another annotation and synchronize it to the remote physician.

```
<message to="JamesBond@server1/JETI" id="JETI_0" type="chat">
  <thread>22b0f010b4a9782fc</thread>
  <body>
    <annotation>
      <point1><x>161.42</x><y>143.64</y></point1>
      <point2><x>81.40</x><y>94.39</y></point2>
      <text>This area seems to be the problem.</text>
    </annotation>
  </body>
  <x xmlns="jabber:x:event">
    <composing/>
  </x>
  <html xmlns="http://www.w3.org/1999/xhtml">
    <body>
      <span style="font-family: Arial; background-color:
    </body>
  </html>
</message>
```

Figure 6 XMPP message for synchronizing annotation.

Besides annotation on images, when the bed-site physician consults the remote physician, the workspace also provides a function for remote physician to document his/her finding. When remote physician submits the findings it will be transmitted in the XMPP message format, as shown in figure 7. When the system parses the finding tag in the message, it will process the

finding message into SR document and trigger the report browser in the bed-site screen to show the finding (Figure 8).

```
<message to="JamesBond@server1/JETI" id="JETI_0" type="chat">
  <thread>22b0f010b4a9782fc</thread>
  <body>
    <finding>
      The examination demonstrates a .....
    </finding>
  </body>
  <x xmlns="jabber:x:event">
    <composing/>
  </x>
  <html xmlns="http://www.w3.org/1999/xhtml">
    <body>
      <span style="font-family: Arial; background-col
    </body>
  </html>
</message>
```

Figure 7 XMPP message for transmitting finding.

| Basic Text SR Document               |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
|--------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Patient:</b>                      | Charles Wilkins (male, #GE0514)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| <b>Completion Flag:</b>              | PARTIAL                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| <b>Verification Flag:</b>            | UNVERIFIED                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| <b>Content Date/Time:</b>            | 2002-05-10, 16:34:56                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| <b>CT Report</b>                     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| <b>Request:</b>                      | CT: Abdomen                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| <b>History:</b>                      | Patient is a 45 year old male with abnormal liver function test. Ultrasound evaluation demonstrated a 3 cm lesion in the medial aspect of the right lobe of the liver. The question was raised regarding potential hemangioma.                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| <b>Procedure:</b>                    | Serial imaging was obtained in the upper abdomen with the administration of oral and intravenous contrast.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| <b>Finding:</b>                      | The examination demonstrates a well circumscribed 3 cm lesion present within the medial aspect of the inferior right liver lobe. Initial evaluation demonstrates lack of contrast enhancement. Subsequent imaging (not shown) demonstrated typical contrast enhancement pattern of a benign hemangioma of the liver. The remaining contrast enhancement pattern in the liver is normal. There is normal appearance of the adrenal glands, spleen, kidneys, and pancreas. There is no evidence of liver metastasis. 3 cm nodule present in the inferior medial aspect of right liver lobe. Contrast enhancement pattern consistent with the diagnosis of hemangioma. |
| <b>Best illustration of finding:</b> | <a href="#">CT image</a>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |

Figure 8 Structured Report Browser.

When the collaborative session is finished, either by remote or bed-site physician exiting the workspace or by network connection exceptions, the annotations and the findings are saved as a SR document and store back to the hospital server. The session can be re-built again when the bed-site physician starts the remote consulting procedure and retrieves the image and document back to the workspace.

## 4.2. Medical Guideline Recommendation

For patients that require complex clinical guidelines to be followed when providing the treatment, the remote





physician can use the workspace to recommend a clinical guideline to the bed-site physician. Guidelines are encoded in the RDF (Resource Description Framework) format [6], and using the Protégé tool [5], guidelines could be viewed in a graphical flowchart way (Figure 8). The RDF is encoded in the body of the XMPP message with the guideline tag. Similar to the annotation and finding tag, the system trigger the guideline viewer to show the guideline to the bed-site physician. When collaborative session is finished, the guidelines could also be saved in a SR document back to the hospital server.



Figure 9 Clinical guideline

#### 4.3. Medical Device Information

Several devices connected to the patient provide useful information when making diagnosis. Most medical devices we integrated are communicated to the bed-site computer through the RS-232 interface. In the collaborative workspace, we integrate the respirator to display the respiratory waveform, the ECG, and also the blood pressure and glucose from the connected instruments. These data are also transmitted with appropriate tags encoded in the XMPP message, making remote physician able to examine real time data from these devices.

#### 5. Conclusion

The collaborative workspace in telemedicine system provides a media to facilitate the consulting procedure between physicians. With the workspace in the system, bed-site physician in rural or less-developed areas can consult more specialized or experienced physicians located at the remote site. Remote physician can

examine the patient's medical images, medical documents, real time data from the device connected to the patient, and provide comments and findings to the bed-site physician. With the integration of clinical guideline system, remote physician can also recommend a clinical guideline to the bed-site physician.

Further development will be focused on building a knowledge base from annotations and documents generated from previous collaboration sessions. With the knowledge base, medical students or junior physicians can study previous cases and gain experience. Since this is a design form our previous works on collaborative session to a real system. In our future work, we would like to make this system works by providing it to the hospitals and evaluated by real users. We hope by the feedback from real users can help us making it a solid and sound system.

#### 6. Acknowledgement

This research was supported by the Service Web Platform Technology Research and Development Project of Institute for Information Industry and sponsored by MOEA, ROC.

#### 7. References

- [1] A. Tsai, H.W. Tung, and S.C. Chung, "Apply Collaborative Sessions with Coordination Rules to Respiratory Care and Monitoring System", *Submitted to MIST 2006*.
- [2] <http://www.xmpp.org/>
- [3] <http://www.jabber.org/>
- [4] DICOM, <http://medical.nema.org/>
- [5] Protégé, <http://protege.stanford.edu/>
- [6] <http://www.w3.org/RDF/>
- [7] CE Chronaki, and et al., "An eHealth platform for Instant Interaction among Health Professionals," *IEEE Computers in Cardiology 2003*, pp.101-104, 21-24 Sept. 2003
- [8] The DICOM Standard, Overview and Characteristics Whitepaper, [http://www.ringholm.de/docs/02010\\_en.htm](http://www.ringholm.de/docs/02010_en.htm)





## Using Hospital Information System as A Basis for Clinical Decision Support in Schizophrenic Patients

Chiung-Hsu, Chen

Taoyuan Psychiatric Center

[forensic@typc.doh.gov.tw](mailto:forensic@typc.doh.gov.tw)

Yu-Chuan, Li

Taipei Medical University

[jack@tmu.edu.tw](mailto:jack@tmu.edu.tw)

I-Ren, Chiang

Taipei Medical University

[ijchiang@tmu.edu.tw](mailto:ijchiang@tmu.edu.tw)

### Abstract

*It was still controversial about the effectiveness of treating schizophrenia patients by second generation antipsychotics(SGAs) or first generation antipsychotics(FGAs). We used database extracted from our Hospital Information System in Taoyuan Mental Hospital as a base for clinical decision support system. Four years' pharmacy prescription data were analyzed. Treatment discontinuation was taken as primary outcome for treating patients with different antipsychotics. The duration of treatment, kinds of antipsychotics, patient gender, age group of patients were collected. Survival analysis found the duration of treatment differed from drug to drug statistically, and the difference persisted even considering the influence of gender and age group by Cox regression. The tool could be applied to the other five major mental hospitals, which had the same HIS, in Taiwan. It was possible to use pharmacy data from pharmacy-prescription data to construct a base for clinical decision support system in schizophrenic outpatients.*

*Keywords: Clinical decision support system, schizophrenia, treatment discontinuation, HIS*

### 1. Introduction

The treatment for schizophrenia had made remarking success after discovering chlorpromazine in 1950'. There were many antipsychotics with dopamine blocking activity in later years. Although they were effective in treating schizophrenic patients, they had many disadvantages. The most often happened side effect was extrapyramidal symptoms. It occurred among about 70-80% patients and hence it made patients' drug compliance poorer. The situation changed after finding Clozapine which caused fewer extrapyramidal symptom side effects. Unfortunately, it might cause white blood cell count decreased and therefore caused more severe complications. Clozapine disappeared later and it was used again owing to its remarkable effect in treating refractory schizophrenic patients in 1990'. There were so-called "Second Generation Antipsychotics" from 1990', including risperidone, olanzapine, quetiapine, ziprasidone...etc. The SGAs had treatment effects and fewer extrapyramidal symptom side effects when compared with so-called "First Generation Antipsychotics". Kane. J reported that SGAs were more effective and safer than FGAs.[1] Others found that it is



not true all the same.[2-6] Besides, it was still controversial whether SGAs could improve patients' cognitive function or not.[7] There were limited reports about longterm effects of SGAs, such as preventing patients' relapse, improving patients' social function, improving patients' quality of life, and lowering caregivers' burdens.[6, 8-9] There were many head-to-head studies comparing effectiveness of antipsychotics between one SGA and one FGA or between two SGAs. However, every study of this kind wasted many resources and a lot of manpower. There were many prospective studies comparing antipsychotics' effectiveness with excellent data.[10-12] Since a decision to continue or stop medications reflected the combined evaluation of efficacy and safety/tolerability of the treatment by the patient and clinician, they used "treatment discontinuation" as primary outcome[10]. This finding made it possible to do large comparison studies by information management and the results would be very helpful in clinical decision support system. This study tried to use outpatients' pharmacy prescription database from hospital information system as a base for clinical decision support system about antipsychotics' persisting use.

## 2. Materials and Methods

There were eight major mental hospitals which cared more than 50% of schizophrenic patients in Taiwan. The study was conducted at Taoyuan Mental Hospital which had 312 acute beds, 420 chronic beds and 280 beds for day-care. It was the largest mental hospital in east-south Asia and it served about one million and eight hundred thousands people. Taoyuan Mental Hospital had the same HIS with the other 43 large official hospitals in Taiwan. Six out of eight major mental hospitals used the same HIS. The developed tools could be applied to other hospitals easily.

Effectiveness for treating schizophrenia patients was hard to define. It included the impact of medication on controlling the various symptoms of illness, as well as the safety and tolerability of the illness.[10] The patients or the doctors would change antipsychotic if the drug is not "effective". It needed a prospective study design and cost a lot of manpower and resources to collect all factors about effectiveness among different antipsychotics. Many of these causes were not available in Hospital Information System. Since "treatment discontinuation" was taken as primary outcome by large scale studies in the USA and the Europe[10-12], we took it as primary outcome in this study. This indicator was easy to define in HIS.

The data base of HIS didn't follow the formal schema of relational database. Many attributes were put together in one column. Besides, it was complied by COBOL language which was different from current popular languages. The supplier of HIS didn't offer the schema of HIS database but only some definitions of columns. To resolve the difficulties list above, we dumped the HIS databases to Microsoft SQL server 2000 at first. Every column in database of outpatient service was analyzed step by step and be separated into different attributes. The HIS was applied to Taoyuan Mental Hospital since July 2001. We focused on pharmacy prescription databases of outpatients service from Jan 1, 2002 to Dec 31, 2005. All schizophrenic patients with the first ICD code of "295". The focus of medications were on major SGAs and some atypical antipsychotics, including Clozapine, risperidone, olanzapine, quetiapine, amisulpride, ziprasidone, and aripiprazole. Treatment discontinuation was the primary outcome.

Every schizophrenic patients prescribed with one of the interested antipsychotics would be counted for its







length of drug usage from the first prescribing date to the date of discontinuation. Demographic data about patients such as gender, age,....etc were also gathered from HIS.

We used SPSS 10.0 Chinese version in statistics. Kaplan – Meier survival curves and life table method were used to estimate the time to the discontinuation of treatment. Treatment groups were compared by using Cox proportional-hazards regression models. Demographic factors such as gender, age...etc were considered in Cox regression models.

### 3. Results

There were 2093 events during 2002 to 2005. The risperidone was the most popular use.(Table 1) We didn't have all interested antipsychotics until 2005. There were 1513 patients in 2093 events and 1100(72.70%) patients were prescribed with only one interested antipsychotic. There were 3(0.20%) patients prescribed with 6 kinds of interested antipsychotics. The gender difference between antipsychotics was noted. Clozapine was more prescribed among males than females. Besides, the females prescribed with ziprasidone and aripiprazole were about twice of males.

**Table 1 First prescription of antipsychotics (2002-2005)**

| Yr | AMI  | ARI  | CLO | LOD | OLA | QUE | RIS | ZIP  |
|----|------|------|-----|-----|-----|-----|-----|------|
| 02 | n.a. | n.a. | 98  | 69  | 68  | 45  | 134 | n.a. |
| 03 | 29   | n.a. | 46  | 40  | 84  | 71  | 142 | 26   |
| 04 | 36   | n.a. | 40  | 85  | 93  | 68  | 182 | 67   |
| 05 | 41   | 141  | 59  | 90  | 66  | 50  | 180 | 43   |
|    | 106  | 141  | 243 | 284 | 311 | 234 | 638 | 136  |

Yr= Year; AMI= amisulpride; ARI= aripiprazole; CLO= clozapine; LOD= lodopine; OLA= olanzapine; QUE= quetiapine; RIS= risperidone; ZIP= ziprasidone; n.a.=not applicable

**Table 2 Discontinuation rate among antipsychotics in 2005**

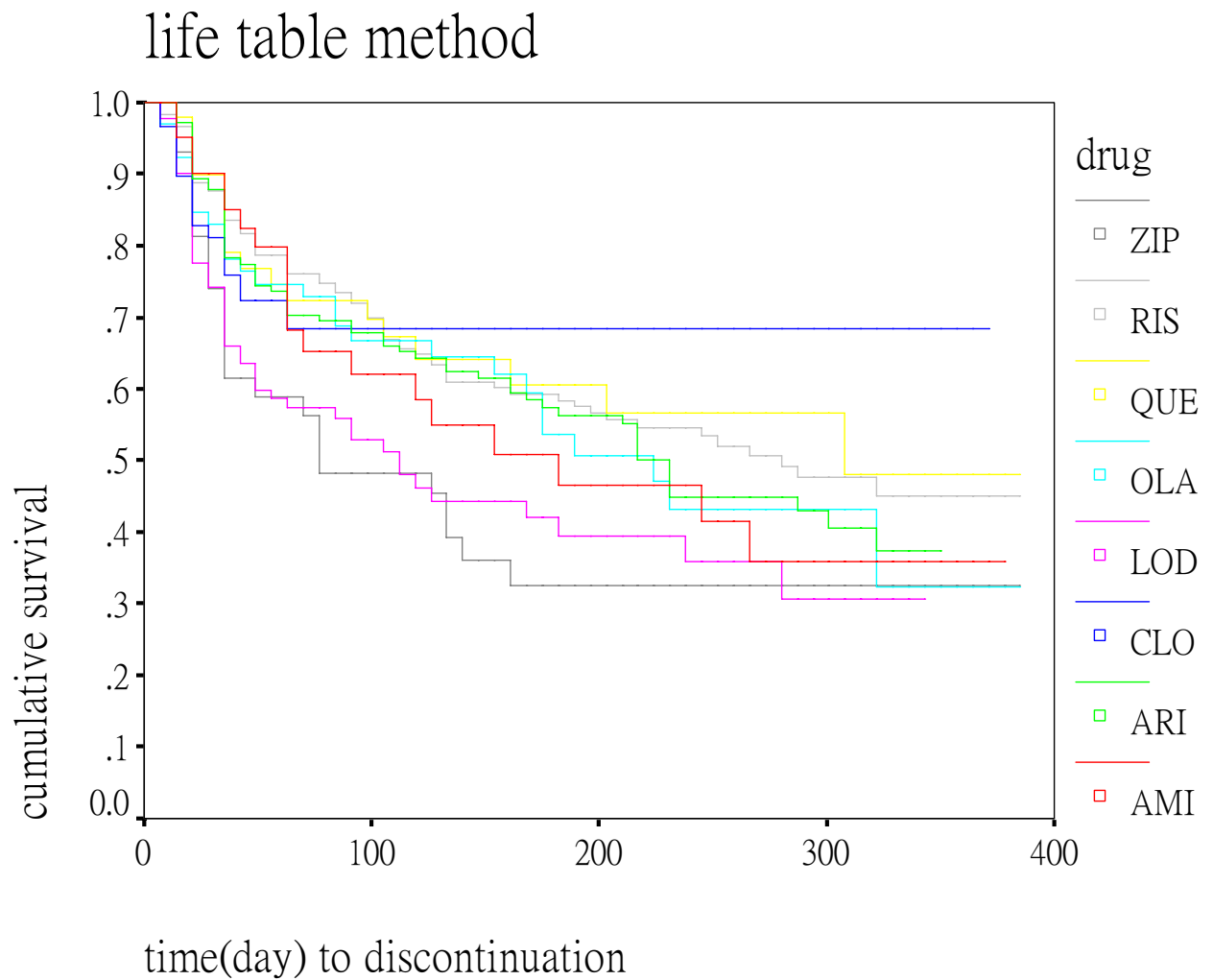
| Continuation(Con.) v.s. discontinuation(Discon.) |      | Con.       | Discon.    | total |
|--------------------------------------------------|------|------------|------------|-------|
| AMI                                              | N(%) | 21(51.2%)  | 20(48.8%)  | 41    |
| ARI                                              | N(%) | 74(52.5%)  | 67(47.5%)  | 141   |
| CLO                                              | N(%) | 41(69.5%)  | 18(30.5%)  | 59    |
| LOD                                              | N(%) | 41(45.6%)  | 49(54.4%)  | 90    |
| OLA                                              | N(%) | 37(56.1%)  | 29(43.9%)  | 66    |
| QUE                                              | N(%) | 31(62.0%)  | 19(38.0%)  | 50    |
| RIS                                              | N(%) | 105(58.3%) | 75(41.7%)  | 180   |
| ZIP                                              | N(%) | 17(39.5%)  | 26(60.5%)  | 43    |
| <i>Total</i>                                     | N(%) | 367(54.8%) | 303(45.2%) | 670   |

$X^2=14.804$  df=7 p=0.039

The later results were focused on 2005 when all interested antipsychotics were included for comparison. There were 670 events by 578 patients in 2005. 501 (86.68%) patients are prescribed with only one interested drug. The age distribution was compatible with schizophrenia patients'. The average discontinuation rate among antipsychotics was about 45.2%. Those with lodopine and ziprasidone had higher discontinuation rate.(Table 2)

The comparison of different antipsychotics' survival curves by life-table method revealed statistic different between drugs. (Figure 1) The Kaplan – Meier survival curves also revealed the same result. Statistical significant among different antipsychotics persisted after considering demographic data such as gender and age group by Cox regression.





Overall comparison statistic= 17.961, df.=7, p=0.0121

**Figure 1 Life-table survival curve of antipsychotics in 2005**

#### 4. Discussion

There were many factors affecting doctors' prescription pattern, such as lack of efficacy, tolerability issue owing to side effect, the official policy by National Health Insurance, and coping strategy developed by hospitals. Patient's condition such as poor disease insight compromised drug compliance and patients dropped out later. Even under the influences of so many factors, it was said

that treatment continuation was a good indicator for treatment effect.[11, 12] It provided a convenient way to handle the complicated effects of antipsychotics' advantages and disadvantages. We tried to use pharmacy prescription database from HIS to compare the outcome of different antipsychotics' effect on patients.

Selection bias was inevitable in such study design and could not be managed by statistics. Those patients prescribed with clozapine were quite different from those



with aripiprazole in clinical setting. There were more males in refractory schizophrenic patients and it might explain the gender difference between clozapine and other antipsychotics. The ratio of female to male was about two in patients prescribed with ziprasidone and aripiprazole. It might be due to the least likely causibility of hyperprolactinemia by these two antipsychotics. However, when stratifying the age group and gender, we could have separated the different reasons for prescription antipsychotics. It could be further developed as a clinical decision support system. For examples: a 27 year-old female diagnosed as schizophrenia for years came to outpatient service for her relapsing psychotic symptoms, which antipsychotic was preferred? With this clinical decision support system, the doctor chose “gender= female” and “age group= 20-30” on the screen and the computer would show him that those with the same characteristic taken different antipsychotics have different survival curves. He could make decision by himself referring this system finally.

Further efforts were needed. We calculated the treatment duration by start-date and end-date. However, there were patients who take some antipsychotic at first, change to another antipsychotic later, and take the original antipsychotic finally. In this kind of situation, we overly estimated the treatment duration and hence this system gave us an optimistic suggestion. It might compromise patients’ best interest. We would try to review every visit of patients and separate those discontinued visits in later studies.

Some might recommend to do the same thing from NHI database. The availability of different antipsychotics of interested differed from hospitals to hospitals. There might be only one choice in one hospital and six choices in the other hospitals. However, it was possible

if we limited our topic on some major mental hospitals. The other five major mental hospitals had the same HIS and this result could be applied to other hospital easily. We could also try to set up a clinical decision support system by using databases of six major mental hospital with the same HIS to get more objective results. .

## 5. Conclusion

We could see the brightening future to use pharmacy database to construct a clinical decision support system in schizophrenia outpatients, disregarding the underdeveloped electric medical records. It would be more promising if we further manage the data and collect databases from the other five major mental hospitals with the same HIS.

## 6. Reference

- [1] Kane J, Honigfeld G, Singer J, Meltzer H. Clozapine for the treatment-resistant schizophrenic: a double-blind comparison with chlorpromazine. *Arch Gen Psychiatry* 1988;45: 789-96.
  - [2] Leucht S, Pitschel-Walz G, Abraham D, Kissling W. Efficacy and extrapyramidal side-effects of the new antipsychotics olanzapine, quetiapine, risperidone, and sertindole compared to conventional antipsychotics and placebo: a meta-analysis of randomized controlled trials. *Schizophr Res* 1999;35:51-68.
  - [3] Geddes J, Freemantle N, Harrison P, Bebbington P. Atypical antipsychotics in the treatment of schizophrenia: systematic overview and meta-regression analysis. *BMJ* 2000;321: 1371-6.
- Wahlbeck K, Tuunainen A, Ahokas A, Leucht S. Dropout rates in randomised antipsychotic drug trials. *Psychopharmacology (Berl)* 2001;155: 230-3.



- [4] Davis JM, Chen N, Glick ID. A metaanalysis of the efficacy of second-generation antipsychotics. *Arch Gen Psychiatry* 2003; 60: 553-64.
- [5] Leucht S, Wahlbeck K, Hamann J, Kissling W. New generation antipsychotics versus low-potency conventional antipsychotics: a systematic review and meta-analysis. *Lancet* 2003; 361: 1581-9.
- [6] Leucht S, Barnes TRE, Kissling W, Engel RR, Correll C, Kane JM. Relapse prevention in schizophrenia with new-generation antipsychotics: a systematic review and exploratory meta-analysis of randomized, controlled trials. *Am J Psychiatry* 2003; 160: 1209-22.
- [7] Keefe RS, Silva SG, Perkins DO, Lieberman JA. The effects of atypical antipsychotic drugs on neurocognitive impairment in schizophrenia: a review and meta-analysis. *Schizophr Bull* 1999; 25:201-22.
- [8] Rosenheck R, Perlick D, Bingham S, et al. Effectiveness and cost of olanzapine and haloperidol in the treatment of schizophrenia: a randomized controlled trial. *JAMA* 2003; 290: 2693-702.
- [9] Csernansky JG, Mahmoud R, Brenner R. A comparison of risperidone and haloperidol for the prevention of relapse in patients with schizophrenia. *N Engl J Med* 2002; 346: 16-22.
- [10] Swartz MS, Perkins DO, Stroup TS, McEvoy JP, Nieri JM, Haak DC. Assessing clinical and functional outcomes in the Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE) schizophrenia trial. *Schizophr Bull* 2003; 29: 33-43.
- [11] Lieberman JA, Stroup TS, McEvoy JP, Swartz MS, Rosenheck RA, Perkins DO, Keefe RS, Davis SM, Davis CE, Lebowitz BD, Severe J, Hsiao JK. Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE) Investigators. Effectiveness of antipsychotic drugs in patients with chronic schizophrenia. *N Engl J Med* 2005; 353: 1209-23.
- [12] Fleischhacker WW, Keet IP, Kahn RS. EUFEST Steering Committee. The European First Episode Schizophrenia Trial (EUFEST): rationale and design of the trial. *Schizophrenia Research*. 78(2-3):147-56, 2005 Oct 15.





# Problems on a transborder EHR system in Asia-Pacific region

Takeshi Tanaka, Kiyomu Ishikawa, Hidehiko Tsukuma, Norikazu Iwata\*, Minoru Ikeuchi\*\*

*Hiroshima University Hospital, Hiroshima, Japan*

*Information Media Center of Hiroshima University, Higashi Hiroshima, Japan\**

*Kure University Junior College, Kure, Japan\*\**

## Abstract

*Using a simple model, the authors show the ethical, technical and administrative problems of trans-border electric health record system in a situation of serious disaster whose victims have various nationalities focused on the problems of languages and identification of dead bodies.*

*As results, the authors show the importance of least data set with multi-language characters and treatment of death to keep the rights of patients and medical professionals.*

## 1. Introduction

On 6th Dec. 2004, the huge earthquake with tsunami from off west coast of northern Sumatra in Indian sea caused approximately 283 thousands of deaths and several kinds of damages in various nations. In pacific region, such an international disaster of earthquake with tsunami from Chili had also caused approximately 22 hundreds of deaths not only in Chili but also Japan, Hawaii and Philippines, and obtained damages to the above nations and USA in 1960. Farther more, all of the largest earthquakes (over M8.5) in the world in the last hundred years happened in this region. Thus, people in Asia-Pacific region may be thought to have such a risk of disaster basically [1].

In the case of Sumatra, the identification of dead bodies was one of significant problems. Dental records were the most useful data for the identification in Thailand, in which about two thousands of foreign bodies from 33 nations were identified [2]. Not only dead bodies, but also identification and protection of children lost their parents were serious problems so that Indonesia government was donated electronic fingerprint child ID systems [3].

As a recent economical development in Asia-Pacific region, many people in this region, people working, living and sightseeing internationally have been increasing. For example, the number of Japanese people of lengthier stay or permanent residents in Asia-Pacific region (except North and South America) were about 303 thousands people in 2004, whose ratio of increase were larger than these in other regions in the world [4].

It is very helpful to make an international electric health record (EHR) system taking into account of multi-nationalities in Asia-Pacific region in the situation of international disaster to provide appropriate medical other services immediately. There are the international

guidelines of transborder flows of privacy data determined by OECD, in which main principles are shown: Several nations in EU and USA have already passed legislation [5], so that a transborder EHR can be constructed between them.

Nevertheless, countries in Asia-Pacific region have more various languages, cultures, religions and so on, than these of EU countries. Thus, we should take into account of these varieties to make an international EHR system in this region.

In this paper, the authors show problems of international EHR system in a case of international disaster by using a simple model. Categories of data and access privileges to them on the model EHR system and sample cases to simulate are shown in Ch. 2, problems picked up are shown in Ch. 3, and they are discussed in Ch. 4.

## 2. Materials and Methods

We suppose an international EHR system between several nations ratified the OECD guidelines, in which EHRs of patients of several nationalities flow.

### 2.1. Players on the model system

We suppose the following players, individual users or officers to use this system as following; a patients, his/her next keen determined by the law, his/her primary doctors in his/her domestic region, medical professionals in charge for him/her, other medical professionals not in charge, the domiciliary registration office of patient, police (both domestic and ICPO) or immigration or consular officers, and system administrators.

### 2.2. Categories of data

Categories of EHR data are classified into three levels; non-sensitive personal data, highly sensitive data, the most sensitive data. The first level is for personal data for an identification of patient but not so sensitive which can be shown by all professionals using the system, name or date of birth of patient, for examples. The second is restricted sensitive data which can be seen by professionals for the patient in charge for their purpose of professions; biometrical or medical data like dental data or fingerprints for examples. The third is restricted highly sensitive data which can be seen by a patient and his/her primary doctor, for an example, DNA information [6].







To control the system safely, access log data must be recorded. And access privileges of all users should be controlled by the administrators of the system perfectly.

### 2.3. Access privileges to EHR data on the model system

The accessibility to EHR data on the model system are defined according to the principle of the guidelines of [5] in the table 1.

Although a patient basically holds a right to control his/her personal information, it does not mean that a patient can write all EHR data at his/her own will, since EHR is an official record of medical practice on him/her. Medical professionals have right to use EHR data of the patient in charge according to their profession, and hold rights to control EHR data of their medical practice with the patient. The primary doctor knows all EHR of the patient.

A role of next keen of the patient in this paper is to help an identification of dead body with police or immigrant or consular officers, who have rights to see EHR data by the domestic law of them. The domiciliary registration office of patient gives only the information whether the patient is dead or alive on his/her domiciliary registration to this EHR system.

System administrators can correct EHR data and access privileges of users by the order of right holders of data, or disclose proper EHR data to appropriate persons by the international guideline or the domestic law. They can also audit log files of EHR data to protect the privacy of patient as professionals to control the system.

Any other intruders to the system are rejected perfectly.

### 2.4. Cases to simulate

Simulating the problems of EHR system between several nations in Asia-Pacific region with multi-nationalities in a serious international disaster, it is very difficult for us to seek keys in solving problems by such a multiplicity. Thus, we will focus on two problems. The first case is due to verbal difference; it is difficult to make sufficient communication between medical professionals and patients. The other is identification of unknown bodies on postmortem investigations.

As a sample of the former is that a person who met a disaster in a foreign country is carried to a hospital: Immediate and continuous medical treatments (ischemic heart disease, diabetes, and so on) should be done: The medical professionals have to obtain his/her minimum health information as following; the diagnosis and drug history at the first step, examination data and operation history at the second, and at the third step, the characteristic of body (length, tooth form, deficit of body, etc...) for identification of unconscious patients and bodies are needed. These steps are determined by the degree of demand on emergency state.

The latter is an identification of an unidentified patient who is dead in a foreign hospital by a disaster: The

police or immigrant or consular officers have the list of missing persons who would be met the disaster.

The identification of victims is an important task in an early phase of countermeasure of disaster. It could be helpful for police or immigration or consular officers to search a victim on EHR with the ID of patient.

On these two cases, we will consider ethical, technical and administrative problems.

## 3. Results

Using the model and sample cases defined in the above section, we will point out problems which are not well discussed up to now.

### 3.1. Problems on a patient using a different language than that used in a hospital

#### 3.1.1. Ethical problem

Generally speaking, there are many problems on informed consents in emergency situations; especially in a case that patient cannot explain his/her will.

In this paper, nevertheless, the authors define the principle of using EHR system is to save a life of patients in any situation, and do not discuss the ethical problems on this case.

#### 3.1.2. Technical problem

Multi-byte characters on the computer are used in Asian countries, for example, China, Japan, Korea and Taiwan; where medical records are written in these characters. If the patient comes from these countries to the country where these characters are not used, medical professionals cannot read the EHR data. Although medical professionals could know his/her disease from EHR by ICD-10 codes, they could not know the drug history since there are no international codes for drugs or medicinal properties. The same thing can also be said on the second step.

#### 3.1.3. Administrative problem

System administrators in every nation should prepare a proper rule and procedure of data disclosure (including access logs) not only for police or immigration or consular officers but also for the patient in a way he/she could understand it to keep transparency.

### 3.2. Problems on the identification of dead

In the second case, the problems are more difficult than the first case since the guidelines are not clearly defined: Many guidelines recommend applying the guidelines for living persons correspondingly to dead [7]. However we don't discuss that problem deeply for a convenient.

In this case, any patient doesn't play a main role: Police or immigration or consular officers play important



roles by the law. According to the list of missing persons, the officers collect the information to identify the dead bodies from medical professionals, the next of kin and EHR data of prospects. In the future, they could obtain DNA information from an EHR system.

### 3.2.1. Ethical problem

After the identification, the reason to present his/her EHR data is lost.

For an example in Japan, the legal death of patient determined by a procedure in his/her domiciliary registration office. Thus, the EHR system doesn't have legal correct data of dead or alive after the system could obtain the "correct" information from the domiciliary registration office. After that the patient loses the right as a living human.

We must have international guidelines of the treatments of dead patient's data.

### 3.2.2. Technical problem

Tooth profiles and surgery reports are useful for identification. Thus, these records should have a data structure to search by a tooth or a part of body.

After the identification, statistic data of dead people are used for public health or preventive medicine or basic data of medical science in many nations. Such a secondary use of medical data for public can be made by anonymization of health records.

In a case of huge disaster, a probability of miss-identifications of dead is not negligible. Thus, we have to anonymize EHR data of dead in a linkable fashion [8] in during some period of time for a "rebirth" of dead on the system, if we take these data from EHR system for a secondary use. To protect the privacy of dead people, we had better to anonymize EHR data of dead in an unlinkable fashion after that period.

Thus, the authors think that we had better two steps of anonymization to use EHR data of dead for secondary use.

### 3.2.3. Administrative problems

Considering on human rights and dignities of dead as discussed above subsections, system administrators should have procedures or data connections to obtain the information of dead or alive from every domiciliary registration office in nations supporting the EHR system timely to protect the rights of living people and dignities of dead people.

They should also take care of disclosing of highly sensitive data as DNA information which may violates not only the dignity of dead patient, but also the rights of his/her living families. Administrators must make strict procedures to disclose highly sensitive data.

## 4. Discussion and Conclusion

### 4.1. Discussion on a patient using a different language than that used in a hospital

For proper first-aid treatments of foreign patients on EHR, it must be necessary the international standard codes of diseases, medicinal properties, operations, treatment, so on. Except ICD-10 for diseases, there are no standard codes enough useful. The standard code for medicinal properties especially should be prepared immediately.

After the first-aid, medical professionals might need informed consent from the patient to obtain the information from EHR or his/her primary doctor. It is helpful and comfortable for the patient that the system shows or prints the information enough to obtain informed consent. In addition, many "Kanji" or Chinese characters have the same or similar sound between them. Thus, use of roman character might bring a risk to mix-up the patient. To solve these problems, it is not enough to use multi-language character code such as UTF-8 in the EHR system. It needs to prepare the informed consent form in both domestic language and common language on the system. POC to his/her primary doctor must be on the system to obtain a letter or summary in a common language. Good software of translation would help not only medical professionals and patients but also administrators. The above factors need to keep the right of users of the EHR system.

### 4.2. Discussion on the identification of dead

For convenience to identify dead bodies, it is necessary to define the structure of second level data in EHR, especially for tooth profiles and surgical reports for searching prospect of bodies. However the treatment of DNA information must be more careful. It also should be defined the procedure to obtain the consents from the next of kin and primary doctor of victim to obtain the DNA information.

The first task of the problem of right of dead person is to determine the guidelines to treat the EHR data of dead which could be obtained consensuses in nations supporting the EHR systems. In Asia-Pacific region, it may be more difficult than that in EU. Thus, this should be considered more actively as a proper problem of this region having various cultures and religion, and connected with seas and disasters.

## References

- [1] U.S. Department of the Interior | U.S. Geological Survey,  
<http://earthquake.usgs.gov/regional/world/byyear.php>  
[http://earthquake.usgs.gov/regional/world/events/1960\\_05\\_22.php](http://earthquake.usgs.gov/regional/world/events/1960_05_22.php)  
[http://earthquake.usgs.gov/regional/world/10\\_largest\\_world.php](http://earthquake.usgs.gov/regional/world/10_largest_world.php)
- [2] P. Sribanditmongkol et al, "Forensic aspect of disaster causality management: Tsunami Victims Identification in Thailand",



[http://www.who.int/hac/events/tsunamiconf/presentation/s/2\\_16\\_forensic\\_pongruk\\_doc.pdf](http://www.who.int/hac/events/tsunamiconf/presentation/s/2_16_forensic_pongruk_doc.pdf)

[3] Cross Match Technologies, Inc., “Cross Match Technologies and ITOCHU Make Joint Biometric Technology Donation to Indonesian Government to Track Tsunami Children”,  
[http://www.crossmatch.com/pr/Indonesia\\_7\\_12\\_05.html](http://www.crossmatch.com/pr/Indonesia_7_12_05.html)

[4] The Ministry of Foreign Affairs of Japan, “The statistics on numbers of Japanese foreign residents in the Heisei 16th year.”  
<http://www.mofa.go.jp/mofaj/toko/tokei/hojin/05/pdfs/2.pdf> (in Japanese).

[5] “OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data”  
[http://www.oecd.org/document/18/0,2340,en\\_2649\\_34255\\_1815186\\_1\\_1\\_1\\_1,00.html](http://www.oecd.org/document/18/0,2340,en_2649_34255_1815186_1_1_1_1,00.html)

[6] K. Ishikawa, “Health Data Use and Protection Policy; Based on Differences by Cultural and Social Environment”, Int J Med Inform. 2000 Nov; 60(2):119-25.

[7] For example, Ministry of Health, Labour and Welfare, Japan, “Ethical guidelines on clinical researches”,  
<http://www.mhlw.go.jp/general/seido/kousei/i-kenkyu/ri-nri/0504sisin.html> (in Japanese).

[8] Ministry of Health, Labour and Welfare, Japan, “Ethical guidelines on research of human genom and DNA analysis”,  
<http://www.mhlw.go.jp/general/seido/kousei/i-kenkyu/in-dex.html#1> (in Japanese).

**Table 1: Access Privileges on the model EHR system**

| Data              | Patient  |           | Medical Professionals |                         |                             | Organization.                   |                                      | System Administrator |
|-------------------|----------|-----------|-----------------------|-------------------------|-----------------------------|---------------------------------|--------------------------------------|----------------------|
|                   | Patient  | Next keen | Primary doctor.       | Professionals in charge | Professionals not in charge | Domiciliary registration office | Police / Immigrant / consular office |                      |
| 1st level         | <b>R</b> | r         | <b>P</b>              | <b>P</b>                | <b>P</b>                    | <b>wr</b>                       | r                                    | <b>wr</b>            |
| 2nd level         | <b>R</b> | r         | <b>R</b>              | <b>R</b>                | No                          | No                              | r                                    | <b>wr</b>            |
| 3rd level         | <b>R</b> | No        | <b>R</b>              | No                      | No                          | No                              | r                                    | <b>wr</b>            |
| System Log        | No       | No        | No                    | No                      | No                          | No                              | No                                   | <b>P</b>             |
| Access privileges | No       | No        | No                    | No                      | No                          | No                              | No                                   | <b>wr</b>            |

**Caption**

**R** : Right holder.  
**P** : Holder of right of proper use by one's professional.  
**wr** : Writable by the permission of right holder or law.  
**r** : Readable by the permission of right holder or law.  
**No** : No privileges



## A Development of Personal Health Exercise Chart System for Home-visit-type Exercise Service

Hirai T.<sup>1)</sup>, Higashi T.<sup>1)</sup>, Fukui M.<sup>2)</sup>, Sagawa S.<sup>2)</sup>, Ohashi Y.<sup>2)</sup>, and Ashida N.<sup>2)</sup>

1) Osaka University, School of Allied Health Sciences, Faculty of Medicine

2) Koshien University, Department of Medical and Welfare Management

E-mail:ashida@koshien.ac.jp

### Abstract

*The super-aging society had come in Japan and healthy improvement support business on exercise instruction has been carried out in trial as care prevention for elderly people. As the problem, it was pointed out that the conventional exercise instruction is limited to continuous use of only the "fine" elderly people who can go to an athletic center such as fitness club. Then, we set up the exercise program at a house. Exercise instructor and dietitian were dispatched to the house of the elderly people who cannot go to an athletic center. The healthy improvement support service which performs exercise and ingestion instruction was developed. Also the computer support system of "Personal Health Exercise Chart" was developed.*

### 1. Introduction

In the current society, it is necessary to raise the awareness of physical activity and exercise while creating the environments and methods for the purpose of promoting the physical exercise on a day-to-day basis. 1) 2) 3). Concerning personal health databases (PHDB), in the past, we studied the integration of medical information and the multicenter collaboration in handling medical information. 4) 5) We have studied the participatory personal healthcare monitoring for the purpose of developing the personal healthcare monitoring system which coordinates with the school health and physical education and the community programs, with the intention of creating the personal healthcare database of the personal physical fitness and healthcare records which suits individual life styles rather than the hospital-centered database of medical records. The elderly do the physical exercise which suits themselves, such as the activities like standing up, walking, and stepping, as well as pedaling a bicycle 1) 2) 3) 6) 7). Their main focus is to compare their newly recorded numeric data with their old numeric data which were recorded before. Therefore, many measurement items are included in a set. In this way, measurement items are being diversified. However, conventional

database systems have difficulties in handling such diversified measurement items. In order to solve this problem, we develop the dynamic database, which converts diversified measurement items into the conventional database configuration and process them.

Since October 2005, we conducted a local-community-based study to find if a home-visit-type exercise guidance program helps elderly people maintain their lives and assure their QOL by themselves, while checking its commercial feasibility. Individual exercise programs were prepared by the fitness exercise instructor who actually provides training in order to support the elderly in their doing exercise at home. This program were started at Sakasegawa district, Takarazuka city, Hyogo, Japan, where has developed from of old as the residential quarter. The subjects were the elderly residents in the Sakasegawa district who were neither covered by long-term care insurance or preventive approaches in long-term care nor participated in fitness club or health club activities. For this program, necessary systems were developed with some modification to the PHDB.

This paper is to introduce the home-visit-type exercise guidance and health management system.

### 2. Method

#### Home-visit-type exercise guidance ("Home delivery" fitness club services)

In forming a habit of doing exercise, what becomes inevitable is not to do exercise blindly, but to have the social support which provides appropriate guidance on correct use of muscles with consideration given to individual physical conditions and checks the results of such guidance. In the case of conventional exercise programs where many gather at a fitness facility and do exercise, it is difficult to practically implement the exercise programs which are tailored to individuals. Therefore, we decided to dispatch fitness instructors to individual homes and provide guidance on exercise. Figure 1 shows the outline of this exercise delivery program.



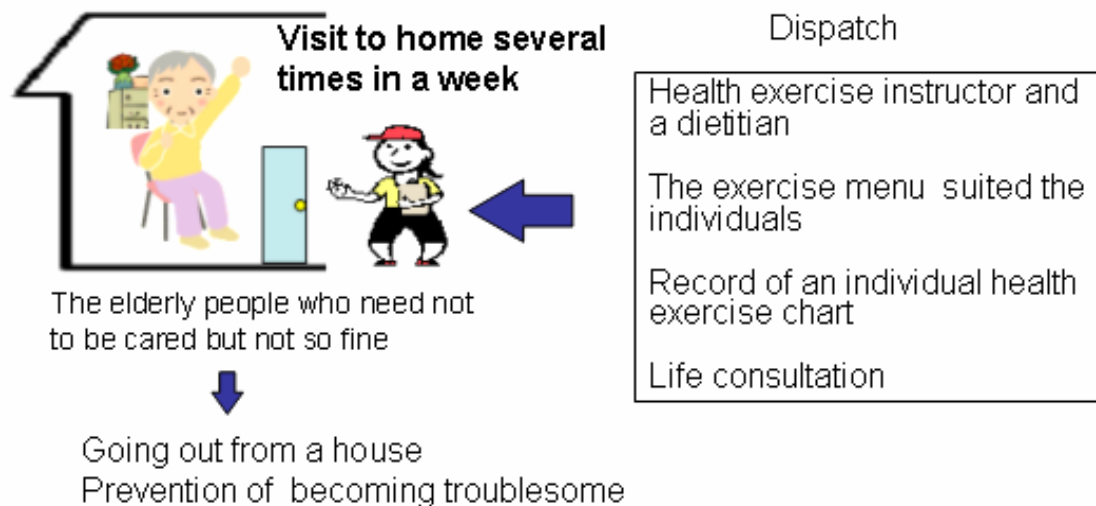


Fig.1. The Outline of Exercise Delivery Program.

### Development of the fitness exercise file for home-visit guidance

In forming a habit of doing exercise, it is important for an individual to set the target by him or herself, improve self-efficacy, and perform self-monitoring. From this viewpoint, an individual was required to keep records on the fitness exercise file by him or herself.

The teaching textbook of main exercise menus at home were made and provided. Figure 2 shows some illustrations how to exercise.. The exercise menu fitted to the individuals was selected. At the beginning, visiting about three times per week instruction and check of whether to perform the right motion are scheduled and later, the visiting became 1-2 instruction per week. Creation of the health exercise chart of home exercise instruction In order to learn a exercise custom, aims are set up by itself, self-effect is heightened and it is important to carry out self monitoring.

Then, health exercise chart was described by himself. It is aware of its health condition by comparing with one's it, and he is conscious of future health. The test-of-physical-strength-and-fitness item (although it has not resulted in what was standardized, and standardization -- a universal thing --) performed now Moreover, the measurement item united with the special feature of the person himself/herself was put in a database, and we registered as many the measurement items and the newly devised measurement items reported by literature as possible, and decided to choose the measurement item which suited itself.

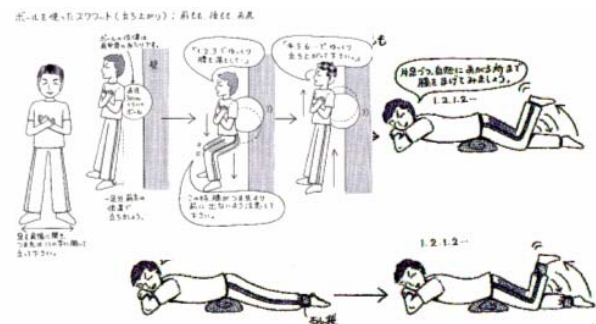


Fig. 2. Muscle training menu at a house: The squat using a ball (standup), leg extension , leg curl, stretching, tiptoe raising, back kick, pose etc.

### System requirement for home-visit type exercise service

To record electronically such various items, which are deferent one by one. Personal health exercise chart (Modified PHDB) was created for the items which are currently measured in physical fitness measurement (those which are already standardized, those which are common but yet to be standardized, and items suitable for individual qualities) and the measurement items which have already been reported in papers and newly developed were added as many as possible, so that the items suitable for individuals can be selected. Also, it contains management system such as scheduling and accounting function for supporting this service.

### 3. Result

#### Community-based experiment for commercialization

The participants of the experiment was eventually comprised of 15 people in total, 14 elderly people ages 53 to 88 and 1 female in her thirties. 5 participants selected (1) to do exercise alone at home and 12 participants selected (2) to do exercise with his or her







friend(s) at the his or her home or his or her friend's home. One participant of each group gave up in midway, and 15 participants continued until the program finished.

### System construction

In this experiment, Excel 2003 VBA (visual basic for application) was used to develop the systems, and all books and sheets could be opened from the sheet 1 of the first book to be used (health management system See figure 1).



fig. 1. Start view of the Personal health exercise chart.

#### (1) Personal list

One book is assigned for a participant. In a book, sheets were prepared for basic information (personal information about a participant, such as the name, age, and address), Self-monitoring (1) (weight, blood pressure, pulse, body fat percentage, and pedometer record), Self-monitoring (2) (training records), physical fitness measurement results, daily reports (written by a trainer who visits the participant), and maps. Trainers check basic information and daily reports to grasp physical condition of participants. Participants enter information about his or her physical information and workouts that they did by themselves for self-monitoring. Physical fitness measurement results include pictures and videos. Pictures and videos are used to help participants grasp their physical changes by checking their pictures and videos. When trainers visit participants, trainers enter information about the workouts done on the day in addition to the muscles which were stretched, so that other trainers can check the content of the previous exercise when they visit next time. From a personal file, plans can be checked and a reservation for a next visit of a trainer can be made. Maps are prepared because trainers need them to visit individual residences. Those maps can be printed out.

#### (2) Training menu

This is to explain the training menu that a participant actually goes through. Videos are prepared so that participants can watch actual workouts.

#### (3) Physical fitness measurement

Periodical physical fitness measurements were carried out to check physical changes of participants. The average results of Japanese people can be checked for measurement items.

#### (4) Schedule

A book was prepared for all trainers to enter their schedules. When a particular date is selected on Sheet 1, it shows the schedule of all trainers on the day. Other sheets except sheet 1 are used for individual trainers so that a trainer can use one sheet to check and update his or her schedule.

#### (5) Commodity control

Some participants purchase balance balls and dumbbells for them to do exercise at home. Such equipment was sold through us, therefore, the sheet to issue a receipt was prepared.

#### (6) Trainer list

The list of the addresses of the participants with the names of the trainers who visit them.

#### (7) Map

The map to show the homes of the participants that trainers visit. It can be printed out, marking individual participants' residents.

#### (8) Accounting

The list to check if a participant pays the fee.



fig. 4 Job flow of the Personal health exercise chart

## 4. Discussion

It seems that precise exercise by instruction and support system corresponding to every elderly people can be taken with this system and the broader customers who need continuous healthy improvement activities are attained. It is said that 10% of local elderly people are "Shutting Oneself Up." Before long, 60 to 70 percent, although everyday life has become independent mostly, it is called what is depended on psychology, society, and an environmental agent. Although this exercise program is not performing health enhancement positively among the elderly people to whom this type shuts itself up and the large majority (90%) of elderly people and elderly people becomes independent, it is aimed at what it is



interested in health and has uneasiness in the future. This social experiment is aimed at investigating how much potential demands such a candidate's excavation method and contents of service have.

Even among the elders, some participate in a full marathon race, and some of those who have neither disturbances nor disorders have problems in walking even several meters. In this way, physical capabilities of the elders vary greatly among individuals, which means age-based criteria are meaningless. As declines in physical strength and capability greatly vary among individuals, they need to be grasped as personal aging phenomena instead of performing a cross-sectional analysis. It is necessary to create the environments where the elders appropriately make physical fitness motions, express them numerically, and compare them with his or her past records, so that they can acknowledge their current health and physical conditions and raise the awareness about future health and physical conditions.

## 5. References

- [1] Ota M., Okufuji T., Ikeda M.; Aged people and physical fitness, JCSM, p549-554, 2003
- [2] Nishijima N.; New physical fitness, JCSM, 20, p53-64
- [3] Tanaka K., Nakamura Y., Akutsu T.; physical fitness for middle and elder, JCSM, 20, p201-210 2003
- [4] 2003 Hasegawa T., Ashida N., et al; Design Philosophy of Personal Health DataBase (First report), 17th JCMI vol17 p422-423 Nov., 1997
- [5] Hasegawa T., Ashida N., et al; Design Philosophy of Personal Health DataBase (Second report), 18th JCMI vol18, p 362-363 1998
- [6] Sagawa S., Ashida N. et al; Follow up the personal changes in women's fitness due to aging (Japanese), 21th Bulletin of Osaka university medical college p83-90, 1993.
- [7] Umeda N., Shigematsu R., et al; Physical fitness in older adult with low, middle and high body mass index J.J.Phys.Educ.Hlth.SportSci.47 p439-450, 2002
- [8] Ashida N., Higashi T., Kubota H. et al.: The Dynamic Database Model Supporting Greatly Varied Items for the Continuous Monitoring of Physical Fitness and Healthcare, JKS Medical Informatics 2003, 9, 2, p327-330
- [9] Ashida N., Kubota H. et al.; Construction of an individual participation type physical strength healthy data management system (Japanese) Japan Journal of Medical Informatics, 23 (Suppl.), p734-737, 2003
- [10] Ashida N., Hirai T., et al.: Educational Campaign of Continuous Exercise from Children to Elderly People. (Japanese). Bulletin of Koshien University College of Contemporary Business Administration, 32, 1-6, 2004
- [11] Okada R., Takashima N., Ashida N. and Higashi T.; Comparison of martial art wrestlers in their shapes -judo, wrestling, and sumo-, 2nd International Judo Federation Symposium, Sept. 2005 Cairo



## Integral Medical Information Interchange with HL7 CDA- A Case Implementation

Chia-Ching Huang<sup>1</sup>, Kuo-Ching Tsai<sup>2</sup>

Cardinal-Tien Catholic Hospital

362, Chung Cheng Rd., Hsintien Taipei Hsien 23137

Taiwan R.O.C.

E-mail<sup>1</sup>: [huang@cth.org.tw](mailto:huang@cth.org.tw)

E-mail<sup>2</sup>: [cth.mis@msa.hinet.net](mailto:cth.mis@msa.hinet.net)

Ruey-Key Chiu<sup>3</sup>, \*Kuan-Chih Lin<sup>4</sup>

Department of IM, Fu Jen Catholic University

No. 510 Chung-Cheng Road, Hsing-Chuang City, Taipei

County, Taiwan

E-mail<sup>3</sup>: [rkchiu@mails.fju.edu.tw](mailto:rkchiu@mails.fju.edu.tw)

E-mail<sup>4</sup>: [jackeir@im.fju.edu.tw](mailto:jackeir@im.fju.edu.tw)

\*(corresponding author)

### Abstract

*This paper presents an experimental system implementation to support patient information interchanging in patient referral between healthcare institutions. The system is implemented based on an agent-based architecture by adopting Health Level 7 Clinic Data Architecture, particularly with the adoption of HL7 v3.0 Standards, and Digital Imaging and Communications in Medicine v3.0 as well as leveraging the use of advanced information technologies. The HL 7 version 2.5 and HL7 version 3.0, respectively are adopted, respectively for message exchange for this research and system implementation.*

*The objectives of this research aim to enhancement the system interoperability and information compatibility among healthcare institutions involved in electronic patient referral.*

*We find HL7 V3.0 provides a thorough component-based reference information model with better transportability, scalability, integration, interpretability but it is also more difficult and complicate to implement but overall it is the trend and unique HL7 standard we need to adopt.*

**Keywords:** Heath Level 7, Clinic Data Architecture, DICOM, Information Technologies, Web Services, Multiple Agents.

### 1. Introduction

The adoption of standards to develop an information interchange system has been regarded as a very important issue to enhance the interoperability and compatibility in all industries in facing this highly competitive age. Health care industry has no exception. Different industries may have developed different standards for different purposes. Regardless of the purposes in adopting standards, they are aimed to improve the efficiency and effectiveness of system development, maintenance, and utilization. In the US, the Administration Simplification Provision of the Health Insurance Portability and Accountability Act (HIPAA) had mandated the use of named health care electronic data interchange standards such as HL7 Standards and its attachment standards such as like CDA for the electronic conveyance of health care data since 1996. Under this addressed act, the professionals of health information systems, providers, payers are requested to adopt the HL7 Standards in sending or receiving detailed claim or encounter information such as patient order data in terms of diagnosis, laboratory results, observations, treatment modalities, and so forth (ASIG at HL7, 2004).

Therefore, in this research we attempt to build a patient information delivery system in patient referral to



enhance the information interoperability and compatibility between health care institutions. The system is implemented by leveraging the use of the standards of medical informatics such as Health Level 7 (HL7) (Hammond, 2003), Health Level 7 Clinic Data Architecture (HL7 CDA), and Digital Imaging and Communications in Medicine (DICOM) and advanced information technologies such as Web services, software agents, and objected-oriented methodologies of system implementation. The HL 7 version 2.5 and HL7 version 3.0 are respectively used to conduct the different message implementation but for the same patient referral application. We use these two different implementations attempting to investigate the advantages, limitation, and constraints of these two versions of HL7 Standards.

## **2. The Standards of Medical Informatics Adopted**

### **2.1 HL7-Health Level Seven and XML**

The HL7 (Health Level Seven, 2005) is developed by HL7 Standard organization which is one of American National Standards Institute (ANSI) accredited standards operating in the healthcare arena. The HL7's domain is clinical and administrative data. Its mission is set the standards for the exchange, management and integration of data that support clinical patient care and the management, delivery and evaluation of healthcare services so that the interoperability, compatibility, and quality of medical data and information can be ensured among healthcare information systems (Health Level Seven, 2006). Its newest version 3, which has been evolving over several years, uses an object-oriented development methodology and a Reference Information Model(RIM) (<http://www.hl7.org/Library/standards.cfm>) to create messages. It will be the most definitive standard to date. It adopts modern class-based concepts to model its information architecture and the use of rigorous

analytic and message building techniques resulted in it has less flexibility with little option and significantly different information model from its past versions. As a result, the downward compatibility of Version 3 becomes very difficult even impossible. In contrast, the HL7 Version 2.5, which was approved as an ANSI Standard in June 26 of 2003 (Health Level Seven, 2003), is the version of HL7 Standards currently most widely implemented in healthcare and medical industries..

As the XML (eXtensible Markup Language) became a single standard for exchanging information, HL7 message is also able to be encoded in XML specification for its version 2.x standards. These XML-specified HL7 Standard version 2.x message is informally named as HL7 v2.xml, were formally approved by ANSI in July 2003 (Health Level Seven, 2003; Health Level Seven, 2004) and has been used by healthcare industry since then. However, HL7 v3.0 RIM is inherited to design in an XML messaging model. It is more consistent and capable of supporting more functionality, completeness, and better for imaging than any of the previous versions but it also looks more complicate in terms of modelling and implementation.

In this research, we attempts to use HL 7 version 2.5 and HL7 version 3.0 respectively to implement the same referral application. Two implementations were conducted in two different stages. We aim to make use of these two different version implementations to assess the advantages and constraints of two versions.

The HL7 Clinical Document Architecture (commonly shorted as HL7 CDA) is an XML document standard that specifies the structure and semantics of "clinical documents" (e.g., admission, discharge summary or progress note) for the purpose of exchange. The CDA can be a payload within an HL7 message, so it is also known as HL7 CDA message. The HL7 CDA is a defined and complete information object which can include text, images, sounds, and other multimedia data.





The Clinical Document Architecture, Release 2 (CDA R2), which became ANSI-approved HL7 Standard in May 2005 (Health Level Seven, 2004), is the most recent version of HL7 CDA. The basic architectural model of a CDA document is shown in Figure 1 (Dolin, et al., 2005).

```
<ClinicalDocument>
  <!-- CDA Header -->
  <structuredBody>
    <section>
      <text> (a.k.a. "narrative block") </text>
      <observation>...</observation>
      <substanceAdministration>
        <supply>...</supply>
      </substanceAdministration>
      <observation>
        <externalObservation>...</externalObservation>
      </observation>
    </section>
    <section>
      <text>...</text>
    </section>
  </structuredBody>
</ClinicalDocument>
```

**Figure 1. Major components of a Clinical Document Architecture document**

As we can see from Figure 1, the basic model of a CDA document, which consists of a header and a body, is rooted by 'ClinicalDocument' tag. In general, the purpose of the CDA header is modeled to identify and classify the document and provide information on authentication, the encounter, the patient, and the involved providers (Dolin, et al., 2005). The body mainly contains the clinical report, organized into sections. The structured body contains one or more 'section' components. The clinical code can be drawn from a LOINC (Logical Observation Identifier, Names and Codes) (Huff, et al., 1998) value set or the clinical data in terms of clinical finding, procedure, observable entity, body structure, and so forth via a code drawn from SNOMED (Systematized Nomenclature of Medicine) (Wang, et al. 2002; SNOMED Clinical Terms, 2006) value set.

In this research, the message include the image is carried by a HL7 CDA message by adopting the Digital Imaging and Communications in Medicine (DICOM) Standard. It is structured as a multi-part document (DICOM Standards, 2004). The image definition,

organizing and transmitting (imaging) adopt DICOM 3.0 Part 3 (NEMA, 2004) and Part 18 (NEMA, 2003). The Web Access to DICOM Persistent Objects (WADO) defined in the Part 18 provides a simple method to carry out the access of DICOM object by using HTML Web page or XML document through HTTP/HTTPS secure network protocol. As a result, in this research implementation, we adopt WADO to carry image object in a HL7 CDA document.

### 3. The Advanced Information Technologies

The modern advanced information technologies are in a variety forms, application domains, and applied techniques. In this research, the major information technologies adopted are software agent, Web services, and object-oriented methodologies for system development.

#### 3.1 Software Agents

The Software agent, one of major studying and applied areas in artificial intelligence, usually bring great benefits in handling routine and repetitive tasks. One of the primary properties of software agents is autonomy, i.e. the ability to continuously monitor and react to changes without the need for human interference or assistance (Shoham, 1993; Jennings, 2000). Other critical properties of software agents are pro-activeness, sociability, and reaction. Software agents exhibiting these abilities can be self-integrated, independently functioning, or inter-coordinated to solve specific problems. They may be easily integrated with an institution's workflow system to play an effective role to seamlessly bridge the gaps between back-end information systems and front-end communication systems (Jennings, 2000).

#### 3.2 Web Services and technologies







Web services are thought to be the paradigm shift of distributed object technology (Booth, et al., 2004). It makes the applications under Web environment proceed to the interoperable services, for examples, description, publishing, searching, and retrieval. Through the use of software components it is possible to loosely couple information systems effectively but irrespective of their platforms or languages as well. The technologies of Web services accelerate the integration of computer systems both in an organization and inter-organizations. It has become a de facto standard under current major IT vendors' supports and advocates.

W3C Web Services Architecture Working Group (W3C WSAWG) (Booth, et al., 2004) defines the architecture of Web services as a layer-plane architecture which can be illustrated as Figure 2. It consists of two layer-planes, which are core service plane and supportive plan. The core service plane provides the service specifications in terms of messages, descriptions, and processes of Web services. The supportive plane contains the rest of service specifications that are defined to support the layer operations of core service plane in terms of the issues of security and management. The core service plane consists of four layer components. The top layer is a process specification layer which is used for the description, aggregation, and choreography of a Web application. In the note released by W3C's WSAWG.

The service specifications of processes layer can be mostly accomplished by the specifications of Universal Description, Discovery, and Integration (UDDI) (UDDI Technical White Paper, 2004). The second layer is the descriptions of Web services language specifications which are commonly named as the Web Services Definition language (WSDL) (WSDL Version 2.0, W3C, 2004).



**Figure 2. The layer-plane architecture of Web services**

The WSDL is an interface definition language based on the XML (eXtensible Markup Language) standard. It provides a method of creating an explicit interface containing the return values, parameters, method names, protocols, and IP addresses so that external applications may remotely access the Web services components published on the web by referring to the information recorded in UDDI. The bottom layer of core service plane is a message layer which is supported by the Simple Object Access Protocol (SOAP) (SOAP Version 1.2, W3C, 2003) to define the protocol for object communications and accessing over the Web. The MIME-based extensive attachment in SOAP Version 1.2 provides a highly secure capability of carrying XML or non XML document in an SOAP object. Thus, HL7 CDA document can be thus securely transmitted and exchange by two communication parties.

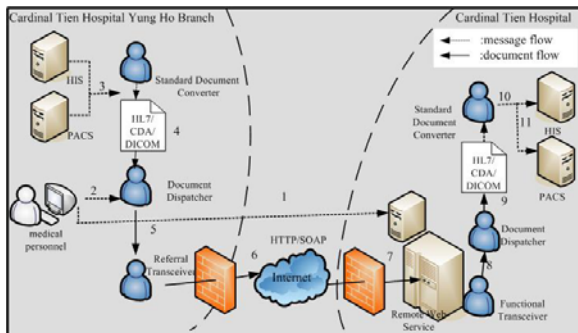
#### 4. Agent-Based Framework and System Implementation

In this research, a multi-agent system framework is designed for implementing the front-end information delivery and interchanging system, which is conceptually illustrated as it is shown in Figure 3. The design of the system framework is aimed to be able to carry out information interchange and transmission for a





medical institution with other ones by the collaborations among different functional role-playing agents.



**Figure 3. The agent-based system framework for information delivery front-end system**

The agent-based system framework is implemented and deployed as a Web services component by leveraging the integrated use of the technologies of multiple software agents and Web technologies. It is mainly composed of three collaborative software agents, which are Standard Document Converter (SDC) agent, Document Dispatcher (DC) agent, and Functional Transceiver (FT) Agent. The front-end system implemented by these collaborative agents can be loosely integrated with institution's back-end systems such as health information system (HIS), radiology information system (RIS), and picture archive communication system (PACS) through SDC, which is another agent designed for data retrieval and conversion.

The SDC may periodically scan and retrieve the data records (e.g., patient data, infectious disease data, or insurance claim data) ready for transferred from the intermediary data repository implemented between the front-end system and HIS, RIS, and PACS. The data records stored in the intermediary data repository are extracted from HIS database (DB) and DICOM database when they are qualified for exchanging and transferring. The intermediary repository is implemented as the apart of databases. The SDC subsequently converts this data record into HL7 CDA document with embedded DICOM image's reference.

The Document Dispatcher (DD) is mainly responsible for activating the Functional Transceiver Agents, which can be Insurance Reimbursement Transceiver to conduct the patient data delivery and interchange to the peer hospital and clinic. The type of message uses LONIC code to identify in the header of HL7 CDA document.

The Functional Transceiver (FT) is mainly responsible for sending message to and receiving the return message (in HL7 CDA) from the external healthcare institutions, which can be another hospital/clinic, CDC, BNHI, and so forth. Again, which destination Functional Transceiver should be activated is dependent upon the type of message to be transferred.

The healthcare standards and advanced information technologies introduced and discussed in previous sections are adopted for the system implementation of this research. Two experimental referral systems are implemented based on this system framework. One is built in HL7 xml.2 standard message and the other is built in HL7 v3.0. Both experimental systems are developed on Microsoft Window 2003 Server running Java Virtual Machine. The databases and intermediary data repository are implemented on Microsoft SQL Server 2000 as well. The object-oriented technology of system development is adopted by using Java programming language with its Web services developed platform. The Web services are implemented by using Java Web Services Developer Pack Version 1.4 (JWSDP 1.4), which is a free integrated toolkit developed and provided by Sun Microsystems (<http://java.sun.com/webservices/jwsdp/>). We take the advantage of this toolkit to build and deploy the Web services and Web applications for the front-end message component of Functional Transceiver between two communication sites. The system experiment environment intends to emulate the real one which includes a main hospital at one city and its subsidiary





clinics located in other cities. The messages are transferred over the Internet between these two sites. The experiment of these two systems and their experimental results are described in the following two sections, respectively.

#### 4.1 Experiment 1- by Adopting HL7 xml.v2 plus CDA 2.0

The sample code for this implementation of using HL7 xml.v2 plus CDA 2.0 is shown as Figure 4.

```
<ClinicalDocument>
  <id extension="G23459876A0000120050714" root="2.16.840.1.113883.933" />
  <code code="28616-1" codeSystem="2.16.840.1.113883.5.25" displayName="TRANSFER SUMMARIZATION NOTE" />
  <title>轉大診所轉診單</title>
  <effectiveTime value="2005/7/14" />
  <confidentialityCode code="N" codeSystem="2.16.840.1.113883.5.25" />
  <author>
    <time value="2005/7/14" />
    <signatureCode code="S" />
    <assignedAuthor>
      <id extension="醫師代碼" root="2.16.840.1.113883.3.933" />
      <addr> <!-- omitted --> />
      <assignedPerson> <!-- omitted --> />
      <representedOrganization> <!-- omitted --> />
    />
  />
  <!-- omitted -->
  <StructureBody>
    <component>
      <section>
        <title>IMAGE_LINK_IN</title>
        <text>
          <paragraph>
            <content ID="Out_image">
              <link>
                <link_html title="CR" href="http://140.136.155.194:8080/FJU_REF/DicomView.jsp?requestType=WADO&studyUID=1.2.392.200036.9125.0.199302241758.16&seriesUID=1.2.392.200036.9125.0.199302241758.16&objectUID=1.2.392.200036.9125.0.19950720112207" />
              />
            />
          />
        />
      />
    />
  />
</ClinicalDocument>
```

Figure 4. The CDA sample codes with HL7xml v2

The CDA Header code element presents this CDA document purpose. In figure 4, the code represent this is a transfer summarization note. The CDA header contains related medicine objects, such author, data enter, patient, represent organization and others. CDA body include related illness description and diagnosis information. The medicine image is indicated by “paragraph” and “link” elements. The DICOM file connected method complied with DICOM part 18: Web Access to DICOM Persistent Objects (WADO).

The selected user interfaces of this system, which are seen at the physician workstations at both sites, are shown in Figure 5. During the patient referral, the physician or nurse can submit and validate the required

data including text, and images to be referred. These data include patient’s demographics, medication summary, laboratory data, relevant images, and physician data from the system installed at the patient referring hospital/clinic site. By the way, the registration for the referred patient to the referred clinic/hospital also can be done during the time of proceeding referral as well. The patient’s medical image files are transmitted simply by giving URL link address so that the referred physician may decide if to refer these images by using download images. The same image viewers are installed in patient’s workstations at both sites. The clinical data are documented as an HL7 v2.xml CDA document.

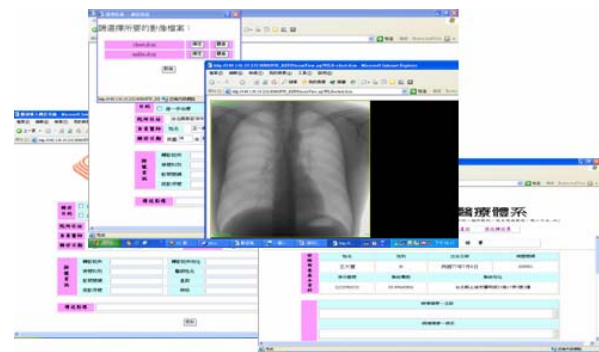


Figure 5. Patient data referred & seen by physicians

The advantages of the use of this system for transferring patient data during patient referral between hospitals/clinics are also very apparently. Without the aids of this system, the hardcopy of the laboratory data, medical image, and hand-writing patient record summary are only carried by human or sent by post mail. It takes from hours or even days to reach the patient referred-in hospital. Now, it can be done within minutes. The transaction data received can also be directly integrated into hospital’s patient record. This can be significantly reduced the time of reproducing the patient data by hand. Meanwhile, it can reduce the cost and the waste of medical resource by sharing the medical record between medical institutions.

#### 4.2 Experiment 2- by Adopting HL7 v3 plus CDA 2.0



Figure 6 shows CDA sample codes transformed by our system. In Figure 6, each medical information object related to this CDA Document is contained in CDA Header, which are referred to HL7 RIM and CDA 2.0. In CDA body, it used <section> and <text> to indicate each medical information, including medication, family history, assessment or physical exam section. The <renderMultiMedia> uses to indicate related multimedia like DICOM file, which is specified in the entry element. Similarly it is implemented in HL7 V2, the DICOM file connected method complied with DICOM part 18: Web Access to DICOM Persistent Objects (WADO).

```
<ClinicalDocument>
<id extension="F12345678920060505" root="2.16.840.1.113883.19" />
<code code="28616-1" codeSystem="2.16.840.1.113883.6.1"
displayName="TRANSFER SUMMARIZATION NOTE" />
<title>轉學醫轉診單</title>
<effectiveTime>20060505</effectiveTime>
<confidentialityCode code="N" codeSystem="2.16.840.1.113883.5.25" />
<versionNumber value="1" />
<recordTarget>
<patientRole>
<id extension="F123456789" root="2.16.840.1.113883.19" />
<patient>
<name><given>林</given><family>大德</family></name>
<administrativeGenderCode code="M" codeSystem="2.16.840.1.113883.5.1" />
<birthTime value="19880403" />
</patient>
<representingOrganization><!-- omitted --></representingOrganization>
<!-- omitted -->
<StructureBody>
<comp>
<section>
<code code="11384-5" codeSystem="2.16.840.1.113883.6.1" codeSystemName="LONIC" />
<title>Physical Examination</title>
<text>頭顱MR檢查表面無明顯外傷</text>
<entry>
<observationMedia classCode="OBS" moodCode="EVN" ID="D166">
<id root="2.16.840.1.113883.19" />
<value type="MR" mediaType="application/dicom">
<reference value="http://140.136.155.72:8080/CTHClinic/imageViewer.jsp?requestType=WADO&
studyUID=1.2.840.113619.2.25.1.1762157631.873231884.123&
seriesUID=1.2.840.113619.2.25.1.1762157631.873231884.162&
objectUID=1.2.840.113619.2.25.1.1762157631.873231884.166" />
</reference>
</entry>
</comp>
</StructureBody>
</ClinicalDocument>
```

Figure 6. The CDA sample codes with HL7 v3.0

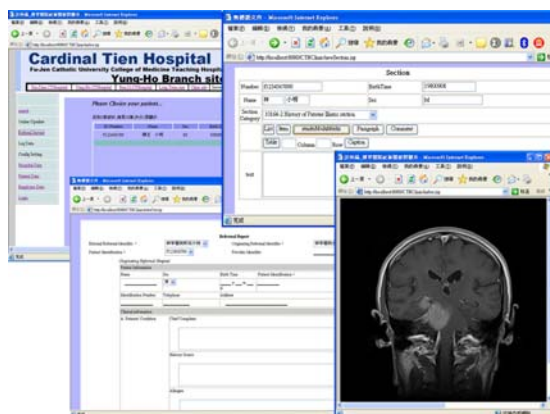


Figure 7. Patient data referred & seen by physicians

The way pf this implementation, not only original system's advantage is retained but also enhances the system expansibility and flexibility. This system's

message transaction uses web services to accomplish. It provides highly transportability, scalability, and integration. The user interface not used original referral from, design involved object concept, which could add section and content flexibility in CDA body. Selected user interfaces are shown as Figure 7.

## 5. Discussion, Suggestions, and Conclusion

In this paper, we present two implementations for patient referral system based on our proposed multi-agent framework to support the information interchange and delivery over the Internet between medical institutions. One is build by adopting HL7 xml.v2 while the other is built by adopting HL7 v3.0 RIM for message transferring and exchanging. Both adopt CDA 2.0 for documentation and DICOM 3.0 for imaging. The advantages and limitations of two implementations are assessed. We found the implementation with HL7 v3.0 RIM could be more interpretable and interoperable than the implementation with HL7 xml V2. More significantly, the use of HL7 v3.0 RIM can transform medicine information into object so that it can be seamlessly integrated into object-based programming methodologies and system implementation paradigm and fully take the advantages of the technologies of Web services for message transferring and interchange. It would promote message interchange effectiveness and efficiency, and enhance system integrity, transparency, and scalability. Presently, the Chinese version of HL7 Standard specification in patient referral is defined based on International standard of HL7 v2.4. The adoption of HL7 3.0 has become a trend and unique HL7 Standards we need to take to develop a standardized medical information delivery system. To promote the use of HL7 v3.0, it is necessary to hasten to define a new Chinese version of HL7 Standard based on current international standard of HL7







v3.0 RIM for the adoption by Taiwan's healthcare to develop their standardized HL7 v3.0 messaging system.

Owing to the resource constraints, this system implementation couldn't take all advantages of HL7 v3.0 RIM. HL7 v3.0 is a very thorough and comprehensive specification but it needs other standardized coding systems, for instances LONIC, SNOMED, and others to fulfil the entire document implementation of an application. Presently, SNOMED is still not licensed in Taiwan. At present, we also don't have a unified and standardized drug coding standard for being applied by local system developers. Our healthcare authorities and business firms have to collaborate to overcome these problems, at the meantime to set up adoptable policies and strategies to speed up the promotion pace, otherwise the standardization of medical informatics might only be an ideal slogan.

## Acknowledgement

We would like to express our thanks with gratitude to the funds granted by the National Science Foundation (2005) of Taiwan and the Cardinal-Tien Catholic Hospital (2006) in Taiwan with the efforts contributed by their domain experts, physicians and professional staffs to make this research success and be unique.

## References

- [1] Attachments Special Interest Group at Health Level Seven (ASIG at HL7), (2004) White Paper: HIPAA and Claims Attachments, Preparing for Relation, <http://www.hl7.org/>. Accessed 2005 May.
- [2] Booth, D., Haas, H. and McCabe, F., Newcomer E, Champion M, Ferris C and Orchard D, (2004) Web Services Architecture. W3C Web Services Architecture Working Group Note 11. <http://www.w3.org/TR/2004/NOTE-ws-arch-20040211/>, pp. 6-88. Accessed 2005 February.
- [3] Chiu, R.K. and Chen, C.H., (2005) A Generic Service Model for Secure Data Interchange. *Industrial Management and Data Systems*, No. 5 and 6, Vol. 105, pp. 662-681.
- [4] Chiu, R.K., Chang, K. and Tsai, K.C., (2005) An Agent-Enabled Message Reporting System over Distributed Open Networks. *Third International Workshop on Supply Chain Management and Information Systems (SCMIS 2005)*, Thessaloniki, Greece, July 6-8.
- [5] Cover Pages Hosted by OASIS, (2004) Health Level Seven Releases Updated Clinical Document Architecture (CDA) Specification. <http://xml.coverpages.org/ni2004-08-20-a.html>. Accessed 2006 February.
- [6] DICOM Standards, (2004) DICOM Supplement 101: HL7 Structured Document Object References. <http://www.dclunie.com/dicom-status/status.html#FinalSupplementsSince2004>. Accessed 2005 November 15.
- [7] Dolin, R.H., Alschuler, L., Boyer, S., Beebe, C., Behlen, F.M., Biron, P.V., and Shabo, A., (2005) The HL7 Clinical Document Architecture, Release 2.0. *Journal of the American Medical Informatics Association*, Vol. 13 No. 1, pp 30-38.
- [8] Hammond, W.E., (2003) The Use of Health Informatics to Improve Health and Healthcare, *The Journal on Information Technology in Healthcare*, 1(5), 321-330.
- [9] Health Level Seven, (2003) Health Level Seven Standard Version 2.5, Accredited by ANSI in June 2003, <http://www.hl7.org/>. Accessed 2005 October.
- [10] Health Level Seven, (2004) HL7 Clinical Document Architecture Release 2.0. <http://xml.coverpages.org/ni2004-08-20-a.html>. Accessed 2004 August.
- [11] Health Level Seven, (2005) HL7 Standards. <http://www.hl7.org/Library/standards.cfm>. Accessed 2005 September.
- [12] Health Level Seven, (2006) What is HL7? <http://www.hl7.org/>. Accessed 2006 March.
- [13] Huff, S.M. Rocha, R.A., McDonald, C.J., De Moor, G.J.E., Fiers, T., Bidgood, Jr., W.D., Forry, A.W., Francis, W.G., Tracy, W.R., Leavelle, D., Stalling, F., Griffin, B., Maloney, P., Leland, D., Charles, L., Hutchins, K., Baenziger, J., (1998) Development of the Logical Observation Identifier Names and Codes (LOINC) Vocabulary, *Journal of the American Medical Informatics Association*, Vol. 5, No. 3, May/June, pp.276-292.
- [14] Jennings, N.R., (2000) On Agent-based Software Engineering. *Artificial Intelligence*, 177: 277-296.
- [15] Jennings, N.R., (2001) An agent-based Approach for Building Complex Software Systems, *Communications of the ACM*, 44(4): 35-41.
- [16] National Electrical Manufacturers Association (NEMA), (2003) DICOM V3.0 standard Part 3 : Information Object Definitions published. [http://medical.nema.org/dicom/2003/03\\_03PU.PDF](http://medical.nema.org/dicom/2003/03_03PU.PDF). Accessed 2005 September.
- [17] National Electrical Manufacturers Association (NEMA), (2004) DICOM V3.0 standard Part 18 : Web Access to DICOM Persistent Objects (WADO). [http://www.nema.org/stds/complimentary-docs/upload/04\\_18p\\_u-2.pdf](http://www.nema.org/stds/complimentary-docs/upload/04_18p_u-2.pdf). Accessed 2005 September.
- [18] Shoham, H., (1993) Agent oriented programming. *Artificial Intelligence*, 60: 51-92.
- [19] SNOMED Clinical Terms, (2006) College of American Pathologists. Available from: <http://www.snomed.org/>. Accessed 2006 March.
- [20] SOAP Version 1.2, W3C, (2003) <http://www.w3.org/TR/soap/>. Accessed 2005 October.
- [21] UDDI Technical White Paper, (2004) Introduction to UDDI-Important Features and Functional Concepts, <http://www.uddi.org/>. Accessed 2005 November.
- [22] Wang A.Y., Sable J.H., Spackman K.A. (2002) The SNOMED clinical terms development process: refinement and analysis of content. *Proceeding AMIA Symposium*, pp. 845-849. [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=pubmed&dopt=Abstract&list\\_uids=12463944](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=pubmed&dopt=Abstract&list_uids=12463944). Accessed 2006 March.
- [23] Web services Description Language (WSDL) Version 2.0, W3C, (2004) <http://www.w3.org/TR/wsdl20/>. Accessed 2005 November..







# Electronic Health Record System Evaluation Based on Patient Safety

Yung Yu (Paul) SU, Khin Than WIN & John FULCHER

Health Informatics Research Centre, University of Wollongong, Australia

Correspondence: john@uow.edu.au

## Abstract

*After describing the general characteristics of Electronic Health Records (EHRs), we consider patient safety. We then proceed to discuss IS evaluation in general terms, followed by HIS system evaluation in particular. The D&M IS Success model, Technology Acceptance model and Task-Technology Fit model are considered as potential candidates for EHR System evaluation. Based upon a critical review of the available literature, we draw some conclusions about the appropriateness of current HIS/EHR evaluation approaches. Finally, we suggest that by incorporating patient safety attributes in any EHR System evaluation framework, then this could lead to improved accuracy, and in turn improved patient care.*

## 1. Introduction

Key considerations for a healthcare organization, apart from the primary one of improved patient care, are cost containment and improving the efficiency of work practices. Health Information Systems have the potential to support these objectives, at the same time as ensuring patient safety. It is therefore imperative that we have means at our disposal to effectively evaluate not only an HIS in a general sense, but more specifically whether the use of Electronic Health Records is helping or hindering the realization of such goals.

## 2. Electronic Health Record (EHR) Systems

Any computer system where failure could have a negative impact on a person's health (life threatening in the extreme), should be regarded as a safety critical system.

Adoption of EHRs can potentially lead to better quality and more efficient healthcare, by way of embedded decision support systems which provide ready access not only to (more complete) information about an individual's state of health, but also to demographic health information.

EHRs contain sensitive health data of individual patients, thus it is essential that these systems are both secure and error free. Inaccurate or insecure information could be detrimental to the individual and subsequently to the responsible company or organization.

EHR systems include different types of data and information for different users. Information included in health record systems needs to be accurate and health

record systems should be safe so that they assist health care workers to improve clinical outcomes.

Typical information management services provided by EHR systems include: (a) Recommendation services, (which determine appropriate activities in specific clinical circumstance), (b) Documentation services (which involve data collection, storage of observations, assessment and interventions), (c) Registration services (which integrate demographic and administrative data), (d) Explanation services (which enhance the credibility of recommendation services by providing supporting evidence), (e) Calculation services (which measure time intervals, medication dosages and other computational tasks), (f) Communication services (which include standards for data transfer and data security), (g) Effective presentation services (which facilitate data visualization;), and (h) Aggregation services (which associate outcome, diagnosis and specific guidelines) [1].

Accordingly, EHR evaluation needs to consider whether or not workflow processes enhance patient safety.

## 3. Patient Safety

The Institute of Medicine (IOM) reports that annual deaths resulting from medical error are of the order of 44,000 to 98,000. The IOM defines patient safety as "the prevention of harm to patients where harm can occur through errors of commission and omission." [2] Now the safety and quality of EHRs are inseparable. This necessitates compliance by all stakeholders with a culture of safety, as well as the development of a HIS with patient safety as a primary focus.

Within the context of patient care, errors of omission, errors of commission, slips and mistakes can be collectively regarded as "human error." [3] Thus, the most important issue with regard to patient safety is considering how to avoid harm by reducing human error. Accordingly, several organizations in the developed countries have taken responsibility for patient safety, including the US Joint Commission on Accreditation of Healthcare Organizations, the UK National Patient Safety Agency, the Australian Council for Safety and Quality in Health Care, the Japan Council for Quality Health Care and the Taiwan Joint Commission Hospital Accreditation.

For instance, the Joint Commission on Accreditation of Healthcare Organizations establishes annual National Patient Safety Goals for reducing medical errors. In 2006, for example, the JCAHO focused on





both improving the accuracy of patient identification and on improving the effectiveness of communication among caregivers [4]. Such considerations must therefore be incorporated in any evaluation of EHR Systems.

#### 4. EHR System Evaluation

HIS evaluation considers the relationship between people, technology, and the environment. [5]. Being a complex organization, a HealthCare System gathers an enormous amount of medical and administrative data on a daily basis. IT system evaluation within healthcare tests not only the IT itself, but also checks the interaction between the IT and the system user. Moreover, systems need to be evaluated on a case-by-case basis, since they invariably cater to individual needs, and are system-specific [6].

Now while IT can be used in theory to improve patient safety, it will fail to do so if it is poorly designed, is unsuited to the clinical task at hand, or uses unnecessarily complex technology for its own sake - these can all lead to increased error [7]. Hence, it is important to evaluate whether EHRs could reduce medical errors and/or overcome barriers.

Since EHR Systems are an integral part of HealthCare systems generally, the risk of incidents due to medical errors can be mitigated by correct identification of initial system requirements. Moreover, the safety aspects of EHRs can be summarized as: identification, system security, privacy, confidentiality, consent, disaster recovery, storage, back up, retention period, data standards, data interoperability, data integrity, medication, alerts, data entry, attributes of data quality, system quality [8]. It is therefore important to consider the risk features of EHRs because evaluation studies need to review an HIS from the perspectives of quality, value, effects, impacts, and applications [6].

Several IS evaluation models have been developed for IS evaluation in the large, and these are summarized in Table 1. The question naturally arises as to how applicable such models are to HIS evaluation? Actually, some earlier researchers have made some tentative steps in this direction (see for example, [12]).

HIS evaluation needs to consider the following: (1) goals (i.e. what to evaluate); (2) methods (i.e. how to evaluate [13-15]. In this paper we focus on so called "goal-based evaluation" [14] in order to assess current-generation EHRs. More specifically, we use safety attributes as estimation criteria, stressing "EHR success", the "Task Fit" of EHRs, and "EHR acceptance" (i.e. what to evaluate) in order to explain how to enhance patient safety by using HIS evaluation models in assessing EHRs (i.e. how to evaluate), as summarized in Table 2.

#### 5. Discussion

Compared with the other two models of Table 1, the Updated D & M IS Success model is more

comprehensive, and incorporates aspects of quality (information, system, and service), user satisfaction, user acceptance (use), and net benefits. It thus could be used to measure the success of EHRs. By contrast, the Technology Acceptance Model focuses on end user acceptance - in the present case, of EHRs. The Task-Technology Fit model attempts to measure the factors which relate the task (i.e. reducing human error) to the underlying technology (i.e. EHRs). TTF could therefore assist in determining whether EHRs assist in reducing human error, and thereby improve the performance and quality of patient care.

Right timing is touted by some researchers as an important consideration, since today's success may fail a decade later due to previously hidden technology limitations, changes in demand, and so on [16]. Now while several researchers have developed frameworks and the like for HIS evaluation [5,6,12,14,17,18], no standard evaluation method exists for evaluating EHR Systems from the perspective of patient safety.

#### 6. Conclusion

We have already seen how safety and quality are intimately linked. The key suggestion which emerges from our critical review of the HIS evaluation literature is the incorporation of EHR safety attributes in the application of existing HIS evaluation models to measure the effectiveness otherwise of EHR Systems.

In short, we believe that the establishment of a "comprehensive", "successful", "high Task Fit", and "highly acceptable" EHR system will help healthcare organizations not only to improve efficiency and effectiveness but also to reduce medical errors.

The next phase of our research will involve the testing and validation of our proposed evaluation framework (Figure 1), focusing on EHR systems within the Taiwanese hospital system. The relevant research questions and appropriate methodology to provide answers to these questions are summarized in Table 3.

#### 7. References

- [1] R.N. Schiffman, C.A. Brandt, Y. Liaw, and C.G. J., "A Design Model for Computer-Based Guideline Implementation Based on Information Management Services", *J. American Medical Informatics Association*, 1999, 6(2), pp.99-103.
- [2] Anonymous, "IOM Report: Patient Safety: Achieving a New Standard for Care", *Academic Emergency Medicine*, 2005; 12(10), pp. 1011-1012.
- [3] V.L. Patel, and L.M. Currie, "Clinical Cognition and Biomedical Informatics: Issues of Patient Safety". *J. Medical Informatics*, 2005, 74(11-12): pp.869.
- [4] JCAHO, 2005 (available online at <http://www.jointcommission.org/PatientSafety/>).





- [5] E. Ammerwerth, J. Brender, P/ Nykanen, H.-U. Prokosch, M. Rigby, and J. Talmon, "Visions and Strategies to Improve Evaluation of Health Information Systems: Reflections and Lessons Based on the HIS-EVAL Workshop in Innsbruck". *Intl. J. Medical Informatics*, 2004, 73(6), pp.479-491.
- [6] E. Ammerwerth, S. Graber, G. Herrmann, T. Burkle, and J. Konig, "Evaluation of Health Information Systems: Problems and Challenges". *Intl. J. Medical Informatics*, 2003, 71(2-3), pp.125-135..
- [7] S.K. Newbold, G.J. Kuperman, S. Bakken, P.F. Brennan, E.A. Mendonca, H.-A. Park, and A. Radenovic, "Information Technology as an Infrastructure for Patient Safety: Nursing Research Needs", *Intl. J. Medical Informatics*, 2004, 73(7-8), p.756.
- [8] K.T. Win, H. Phung, L. Young, M. Tran, C. Alcock, and K. Hillman, "Electronic Health Record System Risk Assessment: A Case Study from MINET", *Health Technology Management*, 2004 33(2), pp.43-48.
- [9] W.H. DeLone, and E.R. McLean, "The DeLone and McLean Model of Information System Success: a Ten-Year Update", *J. Management Information Systems*, 2003, 19(4), pp.9-30.
- [10] F.D. Davis, R.P. Bagozzi, and P.R. Warshaw, "User Acceptance of Computer Technology: a Comparison of Two Theoretical Models", *Management Science*, 1989, 35(8), pp.982-1003.
- [11] I. Zigurs, B.K. Buckland, J.R. Connolly, and E.V. Wilson, "A Test of Task-Technology Fit Theory for Group Support Systems.", *The DataBase for Advances in Information Systems*, 2001, 32(3).
- [12] R. Raitoharju, "When Acceptance is Not Enough: Taking the TAM-Model into Healthcare", *Proc. 38<sup>th</sup> IEEE Intl. Conf. System Sciences*, 2005, p.150c.
- [13] H. Samela, and P. Turunen, "Evaluation of Information Systems in HealthCare: a Framework and its Application", 1997. (available online from <http://www.tukk.fi/tjt/TUTKIMUS/HCISS/Julkaisusivut/dokumentit/pekka/final.pdf>).
- [14] S. Cronholm and G. Goldkuhl, "Strategies for Information Systems Evaluation: Six Generic Types", *Electronic J. Information System Evaluation*, 2003, 6(20), pp.65-74.
- [15] J.C. Wyatt , nad S.M. Wyatt, "When and How to Evaluate Health Information Systems?", *Intl. J. Medical Informatics*, 2003, 69(2-3), pp.251-259.
- [16] T. Burkle, E. Ammenwerth, H.U. Prokosch, and J. Dudeck, "Evaluation of Clinical Information Systems: What Can Be Evaluated and What Cannot?", *J. Evaluation in Clinical Practice*, 2001, 7(4), pp.373-385.
- [17] E. Ammenwerth, F. Kaiser, I. Wilhelmy, and S. Hofer, "Evaluation of User Acceptance of Information Systems in HealthCare: the Value of Questionnaires", *Proc. Medical Informatics Europe Conf. – MIE2003*, 2003, St. Malo, France.
- [18] E. Ammenwerth, and N. Keizer, "An Inventory of Evaluation Studies of Information Technology in healthcare: trends in Evaluation research 1982-2002." *Methods of Information in Medicine*, 2005, 44(1), pp.44-56.





Table 1. IS Evaluation Models, their Dimensions &amp; Main Concepts.

| Evaluation Model                       | Dimensions                                                                                                   | Main concept                                                                                      |
|----------------------------------------|--------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------|
| Updated D & M IS Success Model [9]     | System Quality, Information Quality, Service Quality, Use (Information Use), User Satisfaction, Net Benefits | To measure the factors of a success information system.                                           |
| Technology acceptance model (TAM) [10] | Perceived useful<br>Perceived easy of use                                                                    | To understand the factors of computer-usage behavior that cause users accept or reject an IS/HIS. |
| Task/Technology Fit (TTF) [11]         | Group Task<br>Fit Profit<br>Group Support System<br>Technology<br>Group Performance                          | To match the capability of the technology to the demands of the work environment.                 |

Table 2 Proposed HIS Evaluation Criteria

| <b>Goal-based evaluation</b>                                                                                  |                                                                                                                                           |                                                                            |                                                                                                                                                                           |
|---------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| “The lack of access to information during decision-making” and “need efficient communication” in patient care |                                                                                                                                           |                                                                            |                                                                                                                                                                           |
| <b>Estimation criteria</b><br>(Safety Attributes of EHRs)                                                     |                                                                                                                                           |                                                                            |                                                                                                                                                                           |
| 1.identification                                                                                              | 6.disaster recovery                                                                                                                       | 11.data interoperability                                                   | 16.attributes of data quality                                                                                                                                             |
| 2.system security                                                                                             | 7.storage                                                                                                                                 | 12.data integrity                                                          | 17.system quality                                                                                                                                                         |
| 3.privacy                                                                                                     | 8.back up                                                                                                                                 | 13.medication                                                              |                                                                                                                                                                           |
| 4.confidentiality                                                                                             | 9.retention period                                                                                                                        | 14.alerts                                                                  |                                                                                                                                                                           |
| 5.consent                                                                                                     | 10.data standards                                                                                                                         | 15.data entry                                                              |                                                                                                                                                                           |
| <b>goals of an evaluation</b><br>(what to evaluate/ IT-system as such)                                        |                                                                                                                                           | <b>methods of evaluation</b><br>(how to evaluate/called IT-systems in use) |                                                                                                                                                                           |
| <b>Research Topics</b>                                                                                        | <b>Research Purpose</b>                                                                                                                   | <b>Evaluation Models</b>                                                   | <b>Aims of evaluation</b>                                                                                                                                                 |
| EHRs success                                                                                                  | To find the success and weakness factors, and workflow processes of EHRs in real medical environment.                                     | Updated DeLone and McLean IS success model (Updated D & M model)           | Based on safety attributes to evaluate the success of EHRs with those six dimensions, to enhance the functions of the current EHRs..                                      |
| EHRs acceptance                                                                                               | To recognize, forecast, describe, and recognize the factors of computer-practice behavior that cause users accept or reject current EHRs. | Technology Acceptance Model (TAM)                                          | Based on safety attributes to predict, explain, and realize the factors of the behavior that cause users accept or reject current EHRs.                                   |
| Task Fit of EHRs                                                                                              | It attempts to measure and predict the fit of acceptance and use of technology in user evaluation of EHRs.                                | Task Technology Fit (TTF)                                                  | Based on safety attributes to measure the performance of an current EHRs between technology and task based; to check whether this EHRs could reduce medical error or not. |





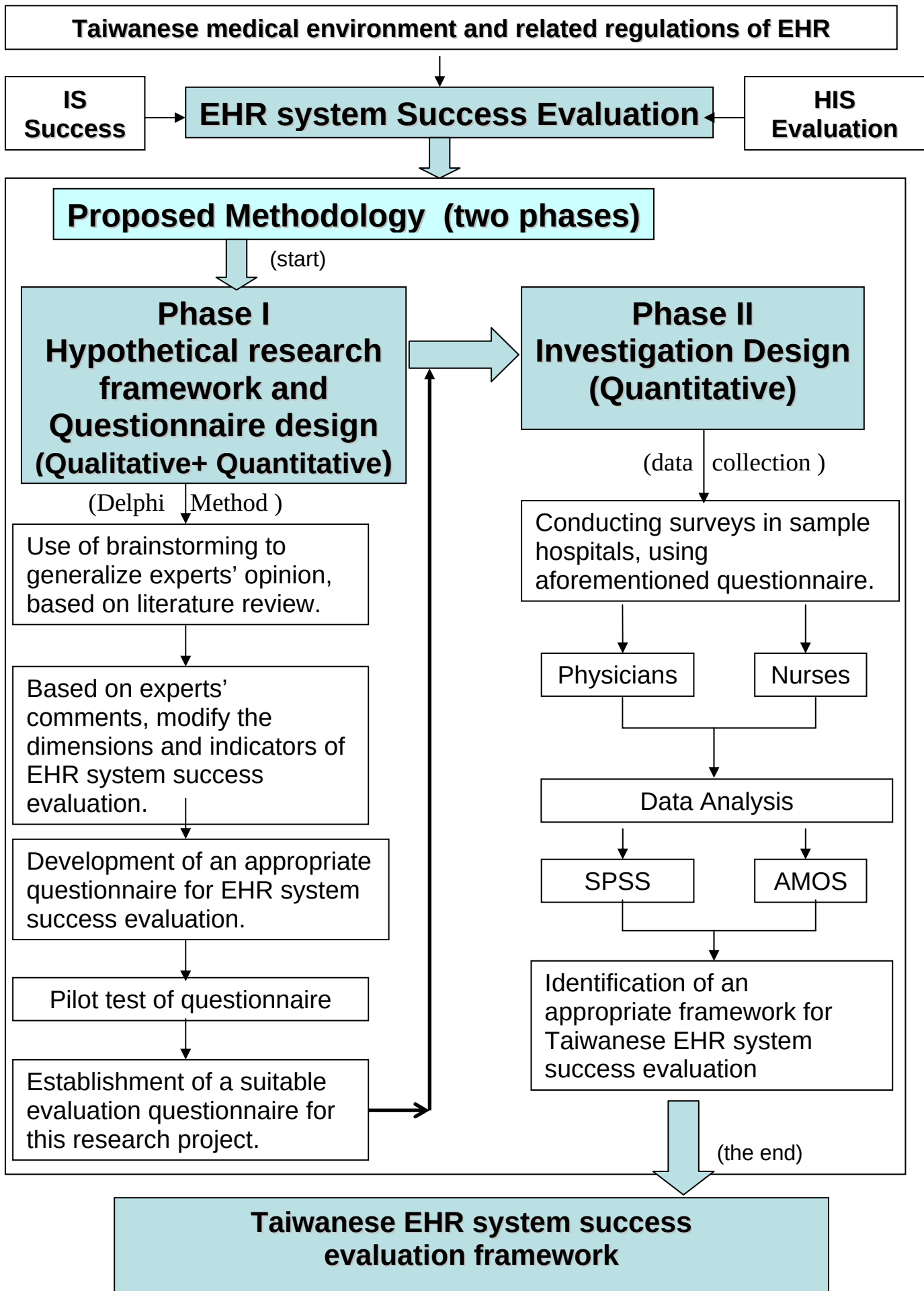
Table 3 Research questions and proposed methods

| Q1: What are the most important characteristics in determining Taiwanese EMR success?                                                                                           |                                                                                                        |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------|
| Specific question                                                                                                                                                               | Proposed methods                                                                                       |
| Objective1: To investigate what the most important dimensions should be in evaluating Taiwanese EMR system success.                                                             | Delphi method<br>Descriptive analysis (SPSS)<br>Reliability test (SPSS)                                |
| Objective2: To investigate what the most significant indicators in estimating Taiwanese EMR success are.                                                                        |                                                                                                        |
| Objective3: Based on the findings of O1 and O2, to develop a hypothetical research framework and suitable evaluation questionnaire for Taiwanese EMR system success evaluation. |                                                                                                        |
| Q2: Does the EMR success evaluation framework demonstrate high content validity and reliability to estimate EMR in Taiwanese hospitals?                                         |                                                                                                        |
| Specific question                                                                                                                                                               | Proposed methods                                                                                       |
| Objective1: To test the accuracy of this EMR success evaluation framework.                                                                                                      | Descriptive analysis (SPSS)<br>Correlation analysis (SPSS)<br>Factor analysis (SPSS)                   |
| Objective2: To test the consistency of this EMR success evaluation framework.                                                                                                   |                                                                                                        |
| Q3: Is this EMR success evaluation framework comprehensive and complete?                                                                                                        |                                                                                                        |
| Specific question                                                                                                                                                               | Proposed methods                                                                                       |
| Objective1: To explore the interrelationship between dimensions and indicators.                                                                                                 | Descriptive analysis (SPSS)<br>Correlation analysis (SPSS)<br>Regression analysis (SPSS)<br>SEM (AMOS) |

Figure 1 Taiwanese EMR system success evaluation framework







## Construction and Evaluation of EHR which Supports Team Practice

Hidehiko Tsukuma<sup>1)</sup>, Takeshi Tanaka<sup>1)</sup>, Kayo Sakata<sup>2)</sup>, Norikazu Iwata<sup>1)</sup>,  
Akie Kawamura<sup>2)</sup>, Kiyomu Ishikawa<sup>1)</sup>

1)Medical Informatics and Hospital Systems Management, Hiroshima University Hospital,  
Hiroshima, Japan

2)Department of Nurse, Hiroshima University Hospital, Hiroshima, Japan

E-mail: tsukuma@hiroshima-u.ac.jp

### Abstract

*Purpose: The new type of EHR is designed to share the information among all medical staff and the patient concerned to promote team practice, which is expected to enhance safety and high-quality medical care. From the viewpoint, the new type of EHR has been constructed and operated in the Hiroshima University hospital since 2003. And the usability of the new type of EHR is evaluated. Methods: The evaluation is done by the questionnaire survey to all of the hospital staff using the HNS. The result of recovering is simple tabulation. Results: About 70 % of the whole answered that the EHR facilitated information sharing among the professions concerned with the medical team. Also, about 65 % of the whole answered that the safety management system was contributing to the improvement of the patient safety. About 66 % of the whole answered that the new EHR contributed to the increase in efficiency of the clinical work. Conclusion: The new type of EHR system is useful for team practice.*

### 1. Introduction

In Japan, information systems in the healthcare field have been developed from the viewpoint of hospital administration for over 30 years. The resulting products off course were still not sufficiently fit for clinical scene. Therefore, making the information system to support clinical staff, over 20 years of the pasts, our interest gradually shifted to EHR.

In the research, we thought that it was important that the future EHR has played a role as a navigator of clinical process in team practice, and a supporting tool for patient safety.

From the above viewpoint, we reexamined the basic design of the conventional information system for the hospital administration, and then the new of EHR for patient-centered team practice has been operated in the Hiroshima University hospital since 2003 [1].

In this paper, the evaluation of the new EHR by the end users is reported, and based on the result, the usability of the system is considered.

### 2. Materials and Methods

#### 2.1. Requirement of the new EHR

The ordering system got excellent results for the abridgment of the patient waiting time and so on, but could not contribute to the efficiency of treatment and nursing process of medical team and to the improvement of the patient safety. Therefore, the authors analyzed the workflow of the ward at Hiroshima university hospital and organized the factor which obstructs the efficiency of them practice and the patient safety as the following.

(1)The medical staff was managing the latest patient information intensively at the staff station, but a lot of posting work was carried out because of the information sharing among the medical staff. The posting work became a new risk factor and also hindered the realtime information sharing among the medical staff. Moreover, the sharing of the information on the urgently changed instruction was impossible.

(2)Because the treatment and nursing care plan were changed frequently, the grasping of the progress of the process was difficult. Therefore, according to the grasping mistake of the latest information, there was possibility of the overlap and non-implementation of medical treatment and nursing care.

(3)The medical staff could not provide the observation and the implementation data occurring at the bedside for the other member of the medical team efficiently until it returned to the staff station.

(4)In the confirmation of the patient who should implement the medical practice, the medical staff sometimes made a mistake.

In order to improve the above-mentioned risk factors, we defined the requirement of the new EHR as the following.: (a)The system should be designed to be able to share the information among the patient and the all of medical professionals concerned to promote team practice, which is expected to enhance patient safety and high-quality medical care. (b)The system should be designed so that the medical professional can simply and surely input the healthcare record in process of medical and/or nursing treatment. Moreover, in the necessary place when necessary, the system should be able to provide an accumulated healthcare record for the medical profession in a form suitable for various scenes of clinical process. (c)The system should have a common browser for all the medical professional of the medical team to grasp the status of progress of medical and/or nursing process. Progress of the clinical process is managed and presented on the browser





chronologically. (d)The system should be designed to lose the posting work and to reduce a work load of the medical professionals. (e)The system should have a function to prevent from the mistaking of a pharmaceutical and/or a patient. (f)The system should be designed that accumulated clinical process data can be used for the rational hospital management.

## 2.2. Outline of the construction system

Being based on the requirements defined in subsection 2.1., we have constructed the EHR with the following functions in the Hiroshima Univ. hospital since 2003.

(a) We introduced 10 laptop computers which the wireless LAN works per 50 beds. Moreover, we developed the special wagon (hereafter, we call it “Bedside Wagon”) to put the laptop computer on the top board. Using the bedside wagon, all medical professionals can carry the laptop PC to wherever the medical practice is done in the ward (for example, patient’s bedside, treatment room, and so on), and can input the clinical data while referring the latest status of the progress of medical practices to be done, or patient’s healthcare information in the appropriate timing and place.

(b) To facilitate safe team practice, we designed and implemented the healthcare navigation system (HNS) to support information management in the clinical process (Fig.1). (A model of the HNS, which is the core of the EHR in the Hiroshima University Hospital, may serve as a practical example of the implementation of HNS in the service of safety management). This tool supports to accumulate the practice data, such as “Watch”, “Think”, “Order”, “To do”, and “Do/Done” to treat patients as team practice, and to use these for decision-making and evaluation. These data are available not only to the physicians but also to all other medical professionals concerned. EHR data are recorded not only by physicians, but also by all co-workers. This may be the basis for team communication and a reliable way of treating in such a way that the evidence of conducted practices is traced.

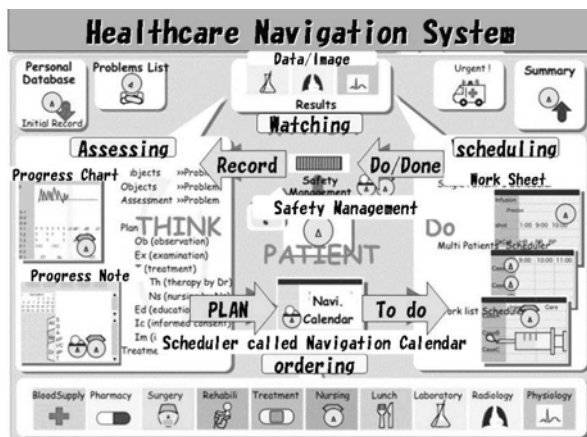


Figure 1. Healthcare Navigation process in EHR

**(b-1) Support of preparing Personal Health Database:** While referring to the anamnesis (support the next item “Watch”), input the basic information to confirm, and make the “Problem List” and prepare the short-term program and the long-term clinical and nursing Policy

**(b-2) Support of “Watch”:** Develop the data chronologically (calendar) to support the clinical process. This facilitates selection, retrieval, and understanding of patients’ data.

**(b-3) Support of “Think”:** This is comprised of Progress Note in accordance with the POS system and Flow Sheet describing the process to be done by the medical professionals. Information can be input byfrom both physicians and nursing staff according to the purpose of practices and is displayed concurrently. DuringWhile inputting of the information, professionals consider and refer to the patient’s condition, process, and the information which other professionals have input. “Input time and date” and “the name of the operator” are automatically attached and accumulated.

**(b-4) Support of “Order”:** This is the function to deal with orders to the other professionals. While referring to the progress, this has the calendar-formed launcher function (called “navigation calendar”; fig.2) to support make future order systematically. Construction of the orders.

Figure 2. Example of the “navigation calendar”

**(b-5) Support of “Do/Done”:**

**a) Work navigation** Orders are made as a list to do and compiled in a “scheduler” (fig.3) by individual patients or physician in charge, and displayed on a calendar form. The process of the practices is displayed in a parallel line along with the actual time until the orders are completed



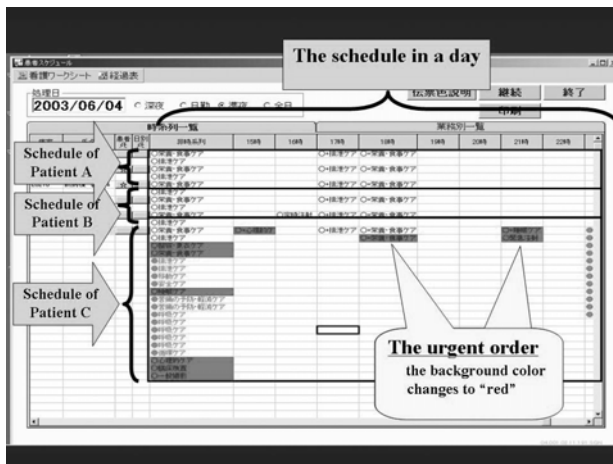


Figure 3. Example of the “scheduler”

**b) Safety management** Throughout the process from the order to dosage, treatment, and other practices, safety management is conducted by CMS to prevent from treating the wrong patient or administering the wrong dosage. Basically, a barcode reader reads “3 items” to confirm “patient”, “practice (including confirmation of medical materials)” and “doer” and logs them with the time (Fig. 4).

The collected data can be accessed on a calendar table, which is easy indicated to recognize when a critical we make close examination procedure is being performed such as comparison between infusion volume and out of water volume for critical cases.

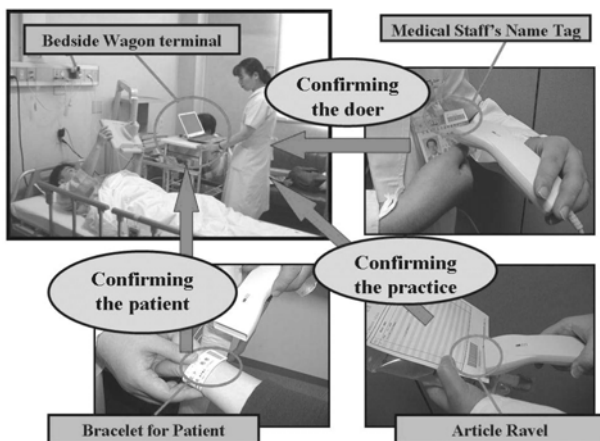


Figure 4. “Three steps confirmation” system both by mobile terminal and barcodes to confirm “patient”, “practice” and “doer”

### 3. Results

#### 3.1 Evaluation Method

By the questionnaire survey to the user of the system, we evaluated the EHR. The questionnaires methods are as follows.

#The period: August 24~30, 2005.

#The object: All of the hospital staff using the HNS (distribution of 1320).

#The answer method: The automatic recording custody method by the questionnaire of the own composition.

#The question contents: “It is possible to do information sharing using the HNS?”, “the contribution to team practice of the HNS?”, “the situation of utilization of the injection safety management system?”, etc..

#The number of answer: Valid answer 784 (Physician 237, nursing staff 411, staff of the central diagnostic and treatment department 136), 59.4 % of answering ratios.

#An analytical method: The result of recovering is simple tabulation.

### 3.2. Result

(a)About 70 % of the valid respondent answered that the HNS facilitated the sharing of medical care information among the medical care staff. Also, about 66 % (82 % if limiting to nursing staff) answered that the HNS was useful for team practice. 30 % of the respondents evaluated the system negatively. The opinion of the respondents was classified into the following three. (a1)The part of the medical record is written using the paper medium, for example the report of the physiology inspection and the pathology inspection and so on. Therefore, the medical staff must refer to both of the PC terminal screen and the paper. This forces a medical staff into the complication and the risk (20 respondents answer). (a2)Because the system response is bad, the system hinders the medical profession works smoothly (19 respondents). (a3)It is sometimes impossible to use a PC terminal when the medical staff needs it because the PC terminals don't exist sufficiently (22 respondents).

(b)About 98 % of nursing staffs answered that they used an injection safety-check system near the patient at the sickroom before the injection. Also, 65 % of the valid respondents (88 % if limiting to nursing staff) answered that this system contributed to the improvement of the patient safety. The opinion of the negative respondents was classified into the following two. (b1)It is sometimes impossible to use the PC terminal when the medical staff needs it because the terminals don't exist sufficiently. (b2)Because the computer doesn't check necessary pharmaceutical individually in this system, the medical staff must confirm with its eyes. If the medical staff mistakes in this step, the patient safety can not be guaranteed.

(c)About 66 % of the valid respondents answered that the efficiency of the medical practice improved by using this system. Specifically, about 60 % of nursing staff answered. that the time spending to create a nursing record decreased. For example, the handwriting job which spends 1 hour reduced to about 10 minutes after the introduction of the system. The reduced time







is effectively reused for patient care at the bedside, conference, and decrease of the overwork time.

On the other hand, the opinion of the negative respondents was classified into the following four. (c1)The system response is bad. (c2)There are few PC terminals. (c3)The medical data which was managed with the paper medium exists. (c4)The perspicuity of the information is bad, because the required data is dispersed on more than one screen.(c1)- (c3) are the same as being pointed out in (a1-3). (c2) and (c3) are the problem which can be solved if there is an enough budget. On the other hand, to solve (c1) and (c4), the improvement of the application must be studied in the future.

From the above result, the EHR system implemented at Hiroshima University hospital can be approximately evaluated as the useful tool for the team practice among the concerned medical staff.

## 4. Discussion

### 4.1. Problem on the system performance

From the result of the previous section, it was clarified that the medical professionals evaluated well the functionality of the EHR installed of this time at Hiroshima University hospital. However, there are many problem on the response time of the system. For example, it takes from about 30 seconds to 1 minute to display medical and nursing treatment schedule of 50 inpatients onto the screen of the patient schedule (referring to fig. 3). And, the retrieval which covers multiple data classification such as medication history and test result also takes considerable time. In the meantime, input and/or retrieval of the ordering data (clinical examination ordering, injection ordering, etc.) finish in the short time (about 2-3 second).

The cause of the response problem in this time is in the data structure of the HER system. As for the system construction in this time, we were based on the package of the system vender which had data structure optimized by the ordering system. Then, we customized the user interface along the concept of the HNS without changing the data structure of the package.

In the scene of team medical practice, the data processing which covers multiple data classification is the most part, therefore the research of the data structure which can efficiently support the clinical work of the medical staff is one of the future problem.

### 4.2. How should IT support medical professions?

(a)When an IT system too excessively supports the human being, the medical professions which can not do act without the system may increase in a healthcare field. In this meaning, we reviewed carefully about how should guarantee safety of the injection by the safety-

check system. The point is confirmation work in the mixing stage of the drug. In the reviewing of the system, we received the request from the physician and/or nursing staff that it wanted to confirm individual drug and consistency of the indication at the mixing stage of the injection with the bar code. If the computer checks individual drug, we think that a new-type risk rises on the contrary. That is to say, the worker of injection mixing should work while confirming the instruction contents of the injection, but when a bar code checking system is once introduced, the worker may omit the confirmation stage of the instruction contents. As a result, the ability of medical professions to check the mistake of the instruction contents probably declines. From that reason, a checking system of individual drag was not installed in our EHR.

That is to say, the IT system accounts for providing the latest (the last) version of the injection instruction in the necessary timing, and the medical staff accounts for evaluating correctness of the clinical contents of the injection instruction and correctly preparing the injection. (b)It is necessary that we consider so that an IT system may not inhibit the direct communication of the medical staff concerned with team practice. For example, we discussed how the nursing staff can certainly receive the urgently inputted information by the physician through IT System. The plan was reviewed, in which the nurse's mobile phone automatically rang when a physician inputted urgent data. But it was not adopted, because we thought that the conversation between the physician and all other professionals concerned with team practice decreased when IT system was overdoing. Instead, it should be covered by the operation as displaying the alert of urgent indication on the terminal screen automatically, so that the physician whom it urgently input may had to contact simultaneously, the nursing staff by the oral.

### 4.3. The problem on the operation

It is necessary that the system operation department would not cause the problem by not expecting system trouble for the EHR system. By the duplication of the system bases such as the disk of the EHR server and the network, we guaranteed a continuous-running to the hardware fault as far as it was possible. As the result, service stop over the whole hospital did not happen by the hardware failure for 3 years since introduction of the system.

If anything, we had another problem: a printer fault (including paper jam, emptying a toner), a mistake concerned with master correction, and so on. These problem occurred frequently. Therefore, we increased about 1.5 employments for the operation of this system.

We consider that the cost decrease by the labor saving of medical staff (see subsection 3.2. (c) in this paper) can sufficiently cover the cost increase by the employment of the operation personnel.







## 5. Conclusion

We have confirmed the effectiveness of the EHR which was implemented according to the requirement defined in section 2. Actually, it is fixed as the tool which is indispensable for the medical team in healthcare field.

Moreover, in the future, using the accumulated data, we want to carry out evaluation of the quality and the process of medical and nursing care.

## References

- [1] K. Ishikawa, N. Konishi, H. Tsukuma, S. Tsuru, A. Kawamura, N. Iwata, T. Tanaka, "A Clinical Management System for patient participatory Health Care Support", *International Journal of Medical Informatics* Vol. 73, 2004 , pp. 243 -249





# System Interoperability Study for Healthcare Information System with Web Services

J.K. Zhang <sup>a</sup>, W. Xu <sup>a\*</sup>, D. Ewins <sup>a</sup>

Centre for Biomedical Engineering, School of Engineering, Surrey University <sup>a</sup>  
W.xu@surrey.ac.uk

## Abstract

*This paper describes the use of a new distributed middleware technology 'Web Services' in the proposed Healthcare Information System (HIS) to address the issue of system and language interoperability raised from existing Healthcare Information systems. With the development of HISs, hospitals and healthcare institutes have been building their own HISs for processing massive healthcare data, such as, systems built up for hospitals under the NHS (National Health Service) to manage patients' records. Nowadays many healthcare providers are willing to integrate their systems' functions and data for information sharing. This has raised concerns in data transmission, data security, and network limitation. Among these issues, system and language interoperability are one of most obvious issues since data and application integration is not an easy task due to differences in programming languages, system platforms, Database Management Systems (DBMS) used within different systems. As a new distributed middleware technology, Web service brings an ideal solution to the issue of system and language interoperability. Web service has been approved to be very successful in many commercial applications (e.g. Amazon.com, Dell computer, etc.), however it is different to healthcare information system. As the result, Web Service-based Integrated Healthcare Information System (WSIHIS) is proposed to address the interoperability issue of existing HISs but also to introduce this new technology into the healthcare environment.*

**Keywords:** Web service, XML, SOAP, WSDL, UDDI, Healthcare Information System

\* Corresponding author: Dr W Xu,  
School of Engineering, University of Surrey,  
Guildford, Surrey, GU2 7XH, UK  
Tel:01483 682368 Fax:01483 689508  
Email: w.xu@surrey.ac.uk

## Introduction

With the evolution of IT technologies and increasing demands from people for computerising medical-related information, healthcare institutions have developed systems to manage and process the large amount of medical information. In order to achieve information exchanging and medical knowledge sharing, hospitals and healthcare providers have intended to integrate their systems' functions and data. This raised some concerns, such as,

data security, data transmission, network limitation and so on [1]. Among these issues, the issue of system and language interoperability are most obvious barrier for the integration of functions and data of different systems. In practice, systems were developed using different languages (e.g. Java, Visual Basic, C++, etc.), different system platforms (e.g. Microsoft Windows operating systems, Linux operating system, etc.) and DBMSs (e.g. Microsoft SQL server, Oracle, Microsoft Access, etc.) [2-5]. These differences between systems are major factors leading to the issue of system and language interoperability. Traditionally, developers use distributed middleware technologies, like CORBA (Common Object Request Broker Architecture)/DCOM (Distributed Component Object Model) to tackle the issue of system and language interoperability. Regarding to the author's research, there have been some of HISs, such as I-HER [3], Hospital Information System [5], BHS [4]. But CORBA/DCOM would become unsuitable while they are applied into those systems based on Internet environment. That is because CORBA/DCOM has the shortcoming in the firewall crossing and wireless environment [6-8]. In addition, CORBA/DCOM is a fairly complex issue, a task requiring special expertise [9]. Nowadays, Web service as a new distributed middleware technology can not only overcome those shortages from CORBA/DCOM but also to successfully address the issue of system and language interoperability. Web service has been successfully applied in many commercial applications, nevertheless, a few of people and organisations formally deployed this technology into healthcare industries. In this paper, a Healthcare Information System named as 'Web Service-based Integrated HIS (WSIHIS)' is proposed and built up with the support of web service and some Microsoft .Net-based technologies to address the issue of system and language interoperability. The main objective of this paper is to illustrate how Web Service and Microsoft .Net technologies are used to deal with the issue of system and language interoperability. Moreover, system interoperability test was carried to assess if whether the web service-based solution is really able to cope with the issue of system and language interoperability.

## The issue of interoperability of current HISs

The interoperability can be defined as "the capability with which two or more programs can share and process information irrespective of their implementation language and platform" [5, 10]. Interoperability concerns also rise as one of the most urgent priorities of healthcare information systems [11, 12]. Connecting for health, as a single IT provider for the NHS defined the interoperability issue 'it is impossible for different healthcare information systems to smoothly, seamlessly, privately communicate information when it is needed to where it is needed' [1]. Most clinical applications are determined by a huge variety of heterogeneous and independent work places, most of them equipped with specialised clinical hardware [12]. Usually, different parts of a large-scale telemedicine system might be from various vendors, who use different standards and information formats, such as, National Program for IT (NpfiT) [13]. These systems are also used by people with different levels of expertises and needs. Once the interoperability problem is solved, the development and maintenance of large telemedicine systems can be





streamlined with data reuse, code reuse, application reuse and choice of an appropriate computing environment using object-oriented technology. The following content would explain the interoperability problem from different perspectives [5, 10]:

**Database System Interoperability:** Patient records are often located in different database systems [14]; however, data from different database systems cannot be exchanged with each other and used by applications based on different DBMSs.

**Language Interoperability:** Usually, different HISs could be developed by different IT providers [15], developers make use of different programming language to built their HISs. This would make the reuse and share of applications between different HISs very hard because of the incompatibility between different programming languages.

**System-platform Interoperability:** System platform interoperability means OS interoperability, but over the last few years the Internet browser has emerged as a platform in itself [16]. As different HISs could be developed based on different development platforms, this feature would make those HISs only work on some certain system platforms.

**Semantic Interoperability:** some interoperability problems are caused by semantic differences. Semantic interoperability assumes that the components of the distributed application will have different meanings. Usually human intelligence is required to solve it, however this problem has been beyond the scope of this paper.

The following section would mainly discuss the issue of system and language interoperability based on those systems delivered by the NHS system as a typical example. Within recent years, there is a huge program being delivered by the NHS – National Program For IT (NpfiT). NpfiT was proposed in February 2002, which is aiming to modernise the NHS system with information technologies [17]. In Oct 2002, this program is formally established, which mainly comprises of various healthcare systems for different purposes. It is no doubtful that the interoperability issue is inevitable because of the following reasons:

**Different business processes in different ways [18]:** There are thousands of hospitals under the NHS, most doing different business processes and having different health systems, which makes NpfiT more complex to interoperate with those various systems.

**The requirement for the interoperability between systems delivered by NpfiT:** For example, Electronic Transmission Prescription, as one of systems delivered by NpfiT was originally envisaged as a separate system [15]. However, this program will also be required to be integrated with another program of NpfiT ‘NHS Care Records Service’ [19].

**Dozens of different patient-record systems in use across the NHS:** There are dozens of different patient-record systems in use across the NHS developed by different software providers [15]. This situation would increase the difficulty for the integration of different records.

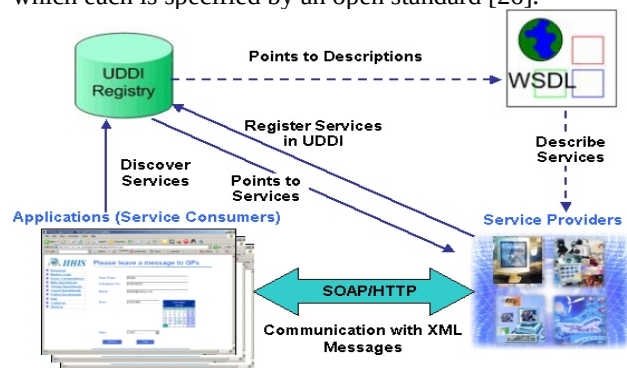
**Need to interoperate with systems on doctors’ desktops [15]:** NHS Care Records Service, as one of services of NpfiT will need to interoperate with systems on doctors’ desktops. This would be challenging enough if the systems were built from scratch. NpfiT must rely on many existing hospital and GP systems being connected to the spine that is a kind of new database system of NHS [15, 20].

**Two separate NHS systems between Scotland and England:** Since Scotland and England are using two different NHS systems [21, 22], both of places are also developing two different Healthcare Information Systems with different IT providers. This situation may lead to the potential system interoperability issue if both NHS systems based in England and Scotland are expected to be integrated for some certain purposes.

Furthermore, under the NHS system, most existing IT systems in trusts are based on either buildings or departments [13]. These systems do not usually support the movement of information between buildings and departments. Consequently, within a single organisation, several records are often created for the same patient [13]. Similarly, in primary care, individual practices have their own IT applications and databases, so patient records are not easily transferred to other practices or care providers [13].

### The Solution of Interoperability – Web Service

Seamless interoperability with applications and data from different systems is a challenging task. Traditionally, solutions of interoperability normally involve developing middleware applications to communicate the non-interoperable applications using the messages, of which the technology is also named as the distributed middleware technology [9]. CORBA and DCOM are two most typical distributed middleware technologies. Nevertheless, CORBA or DCOM is a fairly complex, a task requiring special expertise. Quite often, the achievement of a good interoperability strategy is significantly constrained by many implementation restrictions in CORBA or DCOM [9, 23, 24]. Web services have emerged as the next generation of integration technology [7, 23, 25]. Based on open standards, the Web services technology allows any piece of software to communicate with each other in a standardised XML messaging systems[6, 9, 23]. It solves and eliminates above issues of DCOM/CORBA [23]. As a new type of software service, Web services are modular self-describing, and self-contained applications that can be published, located, and dynamically invoked across the Web. The Web-services technology is built on the foundation of open standards and common infrastructure. The Web-services framework is divided into three areas – communication protocols, service descriptions, and service discovery, of which each is specified by an open standard [26].



**Figure 1. A detailed architecture of Web service**

Figure 1 also indicates how web services works. In general, web services consist of two major technologies (XML – extensible Markup Language and SOAP – Simple Object

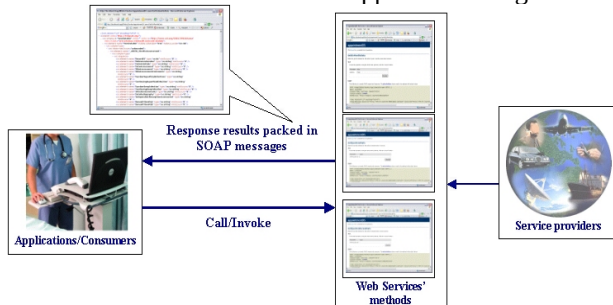




Access Protocol) and two assistant technologies (WSDL – Web Services Description Language and UDDI – Universal Description, Discovery, and Integration).

- Firstly, service providers would make use of WSDL to describe their web services
- Following the above step, service providers would register and publish their services in UDDI
- Applications or service consumers find services via UDDI which would direct service consumers to relevant services according to the description of web services
- Regarding to previous step, applications or service consumers are able to invoke relevant web services using SOAP transmitted via HTTP on the Internet.

Web services encoded in XML, SOAP provides a way to communicate between applications developed with different programming languages and running on different operating systems. In fact, Web services provide a distributed computing technology for integrating applications on the Internet using open standards and XML encoding. The use of standard XML protocols makes Web services platform-, Language- and vendor-independent, thus an ideal solution for use in application integration.



**Figure 2. Interaction between applications/consumers and Web services**

With respect to the above figure, applications send requests and responses to and from Web services via SOAP. When a program invokes a Web-service method, the request and all relevant information are packaged in a SOAP message and sent to the appropriate destination. When the Web service receives the SOAP message, it begins to process the contents (called the SOAP envelop), which specifies the method that the client wishes to execute and the arguments the client is passing to that method. After the Web service receives this request and parses it, the proper method is called with the specified arguments (if there are any), and the response is sent back to the client in another SOAP message. The client parses the response to retrieve the result of the method call.

### **Case Study: Web Service-Based Integrated Healthcare Information System (WSIHIS)**

#### **Background of WSIHIS**

Many people around the world became amputees because of war, traffic accident and some sorts of disease (e.g. Diabetes). Traditionally, these amputees would be treated with conventional socket techniques. With the development of medical technologies, new treatment's techniques for amputees would be based on the Osseointegration that is discovered by Pro. Branenmark.

The first Osseointegration surgery is dental implant, and then he brought this technique further in 1990s [27]. Pro. Branenmark applied the Osseointegration in orthopaedics. In 1997, Branenmark chose Queen Mary hospital as the clinical trail outside of Sweden [27]. Basically, this technique makes use of Titanium implant as the attachment site for the artificial limb attachment. Because this is entire new technique, patients would be selected carefully. A scheme called OPRA (Osseointegration Programme for Rehabilitation Amputee) was developed at Sweden that consists of patients' selection and recruitment, surgical plan after surgeon retirement and rehabilitation. The whole procedure of Osseointegration from the operation to the rehabilitation of patients lasts years. Overall, this process would involve doctors, patients, surgeons, prosthetic clinician and rehabilitation clinician. All relevant information must be recorded and documented for progress review. Patients' files are also required if the infection is developed in later stage. They would require an information system to manage and process these massive health data, as the patient, the hospital and the rehabilitation centre and prosthesis are normally not in same location. The data exchange between them would be very important. Consequently, WSIHIS is proposed to computerise all documents and data, and offer a secure and stable environment for the communication between doctors and patients across the Internet.

#### **Interoperability Issue of WSIHIS**

According to the above statement, it is not difficult to find out WSIHIS would also face the challenging of interoperability issue as other HISs. Principally, interoperability issue of WSIHIS could be concluded as follows:

**System interoperability issue:** Regarding to the description of background, WSIHIS would be required to link with various medical-relevant people (e.g. GP, Surgeon, prosthetic clinician and rehabilitation clinician), who might be using different system on their desktops to access to WSIHIS for different purposes. To achieve this feature, WSIHIS would be required to have the capability to interoperate with those different systems (e.g. Microsoft Windows Systems, Linux Systems, etc). Additionally, some parts of patient records could be produced from some existing systems used by some medical staff (e.g. GP) or even depend on those data from existing healthcare systems. And with the data sharing purpose, WSIHIS would need to interoperate with different DBMSs (e.g. Microsoft SQL server, Oracle, Microsoft Access).

**Language interoperability issue:** WSIHIS is also expected to be integrated with functions of existing and future healthcare systems. Generally, different systems are built up in different technologies, particularly, different programming languages. For example, there are dozens of different patient-record systems in use across the NHS developed by couples of software providers. In order to enable WSIHIS to be compliant with those functions from different systems, WSIHIS is supposed to interoperate with applications built up in different programming languages, such as, Java, Visual Basic, C++ and C#.

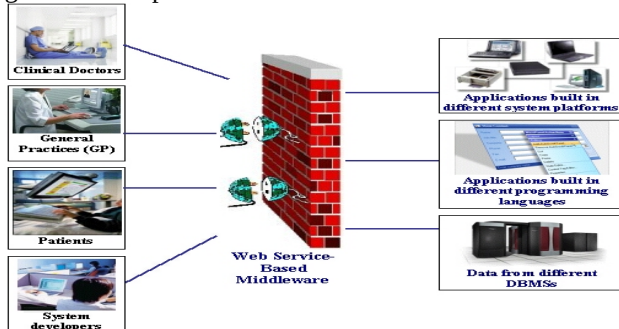
#### **Web Service-Based Solution for Interoperability Issue of WSIHIS**





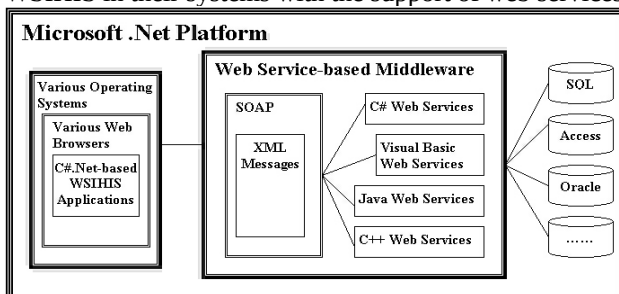


WSIHIS is built up with the support of Web Service to address above interoperability issue. Currently, there are two major application platforms – Microsoft .Net and Java 2 Enterprise Edition (J2EE) being able to create Web services. Microsoft .Net is selected as the major platform for developing the proposed system since Microsoft .Net would be more suitable than J2EE in terms of some features (e.g. overall maturity, interoperability, scalability, cost and the NHS perspective) [5, 28, 29]. Besides, there have also been couples of developers of Healthcare Information Systems expressing the interest to Microsoft .Net for the development of their HIS [2, 5] when Microsoft .Net was not really put into the market. As the result, Microsoft .Net is selected to build up the proposed system. The following figure would illustrate a general concept of Web service-based solution of WSIHIS:



**Figure 3. The general concept of Web service-based solution for the interoperability issue of WSIHIS**

Within WSIHIS, Web service plays as the role of middleware that hides all these differences in system platforms, programming languages and database systems to users and developers. Accordingly, from users' perspective, they would be able to get the access to WSIHIS regardless of their different system platforms. From developers' perspective, they can invoke or reuse applications of WSIHIS in their systems with the support of web services.



**Figure 4. The way of Web service working in WSIHIS with the support of Microsoft .Net technologies**

Figure 3 illustrates a general concept of Web service-based solution for the interoperability of WSIHIS. Figure 4 explains how Web Service and Microsoft .Net technology are applied in WSIHIS. WSIHIS is built upon Microsoft .Net platform, and Web service plays a role of middleware. Web services can be exposed from and consumed by any platform that can format and parse an XML message because using XML for the formatting of requests and responses. This allows XML-based Web services to bring together disparate pieces of functionality – existing or new, internal or external to an organisation – into a coherent whole. Core technologies of Web service are XML and SOAP. Once Web services receive requests from applications, web services would retrieve data from

different DBMSs (e.g. SQL server, Microsoft Access, Oracle, etc.) into datasets according to requirements of applications. All datasets would be written in XML messages, in additions, SOAP would act as an XML envelop to wrap those XML-based datasets into SOAP messages. And then these SOAP messages would be transmitted back to applications via HTTP. SOAP in Web service-based middleware provides a way to communicate between applications developed with different programming languages and running on different operating systems.

Additionally, Microsoft .Net platform is another important part of solution for the interoperability in WSIHIS. It is language neutral. It is best thought of as an open programming platform into which a variety of languages can be plugged. It is achieved by translating all different programming languages into a common language called Intermediary Language (IL). Firstly, source code is translated into Microsoft Intermediate Language (MSIL). This IL code is language-neutral, and is analogous to Java bytecode. The IL code then needs to be interpreted and translated into a native executable. The .NET Framework includes the Common Language Runtime (CLR), analogous to the Java Runtime Environment (JRE), which achieves this goal. The CLR is Microsoft's intermediary between .NET developers' source code and the underlying hardware, and all .NET code ultimately runs within the CLR. Regarding to the announcement of Microsoft company [23], Microsoft .Net platform would also enable the system running under different system platforms and web browser platforms.

#### System Assessment on the interoperability of WSIHIS

The assessment has been done to testify whether Web service and Microsoft .Net technologies could really enable WSIHIS to overcome or eliminate the interoperability issue. The following part would make an analysis and discussion based on the result of assessment. The methodology used for system assessment is shown as the following table:

**Table 1. The methodology of whole system assessment**

| Type of the assessment    | Methodology                                                                             | Based on                                                                                           |
|---------------------------|-----------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------|
| Language Interoperability | Managing to integrate WSIHIS with applications built in different programming languages | Java, Visual Basic, C++, C#                                                                        |
| System Interoperability   | Trying to run WSIHIS on different platforms of web browsers                             | Internet explorer, Netscape, Mozilla Firefox, Deepnet, Avant browser, Opera                        |
|                           | Testifying if WSIHIS is able to work with different DBMSs                               | Microsoft SQL server 2000, Microsoft Access 2000, Oracle                                           |
|                           | Trying to run WSIHIS on different platforms of operating systems                        | Microsoft Windows 2000, Microsoft Windows XP, Ubuntu Linux system 4.02, Knoppix Linux system 5.1.0 |

Regarding to the assessment of language interoperability, four comparable Web service-based applications were built up in different programming languages (C#, Visual Basic, Java, C++). The purpose of assessment is to integrate four applications with WSIHIS to achieve given tasks. The C# application is aiming to validate the reference number entered by the user. With this application, WSIHIS would display the error message for any invalid reference numbers entered by users. The code and resulting screen is shown in Figure 5.





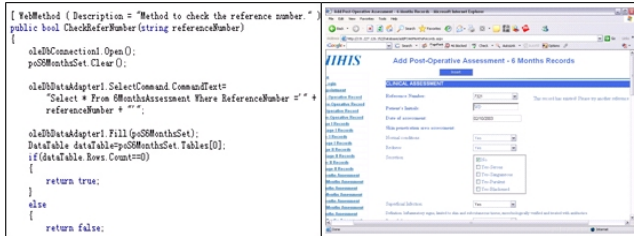


Figure 5. The code and screenshot of application in C#

In WSIHIS, there is a function named as 'Check Appointments' aiming to help doctors search appointments' records made by patients via reference numbers. Nevertheless, if the doctor forgets the reference number of patient, he would be unable to find the appointment record. Regarding to this reason, another function coded in Visual Basic was developed to enable the system display all appointments' records. Figure 6 illustrates the code and screenshot of this application.

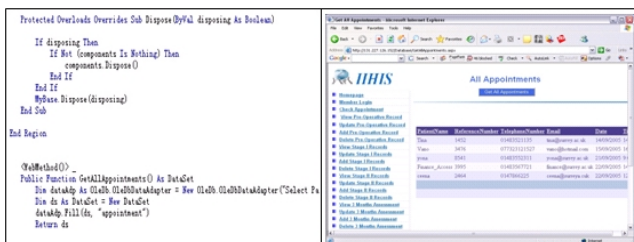


Figure 6. The code and screenshot of application in Visual Basic

Another application is developed in Java that is aiming to assist WSIHIS to find a relevant patient record via a reference number, of which the code and screenshot of application is illustrated in figure 7.

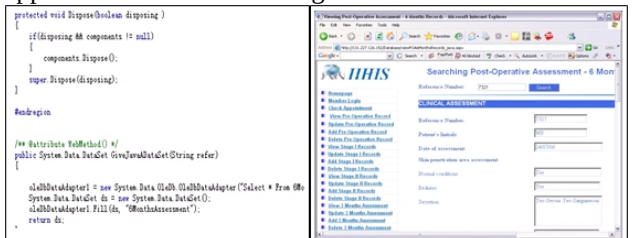


Figure 7. The code and screenshot of application in Java

The function of application in C++ is same as the application in Java that is to retrieve patient records through reference numbers. Its code and resulting screen is shown in Figure 8.

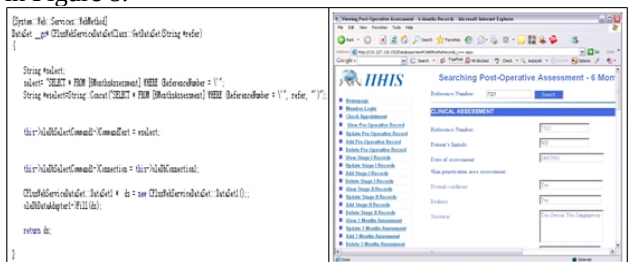


Figure 8. The code and screenshot of application in C++

The assessment of system interoperability comprises of testing the system on different web browsers, different DBMSs and different operating systems. Firstly, this

assessment focuses on testing the interoperability of WSIHIS with six selected web browsers (Internet Explorer, Opera, Mozilla FireFox, Deepnet Browser, Avant Browser and NetScape), on which the author runs WSIHIS to execute specific tasks (e.g. making appointments, changing appointments, etc.) to see if WSIHIS is able to run regardless of different web browsers, of which the result has proved the proposed system has no problems on this point, as illustrated in figure 9.

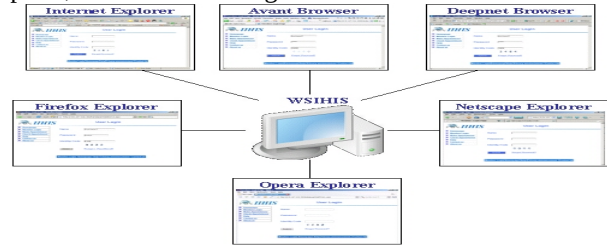


Figure 9. The result of Web Browsers' platform interoperability Testing of WSIHIS

Secondly, the assessment is to testify the ability of WSIHIS for integrating different applications linked with three different DBMSs (Microsoft Access, Microsoft SQL Server, Oracle), and the result of this assessment verified that the system is able to interoperate with applications based on different DBMSs, as shown in figure 10.



Figure 10. The result of DBMS platform interoperability Testing of WSIHIS

Finally, the assessment emphasises on testing the capability of WSIHIS running on different operating systems (Microsoft operating systems and Linux operating systems). WSIHIS was given specific tasks to carry out (e.g. making appointments, searching patients' records). The result of the assessment proved that the proposed system has no problem to work on any different operating systems, as shown in figure 11.



Figure 11. The result of system platform interoperability Testing of WSIHIS

## Discussion and Conclusion

System and language interoperability of HISs is a major concern in this paper. Web services provide an ideal solution for this concern. However, people may question why system and language interoperability is so critical for HISs and why HISs should have the interoperability rather than just stick with the single technology, DBMS and





system platform. Actually, HISs under the NHS in U.K, as an example, are spread over the whole country. Historically, these HISs were developed by different software providers with different programming languages, different system platforms and database systems. Within a single hospital, several records are often created for the same patient; on the other hand, patient records are often located in different database systems. In primary care, individual practices have their own IT applications and databases, so patient records are not easily transferred to other practices or care providers. Today, NHS is developing NpfiT to integrate separated HISs for sharing medical information and functions. The first requirement for this purpose is to enable different HISs to interoperate in a seamless environment regardless of the programming languages, operating systems and database systems. Major benefits from the achievement of interoperability can be concluded from several aspects. Firstly, for existing HISs, medical data and functions can be easily shared and exchanged between different HISs, additionally, systems would be accessible to people using different operating systems. Secondly, for those HISs under the expansion and improvement, new technologies can be easily plugged into original systems. Finally, for those HISs under the development, they can reuse functions from other HISs to reduce the development time scale and cost.

The main contribution of this work is the introduction and use of Web service to achieve system and language interoperability of HISs in a heterogeneous environment. The implementation of WSIHIS has already been carried out based on the background of Osseointegration project in Queen Mary hospital. Within WSIHIS, Web service acts as a distributed middleware to facilitate the interoperability of the whole system with the support of Microsoft .Net technologies. An overall assessment was also executed to test how well WSIHIS to address the interoperability problem with the support of web service and Microsoft .Net technologies. This assessment consists of testing WSIHIS on applications developed in diverse programming languages, system platforms and database systems. Its result shows the application of web service and Microsoft .Net technologies in WSIHIS is very promising and offers more benefits on enhancing the interoperability of system. Another assessment will be carried out based on the wireless environment. This will enable the author to measure how well WSIHIS works and interoperate with other HISs under the wireless environment.

## Reference:

1. DH and NHS, *Connecting For Health - A Public-Private Collaborative*. 2003, Connecting For Health, NHS: England.
2. Matsopoulos, G.K., Kouloulis, V., Asvestas, P., *MITIS: a WWW-based medical system for managing and processing gynecological - obstetrical - radiological data*. Computer Methods and Programs in Biomedicine, 2004. **Volume 76**(Issue 1): p. pp. 53-71.
3. Tsiknakis, M., D.G. Katehakis, and S.C. Orphanoudakis, *An open, component-based information infrastructure for integrated health information networks*. International Journal of Medical Informatics, 2002. **Volume 68**(Issue 1-3): p. pp. 3-26.
4. Wreder, K. and Y. Deng. *Architecture-centered enterprise System Development and Integration Based on Distributed Object Technology Standard*. in *Computer Software and Application Conference*. 1999.
5. Varge, B. and P. Ray. *Interoperability of Hospital Information Systems: A Case Study*. in *Enterprise Networking and Computing in Healthcare Industry, 5<sup>th</sup> International workshop*. 2003.
6. Deitel, H.M., Deitel, P.J., Listfield, J., Nieto, T.R., *C# How to Program - Introducing .Net and Web Service*. 2002, London: Prentic-Hall, Inc.
7. Chester, T.M., *Cross-Platform Integration with XML and SOAP*. IT Professional, 2005. **Volume 9**(Issue 4): p. Page 67-70.
8. Duthie, G.A., *Microsoft ASP.Net Programming With Microsoft Visual C#.Net Step by Step*. 2003, Washington, USA: Microsoft Press.
9. Zhu, J., *Web Services Provide the Power to Integrate*. Power and Energy Magazine, 2003. **Volume 1**(Issue 6): p. Pages 40-49.
10. Pronab, G. and R. Pradeep, *Software Interoperability of Telemedicine Systems: A CSCW Perspective*. IEEE, 2000.
11. Maglaveras, N. and I. Chouvarda, *The Citizen health system (CHS): A Modular medical Contact Center Providing Quality Telemedicine Services*. IEEE Transactions on Information Technology In Biomedicine, 2005. **Volume 9**(Issue 3).
12. Jablonski, S., Lay, R., Meiler, C., *Data Logistics as Means of Integration in healthcare Applications*. ACM Symposium on Applied computing, 2005. **Volume: 1-58113-964**.
13. NpfiT, *Making IT happen - Information about the National Programme for IT*. 2005, Department of Health: England.
14. James, A.E. and Y.H. Wilcox, *A telematic system for oncology based electronic health patient records*. Information Technology in Biomedicine, 2001. **Volume 5**(Issue 1): p. pp 16-17.
15. Cross, M., *In Sickness or In Health?* IEE Review, 2004. **Volume 50**(Issue 10).
16. Albahari, B., Drayton, P., Merrill, B., *C# Essential - A Comparative Overview of C#*. First Edition ed. 2001: O'REILLY.
17. NHS, *History of Connecting for Health*. 2005, NHS Connecting for Health, Department of NHS: England.
18. Guah, M.W.C., W.L. *Logicity of ASP in healthcare: the NHS case study*. in *Proceedings of the 37th Annual Hawaii International Conference*. 2003, Hawaii: IEEE.
19. NHS, *Electronic Transmission of Prescriptions (ETP) Programme*. 2005, NHS Connecting for Health: England.
20. DH and NpfiT, *Choose and Book Service Implementation Guide*. 2005, NHS Connecting for Health: England.
21. Mathews, A.G., Butler, R. *A Vision for the Use of Proactive Mobile Computing Tools to Empower People with Chronic Conditions*. in *Proceedings of the 18th IEEE Symposium on Computer-Based Medical Systems (CBMS'05)*. 2005: IEEE.
22. DH and NHS, *NHS Connecting for Health Fact Sheet*. 2005, Connecting for Health, NHS: England.
23. Banerjee, A., Corera, A., *C# Web ServiceS - Building Web Services With .NET Remoting and ASP.NET*. 2002: WROX.
24. Gokhale, A., Kumar, B., Sahuguet, A. *Reinventing the Wheel? CORBA vs. Web Services*. in *The Eleventh International World Wide Web Conference*. 2002, Honolulu, Hawaii, USA.
25. Umar, A. *The Emerging Role of the Web for Enterprise Applications and ASPs*. in *Proceedings of The IEEE*. 2004: IEEE.
26. Curbera, F., M. Duftler, and R. Khalaf, *Unraveling the Web Services Web - An Introduction to SOAP, WSDL, and UDDI*. IEEE Internet Computing, 2002.
27. Sullivan, J., Uden, M., Robinson, K. P., *Rehabilitation of the trans-femoral amputee with an osseointegrated prosthesis: the United Kingdom experience*. Prosthetics and Orthotics International, 2003.
28. Sessions, R., *Java 2 Enterprise Edition versus The .Net Platform Two Visions for eBusiness*. 2001, ObjectWatch, Inc.: Texas.
29. Vawter, C., Roman, E., *J2EE vs. Microsoft .Net - A comparison of building XML-based web services*. 2001, The Middleware Company.



## Analysis of Misheard Phrases in Nursing Dialogue

Kaoru Sagara<sup>(a)(b)</sup> Hiromi Itoh Ozaku<sup>(b)</sup>  
Akinori Abe<sup>(b)</sup> Futoshi Naya<sup>(b)</sup> Noriaki Kuwahara<sup>(b)</sup> Kiyoshi Kogure<sup>(b)</sup>

(a) Seinan Jo Gakuin University Fukuoka, JAPAN

(b) ATR Knowledge Science Laboratories, Kyoto, JAPAN

sagara@seinan-jo.ac.jp

{romi, ave, naya, kuwahara, kogure}@atr.jp

### Abstract

*We are now collecting, transcribing and analyzing conversation data in medical and nursing facilities. This is part of nursing assistance research being conducted based on ubiquitous computing technologies. In this paper, we focus on communication error analysis. We analyzed 30 hours and 45 minutes of transcribed voice data collected in a hospital. Incorrectly transcribed phrases are analyzed to find features of incorrect transcription.*

### 1. Introduction

In order to reduce accidents and incidents in nursing and to improve the quality and effectiveness of treatment, it is necessary to observe nursing activities and situations and to understand their actual circumstances.

Nanbu et al. [1] pointed out that “Without any communication between various types of staff members, no medical practice can be completed. In addition, ways of information transmission and sharing have close relationship between accidents and incidents. Therefore, they can be important keys to keep safety in medical situations.” A medical treatment evaluation section in the Ministry of Health, Labor and Welfare has collected examples of accidents and incidents from many hospitals in Japan and selected significant examples to upload to their Web site (<http://www.hiyari-hatto.jp/>). In addition, they concluded that:

- 1) It is not easy to propose effective strategies to solve problems in communication among staff members in hospitals, even if we collected a large amount of examples from many hospitals in Japan for analysis.
- 2) In order to fully understand accident examples from the viewpoint of communication among staff members in hospitals, it is necessary to comprehend actual communication situations in hospitals. If we conducted such analyses, it would be possible to develop effective strategies for staff members in hospitals for preventing accidents and incidents. Such strategies could be established not for individual events but for general events. In addition,

communication among staff members in hospitals would be improved, thus helping to maintain safe medical practices.

Kohno[2] pointed out that “In medical places, for information communication media, they use human communication media (e.g. oral communication). Consequently, due to less reliable (forgettable and changeable) human memories and relationships, human errors easily occur.”

We analyzed 221 examples uploaded to the accident and incident Web site (September 2005). We could collect 61 examples (22.6%) involving “conference,” “communication,” “intention transfer,” “handover,” and “oral order.” From the results, we concluded that communication errors could cause accidents and incidents that are classified as human errors. In this paper, therefore, we analyze nurses’ conversation data collected by a wearable recording system.

### 2. Methods

#### 2.1. Voice data collection

**Collection time:** From 13 Dec. 2004 to 18 Dec. 2004. In detail, voice data sets (30 hours and 45 minutes) were collected over those 5 days. They include voice data sets spoken during all of the nurses’ duty time (three shifts).

**Subjects:** Fourteen nurses who have been working from 1 year to 11 years.

**Venues:** The Ophthalmology Department, the Otorhinolaryngology Department, and the Internal Medicine Department of a certain hospital.

**Procedure:** Nurses wore the wearable recording devices shown in Fig. 1. Voice data during “handover,” “oral order,” and “conference” were collected.

**Ethical considerations:** Objectives of the experiment were explained to the subjects. The experiment was conducted with the assistance of those who understood and agreed with the objectives. In addition, we guaranteed that individual information would not become open to the public.





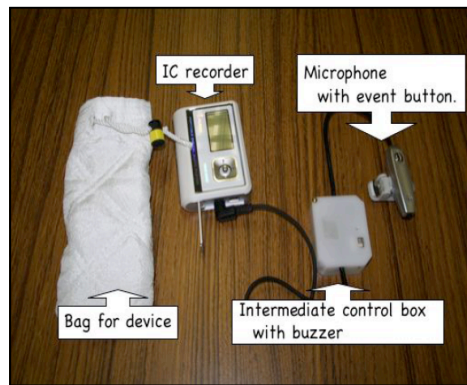


Figure 1. Voice recorder

## 2.2. Voice data collection

- Sets of collected voice data were transcribed by expert transcribers. Then text data sets (12,764 lines) containing unrecognized words were obtained.
- The transcribed text sets were checked by two nurses who had worked more than three years. In addition, the transcribed text sets were checked again by two other nurses who had worked more than ten years. A total of 952 phrases (358 distinct types) were extracted as incorrectly transcribed words.
- The 358 extracted phrases were categorized as follows:
  - 1) Phrases including numerical values
  - 2) Phrases including alphabetic characters
  - 3) Phrases including katakana characters
  - 4) Others
- From the 358 phrases, 319 words excluding numerical values and units were extracted as abbreviations and synthesized words. These were compared with words included in Bunrui Goihyo[3], a reference that includes 100,000 everyday words and 200,000 terminologies for medical and nursing science collected for research (nursing terminology set).

## 3. Results

### 3.1. Analysis of incorrectly transcribed phrases according to forms of characters

As shown above, 358 phrases were incorrectly transcribed. Among them, 16 phrases included numerical values, and in 9 of these the numerical values could not be correctly transcribed (e.g. 眼圧 14 → 眼圧 54). Seventeen phrases included alphabetical words, ten of which were transcribed in the form of alphabetic characters. Among them, only three alphabetical phrases were correctly transcribed. Four phrases that should have been represented in alphabetical words were transcribed in katakana (IVH → アデチ). For one phrase, Chinese characters were used instead of alphabetical characters

(A1C → えー足). Two phrases could not be fully transcribed (HCV → \*ブ).

Table 1. Incorrectly recognized words including numerical values.

| Correct words                   | Incorrect words              |
|---------------------------------|------------------------------|
| 72 の<br>(nanaju ni no)          | 77<br>(nanaju nana)          |
| 眼圧が 14<br>(Gan-ats ga ju yon)   | 眼圧 54<br>(Gan-atsu goju yon) |
| 14 時<br>(ju-yo ji)              | 治療時<br>(Chiryō ji)           |
| 2, 4 分 2<br>(ni, yon bun ni)    | リーヨンブンリ<br>(ri-yon bunri)    |
| 50 ミリ 2 アン<br>(goju miri ni an) | 50 メニア<br>(goju menia)       |
| 1 万円<br>(ichi man en)           | 1 枚<br>(ichi mai)            |
| 1 包<br>(ippou)                  | 1 本<br>(ippon)               |
| 2 直帯<br>(ni chyokutai)          | 2 直隊<br>(ni chyokutai)       |
| 4 検<br>(yon ken)                | ヨンケン (4 件?)<br>(yon ken)     |
| タ R6<br>(UR roku)               | ユーアール (UR) 6<br>(UR roku)    |

Table 2. Incorrectly recognized words including alphabetic characters.

| Correct words                     | Incorrect words                     |
|-----------------------------------|-------------------------------------|
| CPR 前<br>(CPR zen)                | CPR 全<br>(CPR zen)                  |
| CS 入院<br>(CS nyu-in)              | CS にー<br>(CS nii-)                  |
| CT 下の肺生検<br>(CT ka no hai seiken) | CT から肺セウキョウ<br>(CT kara hai seikyō) |
| DM                                | AM                                  |
| PEA                               | PA                                  |
| SMBG                              | SMVG                                |
| SOAP                              | SOP                                 |
| BMI                               | DMI                                 |
| E3<br>(E san)                     | インサン<br>(in san)                    |
| IVH                               | アデチ<br>(a de chi)                   |
| MRSA                              | エマーリセー<br>(ema- rise-)              |
| V line                            | ドエライン<br>(do erain)                 |
| A1C                               | えー足<br>(e- ashi)                    |
| HCV                               | *ブ<br>(...bu)                       |

Incorrect transcriptions would occur due to low-quality recordings. However, even if we could obtain the best quality recorded sound, incorrect transcriptions might still occur. If a transcriber does not know specialized words or does not have background knowledge in the application field, incorrect transcriptions tend to occur. Furthermore, in the case of “眼圧が 14 → 眼圧 54”, “ga ju” was heard as “go ju.”





This type of incorrect transcription cannot be avoided even if a transcriber has sufficient knowledge in the application field, since it is sometimes difficult to distinguish “a” from “o” due to similar pronunciation.

**Table 3. Incorrectly recognized words including katakana characters.**

| Correct words                  | Incorrect words               |
|--------------------------------|-------------------------------|
| モーラスパップ<br>(mo-rasu pappu)     | もう、ラスカップ<br>(mou rasukappu)   |
| ラキシベロン<br>(Rakisoberon)        | だとベロン<br>(dakedo beron)       |
| ヘアバスとベベ<br>(hair bath to bebe) | 部屋バストベで<br>(heya basutobe de) |
| アップしている<br>(appu shiteiru)     | 安定している<br>(antei shiteiru)    |
| タンポン抜去<br>(tanpon bakkyo)      | 担当バック<br>(tantou back)        |
| ワーファリン<br>(Warfarin)           | バファリン<br>(Bufferin)           |

On the other hand, in the case of “V line → ドエライン”, if a transcriber had sufficient knowledge in the application field, such incorrect transcriptions could be avoided.

Considering the case of “2, 4 分 2 → リーヨンブンリ”, the recordings themselves do not seem acceptable. This is because “n” is heard as “r”. Of course, in actual situations, “n” is sometimes heard as “r” if we do not have sufficient knowledge. Therefore, we can easily understand that such incorrect transcriptions can occur if a transcriber does not know the application field.

Among incorrectly transcribed phrases, 111 phrases include katakana. Forty-one of them should be medicine names, but transcribers could not correctly recognize them. On the contrary, in some cases, although phrases or words could be correctly recognized as phrases or words (in the same category), their names are different (ワーファリン → バファリン). In this case, a transcriber seemed to recognize it as a more familiar medicine. If he/she has medical knowledge, such an incorrect transcription should have been avoided, since only one medicine should be applied for a particular medical condition.

### 3.2. Analysis of phrases according to features of included words

Among the 319 phrases that were incorrectly transcribed, 89 words are included in Bunrui Goiho and can be regarded as everyday words. Since they are everyday words, perhaps due to missing context, transcribers seem to make mistakes. In addition, pronunciations between correct words and incorrect words are similar. Thus, this type of incorrect transcription would easily occur (Table 4).

**Table 4. Incorrectly recognized words regarded as everyday words.**

| Correct words      | Incorrect words    |
|--------------------|--------------------|
| 転倒<br>(tentou)     | 点灯？<br>(tentou)    |
| 床上<br>(syoujyou)   | 症状<br>(syoujyou)   |
| 予約<br>(yoyaku)     | 与薬<br>(yoyaku)     |
| 既往<br>(kiou)       | 今日<br>(kyou)       |
| 転科<br>(tenka)      | 検査<br>(kensa)      |
| 前後に<br>(zengo ni)  | 全部に<br>(zenbu ni)  |
| 次回<br>(jikai)      | 時間<br>(kikan)      |
| 病名<br>(byoumei)    | 両目<br>(ryoume)     |
| 気切部<br>(kisetu bu) | 気節部<br>(kisetu bu) |
| 腰痛<br>(youtsu)     | 尿痛<br>(nyoutsu)    |
| 絶食<br>(zessyoku)   | 自食<br>(jisyoku)    |

**Table 5. Incorrectly recognized words regarded as nursing terminology.**

| Correct words            | Incorrect words          |
|--------------------------|--------------------------|
| オペ前<br>(ope zen)         | ブレゼン<br>(pre zen)        |
| ギャッチアップ<br>(Gyacchi up)  | キャッチアップ<br>(Catch up)    |
| 綿球<br>(men kyu-)         | ネンキュウ<br>(nen kyu-)      |
| 輸液<br>(yu e ki)          | デーキ<br>(de-ki)           |
| 眠前<br>(min zen)          | インセン<br>(in sen)         |
| 滲出性<br>(shinsyutsu sei)  | シンシツ性<br>(shinshitu sei) |
| 血糖<br>(kettou)           | きつと<br>(kitto)           |
| 入室時<br>(nyu- shitsu ji)  | 入出時<br>(nyu-syutsu ji)   |
| 蓄尿<br>(tiku nyou)        | 蓄膿<br>(tiku nou)         |
| 舌苔<br>(zettai)           | 絶対 (全体？)<br>(zettai)     |
| 搔痒感<br>(sou you kan)     | ソウヨカン<br>(sou you kan)   |
| 霧視<br>(mushi)            | ムシ (無視？)<br>(mushi)      |
| 保清<br>(hosei)            | 補正<br>(hosei)            |
| 床上安静<br>(syoujyou ansei) | 症状安静<br>(syoujyou ansei) |

As for 230 phrases that do not include words in Bunrui Goiho, 135 phrases included words in the nursing







terminology set. In addition, 53 words were expressed in katakana (Table 5).

**Table 6. Incorrectly recognized words including particles.**

| Correct words        | Incorrect words     |
|----------------------|---------------------|
| 押して                  | をして                 |
| (oshite)             | (wo shite)          |
| リンコデ                 | リンコで                |
| (Rinkode)            | (rinko de)          |
| シックデイ                | シックで                |
| (sick day)           | (sick de)           |
| 切開排膿                 | 切開ハイの               |
| (sekkai hainou)      | (sekkai hai no)     |
| 肝機能障害                | カンキの障害              |
| (kankinou syo-gai)   | (kanki no shougai)  |
| 眼圧が 14               | 眼圧 54               |
| (Gan-atsu ga ju yon) | (Gan-atsu goju yon) |

Other phenomena involve the recognition of part of a word as a particle (押して → をして) and not recognizing particles that should exist (Table 6).

We analyzed 95 phrases that do not include words in the nursing terminology set. Nine words were abbreviations of medicine names, four words were proper nouns, such as the names of companies or products, and 63 words were spoken language (jargon) or abbreviations specialized in nursing activities (Table 7).

**Table 7. Incorrectly recognized words.**

| category                        | Variety<br>(ratio) |
|---------------------------------|--------------------|
| nursing terminology             | 135<br>(42.3%)     |
| everyday words                  | 89<br>(27.9%)      |
| nursing jargon and abbreviation | 63<br>(19.7%)      |
| name of medication              | 9<br>(2.8%)        |
| name of vendor or product       | 4<br>(1.3%)        |
| others                          | 19<br>(6.0%)       |
| Incorrectly recognized words    | 319<br>(100%)      |

Table 7 shows the classification of 319 incorrectly transcribed phrases, from which 39 phrases including numerical values and units were removed. Fully 62% of them were categorized as nursing terminology, nursing jargon, and abbreviations. This, we can assume, would occur because people who had never worked as a nurse transcribed the voice data. Furthermore, 19.7% of incorrectly transcribed words were nursing jargon and abbreviations. Therefore, we can suppose that nurses' conversations include many nursing jargons and abbreviations specialized to the particular hospital that are not included in ordinary nursing terminology

dictionaries. A nurse who frequently uses phrases including such nursing jargon and abbreviations, even an experienced nurse, might mishear or misunderstand a phrase if he/she has just changed his/her working place.

## 4. Discussions

The data shown in this paper are included in nursing conversation during actual nursing activities. Therefore, they have the following features:

- 1) Differently experienced nurses converse together.
- 2) Sometimes, a conversation includes more than one topic.
- 3) Nurses converse in noisy environments due to other conversations and their ongoing activities.

It is not easy to separate each topic in a conversation or to distinguish one speaker from the other ones in such non-ideal situations.

In this study, transcriptions were made by transcribers having the following features:

- 1) Those who could identify topics of conversations and distinguish more than one speaker when only listening to voice data.
- 2) Those who had neither nursing experience nor knowledge of nursing practices.
- 3) Those who knew conversations were made by nurses in nurse stations of the otolaryngology department.

However, as they became engaged in the transcription work, the following features could be observed:

- 1) Their knowledge of nursing practices gradually increased because they carefully listened to nursing conversations.
- 2) Since they used computers for editing with an automatic kana-to-Chinese character transformation system, their knowledge of words that have the same sound but different meaning gradually increased.

The above development of transcribers might be similar to knowledge acquisition by novice nurses. Therefore, analyses of incorrect transcriptions could help us to understand the tendencies of communication errors by novice nurses. In the following, we briefly discuss such tendencies.

### 4.1. Description features

In the case of incorrect transcriptions of phrases including numerical values (Table 1), the following errors were observed:

- 1) Numerical numbers are incorrect.
- 2) Units are incorrect.





Reasons for (1) could be (a) environmental problems such as noisy situations and (b) insufficient knowledge of the proper range of each value. A reason for (2) could be insufficient knowledge of the units used in nursing practices.

As for incorrect transcriptions of phrases including alphabetic characters (Table 2), in a very few cases, transcribers used Chinese characters. They usually use katakana to express objects. Consequently, they seem to recognize that those objects should be alphabetical abbreviations of objects or concept. Therefore, if they had sufficient knowledge of the abbreviations used in nursing practices, incorrect transcriptions would be reduced.

On the other hand, as for incorrect transcriptions of phrases including katakana, many of them were incorrectly separated (Table 3). Expressions in katakana are frequently used not only for medicine names but also for nursing practice names[4]. In addition, expressions in Katakana have problems such as multiple meanings and expressions [5]. As a result, even if those expressions could be correctly heard and transcribed, several communication errors might occur. In addition, in some cases, transcriptions were influenced by the transcribers' own knowledge. For instance, in Table 3, we can see an example transcribed as “バファリン (Bufferin)” instead of “ワーファリン (Warfarin)”. This is because Bufferin is a more famous medicine than Warfarin to laypeople, and thus the transcriber incorrectly recognizes the former after deciding the phrase refers to the name of a medicine.

Therefore, in order to reduce communication errors with expressions in katakana used in hospitals, preparation of a terminology list would be effective. In addition, recording communication results in documents might also be effective for later confirmation.

## 4.2. Specialized and general terminologies

In speech interpretation, we usually use not only grammatical knowledge but also logical, semantical, and contextual implications to reduce ambiguities [6].

Table 4 suggests interesting phenomena on how a normal person identifies rarely heard concepts or objects that have the same or similar pronunciation as general words and thus recognizes them as general words. For instance, 既往 (anamnesis), 転科 (get enrolled in another department), and 床上 (nursery bed soil) can be easily understood if we see them in writing. However, those words are rarely heard in everyday life. Transcribers know that the conversation was done by nurses in the otolaryngology department but guess that the contents do not involve specialized events. As a result, they adopted 今日 (today), 検査 (check), and 症状 (symptom) instead of 既往, 転科, and 床上. Those words are shown as candidates by an automatic kana-to-Chinese character transformation system.

In Japanese, 部 means “part,” 痛 means “pain,” and

食 means “food.” Consequently, we can easily synthesize new word by coupling those words. As for 気切部 (tracheotomy hook), 腰痛 (lumbago), and 絶食 (abrosia), since transcribers guess that those phrases describe part of something, pain in certain parts, and something related to food, they adopt coinages such as 気節部 (coined terms), 尿痛 (urodynia), and 自食 (coined terms) that cannot be easily transformed (Table 4). 腰痛 (lumbago) and 絶食 (fast) are very common words in Japanese. However, they might think that 腰痛 and 絶食 are too common to use in hospitals, so they seemed to use uncommon words such as 尿痛 (urodynia) and 自食. In addition, coinages such as \*\*性 (\*\* feature) and \*\*時 (\*\* time) are also used with certain confidence for transcription (Table 5).

We can observe that they used 与薬 (administration) instead of 予約 (reserve). We assume that transcribers excessively considered the context, i.e. a hospital, and used medical words instead of general words.

Table 5 shows results including nursing terminology. Since, for the transcribers, such words are hardly ever heard, they used katakana to express unfamiliar concepts or objects. Some words such as 保清 (maintenance of cleanness), 床上, and 霧視 (blurred vision) are transcribed as 補正 (correction), 症状, and 無視 (neglect). On the other hand, for 搔痒感 (itching paraesthesia), which is not such an unfamiliar word but uses a difficult Chinese character, they transcribe as ソウヨカン. For 綿球 (swab), they express it as ネンキユウ. The reason for using katakana for 搔痒感 (itching paraesthesia) might be that 搔痒 (pruritus) is not an easy Chinese character. However, we cannot guess why they use ネンキユウ for 綿球 (swab), which does not contain such a difficult Chinese character. If we collect more examples, we might be able to find certain tendencies explaining these phenomena.

Table 6 shows examples of recognizing parts of phrases as particles or failing to recognize parts of phrases. Due to different intonation or accents in particles, parts of phrases might be recognized as particles.

## 4.3. Limitations

We did not analyze the voice data as they were collected but the text data transcribed at first by expert transcribers who have neither nursing experience nor knowledge. Therefore, 358 incorrectly transcribed phrases should include those that can be easily recognized by a nurse working in a hospital, the nurse who took part in the conversation, and a person who has nursing knowledge. On the contrary, the transcribed data also include phrases that cannot be easily evaluated by even a person who has nursing knowledge, if he/she did not take part in the conversation. In addition, as explained in Section 4, since it is not easy to separate each topic in a conversation and distinguish each speaker,



achieving fully correct transcriptions by nurses who are not expert transcribers is also difficult. Therefore, despite the approach taken, it is difficult to correctly analyze actual nursing conversations in hospitals.

## 5. Conclusions

Kohno [2] pointed out that the main reasons for communication errors in medical places were as follows: 1) Medical procedure orders are different depending on the medical ward, 2) Abbreviations are different depending on diagnosis and treatment departments, 3) There is no consistency in descriptions for medical matters, 4) There are problems in communication among specialists, 5) Various types of medicine, apparatus, and papers are placed in a small nurse station, and 6) There are incomplete education programs for new employees and incomplete regular medical competency tests for doctors and nurses. Reasons for communication problems among specialists include not only the difference between their experience and knowledge but also the difference between their culture and dialect. In addition, human relationships, such as preference and hierarchy, cause such communication problems.

In this paper, we do not consider such human relationships. Analyses are only based on transcribers' transcriptional features of voice data. However, we could observe that when they could not catch words or phrases for various reasons, they assigned words or phrases to the uncertain sounds based on their personal interpretations. Accordingly, when human factors are

added to the communication line, communication errors are likely to occur.

## Acknowledgment

This research was supported in part by the National Institute of Information and Communications Technology (NICT).

## References

- [1] Nanbu M., Harada E., Sudou T., Shigemori M., and Uchida K., "Risk sharing communication in medical places", *Japanese Cognitive Science*, 13(1), 2006, 62--79. in Japanese
- [2] Kohno R., "Human errors in medical treatments --- Why do mistreatments occur? How to avoid them? ", Igakusyoin, 2004. in Japanese
- [3] The National Institute for Japanese Language eds., "Bunrui Goiho --- Enlarged and revised edition---, Version 1 (CD-ROM Ver.)", Dainihon Tosyo, 2004. in Japanese
- [4] Sagara K., "Features of Standardized Nursing Terminology Sets", *JAMI-NS: Kango Jouhou Kenkyukai Ronbunshuu*, 2005, 73--75. in Japanese
- [5] Sagara K., "The Syntactic Structure of Sentences in Nurse's Record", *JAMI-NS: Kango Jouhou Kenkyukai Ronbunshuu*, 2004, 85--88. in Japanese
- [6] Takubo Y., Nishiyama Y., Mitou H., Kameyama M., and Katagiri Y., "Discourse and context," *Linguistic Science*, Vol. 7, Iwanami Shoten, 2004. in Japanese



## Context-Based Retrieval System for Similar Medical Documents

Kazuya OKAMOTO  
Graduate School of  
Informatics, Kyoto University,  
Kyoto, Japan.  
[kazuya@kuhp.kyoto-u.ac.jp](mailto:kazuya@kuhp.kyoto-u.ac.jp)

Tadamasa TAKEMURA  
Department of Medical  
Informatics, Kyoto University  
Hospital, Kyoto, Japan.  
[takemura@kuhp.kyoto-u.ac.jp](mailto:takemura@kuhp.kyoto-u.ac.jp)

Tomohiro KURODA  
Department of Medical  
Informatics, Kyoto University  
Hospital, Kyoto, Japan.  
[Tomohiro.Kuroda@kuhp.kyoto-u.ac.jp](mailto:Tomohiro.Kuroda@kuhp.kyoto-u.ac.jp)

Hiroyuki YOSHIHARA  
Department of Medical  
Informatics, Kyoto University  
Hospital, Kyoto, Japan.  
[lob@kuhp.kyoto-u.ac.jp](mailto:lob@kuhp.kyoto-u.ac.jp)

### Abstract

*This paper proposed a context-based similarity measurement model for retrieving discharge summaries and tried to retrieve similar discharge summaries. A discharge summary consists of sentences that have an attribute, one of clinical cycle consisting of "observation", "diagnosis" and "therapy". The model understands a discharge summary as continuation of "observation", "diagnosis" and "therapy", and measures similarities based on the context between the documents. This paper evaluated the model in two retrieval processes, a matching process and a ranking process. There was not so much difference in a matching process between the model and a vector space model that measures similarities by a number of kinds of words included in two compared documents. In contrast, the proposed model was superior to the vector space model in a ranking process. This paper concluded that the context-based model should be adapted with the vector space model executed fast for effective retrieval system.*

### 1. Introduction

The use of electronic patients' record systems has been realized in many hospitals. Such systems make it easier to retrieve information than referencing paper records and medical documents have been pooling.

With the pooling of medical documents, Natural Language Processing has been getting important to analyze the pooled documents in order to produce fruitful knowledge and findings.

One of fruitful adaptations of Natural Language Processing to medical documents is document retrieval system to find out patient records of similar phenomena to a patient under treatment. Clinicians can refer previous patients with similar syndromes by giving a patient record of the patient under treatment as a query

for the system. Browsing similar records, doctors may confirm his medical process to go through, and residents can find overlooked issues.

However similar document retrieval system does not work properly in general. In informatics, relevance feedback was tried to improving retrieval performance [1] and utilizing metadata or annotations as human knowledge has been tried to make an improvement [2]. In contrast, full use of characteristics of medical documents may result in efficient similar document retrieval system for medical documents.

In this article, we exploit characteristics of medical documents for efficient similar document retrieval system. The characteristics are as follows.

- Medical documents mostly consist of medical descriptions.
- Medical documents have clinical cycles.

The first feature gives certain limitation of vocabulary to explain a certain phenomena. Thus, introduction of medical term dictionaries should make similarity analysis easier on the contrast of general documents. So we introduce a medical term dictionary that we developed. Additionally, the first feature gives medical documents a certain manner of writing reflecting clinical process itself. As shown in second feature, clinical process is performed in cycle-wise, the document itself has a certain resemblance to basic clinical cycle. So, we introduce analysis to find out "clinical cycle" from given medical documents to make resemblance more clearly. The clinical cycle is comprised of three attributes: "observation", "diagnosis" and "therapy" [3]. A doctor makes observation of his patient and he uses the observation to make diagnosis. He provides therapy according to his diagnosis. And his therapy leads to next observation. Therefore, similarities of documents are considered related to context of medical documents. We propose a context-based similarity measurement model for retrieving similar medical documents. The proposed model classifies





segments of medical documents into observation, diagnosis and therapy exploiting a statistical learning technique, a SVM (a Support Vector Machine), and a statistical measure, TF (Term Frequency) / IDF (Inverse Document Frequency). And the proposed model finds context of each document based on the classifications and measures context-based similarities between documents.

The proposed model is evaluated in two experiments because similarities between the documents are referred to in a matching retrieval process and a ranking retrieval process. The experiments compare the proposed model and a vector space model that is used generally to measure similarities or is exploited to classify documents [4]. The results validate effectiveness of similar medical document retrieving system that exploits the proposed model.

## 2. Materials

### 2.1. Retrieval medical documents

Retrieval medical documents are medical history data and clinical process data of discharge summaries described by MML (Medical Markup Language) [5], which is an XML-based language and a medical information exchange format. Discharge summaries are for doctors to share patient medical records with other doctors or co-medical staff. Discharge summaries must be summarized frankly and briefly to share essential knowledge. Because of this characteristic, discharge summaries may be regarded as knowledge source. Figure 1 shows an example of a discharge summary.

In this article, retrieval medical documents are composed of `<mmlSm:history>` items and `<mmlSm:clinicalCourse>` items, which are medical history data and clinical process data respectively.

Medical documents are not always described by MML. For example, HL7 [6] is one of the other representative XML-based languages to describe medical records. But diverse data such as medical history data and clinical process data should be described by natural language as elements of each XML-based language because of its diversity. Therefore, the methods in this article do not depend on MML.

### 2.2. Japanese morphologic analysis

Word boundaries are not clear in non-segmented languages such as Japanese and Chinese. In such languages, the morphological analyzer of the corresponding non-segmented language must identify word segmentation.

To identify word segmentation in Japanese, a Japanese morphological analyzer needs dictionaries for guidewords and parts of speech. Furthermore, since medical documents have a lot of medical terms, the analyzer must use medical dictionaries to deal with medical documents. We add medical dictionaries [7] that

have extracted medical terms from some kinds of medical documents such as the Merck Manuals [8] to ChaSen [9], the major Japanese morphological analyzer.

ChaSen exploits dictionaries for guidewords and parts of speech and Japanese syntax. For example, nouns can follow adjectives. In actuality, it is difficult to identify word segmentation because of complexity of Japanese.

```
</mmlSm:patientProfile>
- <mmlSm:history>
  20代より動悸を自覚し、不整脈を指摘されたが加療は行っていなかった。
  <xhtml:br />
  2002年8月3日右片麻痺、左半身の痛覚低下、嚥下障害を呈した脳梗塞(延髄外側症候群)で入院。入院時心電図にて心房粗動を認め自然に洞調律へ回復した。また心原性塞栓の鑑別のため経食道心エコーがなされ卵円孔開存を指摘された。歩行器歩行可能となり同年8月23日リハビリのため転院した。以後後遺症は全快し自宅退院。
  <xhtml:br />
  ...
</mmlSm:history>
- <mmlSm:physicalExam>
  【身体所見】身長178.0cm,体重,BMI,体温37.2℃,脈拍数100/分,呼吸数16/分,血圧(右)132/81(左)/mmHg
  <xhtml:br />
  【輸血情報】輸血なし
  <xhtml:br />
</mmlSm:physicalExam>
- <mmlSm:clinicalCourse>
  <mmlSm:clinicalRecord> 4月7日入院後心電図モニタリングを行い経過観察。動悸やモニター上頻拍発作なく経過。血液検査所見も異常なく4月8日Holter心電図評価でもATなど頻脈発作の記録は無かった。当院不整脈外来でEPS, ablationについて検討予定とし4月9日退院となった。
  </mmlSm:clinicalRecord>
</mmlSm:clinicalCourse>
<mmlSm:dischargeFindings> シネノール300mg3×にて洞調
```

Figure 1. An example of a discharge summary

### 2.3. A SVM

A SVM is a statistical learning technique. Boser, Guyon and Vapnik proposed a SVM [10]. A SVM finds an optimal hyperplane separating training samples each of which is in vector space and has a positive or negative class.

Let us define training data that can be separated (Figure 2). Training samples is  $(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_s, y_s)$ .  $\mathbf{x}_i$  is a feature vector of the  $i$ -th sample and an  $N$  dimensional vector.  $y_i$  is the class label of the  $i$ -th sample and positive (+1) or negative (-1).  $s$  is the number of the training samples. A hyperplane is shown as follows.

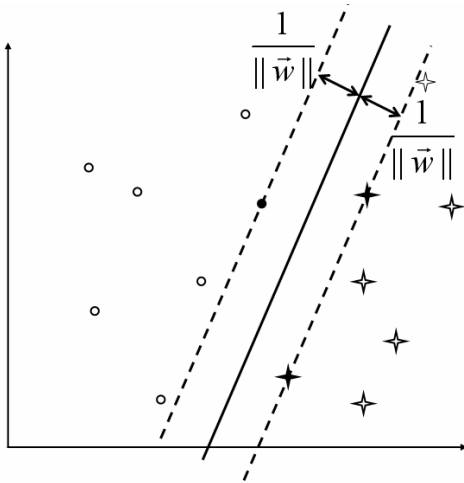
$$\langle \mathbf{w} \cdot \mathbf{x} \rangle + b = 0 \quad (1)$$

$\mathbf{w}$  is an  $N$  dimensional vector and  $b$  is a variable. Under the conditions, a SVM solves the following problem.

$$\begin{aligned} &\text{minimise} \quad \langle \mathbf{w} \cdot \mathbf{w} \rangle \\ &\text{subject to} \quad y_i(\langle \mathbf{w} \cdot \mathbf{x}_i \rangle + b) \geq 1, \quad i = 1, \dots, s. \end{aligned} \quad (2)$$







**Figure 2.** Training data that can be separated

In practice, a SVM is not able to build the separating hyperplane when training samples cannot be separated linearly because of some noise data. In this case, a SVM relaxes the above conditions introducing variable  $\xi_i$  as follows.

$$\begin{aligned} &\text{minimise} \quad \langle \mathbf{w} \cdot \mathbf{w} \rangle + C \sum_{i=1}^s \xi_i \\ &\text{subject to} \quad y_i(\langle \mathbf{w} \cdot \mathbf{x}_i \rangle + b) \geq 1 - \xi_i, \quad i = 1, \dots, s. \end{aligned} \quad (3)$$

$C$  is a constant.

In this article, we use Tiny-SVM [11], one of the most available SVMs.

## 2.4. TF/IDF

TF and IDF are measures of documents' weights to retrieve documents with keywords. TF shows how many keywords are in a document. Behind this weighting is the fact that Luhn has described "a repeated concept is important" [12]. However, a word that is a generally high-frequency word is not important to characterize a document [13] and keywords' weights have been needed.

IDF shows a keyword's weight that is calculated as follows.

$$idf(t) = \ln \frac{N}{df(t)} \quad (4)$$

$idf(t)$  shows a weight of keyword  $t$ .  $N$  shows a number of retrieval documents.  $df(t)$  shows a number of documents that include keyword  $t$ .  $idf(t)$  is largest if only one document includes keyword  $t$ , and  $idf(t)$  is smallest if all documents include keyword  $t$ . TF/IDF is a measure that shows total value of IDFs of keywords included in a document.

## 3. Methods

The proposed model exploits a SVM and TF/IDF to label clauses of medical documents as observation, diagnosis and therapy. Exploiting the labels, the model

can find clauses that have strong anteroposterior relations. Based on the relations, the model measures the similarities between the medical documents.

### 3.1. ODT cycles

The proposed model classifies clauses of medical documents into observation, diagnosis and therapy statistically. Observation, diagnosis and therapy are abbreviated as O, D and T respectively. Each of O, D and T is called an ODT element and a cycle of O, D and T is called an ODT cycle. We count one cycle if each of O, D and T is included only once. Figure 3 shows ODT cycles.

When he was admitted to the hospital, all drugs including rimatil were withdrawn. DOE exacerbation, exacerbation X-ray shadow of his chest are associated with a time when he took rimatil and I doubted his drug associated hepatitis. His DLST of rimatil, and his patch test of rimatil were negative.

入院時より、すでにリマチルを含むすべての薬剤はOFFになっていた。DOEの増悪、胸部レントゲンの陰影の増悪がリマチル内服時に一致していたことよりリマチルによる薬剤性肺炎を疑った。リマチルのDLST、パッチテストでは陰性であった。...



T: When he was admitted to the hospital, all drugs including rimatil were withdrawn.

O: DOE exacerbation, exacerbation of X-ray shadow of his chest are associated with a time when he took rimatil

D: and I doubted his drug associated hepatitis.

O: His DLST of rimatil, and his patch test of rimatil were negative.

...

**Figure 3.** ODT cycles

3.1.1. Training sets. We make two kinds of training sets. A SVM should learn one kind of training sets. The training set is medical documents that are segmented by Japanese pause marks into clauses labeled as O, D and T correctly by personal and one clause may have two or three labels of O, D and T. Figure 4 shows an example of the training set.

When he was admitted to the hospital, all drugs including rimatil were withdrawn. DOE exacerbation, exacerbation X-ray shadow of his chest are associated with a time when he took rimatil and I doubted his drug associated hepatitis. His DLST of rimatil, and his patch test of rimatil were negative.

入院時より、すでにリマチルを含むすべての薬剤はOFFになっていた。DOEの増悪、胸部レントゲンの陰影の増悪がリマチル内服時に一致していたことよりリマチルによる薬剤性肺炎を疑った。リマチルのDLST、パッチテストでは陰性であった。...



T: When he was admitted to the hospital, all drugs including rimatil were withdrawn.

O: DOE exacerbation,

OD: exacerbation of X-ray shadow of his chest are associated with a time when he took rimatil and I doubted his drug associated hepatitis.

O: His DLST of rimatil,

O: and his patch test of rimatil were negative.

...

**Figure 4.** An example of the training set that a SVM should learn

The other kind of training sets leads to TF/IDF. The training set is medical documents that are segmented by





personal into clauses and each clause is labeled by personal to have only one or no label. Figure 3 is also an example of the training set.

3.1.2. Automatic classification utilizing a SVM. A SVM learns a training set that is segmented by Japanese pause marks into clauses each of which may have two or three labels of O, D and T. First, ChaSen, Japanese morphologic analysis, divides the training set into words. And the proposed model puts each clause into vector space each axis of which corresponds one noun or one verb, and vector values are numbers of the corresponding words. If a clause is “He caught a cold”, a value of axis “cold” is one. If a clause is “administering medicine A 50mg and medicine B 100mg”, a value of axis “mg” is two.

The model needs three SVMs each of which determines whether clauses segmented a test set into by Japanese pause marks are O, D and T respectively. For example, a SVM determining whether a clause is O learns a training set labeled O as true data and learns a training set not labeled O as false data. And the SVM finds an optimal hyperplane separating true data and false data. Furthermore, ChaSen segments the clauses of the test set into words and the model puts each clause into the vector space. The SVM determines whether each clause is O or not utilizing the hyperplane.

Three SVMs labeling the clauses of the test set, the clauses have no label, one label or two or three labels. The model erases clauses that have no label because it is not thought to be connected clinical information. Next, the model divides clauses that have two or three labels exploiting TF/IDF.

The model uses the other kind of training sets that is segmented by personal into clauses labeled by personal. ChaSen divides the clauses of the training set into words and the model measures weight vectors of nouns and verbs.

The number of weight vectors’ dimensions corresponds a number of classes. In this case, classes are O, D and T and a number of classes is three. The weight vector for word  $t$  is shown as follows.

$$V_t = (V_{t,O}, V_{t,D}, V_{t,T}) \quad (5)$$

Value  $V_{t,c}$  corresponding class  $c$  is found as follows.

$$V_{t,c} = tf(t,c)idf'(t) \quad (6)$$

$tf(t,c)$  is a term frequency of  $t$  in  $c$  and  $idf'(t)$  is modified version of formula 4 and found as follows.

$$idf'(t) = \ln \frac{\text{NoC}}{df(t)} \quad (7)$$

NoC is a number of classes and is three in this case.  $df(t)$  is a number of classes including  $t$ . If only one class includes  $t$ ,  $idf'(t)$  is highest. If all classes include  $t$ ,  $idf'(t)$  is 0 and  $t$  is not thought to be available to classify

clauses. While  $idf(t)$  of formula 4 shows a weight of a word to identify documents,  $idf'(t)$  of formula 7 shows a weight of a word to identify classes. Table 1 shows words rankings by values of weight vectors. The model divides clauses of a test set that have two or three labels using found weight vectors of words.

The model identifies order of occurrence of classes in clause  $S$ . We assume that  $S$  is up to one cycle. ChaSen divides  $S$  into words and the model sets the nouns and the verbs of the words in order of occurrence in  $S$ . Let  $t_1, t_2, \dots, t_m$  be the nouns and the verbs of the words in order of occurrence in  $S$ . Table 2 shows values of weight vectors corresponding  $t_1, t_2, \dots, t_m$ .

Table 1. Words rankings by values of weight vectors

| O                         |      | D                       |       | T                      |       |
|---------------------------|------|-------------------------|-------|------------------------|-------|
| 認める (recognize)           | 8.82 | 認める (recognize)         | 12.85 | 施行 (do)                | 21.39 |
| 低下 (deconditioning)       | 7.78 | 指摘 (assignment)         | 9.48  | 目的 (purpose)           | 10.13 |
| 自覚 (awareness)            | 7.45 | 異常 (abnormality)        | 9.14  | 紹介 (referral)          | 9.71  |
| した (have be done)         | 6.20 | 疑う (doubt (verb))       | 8.50  | 開始 (begin)             | 8.61  |
| られる ((Japanese suffix))   | 6.08 | 診断 (diagnosis)          | 8.48  | 日当 (daily pay)         | 8.16  |
| 出現 (appearance)           | 5.41 | られる ((Japanese suffix)) | 8.28  | 投与 (administration)    | 7.21  |
| MEQ                       | 5.02 | 疑い (doubt (noun))       | 6.46  | MG                     | 7.16  |
| 頭痛 (headache)             | 4.68 | 考える (consider)          | 6.22  | フォロー (follow)          | 7.06  |
| 全身状態 (general status)     | 4.68 | 明らか (clearness)         | 6.08  | 加療 (treatment)         | 6.87  |
| 改善 (modification)         | 4.54 | 伴う (accompany)          | 5.67  | した (have be done)      | 6.81  |
| 腫脹 (tumor extent)         | 4.49 | 障害 (damage)             | 5.37  | 受診 (consultation)      | 6.68  |
| 軽快 (recovery)             | 4.49 | SLE                     | 5.07  | 中止 (discontinuation)   | 6.34  |
| 症状 (symptom)              | 4.42 | 可能性 (capability)        | 4.91  | 行う (do)                | 6.02  |
| よう ((Japanese suffix))    | 4.38 | 狭心症 (angina)            | 4.74  | 変更 (change)            | 5.89  |
| 転倒 (falling)              | 4.30 | 内膜 (endomenbrane)       | 4.39  | ML                     | 5.77  |
| 続く (continue)             | 4.30 | 肝障害 (hepatopathy)       | 4.39  | 外来 (outpatient clinic) | 5.20  |
| 食後 (after food)           | 4.30 | LESION                  | 4.39  | 経過観察 (follow-up)       | 5.00  |
| 消失 (disappearance)        | 4.30 | CANCER                  | 4.39  | 内服 (administration)    | 4.83  |
| 倦怠感 (fatigue)             | 4.30 | ADENOCARCINOMA          | 4.39  | 受ける (consult)          | 4.71  |
| 圧迫感 (feeling of pressure) | 4.30 | 発症 (occurrence)         | 4.20  | センター (center)          | 4.56  |

Table 2. Values of weight vectors corresponding  $t_1, t_2, \dots, t_m$

|       | $V_{t,O}$   | $V_{t,D}$   | $V_{t,T}$   |
|-------|-------------|-------------|-------------|
| $t_1$ | $V_{t_1,O}$ | $V_{t_1,D}$ | $V_{t_1,T}$ |
| $t_2$ | $V_{t_2,O}$ | $V_{t_2,D}$ | $V_{t_2,T}$ |
| ...   | ...         | ...         | ...         |
| $t_m$ | $V_{t_m,O}$ | $V_{t_m,D}$ | $V_{t_m,T}$ |

We assume that  $S$  is labeled as O and D. If  $S$  has other two labels or three labels, the model can follow the procedure below.

Let  $V_{t_i,O}$  be the highest value in  $V_{t_1,O}, V_{t_2,O}, \dots, V_{t_m,O}$  and  $V_{t_j,D}$  be the highest value in  $V_{t_1,D}, V_{t_2,D}, \dots, V_{t_m,D}$ . The model divides  $S$  in order of O, D if  $i < j$ , and the model divides  $S$  in order of D, O if  $j < i$ . If  $i = j$ , the model compares  $V_{t_i,O}$  and  $V_{t_j,D}$ , and the class of higher value is fixed. We assume that  $V_{t_i,O}$  is higher value. The model erases  $V_{t_j,D}$  from  $V_{t_1,D}, V_{t_2,D}, \dots, V_{t_m,D}$ , and let  $V_{t_j',D}$  be the highest value. The model divides  $S$  in order of O, D if  $i < j'$ , and the model divides  $S$  in order of D, O if  $j' < i$ . If a number of classes is more than a number of words in  $S$ , the model regards  $S$  as having only labels corresponding words.





Fixing order of classes in  $S$ , the model sorts the words into the classes. We assume that  $i < j$ . The model sorts  $t_1, t_2, \dots, t_i$  into class O and  $t_j, t_{j+1}, \dots, t_m$  into class D. And the model compares  $V_{tk,O}$  and  $V_{tk,D}$  corresponding  $t_k$  ( $i < k < j$ ) and sorts  $t_{i+1}, t_{i+2}, \dots, t_{j-1}$  into a class corresponding higher value.

The model made ODT elements each of which has some nouns and some verbs. However, ODT elements are not in strict order of O, D, T, O, D, T, ... . The model modifies ODT elements.

- Continued Os, continued Ds and continued Ts get together.
- If a cycle lacks, for example, T follows O, the model regards D as existing between O and T.

Figure 5 shows modified ODT cycles.

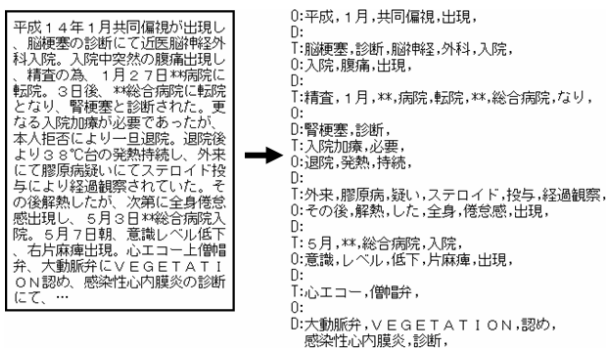


Figure 5. Modified ODT cycles

### 3.2. Measuring similarities between medical documents

ODT elements among one ODT cycle are related strongly because it is thought that O introduces next D and T, D leads next T and pulls out next O, and similarly T pulls out next O and D. Additionally, we can recognize flows of medical documents focusing Os, Ds or Ts respectively. For example, focusing Ds, we can recognize a flow of diseases such as a patient having cataract after having diabetes. So, the model regards anteroposterior relations among one ODT cycle and anteroposterior relations between Os, Ds or Ts.

Let  $D_1, D_2$  be compared documents. Let  $X_1, X_2, \dots, X_{n_1}$  ( $Y_1, Y_2, \dots, Y_{n_2}$ ) be ODT elements in order of occurrence in  $D_1$  ( $D_2$ ). The model measures a similarity between  $D_1$  and  $D_2$  as follows.

$$\sum_{1 \leq i < j \leq m} \sum_{1 \leq k < l \leq n_2} \text{simi}(X_i, X_j, Y_k, Y_l) \quad (8)$$

The model finds  $\text{simi}(X_i, X_j, Y_k, Y_l)$  0 if  $X_i, X_j, Y_k$  and  $Y_l$  do not meet following two conditions.

1.  $X_i$  and  $X_j$  are in one ODT cycle and  $Y_k$  and  $Y_l$  are in one ODT cycle, or  $X_i$  and  $X_j$  have the same kind of labels and  $Y_k$  and  $Y_l$  have the same kind of labels.
2.  $X_i$  and  $Y_k$  have the same kind of labels and  $X_j$  and  $Y_l$  have the same kind of labels.

If above two conditions are met, the model finds  $\text{simi}(X_i, X_j, Y_k, Y_l)$  as follows.

$$\text{simi}(X_i, X_j, Y_k, Y_l) = \left( \sum_{C_m \in C(X_i, Y_k)} \text{idf}(C_m) \right)^2 \left( \sum_{C_n \in C(X_j, Y_l)} \text{idf}(C_n) \right)^2 \quad (9)$$

$C(X_i, Y_k)$  ( $C(X_j, Y_l)$ ) is a set of words included in both  $X_i$  and  $Y_k$  (in both  $X_j$  and  $Y_l$  respectively).  $\text{idf}(t)$  shows a weight of word  $t$  and formula 4 measures it. The less documents include  $t$ , the higher  $\text{idf}(t)$  is.

Finally, the model corrects similarities by dividing them by squared numbers of words of comparative documents because similarities are dependent on sizes of documents. And so, similarities are not absolute values but relative values. A similarity of a medical document  $D_1$  compared to a medical document  $D_2$  is different from a similarity of  $D_2$  compared to  $D_1$ .

## 4. Evaluation

### 4.1. Retrieval processes and evaluation method

Retrieval processes consist of a matching process, a ranking process and a display process. Retrieving models find documents meeting queries in a matching process, and the models rank the found documents in optimality as results of retrieving in a ranking process. In a display process, the models display the results as users can recognize them easily.

We evaluate the proposed model in a matching process and in a ranking process. The model determines whether each compared document is a similar document in a matching process. The model ranks similar documents in a ranking process.

In an experiment 1, we see whether the model can retrieve medical documents of the same disease as each query document to evaluate the model in matching. In an experiment 2, we see whether the model can rank documents that a doctor determines very similar to each query document near the top of documents.

A model compared to the proposed model is a vector space model that measures a similarity by a number of kinds of words included in two documents. The vector space model regards  $\text{idf}(t)$  of formula 4 as a weight of word  $t$ . The vector space model can consider TF, but the vector space model does not consider TF because the model got better results than the model considering TF.

### 4.2. Experiment 1

We evaluate the proposed model in a matching process. The proposed model and the vector space model measure similarities between the 100 reports. We evaluate how much percentage of reports that have the same disease as each query report the two models rank in the top 20. We suppose medical documents having the same disease should be matched by binary decision.

Given 100 discharge summaries consist of 5 sets of 20 reports, where each set corresponds to one of 5 diseases: "type 2 diabetes", "lung cancer", "angina





pectoris", "uterine fibroid" or "cerebral infarction". Because certain diseases, such as type 2 diabets, may trigger other diseases, we selected reports to make given 5 sets independent each other.

### 4.3. Experiment 2

We evaluate the proposed model in a ranking process. We choose 22 discharge summaries that have "liver cancer" as a disease name and have a doctor choose up to 3 reports compared to each report as similar documents available for clinical care. A doctor is able to choose no similar document. In this article, we do not make the doctor rank similar documents as he may not able to apply a single firm standard to judge resemblance of each pair of documents. Furthermore, while a doctor determines that document *A* is similar to document *B*, he may determine that *B* is not similar to *A* because *B* includes contents of *A*, but contents of *A* is not important to *B*.

The proposed model and the vector space model measure similarities between the 22 reports and the models get 3, 2 or 1 points if the models rank the similar documents determined by a doctor first, second or third respectively.

## 5. Results

### 5.1. Experiment 1

Table 3 shows results of experiment 1.

**Table 3.** Results of experiment 1

|                  |          | the vector space model | the proposed model |
|------------------|----------|------------------------|--------------------|
| weights of words | 1        | 66.1%                  | 56.5%              |
|                  | $idf(t)$ | 77.6%                  | 72.0%              |

As a result, the proposed model gave 72.0 % while the vector space model gave 77.6 %. There was not much difference between the models and we saw that a number of kinds of words included in compared documents is important to retrieve documents that have the same disease name.

### 5.2. Experiment 2

A doctor determined that 18 reports of 22 have similar documents in the experiment 2 and a total number of similar documents is 44.

Table 4 shows similarities compared to medical document 2 and table 5 shows similarities compared to medical document 10.

Numbers of the first columns are numbers of compared medical documents and numbers of the second columns and the third columns are similarities given by the vector space model and the proposed model respectively. Similarities were highest when compared documents were query documents. The doctor

determined that documents 9, 10 and 22 are similar documents to the document 2 and documents 2, 8 and 9 are similar documents to the document 10.

**Table 4.** Similarities compared to medical document 2

|    | the vector space model | the proposed model |
|----|------------------------|--------------------|
| 1  | 0.539                  | 0.005              |
| 2  | 2.576                  | 211.710            |
| 3  | 0.246                  | 0.001              |
| 4  | 0.327                  | 0.014              |
| 5  | 0.304                  | 0.001              |
| 6  | 0.511                  | 0.001              |
| 7  | 0.330                  | 0.001              |
| 8  | 0.460                  | 0.088              |
| 9  | 0.282                  | 0.669              |
| 10 | 0.372                  | 0.297              |
| 11 | 0.216                  | 0.000              |
| 12 | 0.257                  | 0.002              |
| 13 | 0.233                  | 0.285              |
| 14 | 0.292                  | 0.003              |
| 15 | 0.452                  | 0.004              |
| 16 | 0.468                  | 0.001              |
| 17 | 0.390                  | 0.015              |
| 18 | 0.582                  | 0.047              |
| 19 | 0.187                  | 0.000              |
| 20 | 0.323                  | 0.012              |
| 21 | 0.605                  | 0.001              |
| 22 | 0.253                  | 0.003              |

**Table 5.** Similarities compared to medical document 6

|    | the vector space model | the proposed model |
|----|------------------------|--------------------|
| 1  | 0.313                  | 0.015              |
| 2  | 0.467                  | 0.468              |
| 3  | 0.237                  | 0.024              |
| 4  | 0.389                  | 0.181              |
| 5  | 0.352                  | 0.060              |
| 6  | 0.627                  | 0.065              |
| 7  | 0.259                  | 0.014              |
| 8  | 0.773                  | 6.565              |
| 9  | 0.299                  | 2.691              |
| 10 | 2.146                  | 415.295            |
| 11 | 0.266                  | 0.020              |
| 12 | 0.206                  | 0.053              |
| 13 | 0.254                  | 0.031              |
| 14 | 0.383                  | 0.208              |
| 15 | 0.334                  | 0.279              |
| 16 | 0.358                  | 0.002              |
| 17 | 0.298                  | 1.411              |
| 18 | 0.519                  | 2.714              |
| 19 | 0.237                  | 0.001              |
| 20 | 0.299                  | 0.316              |
| 21 | 0.576                  | 0.064              |
| 22 | 0.181                  | 0.154              |





As a result of the experiment 2, the proposed model gave 36 points while the vector space model gave 22 points, and the proposed model could rank similar documents first, second or third for 13 documents (72.2% (13/18)).

## 6. Discussion

While the proposed model is not thought to be superior to the vector space model in a matching process, the proposed model is thought to be superior to the vector space model in a ranking process. The proposed model gave a good result in a ranking process. It shows that pursuing clinical processes is important to similar document retrieval for medical documents and the proposed model could pursue clinical processes unlike the vector space model and other models that measure a vector of each document.

To construct similar document retrieval system for medical documents, the vector space model finds similar documents in a matching process because the model is executed fast, and the proposed model ranks found similar documents.

Additionally Japanese different words having the same meaning were found in documents. The models do not consider the words, and because of it, the models were thought not to give good results in any cases. The models need structured medical terminology and need to understand relationships between medical terms.

## 7. Conclusion

We proposed the model that classifies clauses of medical documents exploiting SVMs and TF/IDF into O, D and T and measures similarities between medical documents utilizing the classes. We evaluated the proposed model compared to a vector space model used generally and concluded the proposed model is superior to the vector space model in a ranking process, which is one of retrieval processes. We showed feasibility of similar document retrieval system for medical documents that is executed fast and has high precision when the proposed model is adapted with the vector space model.

## 8. References

- [1] G. Salton and C. Buckley, "Improving Retrieval Performance by Relevance Feedback", *Journal of the American Society for Information Science*, Vol. 41, No. 4, pp. 288-97 (1990).
- [2] F. Ingo, T. Ulrich and K. Thomas, "Annotation-based Document Retrieval with Four-Valued Probabilistic Datalog", *Proc. of the first SIGIR Workshop on the Integration Retrieval and Databases*, SIGIR, Sheffield, pp. 31-38 (2004).
- [3] J. H. Van Bommel and J. H. Musen, *Handbook of Medical Informatics*, Springer-Verlag, the Netherlands, pp. 4-7 (1997).
- [4] M. Benkhalifa, A. Mouradi and H. Bouyakhf, "Integrating External Knowledge to Supplement Training Data in Semi-Supervised Learning for Text Categorization", *Information Retrieval*, Vol. 4, Issue 2, pp. 91-113 (2001).
- [5] J. Guo, A. Takada, K. Tanaka, J. Sato, M. Suzuki, T. Suzuki, Y. Nakashima, K. Araki and H. Yoshihara, "The development of MML (Medical Markup Language) version 3.0 as a medical document exchange format for HL7 messages", *Journal of Medical Systems*, Vol. 28, No. 6, pp. 523-33 (2004).
- [6] R. H. Dolin, L. Alschuler, S. Boyer, C. Beebe, F. M. Behlen, P. V. Biron, and A. Shabo Shvo, "HL7 Clinical Document Architecture, Release 2," *Journal of the American Medical Informatics Association*, Vol. 13, No. 1, pp. 30-9 (2005).
- [7] T. Takemura, H. Matsui and N. Ashida, "Systematization of Medical Terminology Knowledge based on Examples," *Mobile Communications Technology for Medical Care and Triag MCMT 2002*, pp. 92-6 (2002).
- [8] BANYU PHARMACEUTICAL CO., "The Merck Manuals (Japanese)", <http://merckmanual.banyu.co.jp/> (accessed Apr. 2006).
- [9] Y. Mastumoto, "ChaSen's Wiki - Front Page", <http://chasen.aist-nara.ac.jp/hiki/chasen/> (accessed Apr.).
- [10] B. E. Boser, I. M. Guyon and V. N. Vapnik, "A Training Algorithm for Optimal Margin Classifiers", *Proc. of the 5th Annual ACM Workshop on Computational Learning Theory*, pp. 144-152 (1992).
- [11] T. Kudo, "Tiny SVM: Support Vector Machine", <http://www.chasen.org/~taku/software/TinySVM/> (accessed Apr. 2006).
- [12] H. P. Luhn, "A Statistical Approach to Mechanized Encoding and Searching of Literary Information", *IBM Journal of Research and Development*, Vol. 1, No. 4, pp. 307-17 (1957).
- [13] J. K. Sparck, "A Statistical Interpretation of Term Specificity and Its Application in Retrieval", *Journal of Documentation*, Vol. 28, No. 1, pp. 11-21 (1972).





# The Influence that Person's Sensibility Receives from Mixed-Reality based on a *Tanka* (a traditional Japanese poem of 31 syllables)

Hiroshi Juzoji, Muhammad Athar Sadiq, Isao Nakajima, Yasumitsu Tomioka \*  
*Tokai University School of Medicine, Office Tomioka\**  
*juzoji@is.icc.u-tokai.ac.jp*

## Abstract

*We chose a poem depicting the moon and a mountain scene at night, written by Yakamochi Otomo in the Man yo-shu that is thought to represent Japan's oldest diary. The particular Otomo's poem was chosen because it allows easy identification of the where (location) and the when (time) described by the poem. EEG measurement were taken as subjects were presented with Yakamochi Otomo's poem in three different formats: 1) an oral presentation; 2) when viewing a TV screen image; 3) when viewing a three panel wide-vision screen (as a Mixed-reality image). According to results provided by fast Fourier transforms and frequency analysis, the presentation of Mixed-reality displayed on the three panel wide-vision screen was most effective in creating a sense of immersion and a state of relaxation.*

## 1. Purpose

We set out to evaluate the sensitivity and emotional response of subjects as determined by EEG (Electroencephalogram) readings by visually reproducing scenes onto a three panel wide-vision screen adapted from a *tanka* poem written by Yakamochi Otomo. This poem is also thought to represent Japan's oldest diary or journal. For this research, we chose a poem depicting the moon and a mountain scene at night, written by Yakamochi Otomo and recorded in the Collection of Myriad Leaves (*Man yo-shu*, *Man yo* collection of poems). The particular Otomo poem was chosen because it allows easy identification of the "where" (location) and the "when" (time) described by the poem. The night-scene depicted in the poem was visually reproduced onto a three panel wide-vision screen in order to evaluate the effects induced by brain activity when recounting or recalling the past. The subject's EEGs were first converted by a fast Fourier transform, and then evaluated through frequency analysis.

## 2. Background

### 2-1. Research on Sensitivity

Human thinking entails rationalization based on knowledge or intellect. We simultaneously perceive our experiences in both this intellectual way, and in sensuous and emotional terms. Humans recognized the presence of

these two faculties long before Greek civilization. The former has been termed "Logos" and the latter "Pathos." From the earliest times, one way of sharing experience has been poetry – including the *tanka*, a traditional Japanese form of 31 syllables – music, painting, theater, and religion, to some extent. These methods of non-intellectual communication were inherited by contemporary society, and together help define what we know as a specific culture. On an individual and personal level, such works are often inspired by the desire to share various feelings and emotions. Over time, and with advancing technology, this desire to communicate has evolved into media forms such as theater, movies, and television. We regard the psychological extremity of sensitivity sharing this type of communication equivalent to a state of mass-intoxication. Gamelan music indigenous to the island of Bali, resonating over a frequency of 50-100 kHz, is said to induce a neuro-physiological state of intoxication. Similarly, the Kecak dance in Bali, with its hand motions (fever-pitched motion of hands) of a hundred or so naked figures dancing in the dark is said to have the power to induce intoxication in even the audience watching the dance. These states of intoxication occur as the result of an irresistible psychological urge to empathize with the masses. In medical terms, the phenomenon can be regarded as a transient case of reactive psychosis [1].

There are numerous methods by which one can experience the type of emotional communication described above, from *tanka* to Kecak dance. However, we believe a projection of Mixed-reality produced by a synthesis of multiple camera images and computer-generated virtual reality space projected onto a three panel wide-vision screen represents a much more effective form of such communication than reading a book or watching television. We began by comparing the effectiveness of two methods. The first method entails reading of a *tanka* poem, which conveys minimal information, while the second involves recreating a scene mentioned in the *tanka* poem onto three panel wide-vision screen.

### 2-2. EEG and Fast Fourier Analysis

EEGs (electroencephalograms) are recordings of electrical activity of the brain, as detected by silver electrodes positioned on the scalp. The electrical potentials involved fall in the range of 5-300µV, at frequencies of 5-30 Hz. No examination surpasses the EEG's sensitivity





in representing direct brain activity in terms of space and time. However, because EEGs monitor the total electric potential of numerous neuron cells from the surface of the scalp, spatial separation and frequency analysis are critical in evaluating EEG results. We currently use a spectrum analysis method of frequency analysis for the relationship between electric potential and frequency data. This frequency analysis is called a discrete Fourier transform, using algorithms derived from the fast Fourier transform. Here we are most interested in the beta wave's (13-30 Hz) PCA (Principal Component Analysis) and alpha attenuation, the electric potential ratio of alpha waves (8-13 Hz) when eyes are shut and open. Past research has demonstrated that the former (beta wave) indicates concentration/attention span, while the latter (alpha wave) indicates a state of comfort and relaxation [2].

### 2-3. Information on Collection of Myriad Leaves

Ancient Japanese history is classified into the *Yayoi*, *Asuka*, *Nara*, and *Heian* periods. Any writing dating before the *Nara* period, public or private, has been found only in the form of engravings on tombs or stones (a total of 20 artifacts). The oldest extant paper documents are stored at *Shousou-in* in *Nara* prefecture. The number of artifacts totals approximately 12,000 pieces. Valuable historical literature such as "*Kojiki*" and "*Nihon-sho-ki*" are among these documents, as well as Collection of Myriad (ten thousand) Leaves for *waka* or *tanka* poems.

The typical citizens around the time of the *Nara* period could not read or write; the only individuals capable of writing were bureaucrats. Nor was there a tradition of a personal diary or journal, even among the educated class. One of the first such is a compilation of poems edited by Yakamochi Otomo, entitled Collection of Myriad Leaves, which may also be interpreted as a diary or journal written by Yakamochi Otomo [3].

We thought it would be possible to reproduce the night scene described in a poem using a computer if we could identify the location and the date described in the Collection of Myriad Leaves, and if we could find a poem in Collection of Myriad Leaves describing the shape and the location of the moon.

### 2-4. Reproduction of Time

If we enter an arbitrary time in hour / minute / date / month / year and an arbitrary location in latitude / longitude / height, astronomical simulator can reproduce the constellations visible on that particular night. However, this work also involves various adjustments involving date / month settings, since Europe in the Middle Ages relied on the Julian calendar, which set 365.25 days per year. Our current calendar system, the Gregorian calendar, sets 365.2422 days per year. A conversion is needed to accurately recreate the oldest dates. The essential

difference between Julian and Gregorian calendars is the handling of leap years. Europeans added 10 days to the date October 5, 1582, making it October 15, 1582, to make dates consistent with the present system of leap years. Europe converted to the Gregorian calendar after making this adjustment. Japan converted from the old lunar calendar to the Gregorian calendar on January 1, 1873. The Japanese lunar calendar originated in China and is based on the Julian calendar. However, the two may differ by several days. Since dates in Collection of Myriad Leaves are based on the old lunar calendar, we need a formula to convert past dates to the present Gregorian calendar [4].

## 3. Method

### 3-1. System Overview

An overview of the evaluation system is shown in figure 1. A three panel wide-vision CRT screen (hardware configuration providing a 135° angle of view consists of IBM compatible PC with clock speeds of 3GHz) and three LCD projectors (800 x 600 pixels, full-color RGB). The system is designed to examine EEG readings and frequency analysis using oral presentations, single-panel CRTs, and three panel wide-vision screen (135°).

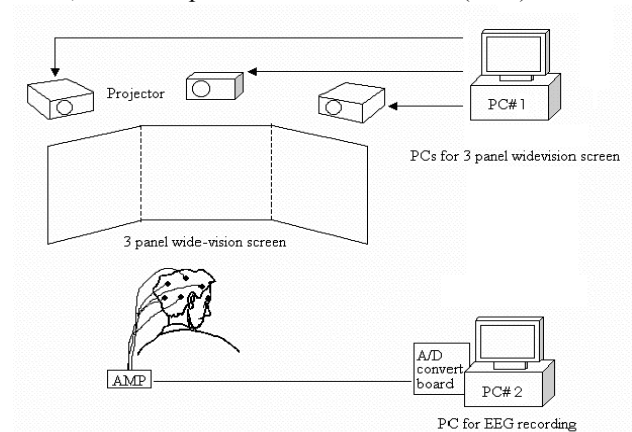


Figure 1. System outline

### 3-2. Examination Method

We chose five subjects whose fields of specialization excluded literature. Subjects were asked to draw straws to determine the order in which they were examined, then asked to stand in the waiting room so that the contents of the examination would remain unknown. Each subject was given a presentation of a Collection of Myriad Leaves poem in the following sequence of presentations, from 1) to 3). The subject was also asked to open and close their eyes from time to time to permit readings to be obtained.

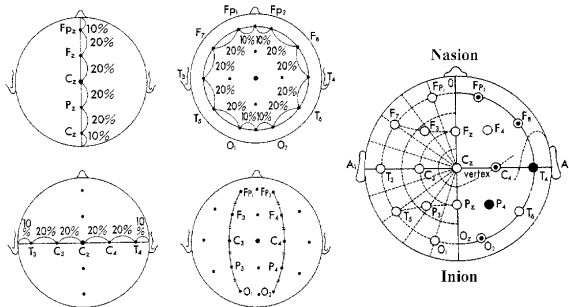
- 1) Listening to an oral presentation
- 2) Viewing a television screen
- 3) Viewing a three panel wide-vision screen





### 3-3. EEG Examination

Silver electrodes were used to measure electric potential, as shown in figure 2. The electrodes were placed in contact with the subject's scalp as single-pole conductors throughout the subject's scalp (evenly spaced and retaining left / right symmetry).



**Figure 2. Electrode layouts**

#### A. PCA (Principal Component Analysis)

PCA is designed to evaluate wave synchronism from a number of electrodes. When synchronism is uniform, the signals are considered to be integrated. When synchronism is off and random, the signals are considered to be non-integrated. An epileptic fit may be regarded as an extreme example of the former. In such instances, the PCA determines that the brain is concentrating on processing information [5].

#### B. Alpha attenuation

When a subject's eyes are shut, the subject's alpha wave activity increases. A comparison of open and shut status implies that levels of awareness [5]. In other words, comparing the alpha attenuation of a subject during testing (test load) can help indicate how well a subject has adapted to our designed environment.

$$\text{alpha attenuation} = \frac{\text{alpha wave: eyes shut}}{\text{alpha wave: eyes open}}$$

### 3-4. Mixed-reality

#### A. Virtual Reality:

An astronomical simulator was used to reproduce the night sky on a three panel wide-vision screen. Software was used on PCs configured for the Windows XP environment, and graphic board that has three outputs were installed to display the image on three screens.

#### B. Reality:

Present night scene actually described in the poem was taken with a triple-lens camera and recorded onto a VCR. The image was then loaded onto PC following editing. The primary reason for this shoot was to capture certain content difficult to reproduce via virtual reality, such as mountains casting a shadow onto foregrounds or grass fields blowing in the wind. These images were then

combined with VR.

#### C. Sound

This sensory component would aid in transporting the listener back to a night sky in the distant past when listening to a poem (factual) written about a certain past time and date (simulated). Some interpretive comments and some sound effects were added to the presentation.

### 3-5. Selected Poem

In recreating the moon and the starry skies of a certain time and date (simulated) in the past on the astronomical software, the following poem has been selected.

"It must be dark, and late at night, because already the moon sinks toward the ridge of Mt. Futagami."

The author is Yakamochi Otomo, born in 718 A.D. The poem was written in the year 746. In the preceding year of 746, in March, he had won a promotion within the government. But after three months, in June 746, he was sent to Ecchu area as Defense Minister, the northern part of Japan. He apparently composed the poem there, while recalling the city far away.

### 3-6. Measurement Results

Measurements were taken for beta wave's PCA and alpha attenuation, the electric potential ratio of alpha waves (8-13 Hz) when eyes are shut and open. Both measurements were taken from the three different presentation formats of Yakamochi Otomo's poem: 1) Listening to an oral presentation (eyes shut); 2) when viewing the TV screen; 3) when viewing the three panel wide-vision screen.

**Table 1. Result by presentation formats**

| EEG<br>(Power Spectrum)              | PCA (Principal<br>Component<br>Analysis) | alpha attenuation<br>(Eyes shut / open) |
|--------------------------------------|------------------------------------------|-----------------------------------------|
| 1) Listening                         | small                                    | small                                   |
| 2) TV screen                         | medium                                   | small                                   |
| 3) Three panel<br>wide-vision screen | large                                    | Large                                   |

## 4. Discussion

### 4-1. Measurement Analysis

In general, alpha waves appear in the parietal region when the eyes are closed, fading when the eyes are opened. However, we also see wave frequencies quite similar to alpha waves when a person watches a dimly-lit screen while deep in a state of delusion. The alpha-wave activity has been observed to increase as the display area of the night sky grows larger (a three-panel screen compared to a single CRT).





Even when we compare the alpha attenuation, the effect of the three panel wide-vision screen is significant, coming close to presenting the subject with the impression that he or she is back in the *Nara* period. The EEG results seem to concur with our findings. On the other hand, the oral presentation indicated low values even in PCA (Principal Component Analysis). The oral presentation alone appears to be too diffuse, failing to capture the subject's full attention.

These findings suggest that the Mixed-reality image projected onto the three panel wide-vision screen provides a deeper sense of immersion for subjects.

#### 4-2. Lunar Calendar in the Days of *Man'yo-shu*

Yakamochi's poem describes the moon nearing the ridges of Mt. Futagami, which also implies that the time is late at night. If we can find a poem with a detailed description of the shape and position of the moon, it would be possible to convert lunar calendar dates into Gregorian calendar dates.

According to research conducted by Japanese literary experts, we can say the following about Yakamochi's poem:

Time: August 27, year 746. Converted in to solar calendar.

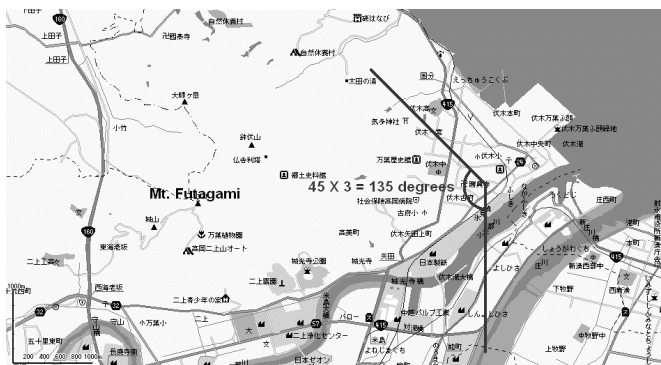


Figure 3. Map around the Fushiki

#### 5. Conclusion

We evaluated the subject's sensitivity and emotional changes through measurements of EEG. EEG measurements were taken as subjects were presented with Yakamochi Otomo's poem in three different formats: 1) an oral presentation; 2) when viewing a TV screen image; 3) when viewing a three panel wide-vision screen (as a Mixed-reality image). According to results provided by fast Fourier transforms and frequency analysis, the presentation of Mixed-reality displayed on the three panel wide-vision screen was most effective in creating a sense of immersion and a state of relaxation.

#### 6. Acknowledgements

Location: Fushiki-furukofu, Takaoka City, Toyama prefecture.

Academic theories point to a location around the present Fushiki meteorological station area.

What the poem fails to reveal:

1) The shape of the moon

2) The time at which the moon sets over the horizon

The position of the moon during this season as it sets falls approximately between west-by-southwest and southwest. If we can assume an angle of elevation of 15°, the moon should be in a southwesterly location one hour before setting.

According to computer simulations, the observation point must be to the north and to the east of Mt. Futagami's peak in order for the moon to be viewed both in a southwesterly direction and over Mt. Futagami itself. The Fushiki meteorological station falls within these parameters. According to our calculations, the moon sets over the horizon at 9:50pm, which seemingly conflicts with Yakamochi's verse, which gives the time as follows: "it must be late in the dark night—." If we can assume a later time for the moon to set – perhaps midnight – it would be reasonable to assume the date is not August 27, but several days later, around September 1.

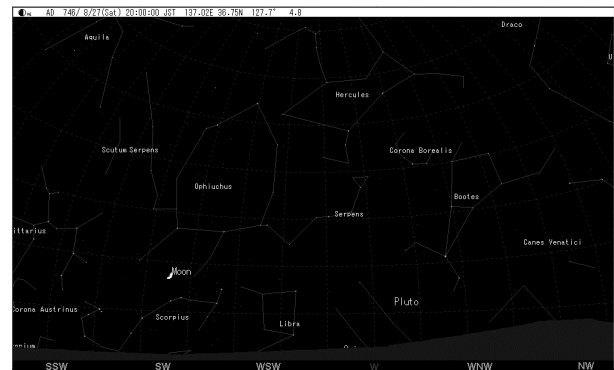


Figure 4. Simulation of starry sky

We would like to express our deep appreciation to the staff curator at the Man'yo Historical Museum of Takaoka City, Collection of Myriad Leaves research authority Souzou Kurokawa, and Takaoka city council member Eiichiro Mukai for their invaluable advice and assistance.

#### 7. References

1. K.Kaketa: Modern Psychiatry Summa 6B, Nakayama Shoten
2. T.Okuma: Clinical Electroencephalography, Igaku-shoin 1991
3. N.Kojima, M.Kinoshita, H.Higashino: Man'yo-shu, Shogakukan 1996
4. M.Uchida: Source book of Japanese lunar calendar, Yuzankaku 1975
5. H.Araga, T.Musha: A new analysis of sensitivity by EEG measurement, NF Circuit Design Block





## Evidence-Based Medicine Study between Bicycle Helmet and Head injury

Tsai Lung Wen<sup>a</sup>, Lin Shyr-Yi<sup>b</sup>, Ing Fang Yang<sup>c</sup>, Ten Fang Yang<sup>d</sup>

<sup>a</sup> EBM Center of Taipei Medical University Hospital

<sup>b</sup> Department of Internal Medicine, Taipei Medical University

<sup>c</sup> Department of Internal Medicine, Jen Chi General Hospital

<sup>d</sup> Graduate Institute of Medical Informatics, Taipei Medical University

E-mail:lungwen.tsai65@msa.hinet.net

### Abstract

*Introduction:* Taiwan has successfully passed the motorcycle helmet law in June, 1997. The head injury rate reduced by 30% after the implementation of law, but the head injury rate of bicycle is increase by 15% proportionately, especially for non-helmet users. It is very important to have the bicycle helmet law to protect teenagers from severe head injury, because it can decrease Years of Potential Life Lost(YPLL) significantly. *Methods:* We used case control study to analysis the study, and the result of GCS to discriminate the severity of head injury. In this study, hospital trauma registry was used for analysis to investigate the trauma patients at six hospitals in Taiwan area from February 2 to November 30, 2001. Telephone interviews were made to confirm whether the injury was caused by bicycle accident, and further to complete the insufficient information. We use frequency and Pearson Chi-Square to test for univariable. For multivariable, we use multiple logistic regression to examine the association between severity and related confounders. *Results:* According to trauma registered data, there were 608 hospitalized patients injured in bicycle related injury, and 600 patients were interviewed. Furthermore, in the logistic regression model, "Male"(odds ratio = 2.13, confidence interval = 1.95-4.38), "Without helmet "(odds ratio=9.34, confidence interval=1.51-62.86), "Helmet is unqualified" (odds ratio = 2.65, confidence interval = 1.23-3.18) and "brakes failure"(odds ratio = 3.08, confidence interval = 2.23- 4.26) had head injury. Furthermore, non-helmet will had serious head injury(odds ratio=4.42, confidence interval = 1.33-65.14) helmet and their severity of head injury in bicycle accident had also been proved significantly lower. *Discussion:* The benefit of helmet is proved in our study. If the bicycle helmet law is implemented, it will significantly reduced the severity of head injury. Besides, gender, helmet quality and status of bicycle need to be considered in preventing bicycle related head injury.

### Introduction

Bicycling is a worldwide activity. In both developed and developing countries it serves as an important means of transportation as well as an enjoyable recreational activity for both adults and children. Thus, injuries related to bicycling are relatively common among them head injuries account for one third of the visits to emergency departments, and up to two thirds of hospitalisations, and three quarters of deaths(1). Head injuries also carry a substantial risk of long term disability. Therefore, preventing head injuries from bicycling is important.

In June 1997, motorcycle helmet law has been successfully passed by the Legislative Yuan of Taiwan. The head injury rate decreased by 30% after the implementation of this law, As for non-helmet user, head injury rate disproportionately increased 15% in bike injury, especially for non-helmet users.

Safety helmets for bicycling have been available for at least 20 years. Although randomised controlled trials have become the gold standard for providing evidence of the effectiveness of clinical interventions, these trials are not feasible in this aspect. Given that the rate of head injury is about 20 injuries per 100,000 people, a randomised controlled trial would need to recruit hundred-thousands of people(2). Evidence for the effectiveness of helmets to prevent head injury has come from two other types of studies: case control studies, in which the proportion of cyclists wearing helmets with head injuries is compared with those without head injuries, and ecological studies examining changes in the rate of head injury over time among populations wearing helmets and those without. The strongest evidence for the effectiveness of helmets comes from case-control studies; this design is one of the cornerstones of modern epidemiology. A systematic review of five case-control studies, published at the Cochrane Library, found that helmets could reduce the risk by 63-88% for head, and brain injury among cyclists of all ages(1). Four of the studies controlled for a series of important covariates(3-6). helmets seemed to be equally effective in reducing injuries in crashes involving motor vehicles and in accidents associated with falls and other etiologies. Cook



and Shiekh reported this using an ecological time series analysis(7). Examining all admissions to NHS hospitals in England over a four year period, the authors found that head injuries as a proportion of monthly admissions for trauma related to bicycles fell from 40% in 1991-2 to 28% in 1994-5 while total emergency admissions for trauma related to bicycles did not change. These changes showed a consistent year to year trend in which the proportion of head injuries related to trauma from bicycles became lower in each successive year. Changes occurred in all age groups and are ascribed to an increase in the use of helmets. Similar findings from ecological studies have also been reported in the United States, New Zealand, and Australia(8-10). These findings were associated with an increased use of helmets occurring as a result of educational and legislative initiatives. Despite this large body of evidence on the effectiveness of helmets in preventing head injuries in cyclists and their beneficial effects for populations of cyclists, critics, especially in the United Kingdom, continue to question the usefulness of helmets. Their criticisms fall into two main categories: "risk homeostasis" and lack of adjustment for other confounders. Hillman has argued that while helmets may offer some inherent protection to cyclists there is no overall benefit because cyclists who wear helmets ride in a less cautious manner so that their overall risk of injury is unchanged(11). This theory of risk homeostasis has been discussed for decades, but the evidence that it applies to helmet use and bicycling is non-existent(12). The other criticism is that case control studies on helmets have not adequately controlled for all potential confounders, especially unmeasured factors such as differential risk taking behavior in cases and controls. Adequate adjustment for differences between cases and controls is important for the validity of any case-control study. Four of the five studies in the Cochrane review controlled for potential differences between cases and controls, such as age and severity of the crash(3-6) Crash severity can be used as a proxy for the hypothesised effects of risk taking behaviour. The magnitude of the protective effect of helmets found by these studies (threefold to eightfold ) makes it clear that unmeasured confounders cannot explain the differences in the risk of injury between cyclists who wear helmets and those who do not.

## Material and Method

The trauma registry is an emergency room based surveillance system(6), hospital is organized and administered by the research group. The companions of the injured patients receive a questionnaire about the injury event, including demographic information, time, location, activities being undertaken preceding the incident, and events that could be the contributing factors to the injury. The attending physician completes a medical questionnaire indicating the nature of the

injuries and the affected body part(s). The information is coded by a trained data entry clerk or nurse, based on a detailed manual. We use case control study to analyze efficacy of bicycle helmets, and GCS score to distinguish the severity of head injuries. In this case control study, we used hospital trauma registered system to investigate the traumatic patients from six hospitals in Taiwan from February 2 to November 30, 2001. In the mean time, telephone interviews were made to confirm whether the injury was caused by bicycle accident, and further to complete the insufficient information. We use frequency and Pearson Chi-Square univariable test. For multivariable, we use multiple logistic regression to examine the association between severity and related confounders.

## Results

Table 1 shows that men tend to have 50 % more severe head injuries than women. The main age group of head injured patient is 10-19 years old, approximately 45%. Under certain circumstances, cyclists may have severe head injuries: speed-change bicyclists (46%), cyclists without helmets (88.9%), large car accidents (45%), bicycle without reflecting objects (86.2%), cyclists carrying heavy items (73.1%), cycling at high speed (86.2%), brake broken system (64.4%). There is no statistically significant difference between wearing safety equipments and cycle with bells. Table 2 of multiple regression shows that the odds ratio in male having severe head injuries is 4.86 (confidence intervals between 2.89 and 5.68); in age group 10-19 is 3.85 (confidence intervals between 1.68 and 4.55); in speed-change bicycle is 2.89 (confidence intervals between 1.48 and 3.13); in cyclists without helmets is 4.64 (confidence intervals between 1.38 and 55.68); in bicycles without reflecting objects is 2.64 (confidence intervals between 1.38 and 4.58); in cyclists carrying heavy items is 5.26 (confidence intervals between 4.23 and 8.98); in cycling at high speed is 2.45 (confidence intervals between 1.65 and 2.88); in bicycles with broken brake system is 2.13 (confidence intervals between 1.25 and 2.69).

## Discussion

Bicycle injuries continue to be a serious public health problem, comprising 20% of all injuries to trauma registry during the study period. The helmet's benefit can be proved from the present study. Speed-change bicycle has been demonstrated to have more severe head injury, this is due to loss of control as has been reported by others(13-14). Thus, specific protection of bicyclists rather than implementation of motor vehicle traffic changes could be the most effective preventive strategy. Although we can not control the type and quality of the helmet, but it is necessary in preventing bicycle related head injury. Besides, comparing what items leads to more severe head injuries would further investigation.

## Conclusion

Many bicycle related injuries can be prevented by changing attitudes and knowledge, and head injuries can be more easily prevented by wearing protective helmets. Legislation,(15-16) together with education(17), should significantly reduce the severity of head injury.

## References

- [1] MacWilliam L, Mao Y, Nicholls E, et al. Fatal accidental childhood injuries in Canada. *Can J Public Health* 1987;78:129-35.
- [2] Brown B, Farley C. The pertinence of promoting the use of bicycle helmets for 8 to 12 year-old children. *Chronic Diseases in Canada* 1989;10:92-5.
- [3] Cushman R, Down J, MacMillan N, et al. Bicycle-related injuries: a survey in pediatric emergency department. *Can Med Assoc J* 1990;143:108-12.
- [4] Laboratory Centre for Disease Control. Children's Hospitals Injury Research and Prevention Program CHIRPP. Technical report #2. Ottawa: Health and Welfare Canada, Laboratory Centre for Disease Control, Bureau of Chronic Disease Epidemiology, 1991.
- [5] Blair L. CMA launches bicycle-helmet campaign hopes to reduce roadside carnage. *Can Med Assoc J* 1991;114:1498-9.
- [6] Shanon A, Bashaw B, Lewis J, et al. Nonfatal childhood injuries: a survey at the Children's Hospital of Eastern Ontario. *Can Med Assoc J* 1992;146:361-5.
- [7] Office for Injury Prevention, British Columbia Ministry of Health. Injury facts and prevention strategies for children and youth in British Columbia. Victoria: Office for Injury Prevention, British Columbia Ministry of Health, September 1993.
- [8] Finvers KA, Strother RT, Mohtadi N. The effect of bicycling helmets in preventing significant bicycle-related injuries in children. *Clin J Sport Med* 1996;6:102-7.
- [9] Schwartz HJ, Brison RJ. Bicycle-related injuries in children. A study in two Ontario emergency departments, 1994. *Chronic Diseases in Canada* 1996;17:56-62.
- [10] Pless IB, Verreault R, Tenina S. A case-control study of pedestrian and bicyclist injuries in childhood. *Am J Public Health* 1996;79:995-9.
- [11] Friede AM, Azzara CV, Gallagher SS, et al. The epidemiology of injuries to bicycle riders. *Pediatr Clin North Am* 1985;32:141-51.
- [12] Thompson RS, Rivara FP, Thompson DC. A case control study of effectiveness of bicycle safety helmets. *N Engl J Med* 1989;320:1361-7.
- [13] Thompson DC, Thompson RS, Rivara FP. Incidence of bicycle-related injury in defined population. *Am J Public Health* 1990;80:1388-9.
- [14] Statts JC, Williamson JE, Whitley T, et al. Bicycle accidents and injuries: a pilot study comparing hospital- and policereported data. *Accid Anal Prev* 1990;22:67-78.
- [15] Baker PB, Li G, Fowler C, et al. Injuries to bicyclists: a national perspective. The Johns Hopkins Injury Prevention Center. *BMJ* 1994;308:173-6.
- [16] National Center for Injury Prevention and Control, Centers for Diseases Control and Prevention. Injury-control recommendations: bicycle helmets. *MMWR Morb Mortal Wkly Rep* 1995;44:1-7.
- [17] Mimaric C, Summers CL, Browning C, et al. Injury patterns in bicyclists attending an accident and emergency department: a comparison of helmet wearers and nonwearers. *BMJ* 1994;308:1537-40.
- [18] Acton H, Nixon JW, Clark RC. Bicycle riding and oral/maxillofacial trauma in young children. *Med J Aust* 1996;165:249-51.
- [19] Scuffham PA, Langley JD. Trends in cycle injury in New Zealand under voluntary helmet use. *Accid Anal Prev* 1997;29:1-9.
- [20] Ginsberg GM, Silverberg DS. A cost-benefit analysis of legislation for bicycle safety helmets in Israel. *Am J Public Health* 1994;84:653-6.
- [21] Nixon J, Clacher R, Pearn J, Corcoran A. Bicycle accidents in childhood. *Br Medical J*. 1987;294:1267-9.
- [22] Sacks JJ, Holmgreen P, Smith SM, Sosin DM. Bicycle-associated head injuries and deaths in the United States from 1984 through 1988. How many are preventable? *JAMA*. 1991;266:3016-8.
- [23] Rosenkranz KM, Sheridan RL. Trauma to adult bicyclists: a growing problem in the urban environment. *Injury*. 2003;34:825-9.
- [24] Eilert-Petersson E, Schelp L. An epidemiological study of bicycle-related injuries. *Accident Analysis & Prevention*. 1997;29:363-72.
- [25] Jacobson GA, Blizzard L, Dwyer T. Bicycle injuries: road trauma is not the only concern. *Australian & New Zealand Journal of Public Health*. 1998;22:451-5.

Table1 Difference of day-time cyclists' head injuries from February 1<sup>st</sup>, 2001 to March 1<sup>st</sup>, 2002

| Character                                  | Severe head injuries<br>(n=90) |         | Mild head injuries<br>(n=234) |         | P value <sup>#</sup> |
|--------------------------------------------|--------------------------------|---------|-------------------------------|---------|----------------------|
|                                            | Number                         | Percent | Number                        | Percent |                      |
| <b>Sex</b>                                 |                                |         |                               |         |                      |
| Female                                     | 45                             | (50.0)  | 147                           | (63.0)  | 0.027                |
| Male                                       | 45                             | (50.0)  | 87                            | (37.0)  |                      |
| <b>Age</b>                                 |                                |         |                               |         |                      |
| 0-9                                        | 32                             | (35.7)  | 91                            | (38.8)  | 0.043                |
| 10-14                                      | 40                             | (45.0)  | 97                            | (41.5)  |                      |
| 15-30                                      | 4                              | (5.2)   | 12                            | (5.3)   |                      |
| 30-45                                      | 6                              | (6.5)   | 24                            | (9.9)   |                      |
| ≥45                                        | 8                              | (7.6)   | 10                            | (4.5)   |                      |
| <b>Bicycle Type</b>                        |                                |         |                               |         |                      |
| Child toy bicycle                          |                                |         |                               |         | 0.040                |
| Child bicycle(with and without aid wheels) | 5                              | (5.7)   | 29                            | (12.5)  |                      |
| Lady bicycle                               | 14                             | (15.5)  | 38                            | (16.5)  |                      |
| Speed-change bicycle                       | 23                             | (25.3)  | 107                           | (45.8)  |                      |
| Mountain bicycle                           | 41                             | (46.0)  | 18                            | (7.3)   |                      |
|                                            | 7                              | (7.6)   | 42                            | (17.9)  |                      |
| <b>Wearing helmets</b>                     |                                |         |                               |         |                      |
| Yes                                        | 10                             | (11.1)  | 10                            | (4.2)   | 0.010                |
| No                                         | 80                             | (88.9)  | 224                           | (95.8)  |                      |
| <b>Wearing other safety equipments</b>     |                                |         |                               |         |                      |
| Yes                                        | 2                              | (2.2)   | 6                             | (2.7)   | 0.823                |
| No                                         | 88                             | (87.8)  | 228                           | (95.3)  |                      |
| <b>Accident causing events</b>             |                                |         |                               |         |                      |
| Large car                                  | 40                             | (45.0)  | 12                            | (5.3)   | 0.043                |
| Compact car or taxi                        | 14                             | (16.5)  | 29                            | (12.5)  |                      |
| Motorcycle                                 | 23                             | (26.3)  | 38                            | (16.5)  |                      |
| Passengers                                 | 4                              | (4.7)   | 107                           | (45.8)  |                      |
| Others                                     | 9                              | (7.6)   | 48                            | (19.9)  |                      |
| <b>Bicycle with reflecting objects</b>     |                                |         |                               |         |                      |
| Yes                                        | 12                             | (13.8)  | 199                           | (85.4)  | 0.004                |
| No                                         | 78                             | (86.2)  | 35                            | (14.6)  |                      |
| <b>Bicycle with bells</b>                  |                                |         |                               |         |                      |
| Yes                                        | 16                             | (17.8)  | 39                            | (14.6)  | 0.754                |
| No                                         | 74                             | (82.2)  | 195                           | (83.4)  |                      |
| <b>Carrying items (bags)</b>               |                                |         |                               |         |                      |
| Yes                                        | 65                             | (73.1)  | 7                             | (2.9)   | 0.000                |
| No                                         | 25                             | (26.9)  | 227                           | (97.1)  |                      |

\*data excluding missing value

\*\*more than one injured body parts on the same patient

# Chi Square or Fisher's Exact Test



Table1 Continue

|                                 | Severe head injuries<br>(n=90) |         | Mild head injuries<br>(n=234) |         | P value <sup>#</sup> |
|---------------------------------|--------------------------------|---------|-------------------------------|---------|----------------------|
| Characters                      | Number                         | Percent | Number                        | Percent |                      |
| <b>Cause of accidents</b>       |                                |         |                               |         |                      |
| Negligence (chat with company ) | 5                              | (5.2)   | 73                            | (31.4)  | 0.031                |
| At high speed                   | 62                             | (68.4)  | 78                            | (33.3)  |                      |
| Cycling at reverse direction    | 23                             | (26.4)  | 83                            | (35.3)  |                      |
| <b>Mechanical problems</b>      |                                |         |                               |         |                      |
| Loss of control                 | 29                             | (31.8)  | 13                            | (5.9)   | 0.024                |
| Broken brake system             | 58                             | (64.4)  | 62                            | (25.9)  |                      |
| Loosing wheel chain             | 3                              | (3.8)   | 159                           | (68.2)  |                      |



Table2 Risk factors of bicycling injured cases having severe head injuries

| Character                                  | Odds ratio | Confidence intervals |
|--------------------------------------------|------------|----------------------|
| <b>Sex</b>                                 |            |                      |
| Female                                     | 1.00       |                      |
| Male                                       | 4.86       | 2.89-5.68            |
| <b>Age group</b>                           |            |                      |
| 0-9                                        | 1.00       | 1.68-4.55            |
| 10-19                                      | 3.85       | 0.18-1.59            |
| 20-30                                      | 1.26       | 0.15-2.99            |
| 30-40                                      | 0.26       | 1.57-2.68            |
| ≥40                                        | 1.65       |                      |
| <b>Bicycle type</b>                        |            |                      |
| Child toy bicycle                          | 1.00       |                      |
| Chile bicycle(with and without aid wheels) | 1.02       | 0.18-1.58            |
| Lady bicycle                               | 0.56       | 0.33-1.98            |
| Speed-change bicycle                       | 2.89       | 1.48-3.13            |
| Mountain bicycle                           | 1.08       | 0.86-2.36            |
| <b>Wearing helmets</b>                     |            |                      |
| Yes                                        | 1.00       |                      |
| No                                         | 4.64       | 1.38-55.68           |
| <b>Wearing other safety equipments</b>     |            |                      |
| Yes                                        | 1.00       |                      |
| No                                         | 1.14       | 0.38-5.68            |
| <b>Bicycle with reflecting objects</b>     |            |                      |
| Yes                                        | 1.00       |                      |
| No                                         | 2.64       | 1.38-4.58            |
| <b>Carrying items (bags)</b>               |            |                      |
| Yes                                        | 1.00       |                      |
| No                                         | 5.26       | 4.23-8.98            |
| <b>Cause of accidents</b>                  |            |                      |
| Negligence(chat with company )             | 1.00       |                      |
| At high speed                              | 2.45       | 1.65-2.88            |
| Cycling at reverse direction               | 1.66       | 1.24-1.98            |
| <b>Mechanical problems</b>                 |            |                      |
| Loss of control                            | 1.00       | 0.18-2.23            |
| Broken brake system                        | 1.21       | 1.25-2.69            |
| Loosing wheel chain                        | 2.13       |                      |



## Posterior walker assisted on gait performance in children with spastic diplegic cerebral palsy

<sup>1</sup>Yu-Ju Chih, yuju519@gmail.com

<sup>2</sup>Chung Huang Yu, chyu@ym.edu.tw

<sup>3</sup>Li-Ru Chen, gracelr@ms1.mmh.org.tw

Institute of Rehabilitation Science and Technology,  
National Yang-Ming University, Taiwan

### Abstract

*To make children who may lose some basic skills to accomplish walking or may have delayed in walking development is a very important issue. Gait in these children may exhibit the following characteristics: poor motor control, poor balance, muscle strength insufficiently, and muscle incoordination. Their mobility may depend upon the use of assistive devices such as canes, walkers, etc. These aids can improve balance and stability or provide body-weight support. They also facilitate forward progression of gait. However, selection and adjustment of walkers are often based on a subjective impression of the children's gait observed by the physician and therapists.*

*There seem not enough information regarding proper selection in choosing the walker design to best suit the children or in comparing the effects on gait load and stability of one design versus the others. Previous researches have studied children with CP and walker usage about lower extremity movement. However, these studies have not quantified upper extremity motions. From the viewpoint of biomechanics, the weight bearing of whole body and the force of upper extremity may influence the usage of walker. So, walker selection requires further study on the full range of upper and lower extremity.*

*The purpose of this study is to analyze the effects of different gait performances of spastic diplegic CP who use different heights of posterior walkers for ambulation.*

*Four spastic diplegic cerebral palsies, of average age 6.75 years, were enrolled in this study. They are all*

*familiarized themselves with posterior walkers. Gait characteristics were evaluated and observed by computer-based kinematics gait analysis using EVaRT 4.6 Motion Analysis, and force analysis by AMTI force plate and force transducers.*

*The conclusion of this study is the different heights of posterior walker would have different effect on gait performance while spastic diplegic cerebral palsy ambulated with walker. In clinic, we should adapt the walkers according to different goals and different abilities of users.*

Keywords: walker, motion analysis, cerebral palsy, gait.

### 1. Introduction

Walking is an important basic skill for human. For the developing children, they can accept more stimuli from environment and other people by walking around. They also can development other skills more maturer and finish many functional activities by walking. Finally, they can be independent in their activities of daily life[1].

Walking is also an important index in childhood development. It means that children from quadruped crawling and biped kneeling to biped walking. It can increase muscle contraction and range of motion of lower extremities, and enhance trunk control[1, 2]. But some children can't do walking well congenitally or acquired.



They may have disability in walking or delay in motor development. Cerebral palsy is a common group in clinic.

Cerebral palsy (CP) is a condition characterized by a motor disorder that is usually diagnosed during the early stages of life. It occurs from brains damage and has symptoms including poor motor control, postural instability, abnormal muscle tone and uncoordinated motor pattern. And it also has some problems in sensation, cognition, speech and psychosocial[3, 4]. Many CP children have difficulties in walking independently because of poor muscle strength, abnormal muscle tone, or poor postural control[5-8]. They often be prescribed with walkers to help stability and support when ambulation, especially spastic diplegic CP[9]. Their cerebral motor cortices have damage and their four limbs are all involved, but the legs more affected than the arms. They have some specific characteristics: increased muscle tone, abnormal motor pattern, joints contracture due to small range of motion, insufficient muscle strength of lower extremity and trunk. In ambulation, they may need some assistive devices to help them. Two types of walker have been advocated for CP in clinic. One is anterior walker, which is positioned in front of the user, the other is posterior walker, which is behind the child. The prescription of the walkers should be well-considered. The ability and limitation of users, the structures and parameters of the walkers, and the usage of the users to walkers are all should be considered. And it should be quantified and objective. However, it is still decided by subjective views of therapists and physician in clinic. So we can't understand how to match the users' goal with the walker parameters and the users' abilities efficiently and currently.

Many researchers have agreed that the posterior walker is better than anterior walker, in the light of upright postures, decreased double support time and increased velocity[10-14].

The purpose of this study is to analyze the gait performance of children with spastic diplegic cerebral

palsy when these children use posterior walker with different heights.

## 2. Methods

### 2.1 Subjects

A convenience sample of four children (2 males, 2 females) with the diagnosis of spastic diplegic CP were studied. Their mean age was 6.75 years, mean height was 116.25cm, their mean height from greater trochanter to floor was 65.5cm, and their mean body weight was 21kg. Inclusion criteria was as follows:

- (1) Spastic diplegic cerebral palsy.
- (2) Use posterior walker as walking aid.
- (3) Can understand verbal commands.
- (4) Be able to ambulate with posterior walker at least 5 meters.
- (5) No orthopedic surgery in 6 months.

Informed consent was obtains from the parents of all subjects before study.

### 2.2 Experiment setup

The posterior walker used in this study had been modified by adding two handle force transducers. The posterior walker had four wheels. The front wheels were limited anterior direction, the back could only roll forwards. The height of handle to the floor could be adjusted from 58cm to 77cm.

The kinematics data of upper extremity and lower extremity was recorded using Evart 4.6 motion analysis system with a 10 cameras system.

### 2.3 Procedures

The posterior walker were adjusted in five different heights: the standard (N)(the height of great trochanter is





equal to the height of walker's handle), one inch higher (H1), two inches higher (H2), one inch lower (L1), and two inches lower (L2). The subject would wear underclothes and couldn't have any object that would reflect infrared in their bodies. There was a 7-meters walkway in this study. Before testing, the subjects would have opportunities to practice walking and the staff would calibrate the hardware and the environment. The subjects ought to have at least three effective trails for every height, once with each height in random order. They walked at a self-selected comfortable speed. They would have rested for 10 minutes after every height. The markers would be stuck on the subjects' bodies by the same staff for higher reliability. The reflective markers on the subjects were attached to the following anatomic locations: head markers at top, lateral, front and back, bilateral shoulder markers at the acromion process, bilateral elbow markers at the olecranon, bilateral wrist markers at the ulnar styloid, sacral marker at the midpoint on the line between both posterior superior iliac spines, bilateral greater trochanter markers at the lateral greater trochanter, bilateral pelvic markers at both anterior superior iliac spines, bilateral thigh markers at the midpoint on the line between the great trochanter and the center of the knee, bilateral knee markers at the lateral knee joint, bilateral shank markers at the midpoint on the line between the lateral knee joint and the later malleolus, bilateral ankle markers at lateral malleolus, bilateral heel markers at , and bilateral forefoot markers at the 2<sup>nd</sup> metatarsal head.

## 2.4 Data analysis

There are high heterogeneities between different CP subjects. So, this study would compare the effect in different heights of the same subject.

The variables of velocity, cadence, and step length were calculated:

(1). Velocity: calculated by subtracting the X coordinate

of the sacral marker at the end of a gait cycle from its X coordinate at the beginning of a gait cycle and then dividing the times of this cycle.

(2). Cadence: calculated by subtracting the times of the end of a gait cycle from its beginning of a gait cycle and then dividing the steps of this gait cycle.

(3). Step length: the distance was from one foot-floor contact to the foot-floor contact of opposite leg. It calculated by subtracting the X coordinate of the second foot-floor contact from the first one.

## 3 Results

### Temporo-spatial variable on gait performance

The results are showed in

Table 1~Table 3

Subject A and D had poor motor control in walking with walker., and subject B and C had good motor control in walking with walker. On gait performance, we could find that the better performances of all subjects were the same.

Subject A: had the faster average velocity, longest step length and highest cadence in L1 walker.

Subject B: had the faster average velocity, longest step length and highest cadence in H1 walker.

Subject C: had the faster average velocity, longest step length and highest cadence in N walker.

Subject D: had the faster average velocity, longest step length and highest cadence in H2 walker.





Table 1 Average velocity

|    | A                       | B                       | C                       | D                       |
|----|-------------------------|-------------------------|-------------------------|-------------------------|
| H2 | <sup>m</sup> 27.74±2.89 | 11.79±1.02              | 19.89±1.23              | <sup>M</sup> 28.91±6.44 |
| H1 | 43.63±16.34             | <sup>M</sup> 14.06±0.70 | 17.69±0.91              | <sup>m</sup> 20.17±1.38 |
| N  | 40.67±6.62              | 12.54±2.57              | <sup>M</sup> 33.31±1.78 | 24.04±5.80              |
| L1 | <sup>M</sup> 47.53±1.70 | 10.99±2.47              | 20.84±2.18              | <sup>m</sup> 20.17±6.28 |
| L2 | 36.69±0.44              | <sup>m</sup> 7.71±2.28  | <sup>m</sup> 17.64±2.50 | 22.36±6.50              |

Values=mean±SD

(m/min)

<sup>M</sup> is maximum, <sup>m</sup> is minimum

Table 2 Step lengths

|    | A                        |                          | B                        |                          | C                        |                          | D                        |                          |
|----|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
|    | Left side                | Right side               | Left side                | Right side               | Left side                | Right side               | Left side                | Right side               |
| H2 | 0.346±0.024              | 0.341±0.013              | 0.274±0.011              | 0.286±0.005              | 0.397±0.034              | 0.400±0.017              | <sup>M</sup> 0.543±0.028 | <sup>M</sup> 0.567±0.028 |
| H1 | 0.347±0.026              | 0.353±0.005              | <sup>M</sup> 0.306±0.009 | <sup>M</sup> 0.302±0.010 | 0.396±0.018              | <sup>m</sup> 0.383±0.014 | <sup>m</sup> 0.438±0.010 | <sup>m</sup> 0.441±0.007 |
| N  | 0.333±0.030              | 0.347±0.034              | <sup>m</sup> 0.273±0.018 | <sup>m</sup> 0.268±0.042 | <sup>M</sup> 0.452±0.031 | <sup>M</sup> 0.473±0.031 | 0.522±0.041              | 0.536±0.057              |
| L1 | <sup>M</sup> 0.357±0.022 | <sup>M</sup> 0.372±0.009 | 0.295±0.008              | 0.284±0.023              | 0.404±0.032              | 0.405±0.034              | 0.449±0.067              | 0.477±0.070              |
| L2 | <sup>m</sup> 0.326±0.005 | <sup>m</sup> 0.332±0.011 | 0.281±0.001              | 0.275±0.022              | <sup>m</sup> 0.385±0.014 | 0.392±0.009              | 0.499±0.046              | 0.488±0.081              |

Values=mean±SD

(m)

<sup>M</sup> is maximum, <sup>m</sup> is minimum

Table 3 Cadence

|    | A                       | B                       | C                       | D                       |
|----|-------------------------|-------------------------|-------------------------|-------------------------|
| H2 | <sup>m</sup> 40.74±5.04 | 27.80±3.91              | 41.45±1.58              | <sup>M</sup> 38.06±6.55 |
| H1 | 59.92±4.65              | <sup>M</sup> 41.09±0.27 | 37.93±3.43              | 36.81±2.26              |
| N  | 58.41±4.56              | 33.82±1.79              | <sup>M</sup> 58.19±2.00 | 37.05±5.30              |
| L1 | <sup>M</sup> 63.88±4.54 | 39.01±0.66              | 40.02±4.71              | 34.60±6.77              |
| L2 | 54.84±1.57              | <sup>m</sup> 19.75±3.26 | <sup>m</sup> 37.64±5.45 | <sup>m</sup> 34.03±4.36 |

Values=mean±SD

(steps/min)

<sup>M</sup> is maximum, <sup>m</sup> is minimum

## 4 Discussion

The height of the walker was one of the convenient way to adjust the walker for different needs or goals and different abilities of users. From the results of this study, we could find that the different heights have the different effects for individuals, although the regularizations between subjects were unobviously, but we could find the suitable height for the subject in different goals.

The samples in this study were a little small, and the subjects' abilities were too divers. The grading of subject's criteria wasn't detail enough, so the results of this study couldn't apply to the subjects of spastic diplegic cerebral palsy.

On gait performance, if the functional performances were our goals, the users would have faster walking velocity, longer step lengths and higher cadence that could let them walk smooth and steady. In our study, we could find that the standard height that used in clinic wasn't the most suitable height for our subjects.

Future works can have more subjects, and the ability of the subject can be more similar and detail to have agreeable results.

### Acknowledgement

This study was aided by department of rehabilitation of Mackay Memorial Hospital.

## 5. References

1. Bax, M., *walking*. Developmental Medicine & Child Neurology, 1991. **33**(6): p. 471-2.
2. Rose, J. and J.G. Gamble, *Human walking*. 2nd ed. Ch.1 Human Locomotion, ed. H.J.R. Veene T. Inman, and Frank Todd. 1994, Baltimore: Williams and Wilkins.
3. Gage, J.R., *Gait analysis in cerebral palsy*. Ch.1 Introduction. 1991, London: Mac Keity Press.
4. Levitt, S., *Treatment of cerebral palsy and motor delay*. 4th ed. Ch.1 The clinical picture for therapy and management. 2004, Oxford [England]; Cambridge, Mass., USA: Blackwell Publishing.
5. Scholtes, V.A., et al., *Clinical assessment of spasticity in children with cerebral palsy: a critical review of available instruments*. Dev Med Child Neurol, 2006. **48**(1): p. 64-73.
6. Leonard, C.T., H. Hirschfeld, and H. Forssberg, *The development of independent walking in children with cerebral palsy*. Dev Med Child Neurol, 1991. **33**(7): p. 567-77.
7. Bottos, M. and C. Gericke, *Ambulatory capacity in cerebral palsy: prognostic criteria and consequences for intervention*. Dev Med Child Neurol, 2003. **45**(11): p. 786-90.
8. Stanger, M. and S. Oresic, *Rehabilitation approaches for children with cerebral palsy: overview*. J Child Neurol, 2003. **18 Suppl 1**: p. S79-88.
9. Sussman, M.D. and M.D. Aiona, *Treatment of spastic diplegia in patients with cerebral palsy*. J Pediatr Orthop B, 2004. **13**(2): p. S1-12.
10. Sutherland, D.H., et al., *The development of mature gait*. J Bone Joint Surg Am, 1980. **62**(3): p. 336-53.
11. Rose, J., J.M. Medeiros, and R. Parker, *Energy cost index as an estimate of energy expenditure of cerebral-palsied children during assisted ambulation*. Developmental Medicine & Child Neurology, 1985. **27**(4): p. 485-90.
12. Levangie, P., Chimera M, Johnson M, Robinson F, Wobesky L., *The effects of posterior rolling walkers vs. the standard rolling walker on gait characteristics of children with spastic cerebral palsy*. Phys Occup Ther Pediatr, 1989: p. 1-17.
13. Logan, L., C.D. Ciccone, and K. Byers-Hinkley,







*Anterior versus posterior walkers: a gait analysis study.* Developmental Medicine & Child Neurology, 1990. **32**(12): p. 1044-8.

14. Greiner, B.M., J.M. Czerniecki, and J.C. Deitz, *Gait parameters of children with spastic diplegia: a comparison of effects of posterior and anterior walkers.* Archives of Physical Medicine & Rehabilitation, 1993. **74**(4): p. 381-5.





## Adaptable Access Control for Electronic Medical Records

Kung Chen<sup>a</sup>, Yuan-Chun Chang<sup>a</sup>, and Da-Wei Wang<sup>b</sup>

<sup>a</sup>Department of Computer Science, National Chengchi University

<sup>b</sup>Institute of Information Science, Academia Sinica & National Yang-Ming University

<sup>a</sup>{chenk,g9305}@cs.nccu.edu.tw; <sup>b</sup>wdw@iis.sinica.edu.tw

### Abstract

*This paper presents an aspect-oriented approach to providing adaptable access control framework for Electronic Medical Records (EMR) on Web-based platform. In our scheme, access control logic is decoupled from the core of application and collected into separate aspect modules which are automatically synthesized from access control rules in XML format and properly designed aspect templates. The generated aspect modules will then be compiled and integrated into the underlying application using standard aspect tools. At runtime, these aspect codes will be executed to enforce the required access control without any runtime interpretation overhead. Future changes of access control rules can also be effectively realized through these mechanisms without actual coding. This will not only improve the system's modularity but also make the task of enforcing comprehensive access control more adaptable. We have built a Web-based EMR prototype implementation using AspectJ to demonstrate our approach.*

**Keywords:** access control, electronic medical record, aspect-oriented programming

### 1. Introduction

There is little doubt that healthcare information systems (HIS) will move towards a fully integrated electronic medical record (EMR). However, as we move closer to a paperless environment and Internet-based applications, we must realize that the risks to privacy and security incurred by using electronic systems are also increased. This is a very complicated issue. Besides social aspects, it also calls for much technological advancement to achieve a proper balance between public interests and personal privacy. In terms of information system development, an effective and flexible access control mechanism is clearly an essential part of a secure and privacy enhanced EMR system [4].

However, it is not easy to develop a comprehensive yet flexible mechanism for access control in HIS with EMR. There are at least two major difficulties. First, like other security requirements, access control is a system-wide concern that permeates through all the major modules of a system. Hence it is very often to see the code for implementing access control scattered over the whole system and tangled with other functional code.

This is not only error-prone but also makes it difficult to verify its correctness and perform the needed maintenance.

Second, access control rules in healthcare domain are inherently fine-grained and dynamic. It is common for information system developers to partition users into different categories, e.g. by roles in an organization, and define access privileges in terms of the application functions that a particular category of users has a right to access, e.g., an administrative clerk is limited to administrative functions and excluded from transaction functions. For HIS, however, an additional level of access control must be defined at the *data field* level. Users may be allowed to access a specific function, but some data elements may be excluded from view. Furthermore, access may be limited to specific patient records or specific elements within a patient record. For example, while a physician can view some fields of a patient's EMR, only the patient's *attending* physician can see the whole record and modify it. On the other hand, we will have to bypass the constraint in an emergency. In addition, changes in legislation or changes in the interpretation of legislation can lead to major revision of the access control rules. All these lead to the needs of frequent changes for access control rules in healthcare domain.

An ideal approach to overcome these difficulties, referred to as *adaptable access control* [9][10], is to provide a framework where the access control logic is separated from the core of application and specified declaratively in a configuration file without actual coding, while still being able to satisfy fine-grained access control requirements and incurring a low runtime overhead. This will not only improve the system's modularity but also make the task of enforcing comprehensive access control more adaptable. Here we show that the emerging technology of aspect-oriented programming (AOP) [6] is a feasible technology to providing adaptable access control. Specifically, we propose an aspect-oriented framework for adaptable access control and present a Web-based EMR prototype implementation using AspectJ [7] to demonstrate our approach.

The remainder of the paper is organized as follows. Section 2 introduces AOP and AspectJ briefly. Section 3 outlines our approach to adaptable access control and overviews our system architecture. Section 4 and 5 explain our design of access control rules and aspect code templates, followed by a description of the main



steps for synthesizing the access control aspects in Section 6. Our prototype implementation of a Web-based EMR is presented in Section 7. Finally, Section 8 concludes and outline our future work.

## 2. Background: AOP and AspectJ

In this section, for the sake of completeness, we highlight the basics of AOP and review the relevant features of AspectJ.

Aspect-oriented programming (AOP) aims at modularizing concerns such as profiling and security that crosscut the components of a software system [9]. In AOP, a program consists of many functional modules and some *aspects* that encapsulate the crosscutting concerns. An aspect module provides two kinds of specifications: *Pointcut*, comprising a set of well-defined points in the execution of a program, designate when and where to crosscut other modules; and *advice*, which is a piece of code, that will be executed when a pointcut is reached. The complete program behavior is derived by some novel ways of composing functional modules and aspects according to the specifications given within the aspects. This is called *weaving* in AOP. Weaving results in the behavior of those functional modules impacted by aspects being modified accordingly.

AspectJ is a seamless aspect-oriented extension to the Java programming language [8]. It provides a rich pointcut language through a powerful join point model. Typical join points in AspectJ are method call/execution and field access. Advice in AspectJ is an anonymous method bound to a pointcut and tagged with one of the three keywords: *before*, *after*, or *around*. The before advice and the after advice are executed before and after the intercepted method, respectively. The case for the around advice is more subtle. Inside the around advice, we can choose to resume the intercepted method by calling the special built-in method *proceed()*, or simply bypass its execution. The following aspect in AspectJ illustrates the power of around advice. It states that, when the update method of class `Patient` is about to execute, control will be transferred to the around advice. If the particular constraint is not satisfied, the intercepted method call will be aborted and the exception handler will take over; otherwise, it will be resumed by calling *proceed()*.

```
public aspect AccessControlPrecheck {
    pointcut pc(Data d):
        execution(public void
            Patient.update(Data d)) && args(d);
    void around(Data d) : pc(d){//advice
        if (constraint(d) )
            proceed(d); // granted
        else forwardToExceptionHandler
            ("AccessDenied");
    } // end around advice }
```

Note that “args(d)” is also a kind of pointcut in AspectJ that can capture the argument(s) passed to the intercepted method.

Furthermore, AspectJ also allows aspect inheritance, abstract aspect, and abstract pointcut. We can write an aspect with abstract pointcuts or abstract methods. A sub-aspect then extends the abstract aspect and defines the concrete pointcuts and methods.

## 3. Our Approach and System Architecture

We propose to apply aspect-oriented technology to design and implement a highly adaptable mechanism for enforcing fine-grained access control in Web-based EMR system. We worked toward this goal from two opposite ends and managed to meet in the middle. At one end, the objective is to accommodate more requirements. We use a flexible modeling scheme based on the tree-structure of EMR and user security attributes that can satisfy a wide range of access control requirements of various granularity levels, including those involving data contents. A high-level form of access control rules is derived from it. At the other end, since access control is a system-wide crosscutting concern, we must impose considerable architectural disciplines on Web applications to layout a good foundation for enforcing the required access control modularly. In particular, we follow the well-accepted Model-View-Controller (MVC) [3] architectural pattern and adopt the popular Apache Struts framework [1] to structure our Web applications.

Next, we apply AOP to devise a flexible implementation scheme that bridges these two ends. We developed our implementation scheme in two stages. In the first stage, we did an in-depth analysis of the structures of aspect code that we developed manually for implementing the forms of access control rule we employed. Such aspect code is classified into a few forms according to their internal structure. Basically, each access control aspect is divided into two parts: a generic part realized by an abstract aspect and a rule specific part realized by a concrete aspect. Indeed, these abstract aspects provide a solid basis towards building an adaptable mechanism.

In the second stage, we focus on how to automatically synthesize aspect code from access control rules. Given the abstract aspects derived in the previous stage, we only need to generate the parts that are rule-specific. Thus we prepared some aspect templates based on the derived aspect code structure to assist code generation. On the source side, in addition to the access control rules, we provide an application specification file that links the logical entities (data objects and operations) referenced in the rules to the real entities defined in the underlying application. Following the current practices of Web application development, we define both of the two input files in XML format and treat them as configuration files, one of each type for every application. Together with the pre-defined aspect templates, the two XML configuration files are processed by a rule translator into concrete access control aspects.



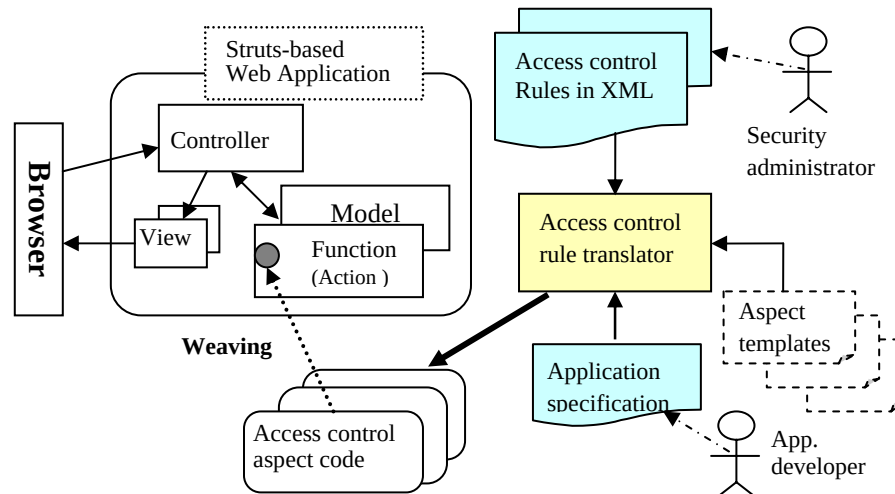


Figure 1. System architecture and mechanisms for adaptable access control

The generated aspect modules will then be compiled and woven into designated functional modules of the underlying Web application using standard aspect tools. At runtime, the aspect code will be executed like common functional code to enforce the required access control. Figure 1 illustrates the system architecture and mechanisms of our approach to providing adaptable access control. Security administrators are responsible for encoding the access control requirements of EMR using our rule format and the system developers need to supply the application specification file for linking the rules with underlying system to facilitate rule translation and aspect code generation.

#### 4. Access Control Modeling and Rules

Access control, also known as *authorization*, is a process by which users, after being identified, are granted certain privileges to information, system functions, or resources. The step to identify a user is usually called *authentication*. Username and Password check is the most basic form of authentication, while digital certificates and biometrics are more advanced ways of authentication.

Role-based access control (RBAC) [8] is the most often cited guiding principle underlying all the proposed solutions to application-level access control. RBAC uses the role abstraction as its primary means of specifying access control requirements. However, as we have argued earlier, roles alone are not sufficient for our purposes; proper access control for EMR usually requires finer grained constraints that are derived from other user attributes and data contents. To handle such fine-grained requirements, we have taken a generic and an attribute-based scheme that is more expressive than RBAC for specifying access control requirements.

We model the interaction between a user and a Web-based EMR application as a sequence of access tuples of three elements:  $\langle \text{user}, \text{action}, \text{data} \rangle$ , indicating a user's request to conduct an action, such as view or

update, on a specific data component of EMR. The access control rules of an EMR system determine which access tuples are allowed and which must be denied. They are derived from the system's access control requirements.

Since EMR is our focus here, the data component of an access tuple plays the central role in specifying access control rules. It must not only facilitate the specification of access constraints of different actions on various parts of EMR but also lend to an easy integration with the underlying HIS. Here we follow the proposal of Taiwan Electronic Medical Record Template (TMT) [5], which aims to provide a suite of standard forms that will become the common basis for developing EMR systems in Taiwan. In TMT, all the data of EMR are categorized and divided into many inter-related documents called *forms*. Each TMT form has a specific name and its data are further grouped into a hierarchy of *modules*. Following TMT, our access control rules are designed to be *form-based* and treat each form as a tree organized into a hierarchy of modules. Specifically, our access control rules take the following abstract format:

```
<FormName, AuthType,
  ModPath1:
    Actions : Constraint;
    ...
  ...
  ModPathn:
    Actions : Constraint;
    ...
>
```

As authentication is required prior to authorization, we also make authentication type (AuthType) part of the rule; the type can be id/password (PWD), health digital certificate (HCA), or any other supported methods of user identification. The *FormName* refers to a specific form of TMT, and a *ModPath* is a tree path specifier referring to a particular module of the form that must be protected against any unauthorized actions. A special





module path called root, represented by “/”, refers to the whole form. The *constraint* is a Boolean expression which must be evaluated to true to grant the attempted execution of the associated actions. Inside a form, the accessibility of any module is defined in a cascading manner according to its position in the tree structured form:

- If the accessibility of the module is explicitly defined in the rule, use the defined accessibility;
- Otherwise use the accessibility of the nearest ancestor that is explicitly set.
- If no ancestor has accessibility explicitly defined, make the module always accessible.

To facilitate the expression of an access constraint without incurring the unnecessary details of the underlying system, we take an object-based approach and supply five generic objects: *User*, *Form*, *Data*, *Cxt*, and *App*, with various attributes that the constraint expression can refer to. Conceptually, the *Form* object and the *Data* object serve as the input and output of an action to execute, respectively. Typical attributes for the *User* object include user's name and roles in an organization. The attributes of the *Form* object include the arguments passed to the protected action, while the attributes of the *Data* object refer to the fields of any modules of a TMT form. The context object (*Cxt*) provides information stored on the server, such as the current time and the list of doctors on duty during off-hours. Like the context object, the application object (*App*) is also a global object which keeps various parameters related to access control yet specific to an application. Some typical attributes of the *App* object include duration of office hours and the IP address of machines dedicated to certain purposes. Clearly, the specific set of attributes for each object depends on individual application's needs. For example, the following rule dictates that, for TMT-Form-N, while all medical staff can read all parts of it, except module /A/B, only doctors can create and update it. Furthermore, for module A of the form, any updates must be done by the patient's doctor-in-charge or by doctors on duty during off-hours from certain dedicated machines. And the data in module /A/B can be read only during normal office hours.

```
<TMT-Form-N AuthType=PassWD
  path=:
    CREATE, UPDATE:
      "Doctor" in User.roles
    READ: "Staff" in User.roles
  path=/A:
    UPDATE:
      User.name == Data.doctorInCharge or
      (User.machineIP in App.dedicatedMachines
      and User.doc_ID in Cxt.docOnDuty);
  path=/A/B:
    READ:
      Cxt.currentTime in App.officeHours
>
```

The above example illustrates the features of our access control rules. In particular, we can associate multiple kinds of attributes with the generic objects, which make it easy to model a multitude of requirements, from simple RBAC to sophisticated instance and module level constraints with little effort.

## 5. Aspect Templates for Access Control

We provide both authentication and authorization aspects. Due to space limitation, we shall focus on authorization aspects which are more complicated than authentication aspects.

Every module action that requires authorization will be associated with an access constraint, which in turn will be converted to real Java code and integrated into a proper access control aspect for enforcing the required control. As stated earlier, to facilitate the generation of access control aspects, an aspect is divided into two parts: generic part realized by abstract aspects and rule specific part realized by concrete aspects. The generic part captures the common code patterns one would develop manually to enforce an access control rule and forms the basis of an aspect template. After some analysis of the code pattern, we identified three most typical generic aspects, namely, *Precheck*, *PostfilterSingle*, *PostfilterCollection*. The availability of the data entities referenced in the constraint expression of a rule, such as user roles, action arguments and data contents, distinguishes these generic aspects. The pre-checking aspect handles the case when the constraint expression involves only the *User* and *Form* objects, whose attributes are all available before executing the protected function. In contrast, the post-filtering aspects are used for the cases when the constraint expression also involves attributes of the *Data* object, which are available only after executing the protected action. Furthermore, as the code structure for handling a single record retrieved by key-based queries is different from that of handling a collection of data obtained by ad-hoc queries, we provide two kinds of post-filtering aspects.

For spaces sake, Listing 1 displays only the code skeleton of the *PostfilterSingle* aspect and its associated template. Each descendant of such an aspect will enforce the access constraints associated with an action on a particular form. The specific association will be established by supplying the concrete pointcut definition in the template. The entry to execute this aspect is the *around* advice. This aspect first checks if certain constraint, *preCondition(..)*, is met before proceeding to execute the designated action. Afterwards, it applies the constraint, *postCondition(..)*, to decide whether the user is authorized to access this particular instance of patient data. Both of these two conditions are derived from the access constraints specified for the *root module* of the underlying TMT form. Finally, before returning it to the user, the aspect needs to filter out any modules of this form that are not accessible to the user. This is accomplished by invoking the *moduleFilter(..)*







method, which will traverse the tree-structured form to perform the filtering task.

Listing 1: The *PostfilterSingle* aspect and its template

```
public abstract aspect PostFilterSingle
    extends Authorization {
    abstract pointcut pc;
    abstract boolean
        preCondition(HttpServletRequest req);
    abstract protected boolean
        postCondition(HttpServletRequest req);
    ActionForward forwardToErrorPage(..){...}
    void moduleFilter(... req){ //template}
    ActionForward around(...)
        :pc(mapping,form,request,response) {
        if( preCondition(request) ){
            ActionForward forward =
                proceed(mapping,form,request,response);
            //instance (form) level filtering
            if( postCondition(request) ){
                // module-level filter
                moduleFilter(request);
                return forward;
            } //end post
        } //end pre
        return //access denied
            forwardToErrorPage(request,mapping);
    }
}
//template to be instantiated
public aspect <aspectName>
    extends Postfilter {
    pointcut pc(..): <pointcut> && args(..);
    <_Library> // library functions
    Boolean preCondition(... req)
    { <Constraint from the root module> }
    boolean postCondition(... req)
    { <Constraint from the root module> }
    void moduleFilter(... req) // visitor pattern
    { <module-level constraints> }
}
```

Since our access control rules are specified in terms of the TMT form, we have a rather stable structure of data to traverse for conducting the module filtering work. Therefore, to facilitate the code generation of the module filter part, we have chosen to implement the module filters using the famous visitor pattern [3], which provides the ability to add new operations to existing object structures without modifying those structures. This advantage of visitor pattern is exactly what we expect our module filter to achieve: traversing various tree-structures and performing different operations, such as masking a module, for access control purpose.

## 6. Synthesizing Access Control Aspects

This section sketches our scheme for synthesizing access control aspects. Figure 2 illustrates the major steps of our aspect synthesizing process. Besides aspect templates, there are two configuration files that serve as the input to our aspect synthesizer: the access control rules and application specification. Following the current practice of designing configuration files for Web-based

applications, both input files are formatted using an XML schema, respectively. After loading these two files and perform the required sanity check, the synthesizer will enter a loop to process every enforce-point specified in the access control rules.

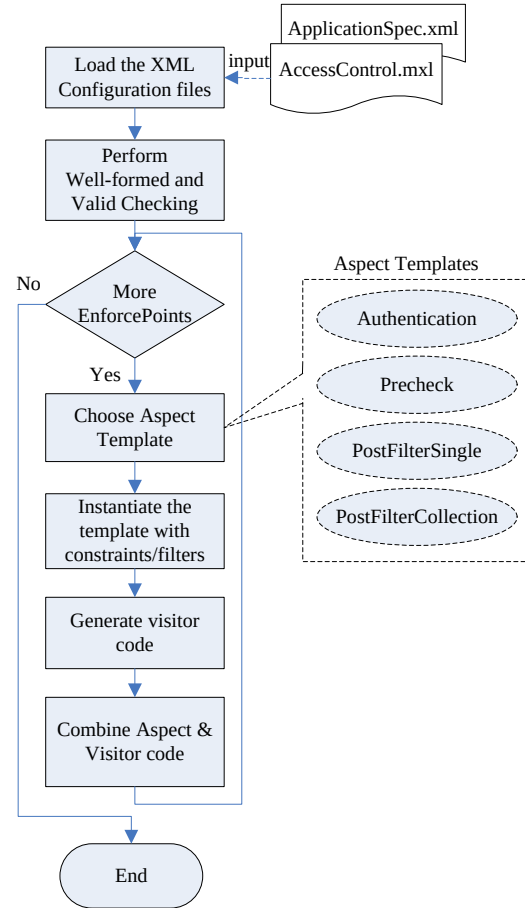


Figure 2. The major steps of the aspect synthesizer

An enforce-point corresponds to an action that needs to be protected for the *root module* of a particular TNT-form. It is the main item specified in the access control XML configuration file. For example, the action of creating a *DischargeNote* form is authorized to only users with a certain role, hence it will be represented as an enforce-point. The constraint associated with that action will become the pre-condition and post-condition of the matching enforce-point. Together with other related module-level actions and constraints, an aspect that inherits from a proper aspect template will be synthesized to realize the access control required by a particular enforce-point.

Figure 3 highlights the major parts of the XML schema for specifying the access control rules. In encoding the form-based access control rules using our schema, enforce-points of the same authentication type requirement will be grouped together under an enforce-domain, as different authentication types imply different security levels. Besides, to facilitate sharing the module-level constraints across module actions, we have prepared a separate group of elements called *formFilters*,







>>> 病歷列表 >> 出院病歷單 >> 編輯出院病歷單

|                  |                |          |           |           |    |
|------------------|----------------|----------|-----------|-----------|----|
| <b>■ 出院病歷單</b>   |                |          |           |           |    |
| * 醫院名稱           | 台北醫學大學附設醫院     |          | * 醫院代號    | 000000001 |    |
| <b>個人基本資料</b>    |                |          |           |           |    |
| * 姓名             | 王明月            | * 病歷號碼   | 000000001 | 年齡        | 30 |
| * 身分證字號          | A123456789     |          |           |           |    |
| <b>出入醫院資料</b>    |                |          |           |           |    |
| 轉入醫院地址           | 台中縣大里市東榮路 483號 | 轉入醫院流水編號 | 000000005 |           |    |
| 入院日期             | 2005/08/12     | 入院科別     | 家庭醫學      | 入院床號      | 20 |
| 轉科日期             | 2005/08/05     | 轉科科別     | 耳鼻喉科      | 轉科床號      | 50 |
| 出院日期             | 2005/08/10     | 出院科別     | 耳鼻喉科      | 出院床號      | 52 |
| <b>■ 疾病診斷模組</b>  |                |          |           |           |    |
| <b>入院時</b>       |                |          |           |           |    |
| 序號               | 疾病名稱           | 編碼系統     | 編碼        | 備註        |    |
| 1                | 耳炎症            | 0001     | 0001      |           |    |
| <b>出院時</b>       |                |          |           |           |    |
| 序號               | 疾病名稱           | 編碼系統     | 編碼        | 備註        |    |
| 1                | 內聽道病變          | 0002     | 0002      |           |    |
| <b>■ 主訴模組</b>    |                |          |           |           |    |
| 描述               | *****          |          |           |           |    |
| <b>■ 出院時管況模組</b> |                |          |           |           |    |
| 出院情況             | 治癒出院           |          |           |           |    |

Figure 5. Screenshot of the prototype implementation for the discharge note of TMT

also plan to develop a rule analyzer for detecting such inconsistencies while synthesizing the aspect code.

## 8. Conclusions and Future Work

In this paper, we have presented an aspect-oriented approach to providing adaptable access control for Electronic Medical Records on Web-based systems. We followed the Taiwan Electronic Medical Record Template to design our access control scheme so that our scheme will integrate easily to future HIS systems based on this standard. Our access control modeling scheme can satisfy a wide range of requirements with different granularity. By employing aspect-oriented technology, we have obtained a highly modular implementation of fine-grained access control and the aspect code for enforcing access control is automatically synthesized. Future changes of access control rules can also be effectively realized through these mechanisms without actual coding. This will not only improve the system's modularity but also make the task of enforcing comprehensive access control more adaptable. We argue that our scheme has achieved an adequate balance between expressiveness of access control and runtime efficiency. We have also built a Web-based EMR prototype implementation using AspectJ to demonstrate our approach.

We plan to further explore this line of study along a few directions. First, besides the authentication type specified for a TMT form, we shall extend our access control rules to enable stricter authentication for individual modules. Second, we plan to design and implement a rule development tool that helps security administrator convert the abstract access control rules into the XML format. Currently, this is done manually, which is not only error-prone but also difficult to maintain. Third, once the numbers of rules increase to a degree, it is very likely that we will mistakenly create some inconsistencies between related rules that may creep into the aspect code without notice. Therefore, we

## 9. References

- [1] The Apache Struts Web Application Framework: <http://struts.apache.org/>
- [2] K. Chen and D.W. Wang. "Toward Configurable Access Control for Healthcare Information Systems", *MIST 2004*.
- [3] Gamma, Helm, Johnson and Vlissides, *Design Patterns*, Addison-Wesley, 1995.
- [4] HIPAA. <http://www.cms.hhs.gov/hipaa>.
- [5] W.S. Jian et al., The Development of Taiwan Electronic Medical Record Template, Draft. Templates download : <http://emr.doh.gov.tw/html/kaisya-annai.html>
- [6] G. Kiczales, J. Lamping, A. Menhdhekar, C. Maeda, C. Lopes, J.-M. Loingtier, and J. Irwin (1997), "Aspect-Oriented Programming," in *ECOOP '97*, LNCS 1241, pp. 220-242.
- [7] G. Kiczales, E. Hilsdale, J. Hugunin, M. Kersten, J. Palm, and W.G. Griswold (2001), "Getting Started with AspectJ", *Comm. of ACM*, vol. 44, no. 10, pp 59-65.
- [8] R. Sandhu, E. Coyne, H. Feinstein, and C. Youman (1996), "Role-based access control models," *IEEE Computer*, 29(2):38-47, 1996.
- [9] T. Verhanneman, L. Jaco, B. De Win, F. Piessens, and W. Joosen (2003), "Adaptable Access Control Policies for Medical Information Systems," *Proc. of Distributed Applications and Interoperable Systems*, 2003, LNCS 2893, pp. 133-140.
- [10] B. De Win, F. Piessens, W. Joosen and T. Verhanneman (2002), "On the importance of the separation-of-concerns principle in secure software engineering," *ACSA Workshop on the Application of Engineering Principles to System Security Design*, pp. 1-10, Boston, Massachusetts.





### Appendix: An example of AccessControl.xml file

```
<AccessControl xmlns=http://www.cs.nccu.edu.tw/plsm/2005/AccessControl
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" >
  <EnforceDomain name="EMR">
    <AuthType value="PWD"/>
    <EnforcePoints>
      <EnforcePoint name="WriteTMTFormN">
        <Rule>
          <Precondition>
            _Lib.contains($User.getAttr("Role"), "Doctor")
          </Precondition>
        </Rule>
      </EnforcePoint>
      <EnforcePoint name="CreateTMTFormN">
        <Rule>
          <Precondition>
            _Lib.contains($User.getAttr("Role"), "Doctor")
          </Precondition>
        </Rule>
      </EnforcePoint>
      <EnforcePoint name="ReadTMTFormN">
        <Rule>
          <Postcondition>
            _Lib.contains($User.getAttr("Role"), "Staff")
          </Postcondition>
        </Rule>
      </EnforcePoint>
    </EnforcePoints>
    <FormFilters>
      <FormFilter form-name="TMTFormN">
        <Module name="A">
          <Rule>
            <Action name="write"/>
            <Constraint>
              _Lib.eq($User.getAttr("Name"), $Data.getAttr("doctorInCharge")) or
              ( _Lib.contains( $App.getAttr("dedicatedRooms"), $User.getAttr("machineIP") and
                _Lib.contains( $Cxt.getAttr("docOnDuty"), $User.getAttr("doc_ID")) )
            </Constraint>
          </Rule>
        </Module>
        <Module name="B">
          <Rule>
            <Action name="read"/>
            <Constraint>
              _Lib.contains($App.getAttr("officeHours"), $Cxt.getAttr("currentTime"))
            </Constraint>
          </Rule>
        </Module>
      </FormFilter>
    </FormFilters>
  </EnforceDomain>
</AccessControl>
```





## MAC Address-Based Position Detection of Illegal Access Terminals on Private Network for Hospital Asset Management

Naoto KUME

*Graduate School of  
Informatics, Kyoto University,  
JSPS Research Fellow,  
Kyoto, Japan.  
kume@kuhp.kyoto-u.ac.jp*

Tomohiro KURODA

*Department of Information  
Processing Science,  
University of Oulu,  
Oulu, Finland.  
tkuroda@kuhp.kyoto-u.ac.jp*

Hiroyuki YOSHIHARA

*Department of Medical  
Informatics, Kyoto University  
Hospital, Kyoto, Japan.  
lob@kuhp.kyoto-u.ac.jp*

### Abstract

*Broadband network infrastructure promotes networked electronic medical record system that integrates whole clinics. And also, progressing of EBM uncovered requirements of comedicals to cite online records for strategic diagnosis and treatment. Advanced medical information system is needed to provide network access service as a necessity. Therefore, conventional network system must arm with several kinds of illegal access protection. Especially, a reasonable solution against access from internal to internal is desired.*

*This paper proposes a method that provides an effective illegal access exclusion solution by easy installation and management. Additionally, custom prevailing of hospital, the system supports asset management by lost terminal detection. The proposed method detects lost and attack terminals by collation of MAC address ACL with ARP log of layer-3 switching router. The proposed system presents the exact position of detected terminals.*

*The developed system was installed on medical information network of Kyoto University Hospital. As a result of real operation, the system was proved effectiveness for easy installation, low load maintenance, and asset management without sacrifice confidentiality, integrity, and accessibility. Every suspicious access was notified by e-mail to make rapid corrective action such as physical exclusion. An illegal access experiment confirmed system performance. Besides, disconnection experiment with registered terminal confirmed effectiveness of the system as lost detection service for asset management.*

### 1. Networked medical information system

In Japan, network connection service is getting established with the expansion of broadband internet as an indispensable infrastructure. At medical workspace, hospital information system (HIS) that is improved on ordering system had been installed since 1980's. After 2005, Prime Minister of Japan and his Cabinet announced promotion policy for acceleration of making medical workspace to online in "IT Policy Package

2005" which took over e-Japan strategy [1]. Besides, the policy aims to popularize electronic medical record system. Now, electronic medical record systems are being installed on a nationwide scale. A promotion policy that defines additional score to a treatment with electronic medical record system was legislated by Japan Ministry of Healthcare, Labor and Welfare at April 2005 [2]. And so, it is easy to presuppose accelerative distribution of electronic medical record system in the next couple of years.

HIS handles such information as hospital management, medicine inventory, diagnosis and treatment and private information that includes past history or corporal characteristics. Therefore, HIS must be kept confidentiality more than ordinary information system. Furthermore, continuance of integrity is strongly required. A lack of diagnosis information or falsification of medical records is critical to cause accident of treatment. For sustaining of confidentiality and integrity, conventional HIS were constructed on closed private network that has no connection to external network such as the Internet. As matters stand now, more and more, diagnosis process is supported by online medical information including highly reliable documents. Online medical services are provided by such as MEDLINE for evidence based medicine. Consequently, the next HIS never be kept behind standalone network.

Besides, stringency of medical payment requires reduction of useless double inspection between two clinics. It makes huge wave to promote electronic chart cooperation system that connects every HIS of hospitals in the same region by global broadband network [3]. In Japan, *Super Dolphin Project* that is organized by Japan Medical Network Association leads incorporated medical record service on united network infrastructure [4]. Therefore, next generation HIS must provide not only access service of external network but also release service of internal information under control.

From the standpoint of hospital asset management, financial administration cannot afford to buy custom-designed expensive equipments that provide easy construction of high security solution. Thus, it is better to construct network infrastructure with general devices and general peripherals. Moreover, not only the cost of installation, but also the cost of management and replacement should be reduced as much as possible.







Especially, the labor and time of network management are critical determinant of quality of service more than money by long-term prospect. On the job of hospital network management, several venders cooperate to activate private network for continuation of HIS activities. Every vender brings and replaces terminals for management and development. Then, network device map state is never settled at the same state. Hence, strict policy of connection makes heavy load for administrators.

Accordingly, for management of medical information system under high confidentiality and integrity, a solution that provides flexible administration is required. Under dynamic circumstance of devices connected to network, a flexible solution should determine valid and illegal access to private network for spill prevention of information. Additionally, detection of lost devices that should connect continuously is desired for asset management. A medical information system is required as follows.

1. Constructed with general devices.
2. Low-cost for installation, management, and replacement.
3. Network access service under confidentiality and integrity with a framework of illegal access exclusion.
4. Detection of network circumstances of connected devices for asset management.

This paper proposes easy administration and effective illegal terminal exclusion system for medical information system.

## 2. Precaution of Illegal Access

Illegal access to local area network (LAN) can be classified into two kinds (from-external network and from-internal network). Several technical solutions had been already provided for access control against from-external network such as control of IP pack filtering, encryption of communication, and IP masquerade. There are many products that can easily exclude several threats from external to internal or from internal to external. What is more in business solutions had been supplied such as network firewall and virus gate-way. In contrast, because most of all current private networks were designed without thinking about internal-to-internal threats, access control of internal connected terminals is not fully supported. Above all, precaution systems against internal-to-internal threats that have customized-design are still expensive. Then, it is not convenient to install precaution of internal illegal access.

### 2.1. Measures for protection of LAN

Generally speaking, permission to users for LAN connection increases idly risks of network trouble caused by such as miss setting, computer virus infection,

and information spill. Not only idly risks but also malicious risks such as misrepresentation of user account, data falsification, data elimination, and information pilferage are also increased. Current general network systems that permit unspecified majority users to connect LAN activate only internet access service controls. On the other hand, since data stream on LAN of HIS is connected directly to clinical services, medical information systems prevent illegal access by limitation of authorized access, for example, a method of terminal limitation by hardware requirement, a method of user account limitation by access control list, and so on.

**2.1.1. Access control by hardware requirement.** A method of hardware limitation by optical cable socket is given. The low diffusion of optical socket makes a terminal equipped only Ethernet socket difficult to connect. Authentication of face recognition requires specific terminal with face recognition peripheral is given [6]. Beyond that in the former case, it is useless after diffusion of optical socket architecture in the near future. And also malicious user needs only to prepare for the optical socket to connect the network. In the latter case, restriction of access control are depends on reliability of authentication device. A kind of access control by hardware requirement is difficult to install because of the huge initial cost. It also costs a lot for replacement. And also it takes a lot of procrastination to use the system. Hence, it is hard to ensure network accessibility.

**2.1.2. Access control by access control list.** It is called access control list (ACL) that coordinates index between user account and terminal network ID, or between terminal characteristic ID and terminal network ID. Access control with ACL guarantees validity of access with one to one collation of IP address and user account on general communication of TCP/IP architecture. There are several methods, for example, a method of fixed IP address distribution to all terminals before connection, a method of dynamic IP address distribution among the terminals that registered MAC address to DHCP server, a method of certification form requires user id and password for permission, and so on [7][8][9]. However, a network that distributes IP address with DHCP only to ACL registered terminals can be accessible by a terminal that has fixed IP address configuration. Moreover, there is no way to determine the segment where the terminal connects to the network. Because a system of information socket with certification needs to establish TCP/IP connection between authentication server and an unknown terminal before transition into authorization form, malicious users are able to find ID and password easily by LAN packet monitoring.

Certification request by every connection is bad for network accessibility of valid users. And also management of ACL is difficult for administrators of a service system including huge accounts. Unfortunately, since interception of anti-illegal connection on network





layer requires high cost for packet logging and access detection, it never works for effective exclusion of illegal terminals.

## 2.2. Measures based on law

Nowadays, in Japan, law maintenance for information security is going to be proceeded a part of *IT policy package 2006*. Above all, according to Provider Liability Law, Unauthorized Computer Access Law, and Act on the Protection of Personal Information, protection by law is activated for illegal access measures by documents and attacked hardware including such as system log, system map, and user account list. Therefore, administrators should organize illegal access detection and hardware retrieval system.

## 2.3. Policy for apprehension

Because of coming and going of unspecified majority at a service station such as hospital, administrators encounter difficulties when they maintain and keep ACL up to date. Even if we construct strict registration system or hardware restrictions system, it is natural to assume that malicious users will finally find a way to evade authentication. Accordingly, administrator's load of ACL maintenance never related to efficiency of apprehension.

Therefore, this study aims to construct a position detection system that works on general network and provide a solution for effective exclusion of illegal access. For a system design of effective exclusion under real operation, system should be support four requirements as follows.

1. Ensuring of network accessibility for users.
2. Easy installation and management.
3. Real-time detection and tracking of unauthorized terminal.
4. Determination of illegal access and position detection for physical apprehension.

Easy installation and low load operation for maintenance makes effective illegal access measures at jobsite for a long time.

## 3. Methods

General network system for large-scale organization, IP routing for VLAN is provide by Layer 3 switching hub(L3 switch) that provides high-speed processing on hardware. L3 switch is a device designed working on network layer defined at third layer of OSI basic reference model for IP routing and switching. IP packet routing is provided with media access control address (MAC address) that is the unique ID of every network devices. MAC address including vender code and unique

address of a network adapter is work for specifying each network devices [10].

This paper proposes a method based on MAC address ACL collation with L3 switch ARP log. This study aims to construct a real-time access detection and position detection system called "*MacMonic*". L3 switch ARP log can be collected from every switch by auto pilot. Any time when registering terminals, MAC address ACL can be updated. Therefore, the method provides easy installation. And it provides easy management that does not require ACL update on demand when user requests authentication for network connection.

### 3.1. Detection system against illegal access

Because L3 switch ARP log includes MAC address and connected port number of the terminal, the proposed method by MAC address ACL collation with L3 switch ARP log can detect the exact position of terminals. To notice illegal access, a list includes L3 switch position, port number, date, and MAC address is e-mailed to administrators as soon as a detection of illegal access. Supposed network topology is shown in Figure 1. Auto pilot program of MAC address grabbing is activated on DHCP server under VLAN. The program goes round all L3 switch configured in network, and retrieves all apparent MAC address. The program and IP distribution control of DHCP server can easily cooperate with each other. After overlap reduction, collation between grabbed MAC address list and prepared ACL detects lost and suspicious terminals.

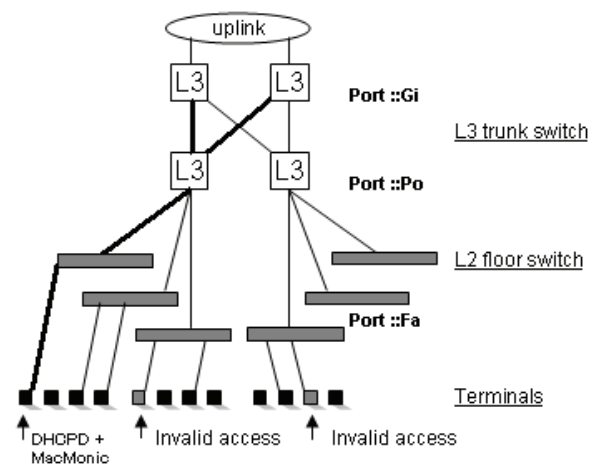


Figure 1. Logical structure of network at Kyoto University Hospital. General L3 switches (CISCO Systems, Inc) is designed with three layers. The top layer L3 switches construct VLAN network.

**3.1.1. Dynamic collation of MAC address.** The general L3 switch supplied by CISCO Systems has a terminal command to check switch report by Telnet protocol. Going round by prepared IP address list of all L3 switch, the grabbing program preserves output results of





terminal command in local. All reports are integrated with L3 switch IP address. The followings show format and a sample of report.

- ✓ --Format--  
<MAC-address>, <type-of-switch>, <earn>, <age>,  
<port-type>/<any>/<port-number>
- ✓ --Sample--  
123.000f.abcd, DYNAMIC, Yes, 155, Fa1/0/4

Report includes MAC address, VLAN number, and port types (trunk = Gi, branch = {path, terminal | Po, Fa}) /port number. Every terminal is assigned logically under VLAN, physically under L2 floor switch (shown in Figure 1). Up link over L3 switches that have cross links is connected to optical fiber network. Port type is defined by the path of switch. *Fa* port is defined as a port of from L2 switch to under layer L3 switch. *Po* port is defined as a port of from under layer L3 switch to upper layer L3 switch. *Gi* port is defined as a port of upper layer L3 switch. Overlap reduction selects only *Fa* port log to determination of terminals.

**3.1.2. ACL with MAC address.** ACL is generated by registered MAC address and physical position. Prepared ACL format is as follows.

- ✓ --Format--  
<MAC-address>, <building-number>, <building-name>, <system>, <type-of-switch>, <model-of-switch>, <number-of-switch>, <comment>
- ✓ --Sample--  
123.00f.abcd, east building 5F, network room, xyz-abc-no007, Catalyst6509, 1, Gi 3/1, any

Item of MAC address is formatted as readable by DHCP servers. The other items are formatted as a comment by DHCP servers.

**3.1.3. Collation of ACL and ARP log.** Collation algorithm of ACL and ARP log is show in Figure 2. Separated ACL and ARP log are integrated and sorted by MAC address by ascending order. Overlap reduction eliminates each log by priority [ *Fa* > *Po* > *Gi* ]. Cleaned ARP log is collated with ACL to classify three states as fine, lost, and attack. A MAC address existing both ACL and ARP log is fine. A MAC address existing only ACL is lost. And a MAC address existing only ARP log is attack doubtfully. In case, attack includes valid access before registration. It ensures accessibility for proper users.

Sorting of MAC list and ACL, the more processing collation decreasing log remains, the more fast collation becomes. At the stage 4 in Figure 2, port number and physical position of the L3 switch is added to output of classified MAC list. Results are notified by e-mail to administrators. This program is implemented by Perl. E-mail format is shown as follows. In addition, the sixth character from the head of MAC address is vender code.

Vender code is replaced to real name to make facility with apprehension.

- ✓ --Format--  
<MAC-address>, <L3-switch-IP-address>, <port-number>, <vender>, <model>, <room>, <ward>
- ✓ --Sample--  
1234.000f.abcd, 192.168.0.254, Po1, CISCO SYSTEMS. INC., xyz-abc-no007, server room, east ward 4F

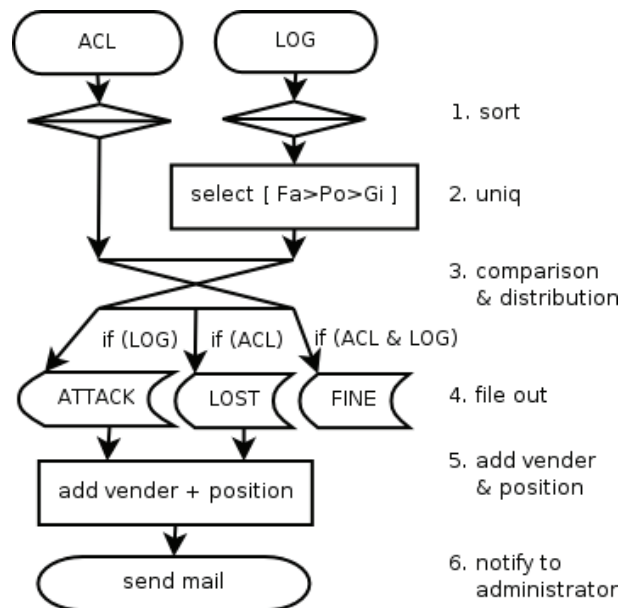


Figure 2. Collation algorithm of ACL and ARP log for access detection.

### 3.2. Detection system operation

This study focuses on real operation and effective solution including physical protection. Intercept on network layer is not enough to protect private network. Both a system working on network system and organization of illegal access apprehension are important. The developed illegal access exclusion system called *MacMonic* is operated by steps as follows.

1. Installation
  - A) Registration of L3 switch (physical position and IP address) with VLAN settings.
  - B) Registration of permitted terminal MAC address (at any time).
2. Operation – activate *MacMonic* at any period
  - A) Determination of suspicious access.
  - B) Receiving notification email.
3. Measures for valid access -- ask users to register.
4. Measures against illegal access -- seize terminals.





## 4. Case study at Kyoto University Hospital

### 4.1. Environment and condition

At Kyoto University Hospital, there are about a hundred floor L3 switches. Total number of port is about 2,500. Switches configure VLAN for every clinical department. About 1,500 terminals includes L3 switch and virtual switch activate network permanently. Number of terminals that change connection state dynamically is about 200. Since January 2005, *MacMonic* was installed on HIS network of Kyoto University Hospital. *MacMonic* was activated every couple of minutes by auto program called *cron*. The detection system worked on a general purpose rack server; Linux installed, Intel Celeron 1.0 GHz, memory 512MB.

### 4.2. Experiment of detection and apprehension

For functional test and performance test, the following approach was employed in the illegal detection experiment. An unregistered terminal was connected at arbitrary selected port. System would detect the terminal and send a notification email to administrator after few seconds. MAC address list registered about 4,000 entries was prepared. Result of online sniffing of L3 switch log, about 1,600 ARP entries after overlap reduction was detected. The results of unregistered connection, it takes only 6 seconds to determinate and locate the access. Administrator could get the email by cell phone to notice the suspicious access quickly. Consequentially, an administrator can call up and check the ward in five minutes at the most. In concluding, the proposed method provided a solution of effective physically exclusion against illegal access by easy operation.

Additional function test of lost detection was designed as arbitrary disconnection of registered terminal. As a result, the lost terminal was listed at e-mail as same as illegal terminals.

### 4.3. Combination with LAN monitoring system

The developed system can implement to combine DHCP server for facileness ACL update. One stop ACL registration at DHCP server works on the developed system. Hence, a registered terminal has no restriction to get free network access wherever it connects. In contrast, unregistered terminal has no way to access without asking identification by manager at the floor as soon as it connects.

### 4.4. Continuous network surveillance

On the job operation, posting a floor switch manager assigned from nurse or staff, administrator only needs to contact them to apprehend illegal terminals. In additionally, even if the malicious terminal accesses with

changing the physical position at every connection, the system can track and record the behavior. Therefore, a terminal once connected, the system can prove illegal access later.

It is concluded that the proposed method that classifies three types of access by MAC address with L3 switch ARP log can be implemented for effective exclusion and asset management system. The developed system provides easy installation and management for administrators. The system serves network accessibility to valid users for freely access at any where without any authentication.

## 5. Discussion

In general, malicious user knows basic knowledge about network especially how-to techniques of IP address configuration. Thus, it can be assumed that the first step of mischievous attack will start at connecting to get IP address by DHCP service. If a system employs MAC address restriction to distribute IP address on DHCP service, attacker try to set fixed IP address with making a conjecture by sniffing network packet at the second step. Conventional system that restricts DHCP IP distribution is not able to defend from intrusion with fixed IP address. On the other hand, the proposed system targets monitoring the connection of the first step to get IP address. Therefore, only an announcement of the system installation could work for preventing attempts of mischievous attacks by users who know, at least, architecture of DHCP and recognizing the difference between fixed and dynamic IP address.

The feature of the proposed system is detection of unique position of every terminal. At the same time, the demerit of the system is useless to counter attack against MAC address arrogation. Close attended malicious user tried to access by a terminal misrepresented with permitted MAC address never be detected. However, because the system never miss the first trial of connection, malicious user should know the permitted MAC address by using a valid terminal or steal original ACL. Consequentially, it is not easy for users to get a permitted MAC address.

A main topic of this study is that the system is useful for asset management. A terminal that should connect to network continuously can be located with lost alert when it disconnected. In case of lost alert, administrator need to reveal the situation is simply disconnection or theft.

It is equally important that the system makes educational affect to improve information literacy. Because idly connection causes to get a caution by a floor manager, clinical staffs will make more attention to access. As it turns out, the ordinary user hesitates to access without necessity.

On the whole, the proposed system made for effective safety information management was designed from three







view points as hardware (system), software (operation), and user (education).

In the near future, it is sure that electronic medical record system will be evaluated as a continuous treatment environment for comedicals wherever they are.

A future use case of electronic medical record system, doctors walk around treating patient every floor with mobile devices. Conventional system for mischievous access to HIS measures to restrict physical site of information socket in a room where allowed only doctors. But for the usability, the developed system provides a solution to install information socket any room and any registered one can use network freely anywhere.

At a viewpoint of post-processing of accident, target terminal that gets troubles such as intrusion from outsiders and computer virus infection can be spotted by the system without any hard investigation on every terminal under network interception. It is necessary to avoid the economic loss by the network interception even if temporary. In the network age, economic effect of a system without network interception for accident measures is inestimable.

It is likely that the worst threat of post network society is "internal users" who sometimes don't understand what they are doing. For guard of information security, it is getting more important that a system design and total network solution should be considered with "internal users".

## 6. Conclusion

This study aimed to provide a solution for information security of networked medical information system. In this paper, we proposed an ACL authentication method based on MAC address collation with ARP record of layer 3 switch. The developed system provided a solution of illegal access detection and lost asset detection. The system also notifies the result by e-mail with exact position of illegal access terminals. The load of maintenance to an administrator is to check email.

The system was easily installed on HIS network of Kyoto University Hospital. The result of real operation, it was confirmed to be able to exclude illegal access effectively without sacrifice network accessibility. And then, for asset management, the system is activated as lost terminal detection system.

The proposed method was confirmed usefulness for a solution against mischievous illegal attack from internal to internal.

## 7. References

[1] IT Strategic Headquarters, "IT Policy Package 2005", *Official Gazette 2005*, Prime Minister of Japan and His Cabinet, 2005,

<http://www.kantei.go.jp/jp/singi/it2/kettei/050224/050224pac.html>.

[2] Standardized Electronic Medical Record Promotion Committee, "Standardized Electronic Medical Record Promotion Committee Final Report", *The Minutes of Council*, Japan Ministry of Labor and Welfare, 2005, <http://www.mhlw.go.jp/shingi/2005/05/s0517-4.html>.

[3] The Higo Foundation for Promotion of Medical Education and Research, "Dolphin Project - Open Network for Kumamoto Local Healthcare and Welfare by Informational Common Electronic Medical Record", 2001, <http://www.kuh.kumamoto-u.ac.jp/dolphin/>.

[4] Kanto Bureau of Telecommunications, "Survey Report of Network Construction at Medical Workplace 2004", *The Material of Survey Study Association*, Japan Ministry of General Affairs, <http://www.kanto-bt.go.jp/stats/data/chosa/chosa02/>.

[5] Office of Information Security Policy, "Outbreak of Illegal Computer Access Acts and Situation of Research and Development of Technology for Access Control Function 2006", *Report Announcement 2006.02.23*, Japan Ministry of Economy, Trade and Industry, 2006, [http://www.soumu.go.jp/s-news/2006/060223\\_1.html](http://www.soumu.go.jp/s-news/2006/060223_1.html).

[6] T. Fukushima, "A Network Socket System with Authentication of Face Recognition", Seminar Report 2002, Research Organization of Information and System, 2002.

[7] S. Maruyama, Y. Asano, H. Tsuji, Y. Fujii and J. Nakamura, "A secure LAN Sockets system for everyone which need not modify existing DHCP clients", *Study Report 1999*, Information Processing Society of Japan, 1999, 99-DSM-14, Vol. 1999, No. 56, pp. 131--136.

[8] H. Ishibashi, A. Sakamoto, N. Yamai, K. Abe, K. Ohnishi and T. Matsuura, "LANA2: An Access Control System for LAN Sockets using VLAN Functions", *Study Report 1999*, Information Processing Society of Japan, 1999, 99-DSM-14, Vol. 1999, No. 56, pp. 137--142.

[9] H. Masuda, M. Suzuki and M. Nakanishi, "Implementation and evaluation of secure access LAN sockets using PPPoE", *Study Report 2001*, Information Processing Society of Japan, 2001, 2001-DSM-021, Vol. 2001, No. 50, pp. 19--24.

[10] Hitachi Systems, "Open Net Guard - DHCP Software using MAC Address Authentication", 2005, <http://www.hitachi-system.co.jp/ong/index.html>.

## 8. Acknowledgement

This research was partially supported by the Ministry of Education, Science, Sports and Culture, Grant-in-Aid for "JSPS Initiatives for Attractive Education in Graduate School".







# Safeguarding clinical software – A managerial case study about project management and oversight

Thomas Wetter

University of Heidelberg, Medical Informatics

thomas.wetter@urz.uni-hd.de

## Abstract

*Clinical software has become such an important factor in health care that government agencies have considered to regulate its release like the one for pharmaceuticals. Scientific organizations have doubted feasibility of all encompassing regulation and rather suggested to establish Software Oversight Committees (SOC's). By using their local expertise they should "detect software problems, analyze their impact, and develop timely solutions".*

*This article suggests that SOC's can be more effective if they focus on both strengths **and** weaknesses of their organizations and when they extend their scope from technology to human resource and to vendor relationships. It presents data gathered around a long lasting failure to integrate critical software components that occurred in a community with 21 hospitals, one being a historical leader in hospital information systems.*

## 1. Introduction

Clinical software has become a factor in health care delivery that may be regarded as essential as medication. Therefore, initiatives to regulate clinical software in the same way as new drugs (i.e. requiring approval before going into nationwide use, requiring ongoing monitoring after introduction, etc) seem to make sense. On the other hand, Miller and Gardner [1] have argued that the attempts to replicate for software what has been practiced for pharmaceuticals for decades would cause major damage to the newly developing field of vendor purchased clinical software. Their major arguments include:

- Medical software often consists of a set of complexly interfaced applications, and that regulating any given individual application misses the essential point that only the integration of that application with the other components provides the required safe clinical data processing functionality.
- Attempting to regulate all possible application combinations, however, would create a number of combinations that is beyond imagination, and

is certainly beyond the number that one could reasonably expect a regulatory agency to handle.

Furthermore, even if the release process for software were twice as fast as the one for pharmaceuticals - although there is no reason to believe this possible -, it would still mean that software would be released when it is technically outdated. In contrast to the situation with pharmaceuticals, where a newly approved drug can be sold profitably for a decade, information technology products can have turn over times as short as 18 months.

As a consequence, Miller and Gardner, working with a variety of professional and scientific organizations in the field <sup>1</sup>, came up with the Software Oversight Committee (SOC) suggestions which can be summarized as follows:

- Regulate high-risk software such as that which controls cardiac pacemakers or insulin pumps. Such software is characterized by its stand-alone operation implying that it can be tested in isolation, and by its immediate impact on patient parameters with virtually no means for a human to intervene in case of malfunction.
- Don't regulate software with only administrative or non-patient specific data (i.e. software that is used in the context of patient care but does not directly influence it).
- Establish local committees that oversee all other software.

For the latter suggestion Miller and Gardner recommended using local expertise to form committees that "detect software problems, analyze their impact, and develop timely solutions". These committees would be analogous to IRBs, but they would focus on issues of software integration including "health care informatics, clinical practice, data quality, biomedical ethics, patients' perspectives, and quality improvement" expertise and to "monitor the procedures for needs assessment,

<sup>1</sup> American Medical Informatics Association, Computer-based Patient Record Institute, Medical Library Association, Association of Academic Health Science Libraries, American Health Information Management Association, American Nurses Association



specification, selection of vendors' products, testing, training, installation, follow-up, users competences, help-desk".

In other words: the "why" and the "what" of the SOC's were clearly specified, whereas the "how" was purposefully left open: it was implicitly assumed that local knowledge and insight would lead to an SOC lineup that makes proper use of its institutions' strengths.

Subsequent to the Miller and Gardner article, research grants were acquired by several institutions to instantiate and evaluate the SOC concept. Now, seven years down the road, it seems worth considering how the concept has evolved, and whether it has delivered to its expectations. A PubMed search did not identify any journal publications on this topic since 1997, although Gardner and Tate [2] have started to teach the topic to a general Medical Informatics Audience. This article can, therefore, be regarded as a first systematic empirical assessment of the concept of SOC's after its inception. It is a detailed yet circumscribed investigation of a software integration issue in a major health-care provider, and considers a problem set typical of those SOC's were designed to address. This article will present the case study by first reporting the few necessary details about the software, its users, vendors, developers and the health care provider environment etc, then will try to convey a more elaborate metaphor for the SOC lineup and how this drove the methods that were used to investigate the case at hand. Next, the methods and results are discussed. They are presented in the order that best makes the case, rather than trying to introduce all subjectivist observation methods first and then proceeding to objectivist data collection [3]. Then the actions that were implemented as a consequence of the investigation are described. Lastly, a wider outlook how the approach developed for this case study can be generalized to other healthcare organizations with different enterprise characters and cultures will be presented.

## 2. Approach

### 2.1. Scenario

Decision makers in the enterprise where the author conducted the study<sup>2</sup> decided to replace the home made blood gas data software that delivered results directly into the respiratory intensive care unit software in a

tertiary care hospital with a vendor based software product. Offers from two vendors were evaluated, and the one with the best adherence to standards and highest responsiveness was chosen. During the negotiations and before a contract was made, the vendor was purchased by another company who, without acknowledging it, backed away from delivering some of the formerly agreed upon functionality specifically required by the tertiary care center.

To the other hospitals in the enterprise the vendor offered former versions of the software with reduced functionalities. While this was satisfactory for these hospitals, it did jeopardize the quality of care in the respiratory intensive care unit (RICU) in the enterprise's tertiary care center and did not succeed for years to upgrade. Workarounds were found in the RICU to assure patient safety, but at a higher manual effort. A common fully integrated vendor made blood gas data system remained an enterprise-wide goal. The enterprise was seeking a way to achieve that goal, and was also seeking recommendations as to how to organize software oversight to prevent such problems in the future.

The author of this investigation came as a visiting scholar to the institution. Preliminary contacts suggested that he qualified as an unbiased third party who might touch upon issues that escaped the attention of all local protagonists. After a brief presentation to the institution's SOC, he was given unanimous consent to investigate the nature of the problem, provide advice regarding solutions, and summarize and publish the findings of the investigation in the scientific literature.

### 2.2. Hypotheses

While Miller and Gardner recommend drawing on the local strengths of an organization in order to lineup an SOC, this article suggests considering the other side of the coin as well.

Hyp.1. Along with its strengths, an enterprise also has its specific weaknesses. Risks, including software related risks, are related to weaknesses in organizational structure and behavior.

Hyp.2. Therefore, it might be as productive as the "strengths" metaphor to look to consider that an SOC has all the specific software related weaknesses of its associated enterprise in its focus.

In a hierarchical enterprise, for instance, an SOC could focus on the weakness that networked operation is not encouraged. As a consequence the SOC should identify those parts of the organization that are not immediately involved in but affected by a software project and get them co-opted in a project team.

Applying the metaphor of limiting the impact of weaknesses of an enterprise entails the following activities:

- identifying the prevailing enterprise culture

<sup>2</sup> Intermountain Healthcare (IHC) is a nonprofit health care system, and is the largest health care facility in the Intermountain West. It provides hospital and other medical services in Utah and Idaho and is headquartered in Salt Lake City, Utah. IHC operates 21 hospitals in Utah and Idaho, among them the Latter Day Saints (LDS) hospital, a major tertiary care hospital, and serves 50% of the state of Utah (55% of Utah inpatient services and 45% of Utah outpatient services) [1]

- finding out weaknesses of that kind of enterprise culture
- trying to find additional evidence that these weaknesses are indeed present locally
- if so: implement or modify the SOC in such a way that the impact of these weaknesses is limited.

Other useful data collected in this process can and should also be used to deliver specific advice concerning the project.

### 3. Methods

#### 3.1. Rationale

In addition to his role as scientific investigator, the author of this study also served as a consultant to the enterprise. As a result, the process of data collection had to be integrated into ongoing, day-to-day peer-to-peer cooperation. Therefore, formal data gathering was kept to a minimum.

#### 3.2. Methods

Partially standardized interviews were conducted which included elements that were intended to identify the enterprise culture and weaknesses entailed by that culture. The interviews explored the interviewees' roles with the software integration situation in the past, his/her assessment of the most serious problems in the present, and his/her expectations about how and with whom the

problem should be solved in the future. One question specifically addressed the role that the SOC should play in this process. As one formal element the interviewees were asked to draw a chart of persons or, if they liked, groups in the enterprise that were related to the problem, and to describe in what way these groups or persons were related to each other. The interviewees were prompted to draw uni- or bidirectional arcs between the entities and name these arcs. They were encouraged to use arc labels they felt were most appropriate. Another quasi formal element was the question whom they regarded as the most appropriate to take the lead in the stalled software integration project. Finally, they were asked to indirectly classify their enterprise's culture. For this purpose, marker sentences were designed, each of which described the central element of one enterprise culture (cf. Table 1) written on cards and offered to the interviewees, prompting them to bring the cards into a rank order according to which best, second best etc characterized how they felt as employees of the enterprise.

#### 3.3. Interviews as data

Handwritten notes of the interviews were transcribed using commercial voice recognition software, submitted to the interviewees for changes, withdrawal of parts they regarded inappropriate for subsequent use and, finally, approval. The interviewees knew that in the case of sufficiently valuable results scientific publication was planned. They were also assured that nothing from our conversations would be used as data unless explicitly approved.

**Table 1**

| <b>Enterprise culture</b>                                                 | <b>Marker sentence</b>                                                                     |
|---------------------------------------------------------------------------|--------------------------------------------------------------------------------------------|
| Hierarchical                                                              | Being part of a precisely defined hierarchy makes life easy.                               |
| Quality driven                                                            | I contribute to a process of care that must be carried out in the best achievable quality. |
| <b>Examples of marker sentences that characterize enterprise cultures</b> |                                                                                            |

### 4. Analytical results

Interviews were planned with 15 individuals, and these occurred (with one exception) one person at a time. Interviewees were selected such that the largest variety of perspectives on the critical software situation that was seemed productive was obtained: Medical vs. information management, managerial vs. technical, local hospital vs. enterprise-wide; as well as human-resources, legal, and risk management.

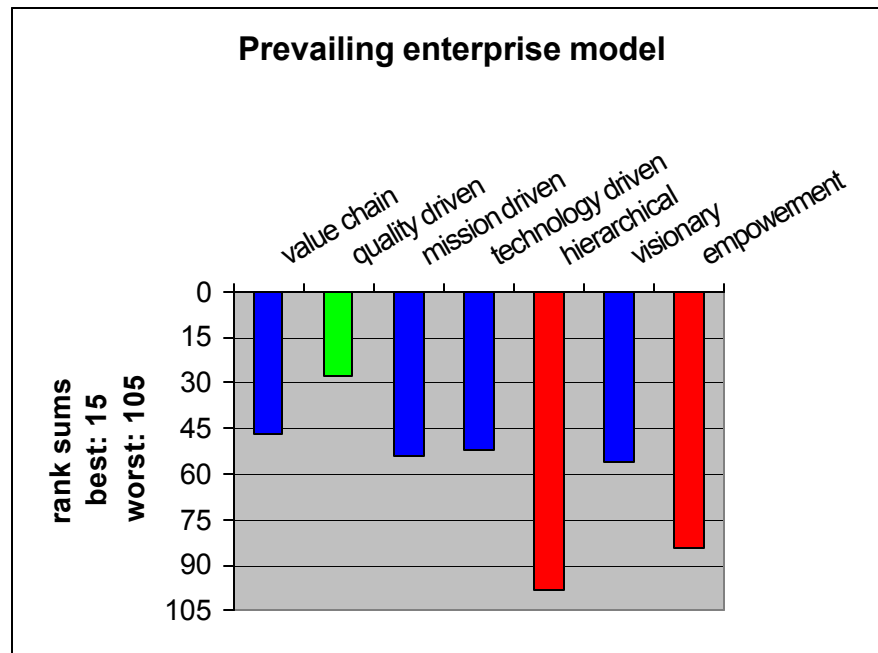
All individuals selected for an interview agreed and were talked to within an interval of 7 weeks, and the interview transcriptions were approved in the subsequent weeks.

#### 4.1. Enterprise culture

The institution is very clearly quality driven. If an enterprise culture were judged equally as top ranking by 15 interviewed individuals, it would collect 15 points. As Figure 1 shows, the "quality driven" marker sentence, reaches 28; meaning that it was ranked top, second, or third out of seven cultures by all employees interviewed. Other enterprise cultures form a middle field which will not be addressed in detail here. However, as clearly as the institution is quality driven, it is clearly not hierarchical. 11 of 15 interviewees rank "hierarchical" as the least appropriate to describe the enterprise culture as they experience it. "Empowerment" is ranked almost as low, which can be interpreted in support of the

conclusions drawn, but will subsequently not be pursued.

### Figure 1



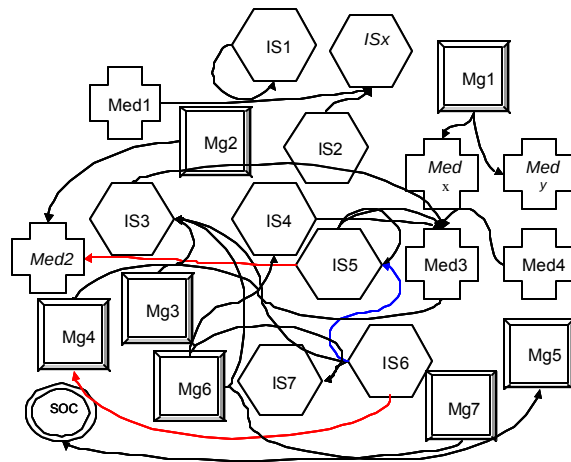
## 4.2. Enterprise culture – a second look

Looking at the details of the interviews and other standardized data reveals that “quality driven” has the more specific meaning of trying to achieve the best for one’s own work context while not covering a wider common good. There was no evidence in the interviews that this localized concept of “quality driven” resulted from an explicit agenda or implementation of an enterprise mission statement. Rather, it reflected a lack of visible structure, power, and managerial tools needed to work towards a widercommon good. The quote in one of the interviews with a senior management medical officer “I have no power” is paramount evidence for an implied “good in my nutshell” set of behaviors. This can also be read from Figure 2 which displays the interviewees’ recommendations when questioned as to who should take the lead to resolve the blood gas data integration problem. In the 14 interviews, 8 different persons and 3 functions within the institution were suggested for this role.

None of these recommendations was out of place. But the fact that this variety existed among people who have been in contact about the problem for up to four years suggested what approaches (other than technical) should the institution start with to address the problem.

## 5. Managerial results

The above and some additional evidence was presented to different bodies within the institution, fol-



**Figure 2.** Each shape represents a person (numbers 1 .. 7) or organizational unit (letters x, y) that was suggested as a candidate to take the lead for the continuation of the blood gas data integration project. Med: Medical; IS: Information Services; Mg: Management. Arrow from a to b: a suggests b. Manager 7 suggested the SOC itself.

lowed by suggestions concerning the project team itself and the SOC lineup.

## 5.1. Project team

Concrete staffing, role allocations and a lead for the project team has been suggested. Required roles to be filled are displayed and briefly commented in the following:

External role: One “face” to the vendor

- In analogy to the “One face to the customer” principle already cherished in service industries: Make sure that the vendor need not negotiate separately with all hospitals and cannot make use of their partially diverging interest.

Internal roles: Make sure that the face represents you

- Define and maintain highest standard for tertiary care hospital
- Collect variety of solutions required by primary through tertiary hospitals

Interface roles: Make sure that vendor can and does deliver

- Make tough contracts (including penalties)
- Control execution
- Offer technical support; make the own asset of the skills of those who developed the former home made system available to the vendor.

## 5.2. Software oversight committee

The recommendation was made that the SOC, in addition to its role of technically monitoring projects, also take on the role of making sure that projects are appropriately staffed. A major strength of this specific committee is that some of the committee members are academic teachers of medical informatics who have an excellent account of the specific qualifications of many of the employees of the institution because they know them as former students. Once they take on the role of monitoring the staffing of projects, they should be very effective.

The recommendation was also made that the SOC widen its scope from a local (the tertiary care hospital only) to a global (all 21 hospitals) perspective. This is intended to limit the impact of the weakness that local quality oriented people miss the global picture.

## 6. Discussion and outlook

### 6.1. The hypotheses revisited

Both formal and informal data collected support the hypothesis that a certain enterprise culture – as positive as “quality driven” may appear initially - comes with weak points. The weak points may be covered or worked around for some time. Since work arounds need extra energy and effort, efficiency may decay unnoticed until a catastrophic failure hits unexpectedly. This has been

discussed for clinical software in [4]. Here the weak point was the lack of managerial tools to handle the project as an organizational and contractual rather than as a technical challenges.

While hypothesis 1 is supported through a small but comprehensive and representative data set, support for hypothesis 2 is weaker. The strongest support for hypothesis 2 comes from the fact that the SOC under study was granted the extension of its scope from “tertiary hospital only” to enterprise wide immediately after presentation of these results. Whether this had longer ranging effects that would statistically support hypothesis 2 could not be investigated because a new software generation was started soon after this study, which entailed so many changes that data cannot be compared.

### 6.2. A broader view at safeguarding clinical software

To generalize this approach, other case studies could be conducted and the instruments used to investigate the enterprise culture could be refined. With some investigations of that type completed, a culture of functional benchmarking could evolve: enterprises could be given advice about how to lineup their SOC's based on how other organizations with similar cultures organize their SOC's. For software engineering in a more restrictive sense the term “experience factory” has already been coined [5]. From an artificial intelligence perspective this strategy can be called a case based approach. In addition to a case based approach, one could also draw on literature from the managerial sciences or organizational psychology to produce intelligent guesses about what weaknesses to expect given a certain enterprise culture. In either case, having the weaknesses of an institution in mind (and not just its strengths) is likely an important aspect to the process of defining the appropriate focus of the organization's SOC.

SOC have been explicitly given the mission to handle multi vendor situations safely [1]. This implies that observations reported in [4] should be taken into account when staffing an SOC and directing its attention. [4] analyzes factors that may lead to late failures of software that has been successfully launched. Among theses factors, vendor strategies, mergers, and bankruptcies are among the hardest to handle because they are outside the sphere where the SOC has a definite role or significant power. If taking seriously this extends the scope of SOC's beyond their mission as of their inception which is still being taught in [2]: to analyze vendor products rather than taking a thorough look at the vendors as well. If neglected one may feel safe until an unattended and hence unexpected move of one or more vendors occurs.





### 6.3. A broader view at human resource issues

The case study also revealed human resource deficiencies which may transfer to other institutions. The hosting institution of this investigation was and still is world renowned for its technical and scientific achievements in both medicine and medical informatics. More than half of the individuals “behind” Figure 2 publish in the best medical or medical informatics journals. However, none of them was an MBA, and the only person with a law qualification did not know whether a contract had ever been written with the defaulting vendor.

This may not be the only case. Health care providers who have been successful with home made systems for decades and eventually decide to purchase the new generation of information systems may be wrong in believing that skillful researchers and developers are also smart purchasers. This has also been neglected in the only major investigation of the past years about skills required in hospital IT service departments [6]. This survey based assessment produces top 10 skills through two different polling methods. Both lists include soft skills such as relational skills or oral and written communication in prominent places but only one specifically purchasing related one (“Administrative and business functions”), occurs near the bottom of the lists. Therefore, skills for successfully purchasing are one issue.

Making efficient use of its asset of technical skills of its researcher and developers and of first of a kind solutions they may have and which the enterprise wants preserved in the to be purchased software is another challenge for an enterprise that converts from home-made to purchasing. The challenge is twofold. The enterprise needs to motivate the respective specialists to alienate what used to be their pride by tolerating that it becomes an undistinguishable part of a vendor software that they will no longer be able to call their own. And the enterprise needs to find policies and business agreements (in terms of license fees, free installations, premier customer status, ..) for having offered the skills of its employers to develop the software which is subsequently purchased and installed.

If the challenges are mastered, all parties can win. The health care enterprise wins because it no longer needs to hold available major human workforce for maintenance and upgrading of software. It also wins because it gets a return on investment on its employees’ skills without having to start being a software vendor.

The vendor wins because he can incorporate best of breed technology without having the respective persons on its payroll.

The experts win – once they decide to like the idea – because eventually they can claim that their efforts bear fruit not only in one hospital but in all hospital where the product is purchased from the vendor. They also win because they no longer need to waste time on maintaining their code but can pick up the next technical challenge once they have helped transfer the former one to the vendor.

Finally, the experts remain experts and can safeguard correct function while the software is installed and during its long term function.

## 7. References

- [1] R.A. Miller, R.M. Gardner. „Recommendations for responsible monitoring and regulation of clinical software systems”. JAMIA (1997) **4**(6), 442-57
- [2] R.M. Gardner, K.E. Tate. „The Software Oversight Committee: Monitoring the Clinical Computing Environment to Reduce Errors and Improve Patient Safety”. Abstract. Proc. AMIA 2003 Annual Symposium, Washington DC Nov. 2003, Tutorial T20
- [3] C.P. Friedman, J.C. Wyatt. Evaluation Methods in Biomedical Informatics (Health Informatics), Springer, New York, 1997
- [4] Th. Wetter, “To decay is system: The challenges of keeping a health information system alive”. IJMI in press August 2006
- [5] V.R. Basili, G. Caldiera, H.D. Rombach (eds) - Encyclopedia of Software Engineering, 1994, <http://ftp.cs.umd.edu/pub/sel/papers/fact.pdf> (last visited May 17, 2006)
- [6] S. Hoffmann, J. Ash, “A Survey of Academic and Industry Professionals Regarding the Preferred Skillset of Graduates of Medical Informatics Programs”. V. Patel et al. (eds) Proc. MedInfo 2001, London, Sept. 2001

## Acknowledgements

Univ. of Heidelberg partially funded the author’s stay in Salt Lake City as a sabbatical leave.

Reed Gardner suggested the project and gave valuable advice throughout its duration.

All interviewees offered an openness to have the case investigated that is beyond imagination. The author was not refused a single interview request and never had the slightest impression that details were tried to be hidden. Withdrawals from the interview transcripts before release for this external study were absolutely minor. All personal concerns that may have been present were put behind a common attitude to use the opportunity for improving on ones quality of service.

Bob Skeate read a former version of the manuscript and provided valuable suggestions on wording and readability.

## Expanding the role of Universal Service Fund to Rural Telemedicine in Developing Countries

Muhammad Athar Sadiq<sup>1</sup>, Isao Nakajima<sup>1</sup>, Hiroshi Juzoji<sup>1</sup>, Kiyoshi Igarashi<sup>2</sup>, Kenji Tanaka<sup>2</sup>

<sup>1</sup> Nakajima Laboratory, Department of Emergency and Critical Care Medicine, Tokai University, 143 Shimokasuya, Isehara-shi, Japan 259-1193

<sup>2</sup> National Institute of Information and Communication Technology, 4-2-1, Nukui-Kitamachi, Koganei, Tokyo, Japan 184-8795

E-mail: sadiq@is.icc.u-tokai.ac.jp

In most developing countries rural population is suffering from the burden of disease and disorders due to resource constraints and severe shortage of health care professionals and facilities. Use of information and communications technologies (ICT) in healthcare system can somehow eliminate the barriers of distance and topography and offer an opportunity to which they have limited or no access. Many developing countries have developed goals to adopt universal service (access) fund to provide affordable access to expand telecommunication and information services to rural and underserved areas to telephones, telecenters, distance learning and to support energy. Communication and information infrastructure offers an opportunity to rural, remote and underserved population to have access to wide range of educational and medical services. To mobilize investment in rural telehealth care systems, there is strong need to broaden the role of universal access funds to rural telemedicine systems. The impact on the education and health care would be dramatic and benefits to rural and underserved communities would be extensive. It is therefore necessary that these populations should have access to information and communication infrastructure. This will require new funding mechanisms and additional government subsidies and new taxes on communication providers.

### Universal Access

Universal access means all the citizens of a country are able to use various services including communication, interaction and transaction services over the ICT infrastructure without too much inconvenience [1]. This does not require computer and internet at every home as it could be achieved through telecenters and through access through schools. Studies have shown that access to ICT can effectively enhance the effects

of other forms of private and public investments in the society.

### Universal Service (Access) Fund

Universal Service Fund (USF) is a mechanism of providing affordable telecom facilities to rural and underserved areas. Rural areas are defined to include villages and communities of 10,000 persons or less. USF is a prominent mechanism of mobilizing investment into rural areas. These funds offer licenses and subsidies to private and public operators to serve designated areas and communities. The idea of creating universal access funds for rural communications has existed since 1994 when Chile in effort to bridge the access gap to rural and underserved areas, established Fondo de Desarrollo de las Telecomunicaciones [2]. Chile, Peru, Colombia, Guatemala and Dominican Republic issued license to rural operators through such funds. Following closely the success of Chile and Peru models other developing countries in Latin America, Asia and Africa are pioneering the implementation of telecommunications development funds to provide communication and information services to reach all segments of urban poor and rural communities. Funds receive their resource from government contributions (Chile, Colombia), levy on telecommunication carriers which are required to contribute small part of their revenues to such funds as in Peru (1%) and Dominican Republic (2%) or radio license fee (2%) as in Guatemala.

### Services of Universal Service Fund

Access to information and communication technology has become important for economic development and poverty reduction yet the scope of universal access in most developing countries, except with few projects, remains concentrated to few regions and population groups and mostly for providing basic telephone





services and telecenters with internet access. Inclusion and expansion of services in universal access needs to be carefully assessed and needs to be adopted in accordance with country specific factors. Developing countries face plethora of problems of funds, expertise, shortage of health care and education facilities and personals. In United States the original purpose of the Universal Service Fund has been extended to support cost communications services for schools, libraries and rural health care providers [3]. Rural health care division of Universal Service Administrative Company (USAC) is responsible for ensuring for health care providers in rural areas obtain the benefits of the internet and current communications technology as provided for by the United States Congress and the Federal Communication Commission (FCC) through universal service support. The FCC established program funded \$ 56 million in 2003 to rural health care providers so that they don't need to pay more than their urban counterparts for the communication services Fig 1.

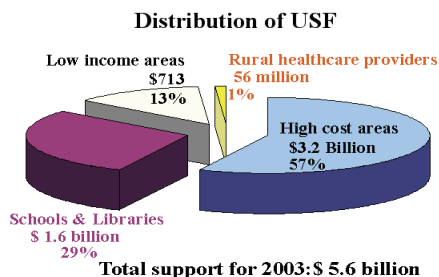


Fig 1. Distribution of Universal Service Fund in USA in 2003.

To expand the distribution of USF for health care provider, FCC in 2004 established a program that will fund up to \$ 400 annually. These measures can be categorized into the subsidies in rural areas and to the people in the lower brackets of income. On the other hand in Japan only universal service carrier, Nippon Telegraph and Telephone Co. Ltd. (NTT-West & NTT-East), has the obligation to provide universal access. The USF was introduced in June 2002 to ensure provision of telecommunications that are essential for national life throughout Japan. Since then the number of mobile phone subscribers has exceeded the number of fixed line telephone subscribers. The use of IP based telephone is also widespread. The competition in telecommunication field has been further intensified due to increased growth of new fixed line telephone services called chokushu-type services in which the subscribers call without using NTT's switching network. This has result in reduction in the basic monthly charges and installation fee of NTT. Currently Universal Service Fund services cover only local access and local telephone services in high cost areas (mostly rural and remote islands). Measures are being introduced to raise incentives in the development of communications infrastructure to high cost areas with tax exemption and low interest or free of interest loans.

However there is strong need to cover the high cost areas with communication services for rural healthcare.

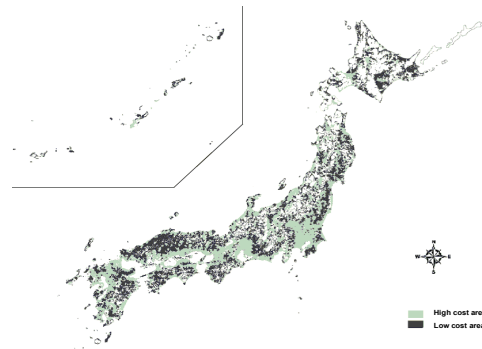


Fig 2. Distribution of high and low cost areas for Universal access in Japan.

Assumptions have been made that narrowing the range of services covered by fund, magnitude of available discounts can be large and broader range of services will allow smaller discounts. Technology is changing very fast and limiting services we might limit our capabilities for integration. Telemedicine is usually considered social objective driven rather than business driven which is not true. The worldwide distribution of mobile phones and internet connectivity has dramatically increased the no of ways doctors and patients can interact. So telemedicine by directly increasing telecommunication traffic indirectly increases business activity. To understand the telemedicine market it is essential to identify its driving forces which are competition within the health care industry, advent of affordable IT solutions especially the internet, insensitivity to distance and increasing functionality and 21st century healthcare consumers [4]. Telemedicine market grows faster than our ability to measure it as business sector. In developed world funding in telemedicine market is being used to stimulate economic growth. Norway funds the largest telemedicine programs in the world. Consultations between remote hospitals and academic centers are being carried out on daily basis [5].

## Why Universal Service Fund for Rural Telemedicine?

The focus of most universal access funds in developing countries is primarily on basic telephony with few projects has focus of promoting access to internet services. A few other funds are planning to include support for points of presence or telecenters (e.g. Bolivia, Bangladesh, Morocco). Peru's FITEL is financing pilot projects involving the internet [6]. In Kenya under InfoDev grant, field trials to improve access to telecommunication and information services for several rural a disadvantaged communities have been completed. InfoDev grants are utilized for



regulatory models for satellite services in Africa and for universal access regulatory harmonization across Africa. Only Dominican Republic has created by law "Fondo para el Desarrollo de las Telecomunicaciones (FDT)" to provide financial incentives to expand telecommunication services to rural and underserved areas that include *telehealth* along with telephones, telecenters and distance learning services. Recently Pakistan also has formulated a Universal Service Fund policy to maximize the commercial availability and coverage of ICT services and applications such as *e-Health*, e-Government, e-Commerce, e-Learning and other ICT services. This has been designed to ensure that un-served or relatively underserved areas gradually receive defined adequate services in a sustainable manner [7]. Telecommunication is a powerful tool in improving both the quality and access to education and healthcare in areas where the medical or educational infrastructure is inadequate or non-existent. In most developing countries the health care and school education system is underdeveloped and is able to provide services to population living in and around big cities. Rural population is suffering from the burden of disease and disorders due to resource constrain and sever shortage of health care facilities and professionals. Many poor patients in rural areas have to rely on basic Medicare services and for specialized care they have to travel to long distances to health care centers concentrated only in big cities. Rural schools lack the funds to attract specialized teachers necessary for courses to increase the competency of students to compete with students of big cities. Considering the rural education and health care status in most developing countries, it is important that the range of services covered by universal service fund in developing countries should expand to include rural schools and telemedicine centers.

The level of services must include bandwidth specification and needs to be adopted in accordance with country specific factors. In developed world where the aim is to have sophisticated health care delivery system, a key demand is to have second opinion and information for citizen to improve quality care and economic constrains. Many medical school and university hospitals provide patient information, course material, up to date medical literature and consultative services via national communication and infrastructure. As a result the two ways, real time video communication is the minimum level of services accepted. On the other hand in many developing countries health care at rural level is relatively simple and mostly consist of performing basic physical examination and dispensing common medicines, performing vaccinations, dressings of simple wounds, providing maternal/child care and family planning advices. These services don't require high band width and use of sophisticated telemedicine equipments. A computer and modem would allow the exchange of electronic messages as e-mails and will also result in

continued medical education of doctors, nurses and other healthcare practitioners in rural areas.

## Considerations

Many countries have adopted universal access to bridge the communication access gap only. Telecommunication in rural telemedicine can bridge the access health and education gap. Governments should expand the original purpose of universal service fund by adapting following measures:

1. Extend reasonably priced telephone services to rural, remote and underserved areas to support the cost communication for schools and rural health care providers.
2. The relatively low income of peoples of rural and underserved areas, it is difficult for the commercial providers of information and communication access to make profits. Therefore it is important to subsidize the provision of ICT services, at reasonably subsidized rates, to all public or nonprofit organizations that interested to provide healthcare and education to rural, remote and underserved areas. There is high potential for abuse of subsidies as have been shown recently in United States in e-Rate subsidy funds [1]. Even greater safeguards must be set in place to prevent abuse of subsidy funds in developing countries where government is seen as a way for politicians and officials to make money.
3. Developing new funding mechanisms that require additional government subsidies, new value added taxes on competitive access providers which are required to contribute a portion of their revenues to universal access fund. USF will make payments from these central funds to support communications services to school and health care programs.
4. Governments must adopt legislative changes mandating the communication and information access providers to include e-health and e-education to rural and underserved areas as a condition of granting any license to provide services to reminder of the population.
5. USF policies should be reviewed regularly as today's best solutions may be the best for tomorrow.

## Conclusion

The governments are the most to gain from promoting the universal access services. In many developing countries education and health care is provided by governments. Telemedicine can reduce pressure on health budgets. It could be an economic means of achieving national health policy objectives. Promoting Universal Service (Access) Fund improve the quality





and efficiency of rural educational system, health care services by providing continued medical education and consultations to remote hospitals thereby reducing the unnecessary referrals, national e-health structure and e-health administration.

## References

- [1] Samarajiva, Rohan. Universal Access Fund: How does it work? How to manage it? Presentation for International Seminar on ICT Policy Reform and Rural Communication Infrastructure 2004.
- [2] ICT Policy Reform and Rural Communications Infrastructure. Institute of International Cooperation, Japan International Cooperation Agency, March 2005: pp. 99
- [3] Heather E. Hudson. Rural Telemedicine: Lessons from Alaska for Developing Regions. *J Telemedicine and e-Health*, 2005; 11 (4), pp. 460-467
- [4] Mittman R, Cain M. The future of the internet in health care. *The internet and health communication: Experience and Expectations*. 2001: pp. 43-73
- [5] Linder, A. "Global Telemedicine and the Future of Medical Science," *Healthcare Informatics*, November 1994: pp. 62-66.
- [6] Dymond, Andrew & Sonja Oestmann. "Rural Telecommunication Development in a Liberalization Environment: An update on universal access funds". December 2003  
[http://info.worldbank.org/ict/WSIS/docs/comp\\_RuralTelecom.pdf](http://info.worldbank.org/ict/WSIS/docs/comp_RuralTelecom.pdf)
- [7] Universal Service Fund Policy, Ministry of Information Technology, Government of Pakistan.  
<http://www.pakistan.gov.pk/ministries/moit/media/USFTask1Report23May06.pdf>







## Safety Management and Crime Prevention by IC Tags –Cases of practical use of IC tag in medical and welfare fields –

Ashida N.<sup>1</sup>, Ashida Y.<sup>1</sup> Sagawa S.<sup>1</sup>, Suto S.<sup>2</sup>, Higashi T.<sup>2</sup>, Makimoto K.<sup>2</sup> and Kawahara T.<sup>3</sup>

1. Koshien University, Department of Medical and Welfare Management

2. Osaka University, School of Allied Health Sciences

3. Matirix inc.

ashida@koshien.ac.jp

### Abstract

*Incidents and accidents are occurring frequently all over the world. The system for keeping safety and crime prevention is required. Also the ubiquitous technology is developing. We tried to construct some systems with IC tag about crime prevention and disaster prevention and put in practical use. We introduce our systems with IC tag for crime prevention and disaster prevention such as monitoring of human behavior, baby abduction prevention, prevention of kidnapping and survivor search etc.*

### 1. Introduction

Terrorism and atrocious crimes are increasing on a world-wide scale. Moreover, natural disasters, such as an earthquake and tsunami, have also occurred frequently. Our safe and comfortable life is threatened by the disaster caused by man and the natural disaster. Even in Japan, the incidents, such as crime targeted the schoolboy and girl, and baby abduction from a hospital, have arisen. On the other hand, an aged society has started and elderly living alone is increasing. The population required care service is also increasing. Elderly people tend to cause accidents, such as a fall.

In order to protect a safe and comfortable life, the policy which prevents an incident and an accident is needed. And the device which prevents an incident and an accident with ICT technology is tried by various matters.

We tried to construct some IC tag systems about crime prevention and disaster prevention and put in practical use, which is pay attention as a fitted technology to ubiquitous society realization among ICT technology. In this paper, we will report our IC tag technology for crime prevention and disaster prevention.

### 2. Method

#### Structure of IC tag

The IC tag is made of IC chip and an antenna. IC chip consists of memory, a radio circuit, a control circuit, and

a decoding circuit. IC tags are divided roughly into a passive type without a battery and the active type having a battery. A passive type IC tag, since energy is supplied from the magnetic field, it is used in communication in a short distance about several centimeters. Since a battery is not necessary, the price down and miniaturizing are possible, and the IC tag which is several millimeters in size is already developed. As it becomes small, including in various goods is possible. In the field of the distribution, traceability is securable by accumulating record in each trigger installation position. It is observed as a system replaced with a bar code.[1][2]

We have developed the battery drive type IC tag of 1 inch size. When the trigger field (RF magnetic field) is installed in a certain point and an IC tag goes into this area, the tag can emit a signal and a tag receiver can catch that signal, we can get tag ID, trigger ID, and time, and those signals transfer the information of "what (who), where, and when." by PC processing. ( Fig. 1)[3]

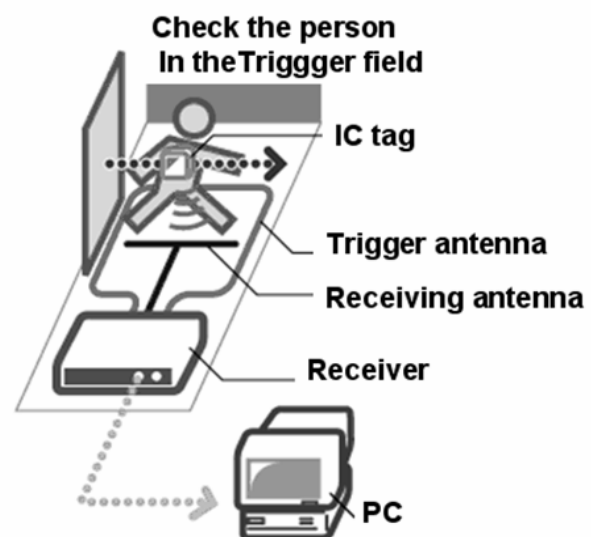


Figure 1. Structure of IC tag system

#### Monitoring of human behavior

The people's movement can also be grasped by an IC tag which acts as a tracer. Moreover, the peculiar service in the place can also be provided to the peculiar people.





For example, an audio assist is performed to people with visual impairment. It seems that it is not suitable for action observation of people by a passive type IC tag, since a strong magnetic field for power supply is needed in order to communicate certainly in several meters distance. If it is an active type, a magnetic field strong of the trigger field will not be required, and the influence on a human body will also become decrease. It seems that there is no influence. The IC tag which we developed is usually in a sleep state in order to prevent battery consumption, when the tag detects the magnetic field of the trigger field, it will be in a power supply ON state, and will send ID code of the trigger field, and its ID code. The signal is received by the receiving antenna attached near the trigger field, and the signal sent from the IC tag is given time information by the receiver with a built-in timer. These data is sent to PC by wireless LAN, and is saved as log data. It is important to give the functions to indicate the information for performing line-of-flow management of people, such as "where is now", "tracking (pursuit)", and "a fixed point observation". And the alarm function is also necessary for the system.

We have developed the software, which shows the trigger area on the map of institutions. According to the stored log data, it displays people with IC tag graphically on a screen. We tried to use and to modify this IC tag system for crime prevention and disaster prevention.

### 3. Result

#### Case reports of construction of IC tag systems

##### 3-1. Monitoring of senile dementia wandering in a nursing home

Since danger, such as a fall, is high, dementia elderly people's wandering has a large burden also for a care worker. The monitoring method using IC tag was developed and dementia elderly people's wandering action was watched by day-and-night in order to grasp the relation between a factor and the pattern of wandering. This method decreased the influence of monitoring on elderly people. The action monitoring of nursing home residents and care workers was performed. In the nursing homes with dementia elderly people's exclusive floor (Osaka, Hamamatsu and Seoul), after obtaining permission of informed consent by dementia elderly people's decision-making representative, action observation was performed for nursing home residents and care workers using IC tags. Trigger areas were installed at some points, such as a bed room, a toilet, a passage and a relaxing place, the continuation monitoring was carried out for day-and-night 24 hours. ( Fig. 2) [4][5][6]

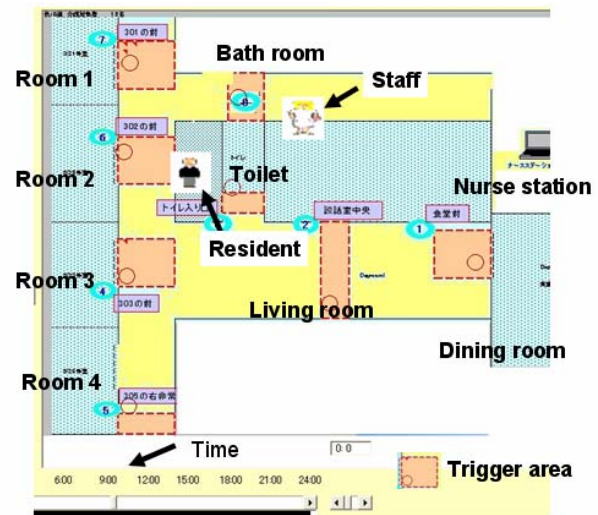


Figure 2. Display of Monitoring system for wandering

##### 3-2. Baby abduction prevention

The wristband type IC tag attached to baby's ankle in order to prevent mistake picking and to prevent kidnapping. This tag is a waterproofed type and always attached from childbirth to the time of leaving hospital. The trigger areas were installed in the entrance of a room, the entrance of a ward, and the hospital door. If an IC tag passes without procedure through these areas (abduction is assumed), an alarm will become and a report signal will send to a building safety control center. (Fig. 3)

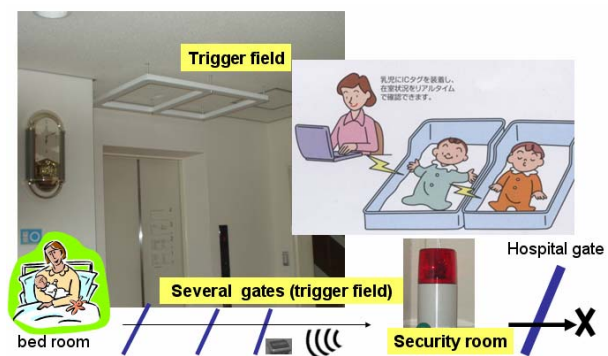
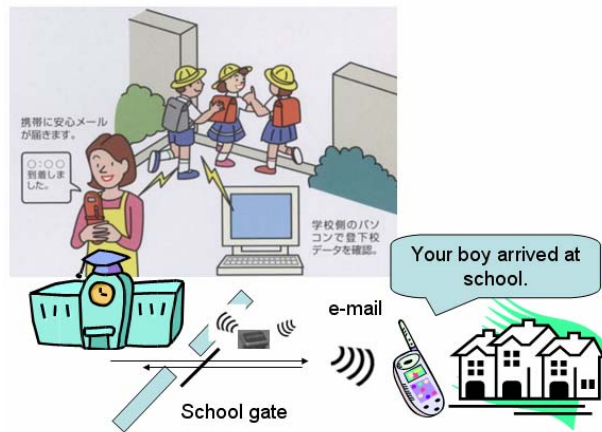


Figure 3. Baby abduction prevention system

##### 3-3. Checking of going to school and coming back to home for schoolboy and girl for prevention of kidnapping.

Trigger areas are set at the gate of a school, or some points of school zone. Passage information is sent to school and parents' cellular phone by E-mail also, and can check safety. (Fig. 4)





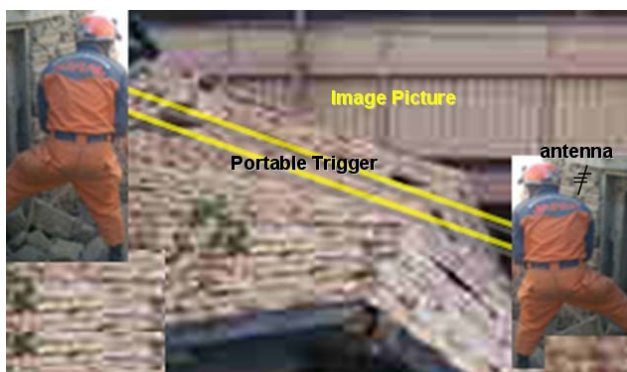
**Figure 4. Checking system for prevention of kidnapping.**

### 3-4. Correspondence to an earthquake disaster of building collapse and survivor rescue

The greatest disaster prevention is to construct an earthquake resistance building. However, if hits by the earthquake beyond an anti-earthquake design standard, it will collapse. Although many victims come out by collapse, there is possibility of survival also under rubble and the pressing subject after suffering a calamity is rescue of survivors. It is important to grasp situation who is surviving in the collapse building.

It can use for management of the entrance and leaving at certain building. (Fig. 5) And additional functions of IC tag system are designed as follows.

1. It will operate automatically when signals from trigger could not catch after disaster.
2. It will operate by the special trigger for a survivor search.
3. Moreover, when the survivor has consciousness, it will operate with the switch for emergencies. [7]



**Figure 5. Survivor search system**

## 4. Discussion

Ubiquitous society has coming. The IC card and an IC tag is convenience to perform individual identification without conscious. It is hopeful in all fields such as delivery, traffic, circulation, and banking. Although ubiquitous technology is the technology which

may make controlled society, the application for "relief and safety" is tried in the medical and welfare field. The service which searches wandering elderly by the position information search engine by GPS or a cellular phone is already commercialized.

The traceability of the IC tag which can be used also in the outdoors and indoor, is effective in the distribution management of a thing. And it is used for proper inventory management also. The IC tag as technology for relief and safety is going to be used for misuse prevention of medical implements and medical supplies

On the other hand, traceability of the IC tag seems to be effective also in action observation for people. Although the method of action observation for people that chronicler follows a candidate and records it serially, and the method of monitoring with a video camera has been used, there is a problem of an operating excess strain or invasion of privacy. IC tag method may be used for monitoring and the job analysis. A passive type IC tag with small size, a lightweight and cheap and is effective in misuse prevention of medical implements and medical supplies. However, it seems that a battery drive type active type IC tag is effective in distance with a detector, or the safety of magnetic field strength when equipping with an IC tag and trying to perform action observation of people.

## 5. References

- [1] ICT Policy Outline in the Heisei 17 Fiscal Year) [http://www.soumu.go.jp/s-news/2004/040827\\_7.html](http://www.soumu.go.jp/s-news/2004/040827_7.html)
- [2]Yoshizaki Masahiro: Japan in U-Japan 2010, New Media (2005. 1),
- [3]Matrix inc. <http://www.matrix-inc.co.jp/index.html>
- [4]Ashida N., Makimoto K. Nozaki C., Nakagawa M., Suzuki M. and Kawahara T. A Monitoring of Wandering Behavioral Pattern of Senile Dementia with IC Tag. 3rd APT TELEMEDICINE Workshop 2005, Proceedings p80-84,2005
- [5] Nobuyuki Ashida, Kiyoko Makimoto, Mizue Suzuki, and kawahara T.; The Monitoring of wandering elderly using a IC tag, the 1st ubiquitous medical symposium (May, Tokyo)
- [6]Ashida N., Makimoto K., and Ashida Y. et al. The action monitoring of the user and care worker in the Nursing home using IC tag ( In Japanese), medical information union convention 2005. November Yokohama
- [7] Ashida N. ,Ashida Y., Takagishi Y. and Kawahara T. :The Survivor Search System using an IC Tag,4th APT Telemedicine Workshop 2006 p220-224,Pakistan 25-26, Jan 2006







# Applying Grid Computation for Evaluating Potentially Inappropriate Medication Use in Elderly Patients

Yung-Tai Yen MS<sup>1</sup>, Yu-Lun Liu MD<sup>2</sup>, Yu-Chuan Li MD, PhD<sup>2</sup>, Chien-Yeh Hsu PhD<sup>2\*</sup>

<sup>1</sup>Graduate Institute of Medical Sciences, Taipei Medical University, Taipei, Taiwan

<sup>2</sup>Graduate Institute of Medical Informatics, Taipei Medical University, Taipei, Taiwan

\*Corresponding Author

[tedyeng@tmu.edu.tw](mailto:tedyeng@tmu.edu.tw), [g158094006@tmu.edu.tw](mailto:g158094006@tmu.edu.tw), [jack@tmu.edu.tw](mailto:jack@tmu.edu.tw), [cyhsu@tmu.edu.tw](mailto:cyhsu@tmu.edu.tw)

## Abstract

**Objectives:** The nationwide health claims data stored in the National Health Insurance Research Database (NHIRD) in Taiwan have been used to analyze potentially inappropriate medication use in elderly patients. To access the huge data of NHIRD, we applied the Grid-computing technology to develop a Grid-based data analysis system in this study.

**Methods:** Globus Toolkit was used to integrate 12 Mac Mini computers and 4 disk arrays for building the Grid system. Prevalence and drug duration score (DDS) of 28 potentially inappropriate medication uses from 2000 to 2002 were calculated.

**Results:** In 2000 – 2002, 68.79% of elderly patients received at least one of 28 inappropriate medications. 14.88% of elderly patients received at least one of 9 drugs classified by an expert panel as “should always be avoided”; 40.45% of elderly patients received at least one of 6 drugs that would rarely be considered appropriate; and 51.70% of elderly patients received at least one of 13 medications that have some indications but are often misused. The DDS of chlorpropamide and carisoprodol was 23.01 and 13.83 respectively, which were the highest scores in the “should always be avoid” and “are rarely appropriate” category.

**Conclusions:** For the task of huge data analysis, the Grid is an appropriate and robust approach to facilitate the analytical process. The inappropriate medication analysis results of this study can be represented as practical observations of potentially inappropriate medication conditions of elderly patients in Taiwan; moreover they can also serve as basic references for further inappropriate medication studies.

## 1. Introduction

Inappropriate medication use is one of major patient safety and public health concerns, especially for elderly patients [1, 2]. Elderly patients who often take polypharmacy are at increased risk of adverse drug effects and drug toxicity [3]. Many researchers studied the prevalence of potentially inappropriate medication use in elderly by analyzing retrospective data [4, 5]. Much of their work has been limited to particular medical facilities or communities.

In this study, we used 3 years (2000 - 2002) complete data of the National Health Insurance Research Database (NHIRD) in Taiwan to extract prescriptions for inappropriate medication identification in elderly outpatients. Nowadays, this database was frequently used to investigate and resolve health and epidemiologic problems [6, 7]. The complete dataset of the NHIRD is too huge to be processed by a personal computer or even a powerful server; therefore most researchers have only used subsets of sampled data for their studies. The two major benefits of using complete data are: (1) the analytical results can be used to represent the entire population with no inferential statistics, and (2) sampling methods may influence analytical results, therefore using complete data can avoid errors due to applying inappropriate or incorrect sampling methods.

Recently lots of claims data, medical records, images, and gene and protein sequence data have been digitized. Traditional data analysis approaches have gradually been rendered incapable of dealing with large datasets. Because the Grid technology [8, 9] is usually used to resolve very complicated problems, which cannot be handled with limited computing resources, this technology has been widely applied in health and biomedical research and projects [10-13]. In general, the Grid is a service for computing resource sharing over the Internet. It can turn computers in different places into one vast computational power to process difficult computational problems.

In this study, we only used part of NHIRD including data of ambulatory care expenditures by visits and details of ambulatory care orders. These data exceeded 200 GB of storage space for 1 year on average. To be capable of analyzing this huge amount of data, we applied the Grid architecture to design the identification procedure for potentially inappropriate medication use in elderly patients.

## 2. Methods

### 2.1. Data sources

In Taiwan, the nationwide health claims data including disease diagnosis, treatment, inpatient and outpatient claims, and details of prescriptions dispensed are stored in the NHIRD hosted by the National Health



Research Institute. The NHIRD for a 3-year period (2000 - 2002) was used to retrieve inappropriate medication information in this study. In the NHIRD, we only used the data of ambulatory care expenditures by visits and details of ambulatory care orders including prescription, patient's ID, gender and age.

An inappropriate medication database for elderly was created based on the study of Zhan et al [14]. Regardless of drug dosage, frequency, and duration of administration, they referred 1997 Beers criteria [15] and restricted their analysis to a subset of 33 drugs. They recruited an medical expert panel and used a modified Delphi method to classify these 33 drugs into 3 categories of drugs that (1) should always be avoided, (2) are rarely appropriate, and (3) have some indications but are often misused. In the Beers criteria, all drugs are classified as high severity or not for elderly persons. That is not sufficient for clinical use. Therefore we followed Zhan's classifications and created the inappropriate medication database to identify potentially inappropriate medications from elderly outpatient's prescription claims data.

## 2.2. Inappropriate medication identification

To identify inappropriate medication use in elderly from prescription data of NHIRD, we have developed an application for inappropriate medication information retrieval. The application procedure is shown in Figure 1. In Figure 1, elderly patient's prescription sheet data is obtained from NHIRD and every drug in a prescription sheet is checked for an inappropriate medication according to the inappropriate medication database. If a drug is identified as an inappropriate medication, the drug name and code are outputted and saved in the result repository. Furthermore, the patient's ID, gender, age, and drug duration of administration are also retrieved from NHIRD.

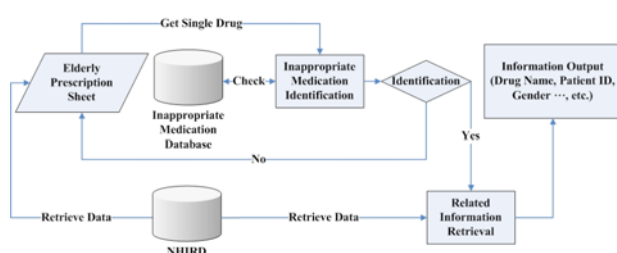


Figure 1: Inappropriate medication identification procedure

## 2.3. Inappropriate medication data analysis

To measure the drug duration of a specific inappropriate medication, we calculated a drug duration score (DDS) for inappropriate medication defined by formula (1):

$$DDS_j = \frac{\sum n_{D_i \in IM_j}}{N_{PS_j}}, \quad (1)$$

where  $IM_j$  is a specific inappropriate medication and  $N_{PS_j}$  is the total number of prescription sheets contained a specific  $IM_j$ , and  $n_{D_i \in IM_j}$  represents the drug duration administration (days) of  $IM_j$ . That is, the DDS is an averaged IM weighted by the drug duration of each administration.

Simply, DDS is the average drug duration of an inappropriate medication for each prescription sheet. For example, the DDS of the inappropriate medication, dicyclomine, is representing the average days that dicyclomine effected an elderly patient when he received a prescription containing dicyclomine.

## 2.4. Grid-based data analysis system

In Taiwan, there are 533 million hospital visits and 529 million prescription sheets every year on average. These data with patient's profile, diagnosis and other related claims data take up over 200 GB of storage space for one year. To efficiently analyze this huge health claims data, we have developed a Grid-based data analysis system for inappropriate medication identification and analysis.

In the Grid architecture, distributive computing resources can be integrated into a single powerful computational unit. Because the number of computers included in the Grid system is unlimited, it is possible to obtain more computational power by just increasing the hardware. In this study, we used Globus Toolkit [16], which is a popular Grid middleware to integrate 12 Mac Mini computers with a total of 720 GB of storage for building the Grid system. In addition, the MySQL DBMS was used to manage the data of NHIRD and to establish the data center. The system architecture is shown in Figure 2.





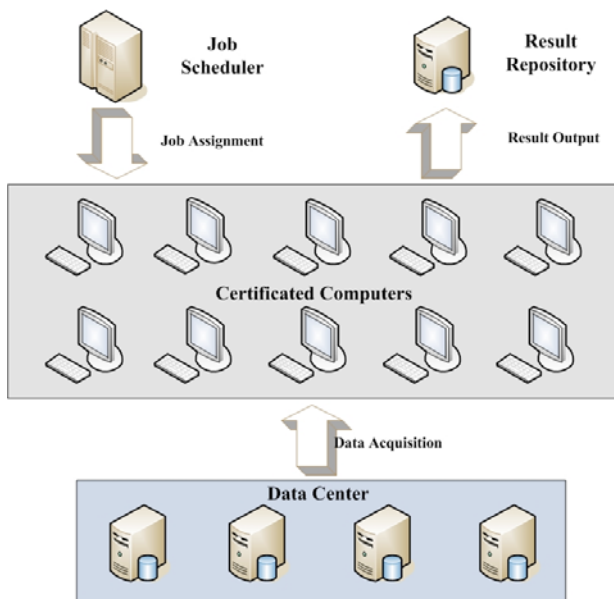


Figure 2: Architecture of the Grid-based data analysis system

In Figure 2, the job scheduler that controls program allocation, job submission, error detection and correction manages the processes of inappropriate medication identification and data analysis. NHIRD data are stored and organized in the data center, which consists of 4 database servers for load sharing. The analytical results calculated by all computers are saved in the result repository.

### 3. Results

We used 12 Mac Mini computers, 4 disk raids to store the NHIRD, and one result repository database to build the Grid system. In this system, it took about 2 hours to retrieve medication information of patients aged 65 years and older from every single prescription sheet and calculate inappropriate medication events for 1 year of NHIRD data, which contained more than  $1.2 \times 10^9$  prescriptions. Table 1 shows the number of prescription sheet received by elderly patients and by the whole population of patient from 2000 to 2002. In Taiwan, elderly people ( $n = 1,975,322$ ) are 8.82% of total population ( $n = 22,401,005$ ) on average, however elderly patient received 17.94% of total prescription sheets on average.

Table 1: Prescription sheet (PS) distribution of elderly patients and the patient population from 2000 to 2002.

| Year | No. of PS<br>received by<br>elderly patient | No. of PS<br>received by all<br>patient | %     |
|------|---------------------------------------------|-----------------------------------------|-------|
| 2000 | 93,236,719                                  | 568,843,314                             | 16.39 |
| 2001 | 93,223,821                                  | 511,190,370                             | 18.24 |
| 2002 | 97,323,170                                  | 506,940,349                             | 19.20 |

Although Zhan et al selected 33 drugs from 1997 Beers criteria as potentially inappropriate medications; we only found 28 inappropriate medications in the NHIRD. Table 2 shows the numbers, rates and DDSs of use in elderly patients of the 28 potentially inappropriate medications in Taiwan from 2000 to 2002. There were 68.79% of elderly patients on average received at least one of 28 drugs identified as inappropriate medications. For the 3 categories of inappropriate medication defined by Zhan et al, there were 14.88% of elderly patients on average received at least one of 9 drugs classified as “should always be avoided”; 40.45% of elderly patients on average received at least one of 6 drugs classified as “are rarely appropriate”; and 51.70% of elderly patients on average received at least one of 13 medications that may “have some indications”. In the “should always be avoid” category, dicyclomine was the drug that the most elderly patients (10.09% on average) received. Diazepam (23.36% on average) and chlorzoxazone (23.05% on average) were the medications that the most frequently used by elderly patients in the “are rarely appropriate” category. Furthermore, 31.11% of elderly patients on average received chlorpheniramine that was the most frequently used drug in the “have some indications” category.

Drug duration information is also shown as DDS column in Table 2. The mean value of DDS for chlorpropamide and carisoprodol was 23.01 and 13.83 respectively, which were the highest in the “should always be avoid” and “are rarely appropriate” category. In the “have some indications” category, the top five medications with higher mean DDS values are methyldopa, ticlopidine, dipyridamole, amitriptyline, and doxepin. The mean DDS value of methyldopa was 24.55, which was much higher than the others.

Overall mean rates of any 28 potentially inappropriate medication uses were higher in women (68.03%) than that in men (65.02%). In addition, overall rates of potentially inappropriate medication uses classified by 5 age subgroups are shown in Table 3. The mean rate of use any 28 potentially inappropriate medication was 66.73% in patients aged 65 to 69 years, 68.51% in those aged 70 to 74 years, 68.66% in those aged 75 to 79 years, 66.41% in those aged 80 to 84 years, and 54.39% in those aged 85 years and older. There was 15.44% of patients aged 65 to 69 years received one of medications that should always be avoided, compared with 14.56% of those aged 70 to 74 years, 14.14% of those aged 75 to 79 years, 13.21% of those 80 to 84 years, and 10.31% of those aged 85 years and older.





Table 2: Numbers, rates and DDSs of use in elderly patients of the 28 potentially inappropriate medications in Taiwan from 2000 to 2002.

| Drugs                     | No. of elderly patient |           |           | % of total elderly patient receiving drugs |        |        | DDS   |       |       |
|---------------------------|------------------------|-----------|-----------|--------------------------------------------|--------|--------|-------|-------|-------|
|                           | 2000                   | 2001      | 2002      | 2000                                       | 2001   | 2002   | 2000  | 2001  | 2002  |
| <b>Always avoid</b>       | 281,178                | 269,523   | 270,395   | 14.91%                                     | 13.83% | 13.43% | 5.96  | 6.45  | 6.55  |
| Barbiturates              | 41,338                 | 31,689    | 26,696    | 2.19%                                      | 1.63%  | 1.33%  | 4.50  | 5.04  | 5.41  |
| Flurazepam                | 11,849                 | 12,641    | 11,593    | 0.63%                                      | 0.65%  | 0.58%  | 15.06 | 14.95 | 14.80 |
| Meprobamate               | 789                    | 776       | 715       | 0.04%                                      | 0.04%  | 0.04%  | 3.12  | 3.61  | 3.92  |
| Chlorpropamide            | 2,006                  | 1,841     | 1,614     | 0.11%                                      | 0.09%  | 0.08%  | 22.72 | 23.03 | 23.28 |
| Meperidine                | 18,963                 | 19,575    | 22,091    | 1.01%                                      | 1.00%  | 1.10%  | 4.83  | 4.86  | 4.83  |
| Belladonna                | 15,455                 | 10,954    | 10,657    | 0.82%                                      | 0.56%  | 0.53%  | 5.21  | 11.00 | 6.62  |
| Dicyclomine               | 195,841                | 195,363   | 198,660   | 10.38%                                     | 10.03% | 9.87%  | 5.20  | 5.53  | 5.58  |
| Hyoscyamine               | 6,652                  | 9,750     | 19,271    | 0.35%                                      | 0.50%  | 0.96%  | 3.10  | 4.51  | 6.13  |
| Propantheline             | 33,617                 | 27,943    | 22,034    | 1.78%                                      | 1.43%  | 1.09%  | 7.65  | 8.89  | 9.06  |
| <b>Rarely appropriate</b> | 786,377                | 780,802   | 797,191   | 41.69%                                     | 40.08% | 39.60% | 8.49  | 9.30  | 9.85  |
| Chlordiazepoxide          | 56,539                 | 45,599    | 40,725    | 3.00%                                      | 2.34%  | 2.02%  | 5.76  | 6.08  | 6.46  |
| Diazepam                  | 468,514                | 447,239   | 448,760   | 24.84%                                     | 22.96% | 22.29% | 9.63  | 10.73 | 11.53 |
| Propoxyphene              | 1,359                  | 887       | 5,553     | 0.07%                                      | 0.05%  | 0.28%  | 11.07 | 5.36  | 5.89  |
| Carisoprodol              | 29,118                 | 31,481    | 33,010    | 1.54%                                      | 1.62%  | 1.64%  | 12.88 | 13.94 | 14.68 |
| Chlorzoxazone             | 441,270                | 448,021   | 458,444   | 23.39%                                     | 23.00% | 22.77% | 6.90  | 7.39  | 7.64  |
| Methocarbamol             | 34,831                 | 37,304    | 39,802    | 1.85%                                      | 1.91%  | 1.98%  | 12.69 | 12.81 | 13.61 |
| <b>Some indications</b>   | 1,004,976              | 1,000,980 | 1,015,627 | 53.27%                                     | 51.38% | 50.45% | 11.07 | 12.00 | 12.48 |
| Amitriptyline             | 17,365                 | 17,191    | 18,012    | 0.92%                                      | 0.88%  | 0.89%  | 15.04 | 15.98 | 17.11 |
| Doxepin                   | 19,200                 | 29,019    | 36,570    | 1.02%                                      | 1.49%  | 1.82%  | 14.95 | 14.47 | 16.61 |
| Indomethacin              | 154,162                | 144,289   | 140,066   | 8.17%                                      | 7.41%  | 6.96%  | 7.79  | 8.32  | 8.20  |
| Dipyridamole              | 340,075                | 329,340   | 316,144   | 18.03%                                     | 16.91% | 15.70% | 18.06 | 19.15 | 20.00 |
| Ticlopidine               | 35,092                 | 37,714    | 38,241    | 1.86%                                      | 1.94%  | 1.90%  | 23.06 | 24.22 | 24.89 |
| Methyldopa                | 4,432                  | 3,914     | 3,302     | 0.23%                                      | 0.20%  | 0.16%  | 23.92 | 24.78 | 24.95 |
| Reserpine                 | 84,787                 | 77,187    | 71,762    | 4.49%                                      | 3.96%  | 3.56%  | 12.13 | 13.31 | 14.38 |
| Disopyramide              | 464                    | 400       | 559       | 0.02%                                      | 0.02%  | 0.03%  | 10.79 | 12.26 | 13.51 |
| Oxybutynin                | 38,233                 | 43,903    | 48,122    | 2.03%                                      | 2.25%  | 2.39%  | 13.31 | 14.61 | 15.97 |
| Chlorpheniramine          | 610,511                | 597,087   | 610,447   | 32.36%                                     | 30.65% | 30.32% | 5.00  | 5.40  | 5.60  |
| Cyproheptadine            | 174,618                | 172,985   | 178,011   | 9.26%                                      | 8.88%  | 8.84%  | 5.87  | 6.10  | 6.39  |





|                 |           |           |           |        |        |        |      |      |      |
|-----------------|-----------|-----------|-----------|--------|--------|--------|------|------|------|
| Diphenhydramine | 2,536     | 2,474     | 2,307     | 0.13%  | 0.13%  | 0.11%  | 3.01 | 2.96 | 2.98 |
| Promethazine    | 2,685     | 1,181     | 879       | 0.14%  | 0.06%  | 0.04%  | 5.92 | 6.83 | 8.19 |
| Any class       | 1,317,155 | 1,330,676 | 1,362,262 | 69.82% | 68.30% | 67.67% |      |      |      |

---





Table 3: Overall rates of potentially inappropriate medication use classified by 5 age subgroups in Taiwan from 2000 to 2002.

| Age   | Always avoid |        |        | Rarely appropriate |         |         | Some indications |         |         | Any class |         |         |
|-------|--------------|--------|--------|--------------------|---------|---------|------------------|---------|---------|-----------|---------|---------|
|       | n / %        |        |        | n / %              |         |         | n / %            |         |         | n / %     |         |         |
|       | 2000         | 2001   | 2002   | 2000               | 2001    | 2002    | 2000             | 2001    | 2002    | 2000      | 2001    | 2002    |
| 65-69 | 105,276      | 99,081 | 98,729 | 269,713            | 258,736 | 260,380 | 327,207          | 312,102 | 309,450 | 443,668   | 432,301 | 434,020 |
|       | 16.14%       | 15.22% | 14.95% | 41.36%             | 39.75%  | 39.44%  | 50.17%           | 47.94%  | 46.87%  | 68.03%    | 66.41%  | 65.74%  |
| 70-74 | 90,917       | 82,363 | 86,490 | 245,403            | 241,805 | 241,431 | 311,023          | 306,545 | 302,602 | 407,480   | 407,561 | 407,427 |
|       | 15.51%       | 13.79% | 14.38% | 41.87%             | 40.50%  | 40.14%  | 53.06%           | 51.34%  | 50.31%  | 69.52%    | 68.26%  | 67.74%  |
| 75-79 | 57,711       | 56,574 | 58,148 | 158,563            | 162,071 | 168,631 | 208,466          | 213,981 | 222,812 | 267,858   | 278,314 | 291,978 |
|       | 14.95%       | 13.90% | 13.58% | 41.06%             | 39.85%  | 39.37%  | 53.98%           | 52.62%  | 52.02%  | 69.37%    | 68.44%  | 68.17%  |
| 80-84 | 26,685       | 28,272 | 29,938 | 75,641             | 80,018  | 85,936  | 104,413          | 111,860 | 119,808 | 133,356   | 143,973 | 155,072 |
|       | 13.82%       | 12.98% | 12.82% | 37.76%             | 36.81%  | 36.81%  | 52.13%           | 51.46%  | 51.31%  | 66.58%    | 66.23%  | 66.42%  |
| >=85  | 14,712       | 14,354 | 14,975 | 38,637             | 39,479  | 4,2082  | 56,730           | 5,8984  | 63,351  | 73,405    | 76,725  | 82,405  |
|       | 10.73%       | 10.18% | 10.03% | 28.19%             | 28.01%  | 28.18%  | 41.39%           | 41.84%  | 42.42%  | 53.56%    | 54.43%  | 55.17%  |

#### 4. Discussion

Many researchers reported their study results about potentially inappropriate medication use in elderly people. Based on 1997 Beers criteria, Rigler et al. [17] analyzed claim data from May 2000 to April 2001 for 3 cohorts of older Medicaid recipients. The rate of any inappropriate medication use ranged from 21% to 48% across the 3 cohorts. Another similar study was done by Simon et al. [18]. They used medication-dispensing data from 10 HMOs to determine rates of potentially inappropriate medication use in elderly persons in the U.S. in 2000 – 2001. Their study was also based on the Zhan's 33 inappropriate medication classifications and found 28.8% of elderly persons received at least one of 33 inappropriate medications. Compared the overall rate of potentially inappropriate medication use estimated in this studies (68.79%) with others, the prevalence of potentially inappropriate medication used by elderly patients in Taiwan was obviously very high.

In Taiwan, it is very easy for patients to seek medical advice and obtain prescription in the outpatient service. From the data of NHIRD in 2000 - 2002, there were 528,991,344 hospital visits and 533,515,227 prescription sheets on average. For a single elderly person, there were 46.87 hospital visits and received 46.80 prescription

sheets on average in one year. In addition, elderly people usually suffered multiple diseases and received polypharmacy at the same time from different doctors or even different medical service units. This may provide a reason for that the prevalence of potentially inappropriate medication use in elderly people was comparatively high in Taiwan.

Drugs with the same chemical components may have different brand names and codes in the NHIRD. For examples, dicyclomine has 69 different names and codes in the database. Other drugs that have similar effects may not be included in Zhan's inappropriate medication categories and Beers criteria. For examples, phenobarbital, pyrabital, and amobarbital etc. have the same effect as barbiturates and only barbiturates was listed in Beers criteria. More examinations are necessary to decide if we should include the other medicines. This would cause errors in our calculation. Thus, more analysis may need to perform for solving these problems.

The updated version of 2002 Beers criteria [19] included several new drugs that were not included in the 1997 list. Further analysis should perform to evaluate these drugs used by elderly people. In this study, we followed Zhan's inappropriate medication categories and didn't assess other domains of potentially inappropriate prescribing, such as drug-drug interactions, drug-disease



interactions, and dosage. According to our previous study result, the frequency of drug-drug interaction in one single prescription sheet for elderly patients was 12.76% in one year, which was much higher than other age groups. Thus, it will be very interesting if we can combine drug-drug interaction into the analysis of potentially inappropriate medication use in the future.

The strength of this study is that we are capable of analyzing the nationwide health claims data by the help of Grid. The study result could be a true representation of the potentially inappropriate medication use status of the whole elderly population in our country.

## 5. Conclusions

The Grid is an appropriate architecture for the task of large-scale data analysis. By the help of the Grid technology, we were capable of using nationwide health claim data to identify and analyze potentially inappropriate medication use in elderly patients. There have been many Grid-based systems and applications for huge data analysis [20-22]. The major benefit of applying the Grid technology is that computational power can be increased without any restriction. Computational units such as personal computers, workstations, or servers are not necessary located at the same place. They can be integrated by the Grid system via Internet. Nakagawa et al. [23] developed a Grid environment includes many PC clusters at different facilities for Magnetoencephalography (MEG) data analysis. Although we only used 12 Mac Mini computers to process the data analysis procedure in this study, we efficiently solved the complicated problem by spending lower costs.

We have successfully applied the Grid technology to analyze the huge health claims data for the task of inappropriate medication identification. The analytical results can serve as a brief representation of inappropriate medication use conditions in elderly people in Taiwan; moreover they can also serve as basic references for further inappropriate medication studies. In addition, although we cannot evaluate the adverse drug event when elderly patients received inappropriate medications, our analytical results can be used as knowledge references to develop a reminding system to improve safety for elderly patient.

## 6. Acknowledgement

This study is based in part on data from the National Health Insurance Research Database provided by the Bureau of National Health Insurance, Department of Health and managed by National Health Research Institutes. The interpretation and conclusions contained herein do not represent those of Bureau of National Health Insurance, Department of Health or National Health Research Institutes.

## 7. References

- [1] S.K. Rigler, C.M. Jachna, S. Perera, T.I. Shireman, M.L. Eng, "Patterns of potentially inappropriate medication use across three cohorts of older Medicaid recipients". *Ann Pharmacother*. 2005;39(7-8):1175-81.
- [2] D. Fialova, E. Topinkova, G. Gambassi, H. Finne-Soveri, P.V. Jonsson, I. Carpenter, M. Schroll, G. Onder, L.W. Sorbye, C. Wagner, J. Reissigova, R. Bernabei, "Potentially inappropriate medication use among elderly home care patients in Europe". *JAMA*. 2005;293(11):1348-58.
- [3] D.S. Chutka, P.Y. Takahashi, R.W. Hoel, "Inappropriate medication use in the elderly". *Essent Psychopharmacol*. 2005;6(6):331-40.
- [4] L. Azoulay, A. Zargarzadeh, Z. Salahshouri, D. Oraichi, A. Berard, "Inappropriate medication prescribing in community-dwelling elderly people living in Iran". *Eur J Clin Pharmacol*. 2005;61(12):913-9.
- [5] S. Niwata, Y. Yamada, N. Ikegami, "Prevalence of inappropriate medication using Beers criteria in Japanese long-term care facilities". *BMC Geriatr*. 2006 Jan 11;6:1.
- [6] H.L. Sun, Y.H. Kao, M.C. Chou, T.H. Lu, K.H. Lue, "Differences in the prescription patterns of anti-asthmatic medications for children by pediatricians, family physicians and physicians of other specialties". *J Formos Med Assoc*. 2006;105(4):277-83.
- [7] H.C. Lin, W.H. Tian, C.S. Chen, T.C. Liu, S.Y. Tsai, H.C. Lee, "The association between readmission rates and length of stay for schizophrenia: A 3-year population-based study". *Schizophr Res*. 2006;83(2-3):211-4.
- [8] I. Foster, C. Kesselman, "The Grid: Blueprint for a New Computing Infrastructure". where: *Morgan-Kaufman*, 1999.
- [9] I. Foster, J. Geisler, W. Nickless, W. Smith, S. Tuecke. "Software infrastructure for the I-WAY high performance distributed computing experiment". *Proc. 5th IEEE Symp High Perform Distrib Comput*. 1997:562-571.
- [10] S. Nakagawa, T. Kosaka, S. Date, S. Shimojo, M. Tonoike, "A grid computing infrastructure for MEG data analysis". *Neurol Clin Neurophysiol*. 2004 Nov 30;
- [11] H.C. van Assen, M.G. Danilouchkine, A.F. Frangi, S. Ordas, J.J. Westenberg, J.H. Reiber, B.P. Lelieveldt, SPASM: A 3D-ASM for segmentation of sparse and arbitrarily oriented cardiac MRI data. *Med Image Anal*. 2006;10(2):286-303.







- [12] Y.M. Teo, X. Wang, Y.K. Ng, "GLAD: a system for developing and deploying large-scale bioinformatics grid". *Bioinformatics*. 2005;21(6):794-802.
- [13] P.H. Beckman, "Building the TeraGrid". *Philos Transact A Math Phys Eng Sci*. 2005;363(1833):1715-1728.
- [14] C. Zhan, J. Sang, A.S. Bierman, M.R. Miller, B. Friedman, S.W. Wickizer, G.S. Meyer, "Potentially inappropriate medication use in the community-dwelling elderly: findings from the 1996 Medical Expenditure Panel Survey". *JAMA*. 2001;286(22):2823-9.
- [15] M.H. Beers, "Explicit criteria for determining potentially inappropriate medication use by the elderly. An update". *Arch Intern Med*. 1997;157(14):1531-6.
- [16] I. Foster, "Globus Toolkit version 4: software for service-oriented systems". *IFIP International Conference on Network and Parallel Computing*. where: Springer-Verlag LNCS. 2005;3779:2-13.
- [17] S.K. Rigler C.M. Jachna, S. Perera, T.I. Shireman, M.L. Eng, "Patterns of potentially inappropriate medication use across three cohorts of older Medicaid recipients". *Ann Pharmacother*. 2005;39(7-8):1175-81.
- [18] S.R. Simon, K.A. Chan, S.B. Soumerai, A.K. Wagner, S.E. Andrade, A.C. Feldstein, J.E. Lafata, R.L. Davis, J.H. Gurwitz, "Potentially inappropriate medication use by elderly persons in U.S. Health Maintenance Organizations, 2000-2001". *J Am Geriatr Soc*. 2005;53(2):227-32.
- [19] D.M. Fick, J.W. Cooper, W.E. Wade, J.L. Waller, J.R. Maclean, M.H. Beers, "Updating the Beers criteria for potentially inappropriate medication use in older adults: results of a US consensus panel of experts". *Arch Intern Med*. 2003;163(22):2716-24.
- [20] Y.M. Teo, X. Wang, Y.K. Ng, "GLAD: a system for developing and deploying large-scale bioinformatics grid". *Bioinformatics*. 2005;21(6):794-802.
- [21] D. Sulakhe, A. Rodriguez, M. D'Souza, M. Wilde, V. Nefedova, I. Foster, N. Maltsev, "GNARE: automated system for high-throughput genome analysis with grid computational backend". *J Clin Monit Comput*. 2005;19(4-5):361-9.
- [22] I. Merelli, G. Morra, D. D'Agostino, A. Clematis, L. Milanesi L, "High performance workflow implementation for protein surface characterization using grid technology". *BMC Bioinformatics*. 2005;6 Suppl 4:S19.
- [23] S. Nakagawa, T. Kosaka, S. Date, S. Shimojo, M. Tonoike, "A grid computing infrastructure for MEG data analysis". *Neurol Clin Neurophysiol*. 2004;2004:36.





## Survey on Medical Error Reporting Systems in Korean Hospitals

Jeongeun Kim <sup>a</sup>

Sukwha Kim <sup>b</sup>

David W. Bates <sup>c</sup>

<sup>a</sup> College of Nursing, Seoul National University, Seoul, Korea

<sup>b</sup> College of Medicine, Seoul National University, Seoul, Korea

<sup>c</sup> Department of Medicine, Brigham and Women's Hospital, and Harvard Medical School

### Abstract

**Background:** Recent data suggest that medical injuries, or adverse events, represent an important international problem, and that many are caused by errors. Spontaneous reporting is the main tool used to detect errors and adverse events in most countries, and reporting systems are believed to be important for improving patient safety. Increasingly, such reporting can be done using information systems, and information systems are widely used in Korea. However, few data are available regarding the use of electronic medical error reporting systems in Korea.

**Objectives:** The objectives of this study were to investigate the present status of reporting system of Korean hospitals, and to compare the current status of medical error reporting systems with that of other health information sub systems.

**Methods:** The chairs of nursing departments of all 283 hospitals nationwide with more than 100 beds were surveyed using a structured questionnaire. The response rate was 35%. In addition, 2 reports on the national use of health information systems in Korea from 1999 and 2003 were analyzed.

**Results:** Among reporting hospitals (n=99), medical errors were reported on paper in 75 hospitals (77%), verbally in 30 hospitals (30%), using word processing in 13 hospitals (13%), and using the hospital information system in only 3 hospitals (3%). In contrast, there was widespread and increasing use of health information technology (HIT) in areas such as medication administration, inpatient and outpatient order entry, and radiology.

**Conclusions:** While HIT is increasingly widely used in Korea in many areas, it is not being used for error reporting. Increasing the use of electronic reporting systems, and systemically evaluating the medical errors and adverse events reported, represent essential steps for reducing systemic errors and improving patient safety.

**Key words:** safety, medical errors, hospital information system (HIS), hospital incident reporting



## Applying Contextual Analysis on Investigating Potential Physical Therapy Contraindications

Wei-Yi Fan<sup>1</sup>, Yung-Tai Yen<sup>2</sup>, Chi-Hsien Chen MD, PhD<sup>1,3</sup>, Chien-Yeh Hsu, PhD<sup>1\*</sup>

<sup>1</sup>Graduate Institute of Medical Informatics, Taipei Medical University, Taiwan

<sup>2</sup>Graduate Institute of Medical Sciences, Taipei Medical University, Taiwan

<sup>3</sup> Taipei Municipal Wan-Fang Hospital

\*Corresponding Author

[m110094008@tmu.edu.tw](mailto:m110094008@tmu.edu.tw), [tedyeng@tmu.edu.tw](mailto:tedyeng@tmu.edu.tw), [chichen@tmu.edu.tw](mailto:chichen@tmu.edu.tw), [cyhsu@tmu.edu.tw](mailto:cyhsu@tmu.edu.tw)

### Abstract

Recent years, several severe medical errors happened in Taiwan. Many of the adverse events were caused by medical errors because of human carelessness. Based on the findings of one major study (in Colorado and Utah), about 44,000 people were killed due to medical errors in U.S. hospitals each year. Another study puts the number much higher, at 98,000. More people die from medical mistakes each year than from highway accidents, breast cancer, or AIDS. Therefore, the patient safety has become a very important topic.

Medical knowledge and information technology grow so rapidly that it is difficult for practitioners to keep up. However we can use them to prevent problems of patient safety and improve medical qualities. Although many patient safety platforms have been developed such as the drug-drug interaction analyzing platform, surgical patient safety platform, adverse drug event reporting platform, and high risk results analyzing platform, very few research reports have been published for the patient safety issue in physical therapy.

Contraindications may occur when physical therapy is intervened especially for some patients who have specific diseases or in special conditions. People who have electronic implants and pregnant woman, for example, are not safe to apply Transcutaneous Electrical Nerve Stimulation (TENS) therapy in some body areas. Several doctors and therapists in the department of rehabilitation participated in this project. After studying many literatures, textbooks, and guidelines, we established two useful tables: the “physical therapies and possible contraindications” and the “mapping table of ICD-9 code to possible contraindication”. The occurring rates for different levels of contraindications were obtained by analyzing the medical records of physical therapy. We found that for the first level contraindication, cardiovascular disease has the highest occurring rate. This indicates that doctors need to pay more attention when ordering the therapy of Cervical/Lumber Traction.

We used Vector Space Model and word segmentation algorithm to analyze the literatures retrieved from PubMed MEDLINE. All the literatures were related to physical therapy and classified according to different contraindications using the information retrieval algorithms. The contraindication analyzing methods for physical therapy was designed by analyzing the patient’s medical records and the ordering physical therapy. The main effects of these analyzing methods can detect possible contraindications and we will build a contraindication reminding system in the future.

**Key Words:** Patient Safety, Information Technology, Rehabilitation, Physical Therapy, Contraindication

### 1. Introduction

After the Institute of Medicine (IOM) published several references like “To Err is Human” and “Crossing The Quality Chasm”, the whole world’s public health bureaus started to take patient safety seriously and the World Health Organization especially paid much attention to patient security as a weight-bearing point. However, in Taiwan, because of several malpractices such as misuse of Hepatitis B vaccine that make patients to suffer severe injure and shock the entire habitants’ confidence in medical center. Therefore, it puts the health bureau to look squarely at this kind of patient safety problems.

From the previous study, 4% patients suffered medical injures and medical errors occupied 66% of all in New York State each year [1]. Furthermore, the book “To Err is Human” estimated about 98,000 people were killed by medical errors in U.S. hospitals each year [2].

“Medicine safely using is the most important factor of improving patient safety in seven targets”, JCAHO indicated that in 2004. According to “India’s First Newspaper for the Healthcare Business” in 2002, there are 180,000 people dead of adverse events in U.S. And other statistics revealed that adverse drug events occupied 20% of medical errors, more than 7% inpatient rate and 0.03% death rate [3, 4].



Moreover, there are 80% inpatients with falling down experience in hospital. Several research results were showed that the medical fee of falling down inpatients each year was almost four billion U.S. dollars in Australia [5] and one hundred billion U.S. dollars in United States of America [6]. These kinds of injuries not only decrease patients' living quality, but also shorten elders' life-span for 10 years [7].

In this study, we classified malpractices into two types: the first type of medical error is external factors such as loud noise or overworked. These kind of external factors are easy to overcome. The second kind of medical error is internal factors such as lacking knowledge or the carelessness of the doctor. Sometimes inadequate therapies also can cause medical error [8-10].

Although problems of patient safety are getting more and more serious, related reports are very rare presently. Clinical doctors indicated that patients' embedded metal devices were not usually recorded in medical records. There will be very serious adverse events such as arrhythmia if patients are treated by electrotherapy or electromagnetic wave without any reminding.

Contraindications of physical therapies are complicated. According to a result of questionnaire survey, the information from contraindication related references is out of date [11]. Furthermore, the report of Consumer Products Safety Commission revealed that there are more than 1500 electric heating pads scalded cases each year [12], however we found these kinds of case were not reported in general journals.

There is no relative research about contraindications in physical therapy in Taiwan. In this study, we have analyzed literatures and medical records for physical therapy and investigated possible contraindications.

## 2. Methods

In order to build a contraindication knowledge base in form medical records, physicians and rehabilitation therapists were invited to attend expert meetings. Those participants have provided lots of medical knowledge and suggested many clinical experiences. The research methods are described as follows:

### 2.1. Literature analysis.

After the expert meetings and literatures reviewing, possible contraindications correlative to physical therapies were determined by the rehabilitation physicians and therapists. This gave us a background of the patient safety problem involved in therapy contraindications. The main resource of medical reference we referred is PubMed MEDLINE of National Library of Medicine (NLM). We used the key words

provided by participating experts to search the related references containing physical therapy contraindications by using the Entrez platform. And we used Vector Space Model to retrieve the most relevant contraindication documents from the references we collected.

In vector space model, document and key word are represented by two vectors and the cosine measure of the two vectors is used to be the similarity basis between a document and key word. Furthermore, this model will use *tf* (term frequency) and *idf* (inversed document frequency) to calculate cosine measure.

*tf* : The frequency of each word appears in each document of all references.

$$tf_{i,j} = 1 + \log(freq_{i,j}) \quad (1)$$

*idf* : The frequency of the key word appears in all references.

$$idf_i = \log((N + 1) / n_i) \quad (2)$$

where N is a total number of all references.  $n_i$  is a total number of a specific word in all references. According to *tf* and *idf*, we can calculate the similarity between documents and key words.

$$SIM(Q, D) = \frac{\sum_{w \in Q \& D} (tf_w \times (tf_w \times idf_w))}{\sqrt{\sum_{w \in D} (tf_w \times idf_w)^2}} \quad (3)$$

### 2.2. Physical therapy contraindication knowledge base creation.

After rehabilitation physicians and therapists attending the expert meetings and referring to lots of medical references, three topics about the physical therapy contraindications were studied:

- (1) Physical therapy contraindications of specific conditions: Rehabilitation physicians and therapists should avoid some therapies when patients have some specific conditions such as patients with dermatitis, hemophilia and acute arthritis should avoid ultra red ray therapy.
- (2) Physical therapy contraindications related to ICD-9 codes: When patients have some diseases treated by physical therapy, contraindications may happen caused by the disease. For example: there is usually a symptom of acute arthritis when patients got gout. This kind of disease should avoid the therapy of hot compress, ozocerite and ultra red ray.
- (3) The severe levels of contraindications: To give us a judgment for the importance and prevalence of a





contraindication, we divided the collected contraindications into two severe levels. The first level included the contraindications which may cause the severest injuries to the patients. The second level contained the contraindications which have the higher rates of occurrence.

### 2.3. Electronic Medical record analysis.

In order to investigate the order of severity and prevalence of potential physical therapy contraindications, we extracted electronic medical records (Jan, 2004 to May, 2005) of the department of rehabilitation from the database of Taipei Municipal Wan-Fang Hospital. According to the physical therapy contraindication knowledge base, we identified potential contraindications by patients' ICD-9 codes and therapy items and calculated prevalence of each contraindication.

## 3. Results

According to the keywords provided by rehabilitation physicians and therapists, we collected 1,330 abstracts from MEDLINE by the following Entrez query:

"rehabilitation"[All Fields] OR "physical therapy"[All Fields] OR "occupational therapy"[All Fields] OR "side effect"[All Fields] OR complication[All Fields] OR "adverse effect"[All Fields] OR contraindication[All Fields] OR precaution[All Fields] AND hasabstract[text] AND English[Lang]

After collecting medical articles from MEDLINE (from 1966 to 2003), we used Vector Space Model of Information Retrieval to calculate the articles, which are most relative to physical therapy and contraindications. Those articles were then reviewed and confirmed by participating rehabilitation physicians. For example, there were 32 articles that were relevant to acute injury.

According to textbooks of physical therapy, medical references and experiences of rehabilitation physicians and therapists, we concluded 110 contraindications. For each physical therapy, the possible corresponding contraindications were identified and were built in the physical therapy contraindication knowledge base. For example, Cervical/Lumber Traction may have 11 contraindications such as claustrophobia and osteoporosis. We also studied the relationship between the ICD-9 codes and the physical therapy items and concluded 724 items related ICD-9 code.

According to the experiences of rehabilitation physicians and therapists, we classified contraindications into two levels in this research.

The first level: cardiovascular disease, over the eye, over open spinal cord, over pregnant uterus, over plastic implants, metal objects, over gonads, contact lenses, over all electronic implants, severe cardiac disturbance, congestive heart failure, acute pulmonary edema, displaced fracture, over the heart.

The second level: spondylolisthesis, cardiovascular disease, osteoporosis, acute injuries, previous spinal surgery, ischemic areas, metal objects, over all electronic implants, over-mobilization, severe diabetes.

We analyzed 4,075,718 order entries from the database of Taipei Municipal Wan-Fang Hospital and collected 5,492 patients' medical records, which may involve in therapy contraindications in rehabilitation department.

Table1 shows that there are 42,521 therapy prescriptions in 25,391 medical orders and we detected 6479 medical orders possible cause contraindications (25.52%). Furthermore, we detected 26,830 contraindications occurred in following 12 therapies.

Table1. Statistics of contraindication for physical therapy.

| Therapy item                                     | Therapy Prescription Frequency (%) | Possible Contraindications Frequency (%) |
|--------------------------------------------------|------------------------------------|------------------------------------------|
| Cervical/Lumber Traction                         | 9013 (21.20)                       | 4158 (15.50)                             |
| Hot/Cold Pack                                    | 11088 (26.08)                      | 11248 (41.92)                            |
| Infrared                                         | 972 (2.29)                         | 1371 (5.11)                              |
| Parafin Bath                                     | 611 (1.44)                         | 602 (2.24)                               |
| Ultrasound                                       | 1765 (4.15)                        | 1312 (4.89)                              |
| Shortwave                                        | 3327 (7.82)                        | 2462 (9.18)                              |
| Microwave                                        | 817 (1.92)                         | 759 (2.83)                               |
| Interferential therapy                           | 10897 (25.63)                      | 4500 (16.77)                             |
| Transcutaneous Electrical Nerve Stimulation-TENS | 3287 (7.73)                        | 181 (0.67)                               |
| Low Power Laser                                  | 574 (1.35)                         | 167 (0.62)                               |
| Magnetic Field                                   | 138 (0.32)                         | 138 (0.20)                               |
| Therapy Circulator                               | 32 (0.08)                          | 32 (0.06)                                |

The highest ratio of contraindication is hot / cold packs (41.92%). The second and third is interferential therapy (16.77%) and traction (15.50%).





Besides, according to 26,830 contraindications, we concluded the occurring times and ratio of the first level and the second level contraindications as following table.

Table2. Prevalence of contraindication classified by severe level.

|                                                     |                     |
|-----------------------------------------------------|---------------------|
| The first level of contraindication:                |                     |
| (Occupied 6.40% of all possible contraindications)  |                     |
| Cardiovascular disease:                             | 846 times (3.15%)   |
| Over the eye:                                       | 12 times (0.04%)    |
| Over open spinal cord:                              | 11 times (0.04%)    |
| Metal objects:                                      | 69 times (0.26%)    |
| Over gonads:                                        | 56 times (0.21%)    |
| Contact lenses:                                     | 4 times (0.01%)     |
| Over all electronic implants:                       | 712 times (2.65%)   |
| Congestive heart failure:                           | 5 times (0.02%)     |
| Acute pulmonary edema:                              | 2 times (0.01%)     |
| The second level of contraindication:               |                     |
| (Occupied 21.27% of all possible contraindications) |                     |
| Cardiovascular disease :                            | 846 times (3.15%)   |
| Osteoporosis :                                      | 257 times (0.96%)   |
| Acute injuries :                                    | 748 times (2.79%)   |
| Ischemic areas :                                    | 2778 times (10.35%) |
| Metal objects :                                     | 69 times (0.26%)    |
| Over all electronic implants :                      | 712 times (2.65%)   |

Cardiovascular disease (3.15%) is the highest frequency of the first level and ischemic areas (10.35%) is the highest frequency of the second level.

#### 4. DISCUSSION

By analyzing the physical therapy contraindications database, diagnosis knowledge base and patient's medical records, we can determine if there is any contraindication may occur according to the continual diagnosis in the medical history and new ordered physical therapies. For example: if a patient was diagnosed to get degenerative spondylolithesis, he may get structural disease due to tumor or infection too. So the treatment of cervical / lumber traction is inappropriate to this patient's situation.

By these analyzing methods, we can detect the possible contraindication according to the diagnosis in patient's medical records. The patient's health status may have been changed, so the result can only be a reference to

doctors for making clinical diagnosis and treatments. In addition, the results show that cold/hot pack, interferential therapy and cervical / lumber traction have the highest probability of contraindication occurrence in the sampled records.

We found that cardiovascular disease has the highest frequency in the first level of contraindications even though the probability is only 3.15%. However, patients' blood vessel of cardiovascular disease is narrow. Patients should be treated carefully as applying cervical / lumber traction to prevent severe injury such as stroke.

At the mean time, according to the regulation of the hospital, some useful information is not included in the medical records such as the detailed data about the applied region of a physical therapy. Without this information, we are unable to identify if the therapy and the contraindication are applied at the same body regions. Further studies are needed to resolve this problem. Another limitation of this study is that the analysis of contraindication is based on a patient's whole medical records 6 months prior to the order of the therapy. However, the patient may have recovered from diseases during the 6-month period. And, this change of status may not be reported in the patient's medical records.

#### 5. CONCLUSION

In this study, we have investigated the contraindication for physical therapy for a medical center in Taiwan and established the mapping table of physical therapies and possible contraindications and the mapping table of ICD-9 code related to possible contraindication detection. These mapping tables were adopted in the analyzing methods to detect possible contraindications according to specific physical therapy.

Retrospective study can help us to discover and establish knowledge base for further contraindication reminding system design. By analyzing the mass clinical data and articles, rules and guidelines can be presented more clearly. According to the study results, we will build a contraindication reminding system and expect the system can provide doctors a better way of detecting contraindications when generate therapy orders.

In conclusion, we expect these analyzing methods can not only save resources of the hospital, but can also increase the medical quality and improve patient safety in the future.

#### 6. REFERENCES

- [1] Ruei-Yung, Yang. Why Doctors Make Mistakes?. *Chang Gung Medical Journal* 2003;24:2-4.
- [2] Chau-Hsin, Chang. Discuss malpractices of inpatients in U.S.A. and Taiwan. *Medicine Journal*





- 2003;131.
- [3] Robert A. Raschke, MD, MS; Bea Gollihare, MS, RN; Thomas A. Wunderlich, RPh; James R Guidry, PharmD, BCPs; Alan I. Leibowitz, MD; John C. Peirce, MD, MA, MS; Lee Lemelson, RPh; Mark A. Heisler, PharmD; Cynthia Susong, RN, MS.. A Computer Alert System to Prevent Injury From Adverse Drug Events. *JAMA* 1998;280(15):1317-1320.
  - [4] Blum KV. Abel Sr. *et al.*. Medication Error Prevention by Pharmacist *Am. J Hosp pharm* 1988;45:1902-1903.
  - [5] Mathers C, Penm R. Health System Costs of Injury, Poisoning and Musculoskeletal Disorders in Australia 1993 – 1994. Canberra: Australian Institute of Health and Welfare. *Health and Welfare Series* 1999;No 6.
  - [6] Sattin RW. Falls among Older Persons: A Public Health Perspective. *Annu Rev Public Health* 1992;13:489-508.
  - [7] Berg RL, Cassells JS, eds. The Second Fifty Years: Promoting Health and Preventing Disability, Falls in Older Persons: Risk Factors and Prevention. Washington, D.C.: *National Academy Press* 1990:263-90.
  - [8] Jau-Jiun, Ji. *et al.*. Discuss the cause of malpractice with clinical pharmacologist. *Hospital Pharmacology Journa* .1994;11:189-192.
  - [9] E.L. Allen, K.N. Barker, M.R. Cohen, N.M. Davis, R.E.Pearson. Draft Guidelines on Preventable Medication Errors, *American Journal of Health-System Pharmacy* 1992;49:640-648.
  - [10] Leape, Lucian L.;Bates, David W.;Cullen, David J.;Cooper, Jeffrey *et al.* Systems Analysis of Adverse Drug Events. *JAMA* 1995;274(1): 535-543.
  - [11] Scott F. Nadler, DO, Michael Prybicien, ATC, Gerard A.Malanga, MD, Dan Sicher, ATC, MD. Complications From Therapeutic Modalities: Results of a National Survey of Athletic Trainers, *Arch Phys Med Rehabil*. 2003 Jun;84(6):849-53.
  - [12] Bill TJ, Edich RF, Himel HN. Electric Heating Pad Burns. *J Emerg Med* 1994;12:819-24.
  - [13] R. B and B. RN, “Modern Information Retrieval” , *Addison Wesley Longman* 1999.
  - [14] Kenney Ng. Survey of Approaches to Information Retrieval of Speech Message. *Draft SLSG MIT*, 1996.





# Handheld RFID Patient Safety Improvement – Case Study in Wan Fang Hospital

Xu Li-Xin, Luo You-Sheng\*, Zhou Zi-Xiang, Zhou Zi-Xiang, Xu Ming-Hui\*, Kao Cheng-Yan  
*Department of Computer Science and Information Engineering, National Taiwan University, Taipei, Taiwan*

*Taipei Medical University-Municipal Wan Fang Hospital Departments of Information \**  
*r94922095@ntu.edu.tw*

## Abstract

Nowadays, medical centers are working on setting a safe medicating environment by using information technology. This research is to build up a wireless handheld RFID patient safety enhancement system for medical personnel, which combines Radio Frequency Identification (RFID) technology, IEEE 802.11 (Wifi) Locating Engine, Personal Digital Assistant (PDA) and Front-Monitoring System. The goal is to improve patients' safety, communications among medical staff and medical staff and patients, medicating quality, and efficiently practicing in medical processes.

Medical staff experiences a high level of mobility. This phenomenon makes hospitals ideal application environments for pervasive or ubiquitous computing technology. Through the combination of information techniques, we can identify patients and medical staff by non contact identification, do real-time verification, getting medical record and colleagues' location immediately and provide context-aware service. Meanwhile this system can record the contact history of the interaction between medical staff and patients. It can also give alarm when clinical decision engine reporting a high risk testing result, notify of new result, provide medicine safety suggestion, improve the mobility and instantaneousness of patients' medical information and create a safe environment for patients.

**Keywords:** Patient Safety, RFID, Wireless Locating Engine, Context-Aware Service

## 1. Introduction

Taiwan Health Reform Foundation also point out that such research illustrated that if the number of patients cared by one nurse increase to 8, the mortality of patients in 30 days increase 31%.

The U.S. healthcare evaluation organization Joint Commission on Accreditation of Healthcare Organization (JCAHO) [10] proposed seven goals for all healthcare organization to improve patient safety, including increasing the accuracy of patient recognition, improving communications among healthcare service providers, elevating the safety in administration of high risk drugs, avoiding mistakes in operation procedures and operation sites in patients, enhancing the safety of

infusion pumps usage, reducing the risk of infection through healthcare [9].

Also, in 2004, IOM [11] had proposed that in order to improve patient safety during healthcare procedures, and one of major tasks is to use medical informatics techniques actively. In one way, real-time, accurate information are provided to the healthcare personnel; whereas information techniques are used to unveil adverse or near mistaken operations during healthcare processes. Healthcare procedures can be improved or redesigned based on this point of view.

Wanfang Hospital is a medical center hospital which works actively in integrating patient-oriented healthcare service system. Several goals are expected: 1. secure the patient safety, 2. develop patient safety information system and building a complete hospital-wide patient safety report system, 3. enhance patient-oriented healthcare quality.

In 2004, the Center for Patient Safety Informatics (PSI) was formed to promote the construction of infrastructures and applications of patient safety information system in Wanfang hospital.

In early stage, a medical information interaction model was built based on interactions among patient basic information, pharmacy, laboratory examination, diagnostics, and medical devices. This model is further combined with communication techniques and plans a series of patient safety information infrastructures and applications. The patient safety information system is partly described below.

Currently, Wan Fang hospital will issue a report while examination and radiology results are judged as high risk cases. These notifications will be sent with instant message or e-mail actively as soon as possible. The healthcare personnel are informed of the state of patients in real time, and take suitable procedures to avoid any possible delay.



**Figure 3** HRR SMS notification





Adverse Event Reporting System (AERS) provides an informatics interface to improve communications among healthcare service providers. Also, the records that are stored in reports can be fed to Computerized Physician Order Entry (CPOE) and produce analytical reports, avoiding adverse responses of patients from recurring (Figure 4).



**Figure 4** Adverse Event Reporting System (AERS)

In the adoption of novel technologies, Wan Fang Hospital introduces radio frequency techniques in the emergency area to verify patient identity through non-contact mean. Afterward, the integrated patient medical information can be obtained; interactions among healthcare personnel and patients are recorded; high risk laboratory examination results are highlighted actively; hints for drug use safety are provided (Figure 5). All in all, these procedures improve the accuracy of patient healthcare information promptly and construct an environment for patient safety.

Wan Fang Hospital has constructed a patient safety information infrastructure, and uses Web Services (developed by Visual Studio.NET platform) to integrate related patient safety information within the hospital. Wan Fang Hospital also attempts to build an information exchange platform for data integration and program component sharing. Each functionality is highly flexible and innovate applications are possible. Also, through the dynamic combinations of services, hospitals and clinics can respond to specific services or projects rapidly.

Hospitals are complex information-rich environments including a significant technique and computing infrastructure, the need for coordination and collaboration among specialists with different areas of expertise, an intense information exchange, and the mobility of hospital staff, patients, documents and equipment [2].

Unfortunately, the personnel usually have to use specific fix computers to access hospital information system, not to get immediately his/her colleague's location and need a convenient interface to get the patient's information.



**Figure 5** Emergency RFID system

For highly mobile healthcare personnel, unnecessary efforts and times are wasted with these inconvenient. In this work, we propose a concept that front-end system can be carried around by healthcare personnel, access hospital information system (HIS) everywhere, get the location information of other colleague anytime, and use RFID module to improve patient identification procedure.

In Section 2, we introduce an overview of RFID application in healthcare systems. Section 3 we describe the research method we use. Section 4 presents the overall architecture, including the wireless locating engine, of our system. Section 5 is the result and discussion. Finally, Section 6 we present the conclusion.

## 2. An Overview of RFID applications in healthcare systems

Radio Frequency Identification (RFID) has been applied in various fields, including libraries, malls, and MRT. RFID has following characteristics: 1. non-contact data accessing; 2. miniature and variation of tags; 3. well adapting to various environments; 4. high data security; 5. reusable tags; 6. no restrictions on directionality; 7. large data contents [2]. With These seven characteristics, RFID is well suited to perform identification in hospitals.

Last few years, most studies on the applications of RFID are focused on improvement of supply chain management. Seldom did they apply RFID to healthcare. In year 2005, the Compound Annual Growth Rate (CAGR) of RFID in healthcare is 54.3%, which is the highest CAGR among other application domains.

Several kinds of RFID application in healthcare is carry out such as tracking the movements of patients in ER, recording the time spend in ER, and using RFID information to improve patient waiting time [4].







In nurse station, using identification technology to authenticate healthcare personnel and build real-time shift database, to provide information on patient clinical treatment and queries on shift data. It also provides more basic patient information for later check and applications. These procedures can replace paper works for confirmation and control in healthcare and shift procedures. This could further reduce manual operations and improve the accuracy of patient medical information promptly [4].

During the outbreak of SARS, RFID was also applied to related epidemic prevention tasks. In the R&D program supported by MOEA, ITRI and Ton Yen Hospital developed “RFID tracking and control system” in collaboration, this system can track and segregate possible infected personnel rapidly. Later on, they developed a “RFID medical instruments management system”. This system can manage the access personnel, quantity, and locations of instruments. It can also monitor the on shelf periods.

Sun Yat-Sen cancer center applies RFID technology to pharmacy matching system, to avoid controversies in misuse of drugs.

For the improvement of patient safety, Taipei Medical University Hospital puts identification RFID tags on patients in operating room and provides related information, images, and audio to help medical personnel identify the patients. This information could further confirm the sites to be operated [6].

### 3. Methods

In the planning phase of the study, numerous interviews have been held among medical personnel from Wan Fang Hospital (including physicians and nurses) and the development team. During interviews, we discuss system requirements and design the system architecture and operating procedure. Later on the development team implements the practical system. In Table 1, we list major requirements from the medical personnel and our corresponding improvements.

### 4. Developmental environment and system architecture

We introduce a simple scenario of pervasive computing environments in hospitals.

While Dr. Lin is making his regular round in the hospital, his PDA allows him to review the patient's healthy information or various kinds of reports wherever he goes, avoiding moving to the specific computer.

**Table 1** Strategy for Improvements

| Requirements                                                                  | Improvements                                                                                                                                                                                                                           |
|-------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Identification                                                                | RFID and authentication                                                                                                                                                                                                                |
| Improve interactions between medical personnel and patients                   | Medical personnel can use PDA to display patient information, including reports, diagnostic descriptions, drugs, etc. Through this information, personnel can understand the status of patients rapidly and discuss with the patients. |
| Improve clinical notification system                                          | High risk report (HRR) from expert system displayed in highest priority, notify medical personnel to check these reports                                                                                                               |
| High risk drug use safety                                                     | Queries on patient drug-usage and online drug dictionary                                                                                                                                                                               |
| Reduce risks of infections from healthcare                                    | Through usage and RFID reading log, once the outbreak occurred, the system can help medical personnel to understand the contact history between medical personnel and patients.                                                        |
| Hospital personnel wants to check reports on patients anytime and everywhere. | Personnel can use this system to download all kinds of patient reports from HIS through hospital local wireless network.                                                                                                               |

Upon checking the patient in bed 222, he downloads the bed 222 patient's data through reading the patient's RFID tag. At the time his PDA alarms that this patient has High Risk Reminder (HRR) reports, judged by clinical decision engine. Dr. Lin must check these reports immediately to give a proper treatment. While analyzing these data, Dr. Lin has a little doubt and needs to discuss with the resident physician, who is also informed that HRR report exists by the message on his cell-phone. Dr. Lin's can use his PDA to find a nearest physician's location on the map and call the physician over. After a brief discussion and giving proper treatment, Dr. Lin moves to bed 223. He finds a notification on his PDA alarming him that this patient's new examination is on the line to review. Dr. Lin can use his PDA to these examinations without moving and discuss with the patient. This scenarios bridge the gap between medical staffs and patients.

There are five major components in this system, including “PDA front-end monitor system”, “Wifi location engine”, “medical information service platform”, “clinical decision engine”, and “medical databases” (Figure 4).





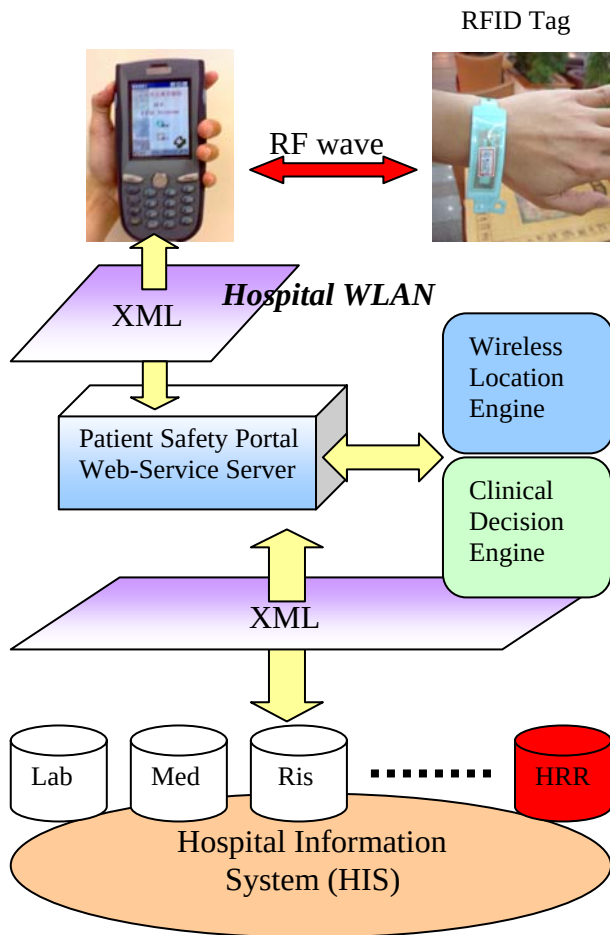


Figure 4 System Architecture

#### 4.1 Medical database

“Medical Database” is the relational databases storing patient-related data. The databases keep track of all histological data regarding patients, and provide all sorts of information required by various software system built for the hospital, this data, including basic information, registration time, drug usage, diagnosis, treatment, examination results, and information from shifts. These databases can speedup the information exchange procedure in the hospital.

#### 4.2 Medical information service platform

The major goal of medical information service platform is to integrate different database tables in the hospital, and extract medical information required by medical personnel. These services are packaged into many Web Services and medical system programmers can use these Web Services as components to develop novel applications (Figure 4).

The advantages for constructing this platform include:

1. programmer do not need to directly operate on numerous large databases distributed in the hospital.
2. Indirect accessing to databases avoids human-made damage to databases, also avoids database

synchronization problem. 3. The extensibility of such hierarchical platform is pretty well. Different services can be constructed for different systems. This facilitates the introduction of new techniques to the hospital.

#### 4.3 Clinical decision engine

“Clinical decision engine” is an expert system combine expertise of professional medical researchers. These professionals collect and build all sorts of clinical information to be used in the clinical system. Possible mistakes in operations are included, so as the system can provide hints to medical personnel at any time, which in turn improve patient safety. Related decision information includes clinical directions, drug usage directions, and examination reports thresholds. The decision engine can provide clinical directions to medical personnel. Also, clinical reports on patients can be compared with related knowledge stored in decision engine, and the data of patients exceeding threshold will be passed to medical personnel immediately.

#### 4.4 Wifi location engine

As a common wifi location mechanism, we divide whole indoor space of each floor into several small cells and then choose some reference points to gather the signal strength of access points of the hospital WLAN in these cells. Since these rooms have concrete walls, the radio signal strength of each room is quite different. We can construct a radio map for every floor with these power signatures. These cells are distinguished by the mean and standard deviation of their signal strength [9].

To reduce the computing burden of the PDA, all computing jobs is done by the server and data for location estimating, such as signal strength, is stored to the server-side storage. The front-end monitor system on the PDA just needs to send the gathering power signature of near access points to the server and then gets the returned estimated location from the location engine. After getting the estimated location, medical staffs can see the location of colleagues and specific medical devices on the map showed on the PDA. Through this locating mechanism, we can keep track of personnel's location information to enhance the communication procedure among personnel.

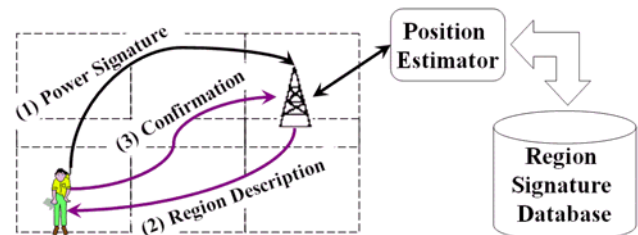
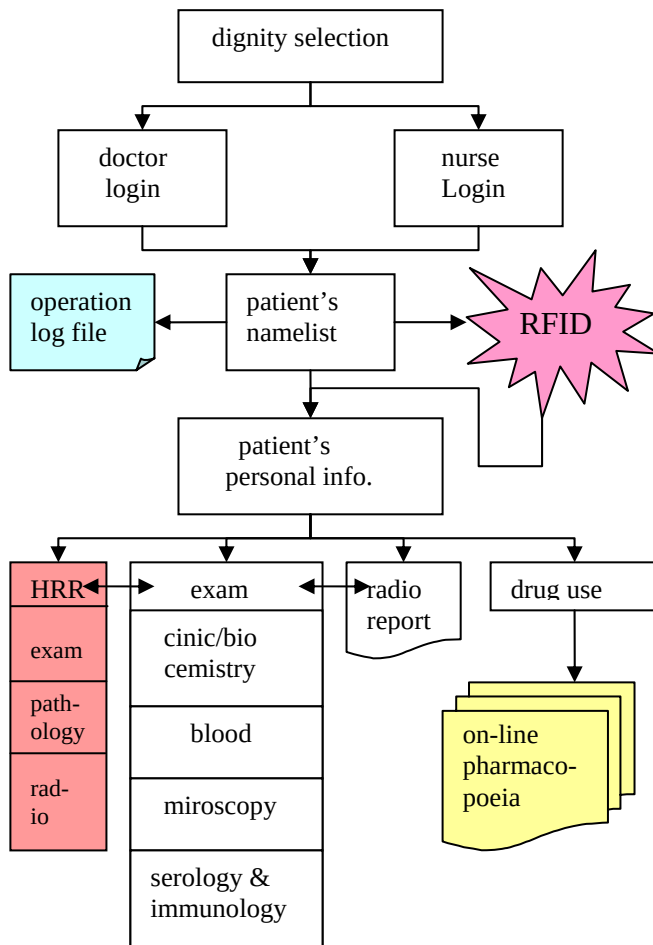


Figure 5 Wifi location engine

#### 4.4 PDA front-end monitor system





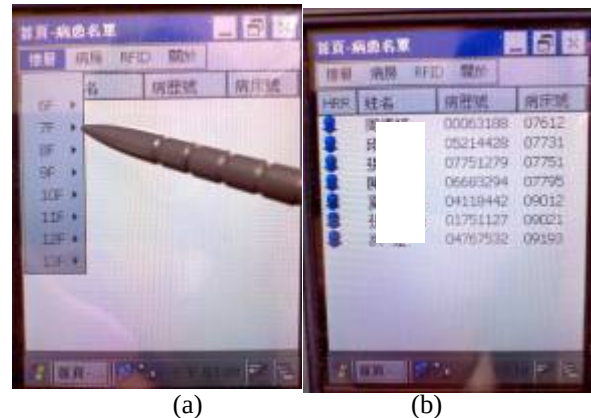
**Figure 6** PDA interface system diagram

The development environment of this interface is Microsoft Visual Studio .NET 2003. The designed operation flow of the system is illustrated in Figure 6.

To physicians, the patient list only includes information of patients handled by the physician (Figure 7b) to personalize the presentation of information to the doctor.

To nurses, when switching to patient list, they can select all rooms in the hospital and view patient information based on rooms and floors to adapt to the mobility of daily work flow of nurses (Figure 7a).

The icon in the left (Figure 7a, 8a) is used to distinguish high risk patients, which is judged by the clinical decision engine. If the icon is in red, the patient has at least one high risk reports; otherwise, the patient is in the normal group. Such design helps personnel to identify high risk patients and view their room numbers and bed numbers as soon as possible. This design enhances the performance of high risk notification system. User can press the column head of each column, e.g. serial number, the patient name list will be sorted based on the pressed field to provide multiple ways to review

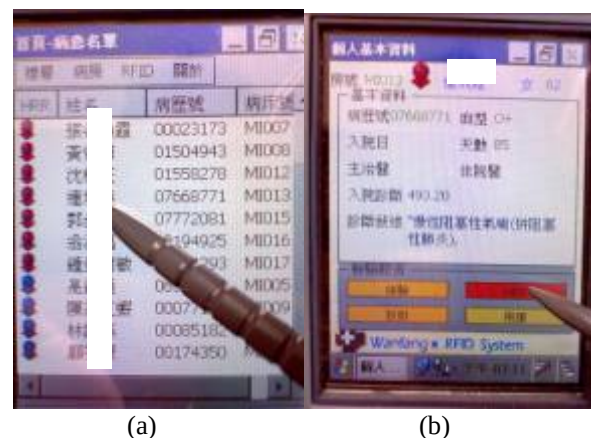


**Figure 7** Front page – patient information

The system can also keep track user's operation flow (Figure 13). User can check their own usage log, including every action and respective time after login, on the PDA. Medical personnel can rapidly lookup which room, patient, or RFID tag has been inspected. Also, analysis can be performed on these usage logs to improve system performance and work flow. Hospital can also track the contact information between physicians and patients through these logs.

If the medical personnel want to obtain more detailed information about a specific patient, s/he can double-click on that patient or read patient's RFID tag to bring up detailed patient information (Figure 8a, 8b).

Detailed patient information (Figure 8b) displays patient name, room number, HRR icon, gender, age, days in the hospital, date of admission, physician, resident physician, diagnosis at admission, and diagnosis description. Also, through clicking the four buttons at the bottom, one can view the examination report (Figure 9a), radiology report (Figure 9b), HRR report (Figure 10), drug use, and online pharmacopoeia.



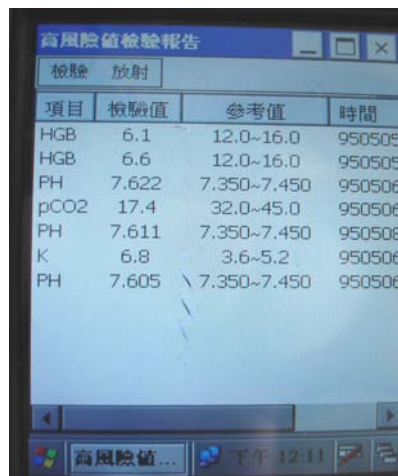
**Figure 8** Patient list and basic information

Examinations include biochemistry, hematology, micrology and serum immunology.





(a) (b)  
**Figure 9** Examination and Radiology Report



**Figure 10** High Risk Reminder Report

To date, the most popular medical applications on handheld devices are the ones that provide access to reference material such as drug information database [3]. The use of drugs change frequently and more than 20,000 drugs are available in the hospital. To avoid mistakes in drug use, we also provide query and display of patient drug use. We divide each day into 24 hours, when users select one of these time slots, the statistics of drug use during the period will be presented, including drug code, drug name, dosage, status, route of administrator, frequency, and unit (Figure 11a).

Medical personnel can query the drug in the online drug dictionary based on the selected drug number (Figure 11b). More information about this drug can be obtained, including: image of the drug, indications, contraindications, pharmacology, etc. These information help medical personnel to explain these drugs to patients, also reduced mistakes during drug using

#### 4.5 PDA

Through a series of test and compare, we adopt Unitech PA960 (Figure 12) as our handheld device. This is an industrial PDA with a built-in RFID reader, a barcode reader, large buttons and a 802.11 wireless network adapter. It is more suitable for medical

personnel, with higher mobility, to carry around than common commercial PDA devices

### 5. Result and Discussion

We have showed our system prototype described above to the hospital personnel and teach them how to use. We get good responses on this procedure. There are still some points we must take carefully. The computing power of a PDA is weaker than that of a common desktop computer. To make the operation smooth, responsive, and user friendly, special cares are taken in the design of work flow and switch of forms.



(a) (b)  
**Figure 11** Drug use and only drug dictionary



**Figure 12** Unitech PA960 PDA

Our tests and discussions have revealed the following points:

1. Simple operation flow. For example, in the programmer's view, the activation of RFID reader requires 5 or 6 instructions and several parameter initializations. For the common users, these steps should be combined into one button-click to reduce the manipulation burden. For heavy loaded medical personnel, simpler operations are better. Complex procedures will lower the acceptance of the system.

2. The presented medical information format should be simple and clear. The medical personnel face multiple patients and various reports everyday. They have to switch their attentions between patients and reports. The theme and functionality on each form should be

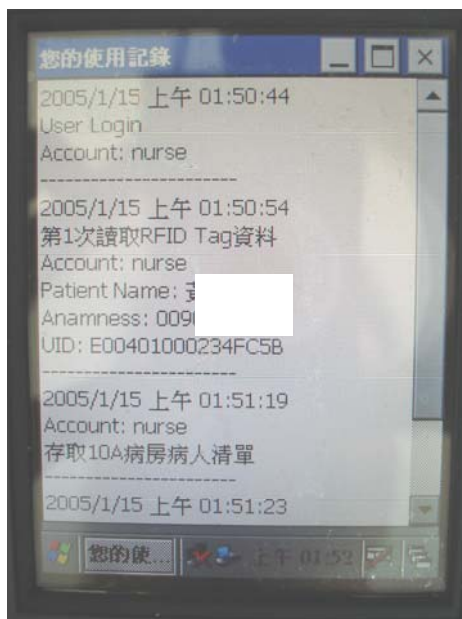






presented clearly to save operation time. In early design phase, we tried to display all information in one form on the screen. However, when tested by the medical professionals, it was complained that the forms are too complex and not easy to extract required information. Therefore the information should be prioritized and the forms should be tabularized, to avoid display of too much information or texts. The programmer might search non-official SDKs [12] from Internet to present data in a more complete fashion.

3. The accession of Web-Service platforms should be combined to speed up the data gathering procedure. Our front-end monitor system on the PDA accesses medical database through IEEE 802.11 wireless networks. The connection procedure of IEEE 802.11 wireless network is much slower than wired Ethernet. Based on our experiments, access to server and call a Web Service takes 3 to 5 seconds (based on the amount of data). If continuously access database in a single operation, users have to wait a long time and might misinterpret the system crashed. This situation should be avoided in both the client and server side.



**Figure 13** Log of user operations

4. Log of user operation flow. Tracking and analysis of user habits is critical to system development and research. With mathematical analysis of this information, we can not only improve system performance, but also provide more active services in a context aware manner [5]. Also, these log information are helpful in debugging the system. Furthermore, medical professional can get a grip of their work progress and contact history through the inspection of these information (Figure 13).

This system has been transferred to Wan Fang Hospital for testing. Therefore the medical personnel can operate the system after short educational training. They are also satisfied with the system.



**Figure 12** Educational training

## 6. Conclusion

The major goal of this study is how to use informatics techniques to help medical personnel enhance their daily work procedure, improve patient safety and spread properly the use of hospital information system.

There are seven goals for us to improve, including: 1. identification mechanism, 2. interactions between medical professionals and patients, 3. clinical notification system, 4. high risk drug use safety, 5. lower risk of infection through healthcare, 6. inspection of all kinds of reports, 7. inspection of basic patient information.

We have presented our efforts on this handheld interactive medical information system prototype, and it is operating online in hospital.

Hospital is a technical and context rich environment to apply mobile, ubiquitous, and pervasive computing. We hope this system to be a good reference for development and distribution of handheld medical information system in a hospital. We will continuously add more context-aware services to this system and make it powerful in the future. as a direction toward the development of a pervasive computing environment.

## 7. References

- [1] Yu-Sheng Lo, Cheng-Chung Yen\*, Min-Huei Hsu、Chien-Yuan Chen, Lin Yen-Yu, Chieh-Ming Huang, Da-Wei Liao, "The Research For Integration Framework of Hospital Patient Safety Based On Web-Services", Taipei Medical University-Municipal Wan fang Hospital Departments of Information, National Chung Cheng University Department of Information Management\*.
- [2] Jesus Favela, Marcela Rodrigues, Alfredo Preciado, and Victor M. González, "Integrating Context-Aware Public Displays Into a Mobile Hospital Information System", IEEE Transactions on Information Technology in Biomedicine Vol.8, No.3 2004.





- [3] Marcela D.Rodrigues, Jesus Favela, Edgar A.Martinez, and Miguel A.Munoz "Location-Aware Access to Hospital Information" IEEE Transactions on Information Technology in Biomedicine Vol.8, No.4 2004.
- [4] Li Hui-Ru, Cai Rong-Long, " The Application of RFID in Medical Industry in Taiwan".
- [5] Anind K. Dey , " Understanding and Using Context". College of Computing & GVU Center, Georgia Institute of Technology, Atlanta, GA, USA. Personal and Ubiquitous Computing 2001, pp. 4-7.
- [6] Wang Yue-Chuan, Li Liu, "Applying RFID to Improve the Patient Safety and Privacy ", MIST2004.
- [7] Yu Jain-Cheng 、 Zhang Bo-Lun, "The Development of Clinical Hand-Over System on PDA ", MIST2004.
- [8] Li Jia-Wei, "Re-examination the Relationship of Medical Staffs and patients, from the drama -White Tower ", NFL Research, Report National Policy Foundation , Oct. 8 , 2004.
- [9] Ming-Hui Jin, Eric Hsiao-Kuang Wu, Yu-Ting Wang and Chin-Hua Hsu, "An 802.11-based Positioning System for Indoor Application ", Communication Systems and Applications 2004.
- [10] Joint Commission on Accreditation of Healthcare Organization, <http://www.jcaho.org/>
- [11] Institute Of Medicine, <http://www.iom.int/tsunami>
- [12] The Premier .NET Compact Framework Shared Source Site, <http://opennetcf.org/CategoryView.aspx?category=Home>
- [13] <http://www.w3.org/XML/>







# The Requirement of Health Record to satisfy the patients' safety management

Kiyomu Ishikawa<sup>1)</sup>, Hisashi Ohmichi<sup>2)</sup>, Yoshimasa Umesato<sup>2)</sup>, Hitoshi Terasaki<sup>2)</sup>, Hidehiko Tsukuma<sup>1)</sup>, Norikazu Iwata<sup>1)</sup>, Takeshi Tanaka<sup>1)</sup>, Akie Kawamura<sup>3)</sup>, Kayo Sakata<sup>3)</sup>, Teruko Sainohara<sup>3)</sup>, Miyuki Sugimura<sup>3)</sup>, Nakao Konishi<sup>4)</sup>, Masashi Tooya<sup>5)</sup>

1) Medical Informatics and Hospital Systems Management, Hiroshima University Hospital, Hiroshima, Japan

2) Hospital Administration Science, Nihon University, Tokyo, Japan

3) Maternity and Perinatal Center, Hiroshima University Hospital, Hiroshima, Japan

4) Department of Nurse, Hiroshima University Hospital, Hiroshima, Japan

5) Department of Safety Management, Japan Council for Quality Health Care, Tokyo, Japan

E-mail: [humind3@hiroshima-u.ac.jp](mailto:humind3@hiroshima-u.ac.jp)

## Abstract

**Purpose:** As a basic medium of team medical care process under multi-specialties' cooperation, EHR is required to secure patient safety and meet the information disclosure demand. Based on the reassessment of the structure and description requirement of EHR, its next generation construction policy is proposed.

**Methods:** Surveys as the basis for its work were carried out by the authors organized and headed task forces in the hospitals certified by Japan Council for Quality Health Care (JCQHC). One of surveys was the assessment of the current state Health Record computerization. In addition, The Guidelines for healthcare record (HR) Structure and its computerization policy based on two questionnaire surveys were executed. Practically, they were used as criteria in evaluating handwritten health record and traditional EHR description.

**Results:** The survey of the above hospitals showed that most have computer supported administrative procedures, but only few computer-based health records. Besides, in terms of healthcare record which could be disclosed to patients, the attitudes of the Hiroshima EHR users showed that while they expected efficiency and quality improvements, there were considerable gaps among their evaluations based on the new guidelines. Especially, the lack of description items, system imperfection, and description contents varied conspicuously.

**Conclusion:** Systematically, EHR is available for integrated description. However, this result was caused by the subsystem construction for different users without sufficient understanding of medical care process. These insights led to the initiation of curricula educating "Health Information Technologist" which is promoted by Japan Association Medical Informatics and the criterion of Chart Review Promotion of JCQHC. They will also contribute to improved and advanced EHR.

## 1. Introduction

In order for the standardized HR to become the basic medium for the interaction of specialists and patients in the practice process, we have to comprehend the HR from the perspective of safety and quality of healthcare and reevaluate its computerization to arrive at a better blueprint for the next generation health system.

Since the Ministry of Health and Labor issued the basic design of the EHR in 2001, the computerization of the health record has met with increasing public interest. Additionally, the "IT policy objectives package 2005" released by Japanese government [1] will accelerate EHR development. As a next step, we plan to develop a Healthcare Navigation system over the next 5 years. The HR should serve as the basis for process and outcome evaluation. However, currently HR quality varies so widely that it is unlikely to serve as the basis for developing patients' confidence in the healthcare process.

Though the EHR is expected to decrease the variation in record quality, there is still no consensus on its social role. The idea that it is just the computerized version of the paper record still prevails. In our view, it will be necessary to develop it beyond that into the key medium for interaction of specialists in medical care. Despite our effort to reflect the care process in the system operation, the current system in the opinion of medical staff is so complicated and unusable that it might impinge on operational efficiency and threaten patient safety.

## 2. Research organization and activity

The Chart Review Committee (CHC) headed by the first author consists of 35 appointed members: healthcare system investigators (5), chart review specialists (15), and clinical specialists (7 medical doctors and 8 nurses)



of member hospitals of the Patient Safety Promotion Organization of Hospitals certified by the Japan Council for Quality Health Care (JCQHC).

The JCQHC is the nationally authorized organization, which accredits hospitals based on neutral assessment with academic perspectives in order to improve the quality of healthcare. In March 2006, there were over 1997 accredited hospitals, and 1050 hospitals were also members of the Patient Safety Promotion (PSP) council, an independent organization. Since PSP established CHC in 2004, its projects have focused on the EHR structure as a strategic tool to improve medical care safety through the following initiatives:

- 1) A survey of the actual status of inspection and evaluation of HR structure in the participating hospital,
- 2) Development of guidelines for HR structure with emphasis on clinical safety management and on making these guidelines available to the hospitals,
- 3) In this context, the advantages and disadvantages of HR computerization were analyzed in relation to their potential role in evaluating the quality of care process.

### 3. Survey of the health record inspection and evaluation in PSP member hospitals

A questionnaire was mailed to 1370 PSP member hospitals by CHC in 2004, and 972 hospitals (70.9%) replied. The number of beds under operation ranges most evenly from across 100 beds or lower to over 500 beds or more.

#### 3.1 Current standardization of EHR recording

There are 569 hospitals (58%) with individual guidelines and manuals for recording, and 341 hospitals (35%) do not have. In terms of summary forms for post-evaluation of medical care process, 654 hospitals (67%) use the unified forms in every department, but in 256 hospitals (26%) only main departments follow them.

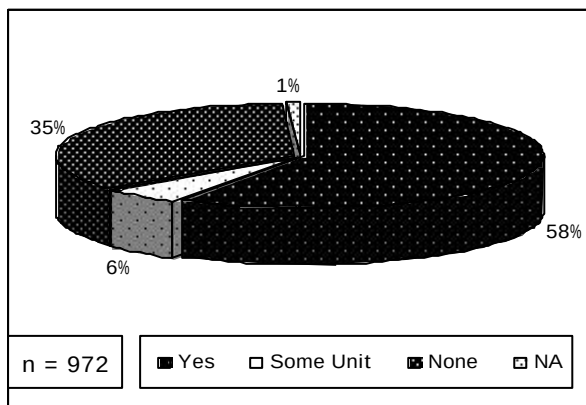


Fig.1 Current standardization of EHR recording

#### 3.2 Inspection condition of the HR

Sixty three percent of the hospitals serve as the center of clinical education and training for interns.

Among these hospitals, 82% have the management sections of HR, and 86% of them have review or inspection committees for the HR.

#### 3.3 Computerization circumstances

In 92% of hospitals, Patient Data were computerized, but health data documentation systems as "official document" is computerized in only 21%.

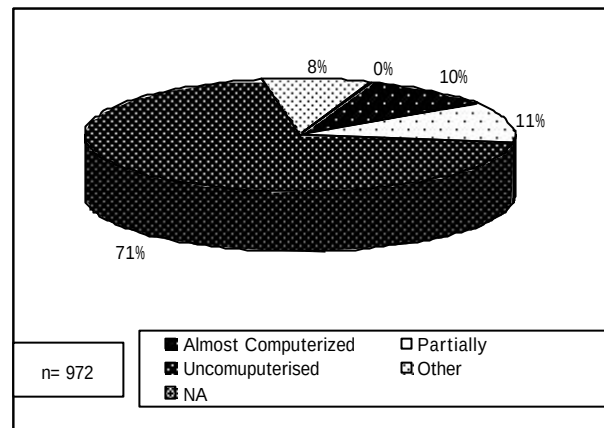


Fig.2 Japan current situation of EHR recording

### 4. Recommended Health Record Structure

In order to make the HR the key medium for organized medical service involving a variety of specialists, the HR Structure Guideline as composed by CHC is based on the following seven principles.

#### 4.1 Compliance with Personal Health Information Act

Healthcare records must be structured in compliance with the Personal Healthcare Information Act and related laws so that inpatients can control their personal information flow and at the same time rely on;

- a) Confidentiality
- b) Disclosure of personal information to patients
- c) Correction of errors
- d) Restricted use of personal information.

Careful documentation is required in order to protect the dignity of the inpatient and his/her family even after his death.

The primary use of information for the medical care of the patient has to be strictly distinguished from the secondary use, including social availability of information. In addition, the patient must be informed of



the purpose of secondary information use in advance in order to protect his/her dignity.

Appropriate information use should be guaranteed by appointing a managing specialist of healthcare information in every facility.

#### **4.2 Healthcare records as communication media in team practice**

In addition to records on medical treatment, operation, and anesthesia produced by doctors, organic coordination of documentation is also required for nursing records, rehabilitation records, records of pharmaceutical management and conduct, nutrition conduct, and medical consultation records.

In order to show the interrelation in team medical care, the form of document organization and filing should be documented.

The organization of the medical service is considered to be appropriately reflected by team-sequenced form of records, which integrates prioritized and ordering issues, setting targets, treatment plans, recording requested orders and their results, as well as conduct confirmation, and outcome evaluation.

Especially, the problem list and the care follow-up are expected to be integrated, precisely indicating the specialists' reasoning.

#### **4.3 Target-oriented description**

The target of medical care, which ranges from short and mid to long term goals in physical, psychological, and social life environment, should be defined and documented.

The action plan for achieving the target must be specified.

The target and action plan need to be adapted to changes in the patient's physical condition and circumstances.

#### **4.4 The guarantee of generalization**

Since various events and episodes occur in parallel during the care process, and the general condition of a patient and care process should be easy to grasp, it is necessary to arrange and specify medical care conducted and documented by other medical staff, so that specialists with diverse viewpoints can cooperate in achieving the same treatment goal.

The lists of problems, targets, and the plan of action are to be placed at the beginning of the file, and an interim summary is to be included in the process records. It is crucial that each episode consists of the compact and clear description on facts, events, subjective/objective problems, assessments, plans, interventions, outcome evaluation, and a plan to the next stage.

In reappearance of problems, the previous listed

number is reused. While, in the case of unlisted episode the new number is added for achieving continuity.

In order to promote information availability of the patient and his/her family, the use of jargon or abbreviations is discouraged in favor of understandable terminology.

#### **4.5 The guarantee of continuity**

For newly emerging conditions, whether related to medical or nursing treatment, including interview, observation, measurement, evaluation, diagnosis, planning, procedure, treatment, and evaluation of outcome, recording of the date of occurrence is required.

Recording of time is desirable when complex medical treatment is conducted during a single day. The start of an episode of the disease process, the time relation of change, the service sequence among several attending specialists should all be systematically described for easy understanding.

When reference is made to past episodes of medical or nursing care, the date and time of episode and those of the recording should both be documented.

#### **4.6 The guarantee of authenticity**

In order to make certain and correct communication among specialties, logging of the following items is indispensable with practical date and time. (Fig. 3)

- a) Orderer
- b) Subject and object
- c) Receiver
- d) Check the order correctness to avoid mismatching with previous information
- e) Doer
- f) Confirmation of action

Objective description of events is required. Accurate verbs should be used in expressions of speculation and inference. The newest data should be shown. Basically, an episode must be recorded as soon as it is recognized while the staffs concerned are on duty. (At latest within the same day)

In the EHR, a permanent preservation file is required in order to avoid replacing the original records. "Documentation errors" of care should be positively corrected. The date and time of correction, those of the recording of the outdated information, the content before and after correction, and the name of corrector all have to be documented.

Basically in correcting handwritten healthcare records, after double lines are drawn to show the incorrect content, new content is added. Use of correction fluid and erasers is not allowed. In documentation, simple writing errors, which do not affect recorded context, serial correction stamps, and signatures, are unnecessary.

#### 4.7 The guarantee of original records

Information meant to serve as official records should be shown individually according to forms such as the first record form for public insurance healthcare, patient

charts, ordering forms, process list, operation and anesthesia reports, summary, and test results.

Medical facilities, which use both handwritten records and EHR, need to define the official form for each healthcare record.

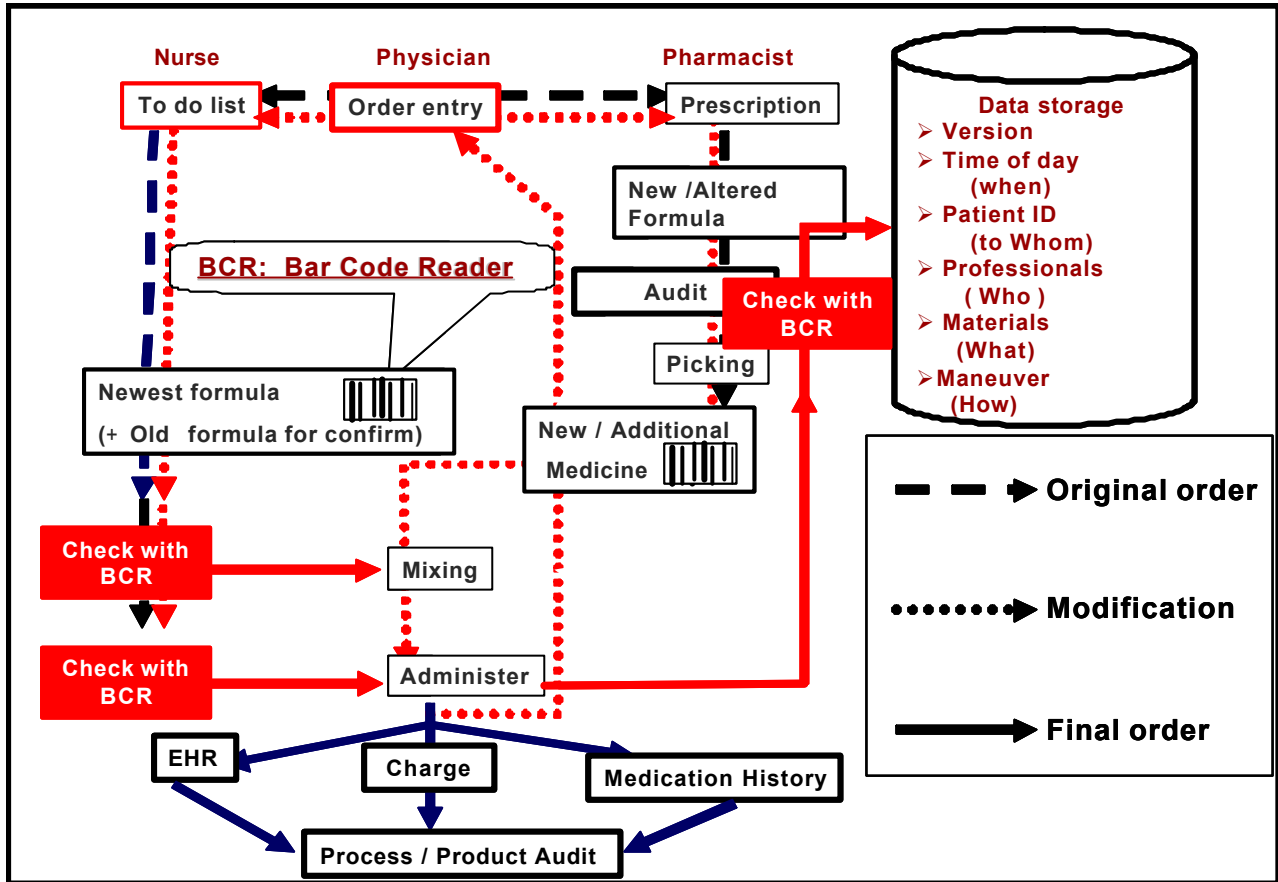


Fig. 3 Business flow for guarantee of responsibility of action and authenticity (Ex. the data accumulation and storage process model in injection)

#### 5. Audit of handwritten HR and classical EHR on safety management

Based on the recording guideline mentioned above, the recording evaluation guideline was established to secure patient safety.

In order to obtain the objective outcome, it consists of 15 items from the view point of “Formative Issues” and “Narrative Issues”. The one focuses on the readable description without misleading expression, and the other team practice. In hospitals, diverse specialists are involved in evaluating operation. Tentatively, 12 reviewers (3 doctors, 3 nurses, and 6 health record inspectors) made the evaluation by means of 3 handwritten healthcare record files and 3 provided print out EHR files.

According to its outcome shown in Table 1, there was a gap among individual evaluation scores though

cases concerned were insufficient for making statistics. Specifically, in terms of “Formative Issues”, EHR got higher score than handwritten HR. In contrast, in “Narrative Issues”, they were almost equal, or handwritten HR was superior

#### 6. An example of the EHR to the safety of medical care

##### 6.1. To guarantee the generalization and continuity of practice

The Health Record (HR) should serve as the basis for process and outcome evaluation. Here, as a practical case of safety management, Healthcare Navigation System is presented. (A model of the HNS, which is the core of the EHR in the Hiroshima University Hospital, may serve as a practical example of the implementation of HNS in the



service of safety management. [2]) The HNS has to facilitate safe team practice see. (Fig.4)

Data through patient healthcare (PHD) are accumulated not only by physicians, but also by all

co-workers. This may be the basis for team communication and a reliable way of treating in such a way that the evidence of conducted practices is traced.

**Table 1 Outcome of Recording Guideline-based Trial Evaluation to Secure Patient Safety**

| Issues                                                               | ***Reviewer |   |   |   |   |   |   |   |   |   |   |   | total | average |
|----------------------------------------------------------------------|-------------|---|---|---|---|---|---|---|---|---|---|---|-------|---------|
| 1) Is lettering legible ? *                                          | 2           | 2 | 3 | 3 | 2 | 3 | 2 | 2 | 2 | 2 | 2 | 3 | 28    | 2.3 #   |
|                                                                      | 4           | 4 | 3 | 3 | 4 | 3 | 4 | 4 | 3 | 3 | 4 | 4 | 43    | 3.6     |
| 2) Is there any lack of description ? *                              | 2           | 3 | 3 | 3 | 2 | 2 | 2 | 2 | 3 | 2 | 2 | 3 | 29    | 2.4     |
|                                                                      | 2           | 4 | 3 | 3 | 2 | 2 | 3 | 2 | 3 | 3 | 2 | 3 | 32    | 2.7     |
| 3) Is there any recording form ? *                                   | 2           | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 2 | 2 | 2 | 3 | 32    | 2.7     |
|                                                                      | 4           | 4 | 4 | 3 | 3 | 3 | 4 | 4 | 3 | 4 | 3 | 4 | 43    | 3.6     |
| 4) Is there any unsuitable term ? **                                 | 1           | 2 | 3 | 3 | 2 | 2 | 3 | 2 | 2 | 2 | 2 | 2 | 26    | 2.2     |
|                                                                      | 1           | 2 | 3 | 3 | 3 | 3 | 2 | 1 | 2 | 2 | 2 | 2 | 26    | 2.2     |
| 5) Is filing understandable for healthcare process ? **              | 2           | 2 | 2 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 2 | 3 | 32    | 2.6     |
|                                                                      | 2           | 4 | 2 | 4 | 3 | 3 | 2 | 2 | 3 | 3 | 2 | 3 | 33    | 2.8     |
| 6) Is there any misleading expression ? **                           | 3           | 3 | 2 | 3 | 3 | 3 | 3 | 2 | 2 | 2 | 2 | 2 | 30    | 2.5     |
|                                                                      | 2           | 2 | 2 | 3 | 2 | 2 | 3 | 2 | 3 | 2 | 3 | 2 | 28    | 2.3     |
| 7) Is team practice expressed ? **                                   | 2           | 2 | 1 | 2 | 2 | 2 | 2 | 1 | 2 | 2 | 2 | 2 | 22    | 1.8     |
|                                                                      | 2           | 1 | 1 | 2 | 2 | 2 | 2 | 1 | 2 | 2 | 2 | 3 | 22    | 1.8     |
| 8) Is problem clearly indicated ? **                                 | 2           | 2 | 2 | 2 | 2 | 2 | 2 | 1 | 2 | 2 | 2 | 3 | 24    | 2       |
|                                                                      | 2           | 3 | 2 | 2 | 2 | 2 | 2 | 1 | 2 | 2 | 2 | 3 | 25    | 2.1     |
| 9) Is the goal of healthcare specific ? **                           | 2           | 2 | 2 | 2 | 2 | 2 | 3 | 3 | 3 | 2 | 2 | 3 | 28    | 2.3     |
|                                                                      | 2           | 2 | 3 | 2 | 2 | 2 | 2 | 3 | 3 | 2 | 2 | 3 | 28    | 2.3     |
| 10) Is solution procedure clearly indicated ? **                     | 2           | 3 | 2 | 2 | 2 | 2 | 1 | 2 | 2 | 1 | 2 | 1 | 22    | 1.8     |
|                                                                      | 2           | 3 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 1 | 2 | 2 | 24    | 2       |
| 11) In case any care process change occurs is the fact recorded ? ** | 2           | 2 | 2 | 2 | 2 | 2 | 3 | 3 | 3 | 2 | 3 | 3 | 29    | 2.4     |
|                                                                      | 3           | 2 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 2 | 3 | 2 | 30    | 2.5     |
| 12) Is outcome expressed ? **                                        | 1           | 2 | 1 | 1 | 1 | 1 | 3 | 3 | 2 | 2 | 3 | 2 | 22    | 1.8     |
|                                                                      | 1           | 2 | 2 | 1 | 2 | 2 | 2 | 3 | 2 | 2 | 3 | 2 | 24    | 2       |
| 13) Are evaluated process and outcome ? **                           | 2           | 2 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 2 | 19    | 1.6     |
|                                                                      | 2           | 2 | 2 | 1 | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 2 | 19    | 1.6     |
| 14) Is there any inconsistency of fact among recorders ? **          | 1           | 1 | 2 | 1 | 2 | 2 | 1 | 2 | 1 | 2 | 2 | 1 | 18    | 1.5     |
|                                                                      | 1           | 1 | 1 | 2 | 2 | 2 | 1 | 2 | 1 | 2 | 2 | 1 | 18    | 1.5     |
| 15) In case any adverse event occurs is the fact recorded ? **       | 2           | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 1 | 2 | 2 | 2 | 21    | 1.7     |
|                                                                      | 2           | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 1 | 2 | 2 | 2 | 21    | 1.7     |

1)-3) Formative Issues\*, 4)-15) Narrative Issues\*\*, A – L: Reviewer A-C= MD, D-F= Nurse, G-L= Health Record Inspectors\*\*\*In each field Upper Row: handwriting HR, Lower Row: EHR, Score #: suitable < 4, 3, 2, 1 > unsuitable

## 6.2. Contribution to safety healthcare

The system is designed to share the information among all attendants and the patient concerned to



promote team practice, which is expected to enhance safety and high-quality medical care as well as rational hospital treatment. Modules of HNS are available not only to the physicians but also to all other health professionals concerned. [2]

A physician may give orders to other professionals

who show order confirmation signed on the same browser. The interactions of information management in the clinical process are implemented and presented on a browser chronologically and iteratively.

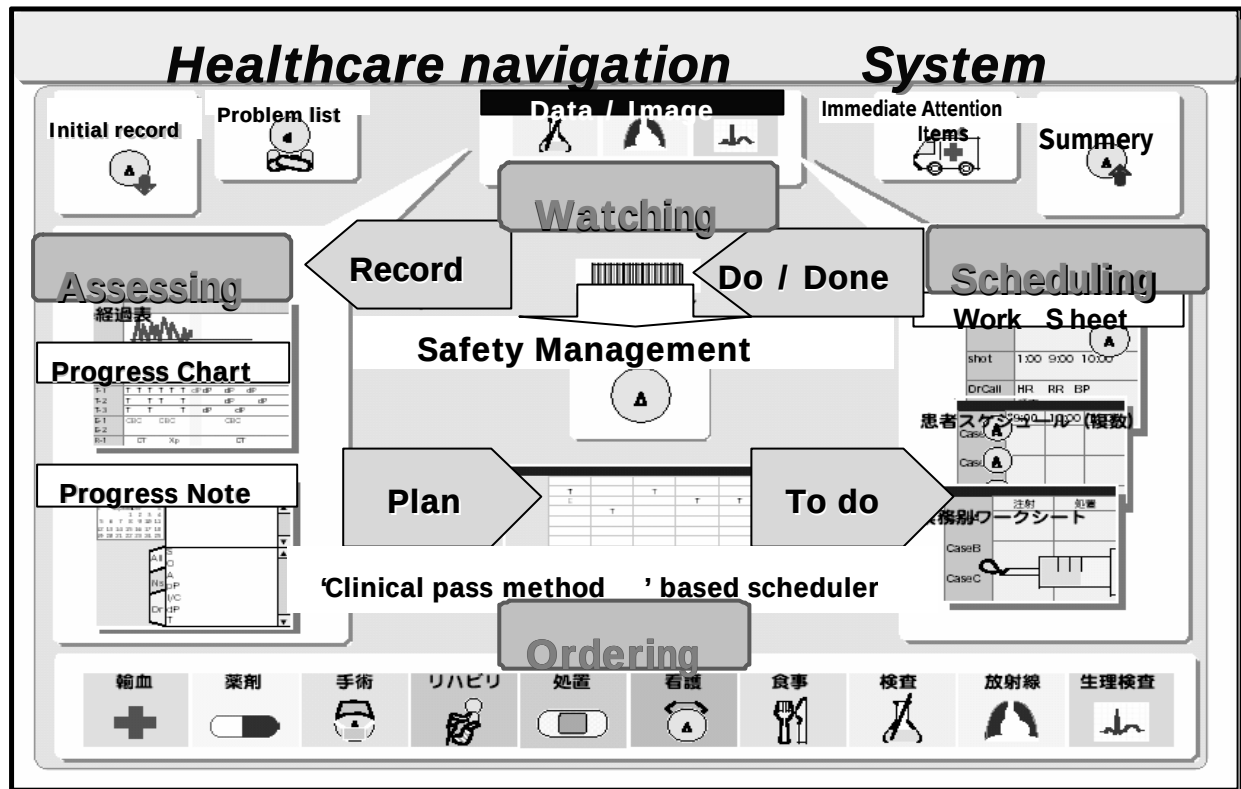


Fig4 the Index browser of Healthcare Navigation System; Specialists are able to access suitable function module on demand with this integrated browser.

### 6.3. Responsibilities for data users

Using conventional technology, it is difficult to fit our traditional handwriting with the clinical process. (In Japan, the common procedure for “official document” is simultaneous recording by physician himself with conducting medical treatment except operation or invasive treatment record. It is unusual that physicians make oral record of the findings of his patients’ condition, results of examination and treatment and so on.)

In the case of the Hiroshima University Hospital and associated hospitals, the process of problem identification, target setting, planning, ordering, order confirmation, result reporting, and evaluation are carried out with a combination of computerized and manual documentation.

A questionnaire survey of all healthcare staff of several hospitals in Hiroshima was conducted, focusing on assessing the perceived advantages and disadvantages of the EHR on medical care safety and operation efficiency.

The results are the following:

**6.3.1. Advantages of the EHR.** The perceived advantages of the EHR include:

- Availability of objective information; the medical and nursing care process are documented in line with the organizational rules such as operation management, procedures, and recording forms.
- The need for repeated copying is reduced, which potentially increases operation efficiency. Furthermore, some specialists answered that computerization was effective in overcoming the drawbacks of handwritten records.

**6.3.2. Disadvantages of the EHR.** These contrast with the following perceived disadvantages:

- Security assurance is difficult; Accidental system troubles disrupt recording or daily operation. Sometimes, a stored record is deleted.
- It is difficult to draw attention to specific issues in the clinical process; with handwriting, it is easy to emphasize important parts of a document.

- c) Lack of flexibility in operation; specifically, slow response. Due to the complicated screen layers and monotone screen layout and presentation, the recording operation is made difficult.
- d) Time sequences are difficult to represent; information retrieval is difficult, especially across different data classes, such as test results and pharmaceutical history (Database structure has to be improved.)
- e) Interactive use involving several specialties is particularly inconvenient. This impacts team care negatively. (Traditionally, specialized sections have evolved independently, which now leads to the need to integrate these separate systems.)

## 7. Discussion

Recently, Japan is facing the critical issue of accountability to secure patient safety. Then, the appropriate recording and post-evaluation of medical care process is expected as effective solution measures. (Motive for the increased attention to quality is a number of adverse or incorrect events that were identified in the clinical process. As a result, the management of the safety of the healthcare process and the manner of disclosure of patients' information have become high priority issues [3]-[6].) This paper focuses on the HR architecture

### 7.1. Confidentiality and Data disclosure are skeletal essentials of patient oriented care

The Japan Personal Data Protection Act was established in May, 2003. It is based on OECD 8 principles which practically became the major pillar of US HIPPA-led worldwide legislation. After adding new clauses such as the obligation of organizations dealing with personal data (Hospitals, healthcare facilities, nursing homes and so on in the medical field), it was put into effect in April, 2005. [7]

Since the act urges that ministries which manage the project "should develop the guidelines for the right" use of personal data, the National Life Council issued a guideline on personal data protection based on the fundamental principles and anticipated future direction of the act. [8]

Under the ongoing sophisticated computerization, and with consideration of the usefulness of personal data to both the public and private sectors, the act aims to protect personal right and benefit, including individuality, dignity, and property. This is expected to lead to well-balanced data access. [9]

It is not too much to say that medical services rely on the accumulation and use of PHD. Besides, without doubt, "right" data access is the key to the improvement in public health and the development of science. Here, the Personal Data Protection Act will be

refocused in order to patient oriented healthcare.

## 7.2. Requirements for Safety Management of Healthcare Record

### 7.2.1. Securing accountability and Authenticity.

Healthcare record originates from patients and their families who are benefited by medical service. Then, medical care record is described by diverse medical specialists who extract information from patients based on their individual viewpoints in the process of medical examination, treatment, and nursing. Furthermore, the recording responsibility is certified as original sources.

**7.2.2. Securing integrity.** Team medical practice concerns diverse specialists. In addition to the conscious manipulation, the content change caused by copying errors during information transmission is unacceptable.

**7.2.3. Securing availability.** Securing availability is the top priority in PHD computerization. It is essential that the medical care beneficiaries and the authorized specialists for information use can refer PHD as needed. Consequently, the following requirements are the major pillar.

- a) "Correct information" is recorded in "understandable expression."
- b) "Integrated information sharing" is carried out between medical care beneficiaries and specialists concerned.
- c) "Information flow" is "shown in record."
- d) Not only personal record but also "information of similar cases" can be referred as needed.

## 7.3. The requirements for EHR functionality

The following principles have been defined for EHR functionality:

- a) Security, authenticity, accessibility on demand, and preservation
- b) Convenient and reliable documentation in accordance with the need of the medical staff during the team treatment process
- c) Support for prompt overview and easy generalization of medical treatment process on demand
- d) Information sharing by all staff concerned

On the other hand, despite our effort to fit the care process to system operation, the current imperfect system has, in the opinions of the healthcare staff, the potential to become so complicated and unusable that it might decline the operation efficiency and threaten safety.

We therefore have to re-evaluate the structure; functionality and architecture of our HR system from the perspective of safety and healthcare quality in order to arrive at a better blueprint for the next generation system.

## 8. Conclusion

In Japan, majority of medical facilities have aimed at constructing the integrated-typed EHR, whose current function is still imperfect in terms of information sharing in local healthcare network. Specifically, despite the need of common data item set, Japanese EHR system is facing the following issues. 1) Due to the individual data item sets expression in each EHR, it is difficult for medical specialists to understand the record description of other medical teams. 2) Under the popular time-sequenced recording, medical process is not easily understandable. The reasons are; 1) Primarily, the construction of EHR was based on hand-written HR of individual medical facilities. 2) HR description methods did not focus on the concept that recording of medical process would lead to medical safety. Consequently, based on the preceding points, this study drew out the new guidelines from 7 viewpoints.

Guidelines for healthcare record description will enable medical professionals to record systematically the common data on patients concerned during medical practices, which aims at the further efficient and continuous patient-oriented medical care.

Self check sheet for safety medical care: Based on the above guidelines, this sheet will be available in checking and evaluating the recording description. At the same time, nation-wide network systems of their procedures and standards are its ongoing project. Furthermore, in 2005, we began the related research which focuses on EHR construction used as strategic tool to improve the safety medical care. Traditionally, in terms of healthcare record, disclosure to patients has not been the priority. At the same time, medical professionals have been heavily tasked due to several data input including ordering of medical practices, assuring orders, and recording practices. Conclusively, current EHR cannot cope with thinking and practicing process of doctors and nurses in the innovative medical playing field. Here, this workshop will discuss its future direction based on the working hypothesis that the conception of "simple recording" will give a negative impact to its development.

Information technology guided communication should be the priority for the future in order to enhance the patient's participation in his healthcare, as well as safe management of team practice. In addition, our concept for the next generation EHR emphasizes safety management and patient participatory health care.

Though the EHR is expected to render documentation quality more comparable, its social comprehension has not yet reached a consensus, and the idea that it is just the computerized version of the traditional paper record is still prevalent.

In order to develop information basis which contributes to securing higher quality, efficiency, and transparency of healthcare, not only technical innovation,

but also health information officers with global standard for healthcare play the leading role. Therefore, fostering such specialists in information venders and management hospital staffs will be the key.

Japan Association of Medical Informatics (JAMI) launched the human development program to educate "Healthcare Information Technologists (Health IT)" since 2003. Now, at the end of 2005, there are 3843 Health Information Technologists. This is expected to contribute to a solution of the identified issues.

## 9. References

- [1] IT policy objectives package 2005, <http://www.kantei.go.jp/jp/singi/it2/kettei/050224/050224pac.html> (Visited at 03.19.2006)
- [2] K. Ishikawa, Nakao Konishi, Hidehiko Tsukuma, Satoko Tsuru, Akie Kawamura, Norikazu Iwata, Takeshi Tanaka. A Clinical Management System for patient participatory : Health Care Support, International Journal of Medical Informatics 2004; 73:243 -249
- [3] Kohn LT, Corrigan JM, Donaldson MS. To err is Human: Building a safer Health System. Washington, DC: National Academy Press; 1999
- [4] Brennan TA, Leap LL, et al. Incidence of Adverse Events and Negligence in Hospitalized Patients: Results of the Harvard Medical Practice Study I. N England J Medicine. 324:307-376, 1991.
- [5] Thomas EJ, Studdert DM, et al. Incidence and types of adverse events and negligent care in Utah and Colorado. Med. Care. 38:261-271, 2000.
- [6] Wilson RM, Runciman WB, et al. The quality in Australian Health Care Study. Med J Australia. 163:458-471, 1995.
- [7] Japan's personal information act, [Law No.57, 2003] [completely enforced on April 1, 2005] [http://www.privacyexchange.org/japan/JapanPIPA2003v3\\_1.pdf](http://www.privacyexchange.org/japan/JapanPIPA2003v3_1.pdf) (Visited at 03.19.2006)
- [8] The guideline for the appropriate handling of personal health data <http://www.mhlw.go.jp/shingi/2004/12/dl/s1224-11a.pdf> (Visited at 03.19.2006)
- [9] The guideline for safety management of the health information systems <http://www.mhlw.go.jp/shingi/2005/03/dl/s0331-8a1.pdf> (Visited at 03.19.2006)

### (Acknowledgements)

I would like to thank Jochen Moehr, Francis Roger France, Albert Bakker, and Barry Barber for giving advice for this article. This research was supported by the Japanese Ministry of Education 2005 and Sciences and Ministry of Health Welfares and Labors 2004 -2005.



## A Study on a Nationwide Ambulance Communication System via HEOs to Support Telemedicine

Isao Nakajima, Hiroshi Juzoji, Muhammad Athar Muhammad,  
Toshihiko Kitano, Naokazu Hamamoto\*  
Tokai University School of Medicine  
National Institute of Information and Communications Technology\*

### Abstract

The Government of Japan is investigating the ability of HEOs( Highly-inclined Elliptical Orbits ) to cover nationwide without blocking and shadowing, especially application of mobile communications during disaster and emergency situations. We have studied on a nationwide emergency medical communication system via these satellites to support telemedicine during the transportation.

### 1. Purpose

The Emergency Medicine has the requirement to transmit real-time video from a moving ambulance to an Emergency Medical Facility (Triage Center) to enhance efficiency. Real-time video reception will allow EMS staff to assess patient condition, perform triage, make preparations, and provide critical guidance to ambulance personnel while the patient is en route. Tokai University and JAXA have developed a state-of-art satellite tracking and transmitting system on Ku, or Ka-band with real-time MPEG-2 and MPEG-4 video up to 6 Mbps for HEOs system. The outline of the HEOs system( 3 satellites constellation, 8 hours for each ) shown on the Figure 1 and on Figure 2.

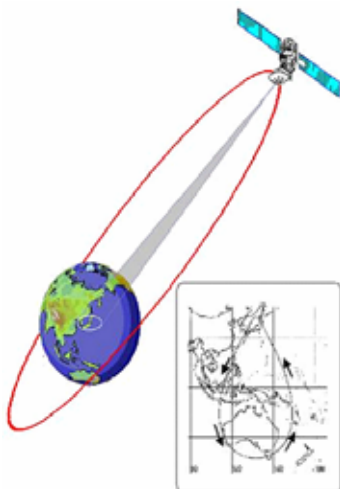


Fig. 1 HEOs system

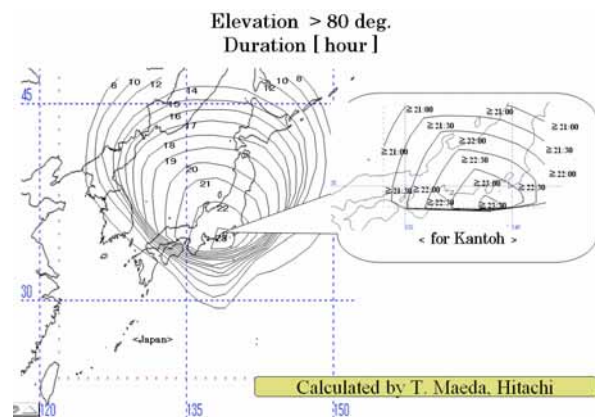


Fig. 2 Total duration with elevation angle over 80 degree

### 2. Back ground

ITU-D SG-2 Q14/2 "Telemedicine" has been study to telecommunication application for medical field, "Telemedicine Lessons Learned" and "Telemedicine Directory" ( Documentation 2/195 ) were submitted by Q14/2 to SG-2 in September 2001 and 2005 respectively, and are now available on the Web. These activities supported by Prof. Leonid Androuchko; Rapporteur, Isao Nakajima; Vice-Rapporteur of Q14/2, and they have described how to establish a commercially successful telemedical service on the ITU publication. Based on these activities, Tokai University School of Medicine; the first sector member of ITU-D from medical field, NICT, and JAXA have been studied on "Mobile eHealth"; next generation ambulance communication system to support telemedicine in Japan.

### Interpretation of Article 20 of the Medical Law

As of December 24, 1997, the Director of Health Policy Bureau sent to all municipal governors nationwide a notification (Health Policy Bureau No.1075) identifying the principle of face-to-face diagnosis in Article 20 of Medical Law and the basic ideas of remote medicine. This notification comprised ten items. The major items can be summarized into the five points, as follows in





Table 1. Of particular importance, we can note that the first medical diagnosis performed through a communications channel has been permitted, and we presume that it includes the diagnosis of a patient in an ambulance through high-quality communications channels. Most healthcare controls have been implemented on the basis of such notifications in Japan. Legal experts thus believe that the remote medicine employing image data transmission from ambulances has a solid legal foundation.

1. Remote medicine using information communications tools does not violate Article 20 of the Medical Law.
2. Remote medicine is to be regarded as a tool supplemental to face-to-face diagnoses, and its application is to be limited to patients in stable condition – for example, patients returning for a second or subsequent visit for the diagnosis of a chronic disease.
3. Remote medicine is to be performed at the request of the patient and must be beneficial to the patient.
4. Remote medicine may be used in the cases where there is difficulty gaining access to in-person medical examinations such as patients in remote areas and bedridden senior citizens (in this case, primary diagnosis by remote medicine is permitted).
5. The accountability of doctors and the responsibility of patients are clearly stipulated.

Table 1. Summary of the official notification on remote medicine in Japan

#### Ambulance communication in Japan

In 1983, the current emergency radio system was increased to 42 frequencies (21 opposing) by the Ministry of Posts and Telecommunications Radio Frequencies Administration Head. Mobile phones are used alongside this system to relay patient information, but they suffer the following problems:

- . Small service area
- . Limited number of frequency channels.
- . Interference from adjacent areas
- . "Shadowing" effect of buildings and trees.
- . Insufficient bandwidth for image transmission.

### 3. Study

#### 3.1 Necessity channels

Year after year, the total number of transported patient have been increased all over Japan. Statistics from the Tokyo Fire Defense Agency were the only large-scale data available for our study in Japan. Thus, Tokyo was unable to coordinate medical control for ambulance.

As of 1999, Tokyo had a population of 12,059,237, served by 201 ambulances. The ambulances were operated a total of 537,416 times, with 504,675 people transported(Fig. 3). The average transport time was 1,002 seconds and victims who were seemingly dead or

victims in critical or serious condition accounted for 8.4% of all those transported. We thought that this 8.4% of victims would be the target of the patient data transmission system. With Erlang's B-formula(Figure 4), the minimum number of channels required was calculated as six ( for daily operations in Tokyo) for a call loss probability <0.02. So, 6 channels for Tokyo, then total 64 channels should be needed nationwide with 4800 ambulances.

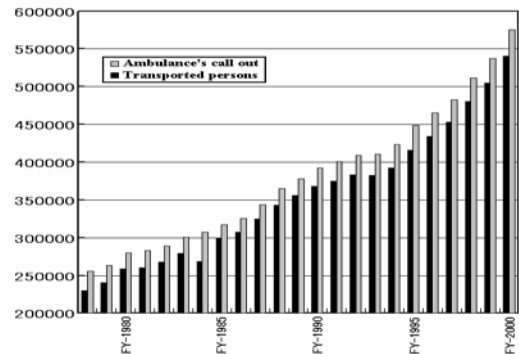


Fig.3 Emergency transportation by ambulance in Tokyo

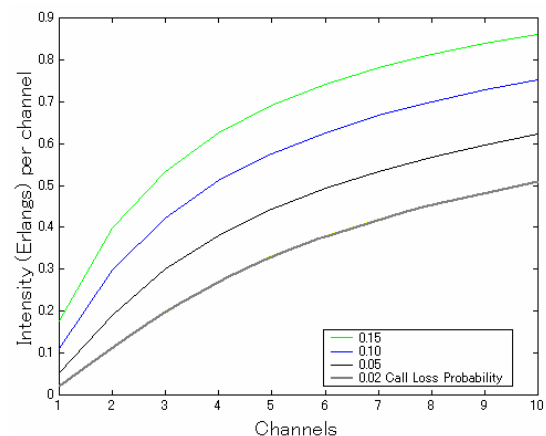


Fig. 4 Results of Call loss probability; Erlang's B-formula

#### 3.2 Proposed new service

We have designed and made a prototype of an emergency ambulances mounted satellite tracking system and a IP-based wireless radio system to support telemedicine. The roof top drive system features a compact, simple design, and mechanically controls a dish antenna 40 cm in diameter(Fig 3). With GPS continuous kinematic positioning, a fiber-optic 3D-gyro, and a quadrant detector, DC motor systems for azimuth and elevation control are installed, and the reduction gears have a harmonic drive mechanism with non-backlash gears aligned on a single input/output line.

With this system we can send up to 2.0-6.0 Mbps (MPEG-2, MPEG-4 ) motion picture from ambulance via HEO satellites, and/or wireless radio on S band on the ground. Telemedicine, especially triage for emergency transportation must encompass a wide range of patients, from children to the elderly, not only developed country, but also developing countries. The transmitted data should







be IP-based multimedia data that provides diverse information such as 12-lead ECG,  $\text{SaO}_2$ , and digital video of cardio-echogram of myocardial infarction and the affected area of trauma victim during disaster.



Fig. 5 Prototype super ambulance  
Two diversity antenna mounted on the roof

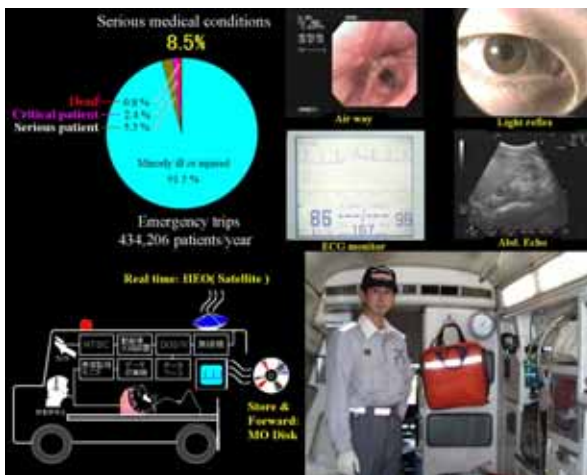


Fig.6 Medical data for transmission; Light reflex, Vocal cord, Echogram



Fig. 7 Tracking test in urban area Japan

#### 4. Considerations

Tele-education and telemedicine are regarded as one of the triggers to spread ICT. However, is it appropriate to follow the same extension policy as the other ICT application? Preparation for communication infrastructure such as PSTN, mobile communications, satellite communications, and optical fiber communication networks differs according to each user's circumstance. Furthermore unlike other needs whose markets change on the principle of free competition, medical service is provided by the managed economy and administered based on a government's powerful support policy in each nation. Therefore, it is indispensable to implement not only hardware but also software suitable to each communications environment.

In case of Acute Myocardial Infarction, thrombolytic agents are reportedly effective when injected into vein within one hour after the attack. The patients who are lucky enough to be transported to a CCU (coronary care unit) in emergency centers are in most cases given thrombolytic agents while undergoing PTC (Percutaneous Transluminal Coronary) operations to remove the coronary thrombus. Thrombolytic agents are reportedly effective even when injected into a vein, if injected in the early stages. In fact, some trials of doses in ambulances have been initiated. We estimated that over 50% of patients will benefit from this satellite system. One solution for curbing medical expenses in Japan, which is currently growing 5% annually, is improving pre-hospital care. The high-speed data channels will enable doctors to assess the conditions of the patient in an ambulance regardless of the distance.

#### 5. Acknowledgement

This study was performed through Special Coordination Funds of the Ministry of Education, Culture, Sports, Science and Technology, the Japanese Government. We would like to offer our thanks to Prof. Kiyoshi Kurokawa at Tokai University for his many useful suggestions.

#### 6. References

01. E.J. Dutzi, G.H. Knouse. Mobile satellite communications technology: A summary of NASA activities. The 37<sup>th</sup> Congress of the IAF 1986.
02. M. Richharia. Mobile Satellite Communications. Published by Addison-Wesley, ISBN 0 20133142-X
03. I. Nakajima, Y. Yagi, H. Juzoji, et. al. Highly Inclined Elliptical Orbiter for Motion Picture Transmission from Ambulance, WPMC01:pp.1585-1590, 2001.[ISBN 87-988568-0-4]
04. I. Nakajima, Y. Yagi, y. Zhao, et. al. Cost Comparison between HEO Satellites and 3G Wireless Mobile Phones for Video Transmission from Ambulances, WPMC01:pp. 1573-1578, 2001.[ISBN 87-988568-0-4]





## **A Pilot Learning/Training/Teaching System of Mammography Screening for Radiologists**

Mann-Li Chang, Chien-Peng Chuang, Tai-Long Kuo

*Department of Industrial Education, National Normal University, Taipei, Taiwan, R.O.C.*

Ching-Chi Wu

*Bureau of Education, Taipei City Government, Taiwan, R.O.C.*

Giu-Cheng Hsu

*Department of Radiology, Tri-Service General Hospital and National Defense Medical Center  
Taipei, Taiwan, R.O.C.*

Chia-Pei Huang, Yang Lang Chang

*Department of Electrical Engineering National Taipei University of Technology, Taipei, Taiwan  
San-Kan Lee*

*Yuan-Shan and Suao Veterans Hospitals, Yilan, Taiwan, R.O.C.*

### **Abstract**

Mammography is by far is the only effective means to be able to detect breast cancer in early stage. It has been widely used as a screening procedure to provide radiologists with a second opinion and further help radiologists improve their diagnosis. An experienced and well-trained radiologist in mammography can significantly reduce fatality rate in breast cancer by 20-30%. As a result, how to effectively use mammography to detect early breast cancer becomes increasingly important in women's health screening procedure. In order to meet this need, this paper presents a learning system that can be used to train and teach radiologists to improve their diagnosis. It adopts Microsoft Visual Studio Net 2003 to develop a new software viewing tool, called ImageViewer that allows radiologists to self learn and teach themselves through biopsy-proven mammograms in a well-designed data base via easily used commands such as zoom-in, zoom-out, contrast enhancement, filtering in a friendly

environment so that radiologists can manipulate functions provided by the ImageViewer at their discretion to increase their experience and further improve their diagnosis.

### **1. Introduction**

Breast cancer is a second leading cause of cancer in American women. Although breast cancer can be fatal, it has been reported by many study groups that its survival rate can be very high if radiologists can detect the cancer at very early stage. Therefore, early detection and treatment plays a key and critical role in increasing the survival rate. The mammography screening has been recommended as the most effective method for breast cancer early detection. A screening program generally involves a huge number of mammograms, which require a limited number of radiologists for examinations. As a result, misdiagnosis often occurs due to human errors and visual fatigue. To address these issues, two





approaches are usually taken. One is to train more radiologists ready for the job. The other is to develop a Computer Aided Diagnostic (CAD) system to assist radiologists in providing a second opinion.

According to the randomized controlled trial: RCT mammography in breast cancer detection has been widely considered to be very effective via. Due to high-rise on incident rates the American Cancer Society has recommended that all women between the ages of 40 and 49 should get their mammograms every two years and every year thereafter. In the Northern America and Europe mammography has become a standard screening procedure for women with high risks and women whose ages are over 50. In particular, since 1995, the Department of Health in Japan has been pouring government funding to support research and conduct studies in breast cancer and has also implemented the same policy which also recommends that women with ages older than 50 should have annual mammography screening and physical examination. [1]

Like western world, the incident rate on breast cancer has grown rapidly in recent years in Taiwan. In the past 20 years, this number has doubled and tripled, and the tendency has not been slowed down, rather the number is continuously growing. [6] In order to cope with this increase many hospitals have included mammography screening as a routine check-up in women's physical examination. When the mammography is used as a screening method to detect breast cancer, radiologists are often looking for small, low-contrast objects in breast images which are likely microcalcifications or masses related to cancer, particularly small tumors in dense breasts. Statistics have demonstrated that 70% of cases found by the mammography screening turned out to be suspicious lesions and required for further follow-up. Despite that not all small objects detected by the mammography are breast carcinomas, clinical studies

have shown that the mammography remains the only effective means in early detection of breast cancer. [8]

Bureau of Health Promotion Department of Health in Taiwan began its project of mammography screening in July 2002 for women with ages between 50-69 and received very good results. As a consequence, it decided in July 2004 to make free offer of mammography screening to women whose ages are from 50 to 69 every other year. Based on the finding from the survey conducted by Bureau of Health Promotion Department of Health for more than 80 hospitals that participated this project, the image quality of mammograms is not up to expectation. If the standard set by American College of Radiology is applied, many hospitals would be disqualified to perform mammography screening. [2] Most importantly, due to poor image quality of mammograms, the ability of radiologists' diagnosis can be significantly impaired. Therefore, misreading and poor diagnostic accuracy occur in many cases of suspicious abnormality. In order to resolve this issue, the work presented in this paper is to develop a system with two goals in mind. One is to provide radiologists with learning and training environments from which they can self improve their diagnosis experience. As indicated, due to shortage of experienced radiologists in mammography many radiologists must learn their skills and knowledge from various sources such as taking classes. However, due to nature in their job, one of most effective and easy ways to accomplish this goal is to provide radiologists with easy access and learning tools, which allow them to do anytime and any place whenever they would like to do it. Our system is developed to meet such need. The other is to provide a vehicle of conducting survey for the purpose of teaching and education and a forum of exchanging opinions among radiologists for the purpose of research.





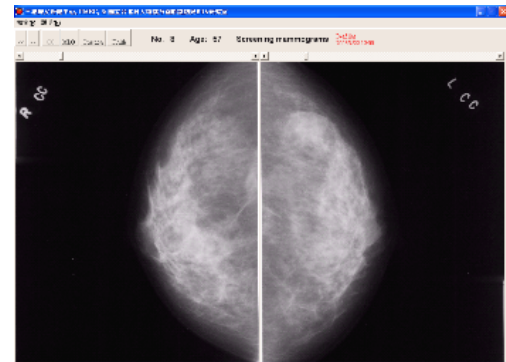
## 2. Materials and Methods

According to reports by the American College of Radiology and its 2003 published Breast Imaging Reporting And Data Systems (BI-RADS), the classification of mammography screening can be categorized into 6 categories, Category 0: No findings, Category 1: Negative, Category 2: Benign Finding, Category 3: Probably Benign Finding, Category 4: Suspicious Abnormality, Category 5: Highly Suggestive of Malignancy. This 6-category categorization described below will be used as a guide to develop our system. [3-5]

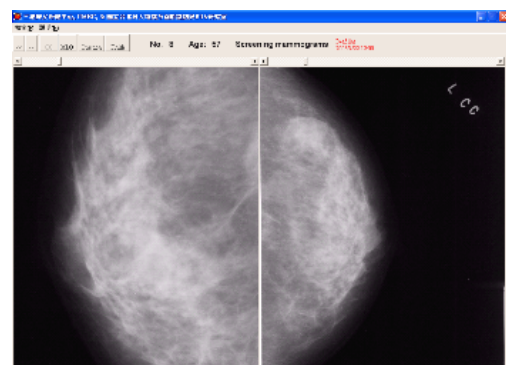
The system to be developed comprises of three component modules, a software module, called ImageViewer which includes many functions to allow users for manipulation of mammograms, a computer-aided diagnostic (CAD) module which enables users to learn and educate themselves via a provided CAD system through custom-designed data bases and tele-mammography module which enables users to communicate and consult at remote sites via internet. Each of these three modules can be developed separately and jointly used for various applications. This paper only presents the preliminary report of the first module, ImageViewer in our currently developed system. It is designed based on the Microsoft Visual Studio.NET 2003 coupled with the BI-RADS guideline to provide an environment with which radiologists can learn various cases from a data base which can be either obtained from the public domain or selected from well-documented cases.

The current system, ImageViewer is basically a software-driven tool. It allows users to perform various functions such as zoom-in, zoom-out, contrast adjustment, Region of Interest (ROI) finding, position adjustment of mammograms, etc. via maneuvering a mouse. Fig. 1 provides an illustrative example to

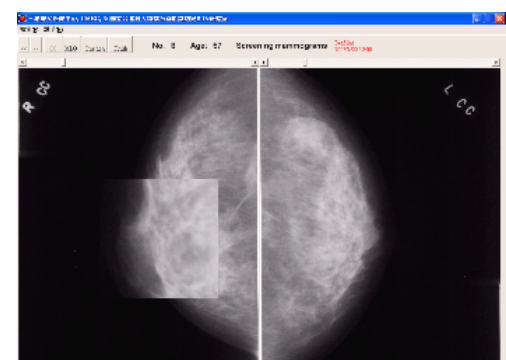
demonstrate some of its functionalities. Fig. 1(a) shows two views of left and right breasts brought up from the database. Fig. 1(b-d) shows the zoom-in view, ROI view and contrast enhancement view of the breast image at the left of Fig. 1(a) respectively.



(a) Two views of left and right breasts



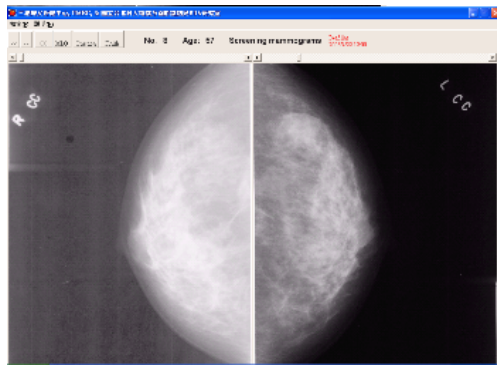
(b) Zoom-in view



(c) ROI view







(d) View of contrast enhancement

**Figure 1. Illustration of ImageViewer**

### 3. Results

In order to justify its utility of our developed ImageViewer, it has recently been used for certification of radiologists in mammography where several requirements must be satisfied. First of all, a radiologist to be certified must read mammograms at least for 240 cases within most recent 6 months. Second of all, he/she must read mammograms at least 1000 cases within most recent two years. However, due to various reasons, radiologists may go aboard for trips or studies or small hospitals that radiologists work for may not have sufficient mammograms for radiologists to read. Under circumstances, the requirements for 6-month and two-year mammogram reading cannot be met. On way to mitigate this dilemma is to use the ImageViewer to train themselves by reading mammograms custom-designed and collected for this purpose so as to increase their experience in clinical diagnosis. In the mean time, the ImageViewer can facilitate radiologists to be familiar with the guideline provided by the BI-RADS. [9]

In the initial phase, 77 radiologists who request to be certified were provided with 12 custom-designed cases for examination. Two scenarios were conducted. The first scenario was conducted twice and required the participants to only use a given set of magnified views to answer questions according to the BI-RADS's guideline.

The second scenario was provided with the ImageViewer to examine cases of mammograms. Since the ImageViewer offers many functionalities that allow the participants to make various maneuvers at their discretion to answer the questions, it was expected that the diagnosis accuracy of the second scenario was 79% on average which was significantly better than 61% and 63% using the first scenario, while the two results using the same first scenario, 61% and 63% did not change much. This experiment demonstrated two key improvements of using the ImageViewer over the traditional method. It was clear evidence that the ImageViewer provides great benefits to radiologists who can take advantage of the software to improve their diagnosis. This was witnessed by the fact that the participants could not improve their diagnosis accuracy very much using the same given set of magnified views very much even if they were another chance to do it again and the benefit gained from repetitions was very limited. The second improvement is the versatility of the ImageViewer, which can be very flexible to be used by radiologists without time and location constraints. Additionally, the database will be also updated and custom-designed to meet different criteria and needs.

Finally, two concluding remarks are noteworthy.

1. In the current developed certification system, the criteria to measure accuracy of diagnosis are based on the six categories recommended by the BI-RADS. The ground truth will be used to verify participants' reading skills on the mammograms provided by the program authority.
2. Due to unavailability of digital mammograms this time, all the mammograms to be used for the test were digitized to 0.05mm (or 50 $\mu$ m) per pixel. The ImageViewer allows participants to adapt their viewing machines up to the image resolution of 0.05mm/pixel. It is believed that this year may be the last year to use digitized mammograms. In the







future, digital mammograms will be used by this program.

#### 4. Conclusions and Future work

Despite the fact that mammography provides an effective screening method to detect breast cancer. It has been shown that it can be effective in detection of microcalcifications, but may not effective in detection of masses. In this case, the ultrasound must be used in conjunction with mammography. [7]

Due to differences in the breast structure such as size, density, shape and a variety of reasons such as, giving a birth, breast feeding, eating habits, breasts appear in many different forms and no gold standard for reading mammograms. Furthermore, there are no clear mammographic differences for radiologists to discriminate malignant and benign lesions, which make extremely difficult for them in their diagnosis. However, the accuracy of diagnosis in mammography can be improved if radiologists are well-trained by their experience or learning and studying more cases. The ImageViewer is developed to help less-experienced radiologists learn and teach themselves by working on many cases from data bases to accumulate their knowledge and increase their experience.

The mammography in Taiwan is still in its infancy. Many resources have not been allocated to implement the mammography screening as a routine procedure. One of most important issues is the shortage of experienced radiologists. In order to correct this problem, one most effective way is to train new radiologists through a learning process that is well-designed by experienced radiologists in a form of teaching classes or offering short courses. Our proposed ImageViewer can serve this purpose and provide access to users who can self learn and self teach themselves using their own time at free of locations.

The current ImageViewer is a preliminary pilot system that is developed to work on off-line to evaluate its efficacy and utility for clinical value. Several improvements are currently being investigated. Due to enormous image files, the image transmission is a major issue if the ImageViewer is implemented for the Internet use. The data storage and transmission speed are key issues that must be resolved. Second, due to filmless requirement for radiology department, the Picture Archiving and Communication Systems (PACS) has been used for radiologists to make diagnosis and writing reports. The ImageViewer has great potential integrated with the PACS. Third, it can use in conjunction with a CAD system that can be used as a teacher while the ImageViewer is used as a platform that allows radiologists to interact with the CAD-based teacher. Finally, the ImageViewer can be implemented as a vehicle for tele-mammography to make tele-consultation possible.

#### References

- [1] [http://www1.cgh.org.tw/content/healthy/magazin/e/medi/01-04\\_04.htm](http://www1.cgh.org.tw/content/healthy/magazin/e/medi/01-04_04.htm), “乳房X光檢查”, 輻射防護委員會.
- [2] <http://www.imaging.com.tw/center15.htm>, “乳房X光可疑異常疑似癌症誤判率七成八”, June 7 2004聯合報.
- [3] C. Zhou, H.P. Chan, N. Petrick, M.A. Helvie, M.M. Goodsitt, B. Sahiner and L.M. Hadjiiski, “Computerized image analysis: Estimation of breast density on mammograms,” *Med. Phys.*, vol. 2, no. 6, pp. 1056–1069, 2001.
- [4] K. Bovisand S. Singh, “Classification of mammographic breast density using a combined classifier paradigm,” in *Int. Work. On Dig. Mammography*, 2002, pp. 177–180.
- [5] M. Heath, K. Bowyer, D. Kopans, R. Moore, and P. J.





Kegelmeyer, “The digital database for screening mammography,” in Int. Work. On Dig. Mammography, June 2000.

[6] 李三剛, “乳房攝影檢查於乳癌篩檢的重要性與現況”, 95年度防癌防治「蘭陽護波月」活動

[7] 李三剛、李覃、劉自嘉, “乳癌病例女性親屬的早期乳癌偵測：臨床乳房檢查、乳房攝影與乳房超音波的比較”, 世界醫用超音波大會2000.

[8] 許居誠, “乳房X光攝影 偵測細微鈣化利器”, 中華民國癌症希望協會, June 2004.

[9] 許居誠, “乳房X光攝影醫療機構認證制度”, 中華民國放射線醫學會舉辦「95年度乳房X光攝影品質提升教育訓練課程」, May 2005.



## Artificial Neural Networks: a Supporting Role in Improving Prediction of Benign Prostatic Obstruction Evaluated by the International Prostate Symptom Score

Chih-Cheng Lu, Division of Urology, Department of Surgery, Chimei Foundation Hospital, Liouying,  
Department of Information Management, National Chung-Cheng University

Jinsheng Roan, Department of Information Management, National Chung-Cheng University

Yu-Jen Lee, Division of Urology, Department of Surgery, Chiayi Veterans Hospital,

**Introduction and Objective:** The International Prostate Symptom Score (IPSS) is used exclusively for evaluating patients with a prostatic condition and for following various treatment modalities. This non-invasive tool is insufficient for final clinical diagnosis. To evaluate the performance of an artificial neural network (NN) in the diagnosis of men with lower urinary tract symptoms and to compare its performance to that of a traditional linear regression (LR) model for evaluating prostatic obstruction. **Methods:** 331 qualified patients visiting outpatient clinic at the hospital between 2001 and 2003 received routine investigation, consisting of transabdominal sonography of the kidney, urinary bladder, and prostate, serum prostate specific antigen measurement, assessment of urinary symptoms and quality of life by the IPSS, urinary flowmetry with determination of maximum flow rate, and voided volume estimated based on all available non-invasive diagnostic test results plus patient age. Total patients were randomly assigned 50 times into 2 groups, training and test groups. The performance of the NN was quantified by using hit ratios. The 50 results of the NN approach were compared with those of LR analyses by independent t-tests. **Results:** There were five different sets of results for test groups: IPSS-7 (urinary symptoms in IPSS), Logi-7 (urinary symptoms in IPSS by LR), Logi-8 (IPSS by LR), NN-7 (urinary symptoms in IPSS by NN), and NN-8 (IPSS by NN).

**Results:** The results of the NNs showed better predictive values concerning the outcome of final diagnosis. Use the same seven urinary symptom questions, the mean estimated hit ratio was 69.9% by NN (NN-7) and that by IPSS (IPSS-7) was 65.9% ( $p < 0.01$ ). The eighth question in the IPSS, which is assessment of quality of life, showed its critical role as an important factor in the NNs (hit ratio 74.1% in NN-8 and 69.9% in NN-7,  $p < 0.01$ ).

**Conclusions:** In the computerized medical environment, this study shows that the neural network could be an acceptable and useful tool for the clinical help in the prediction of prostatic obstruction.

# A study on the Metaheuristics - apply biologically inspired dynamic formation to telecommunication network -

I. NAKAJIMA, H. JUZOJI, M. A. SADIQ  
N. HAMAMOTO\*, S. TAIRA\*

Tokai University School of Medicine  
National Institute of Information and Communications Technology\*  
E-mail: js2hb@ets8.jp

## Abstract

A Metaheuristics is a set of concepts that can be used to define heuristic methods that can be applied to a wide set of different problems. In other words, Metaheuristics can be seen as a general algorithmic framework which can be applied to different optimization problems with relatively few modifications to make them adapted to a specific problem. We apply the ant's Metaheuristics for Search and Rescue System with ETS-VIII satellite.

### 1. Purpose

We have started this Metaheuristics since 1997, and our paper was submitted in IEEE<sup>[1]</sup> 1998. We have recorded and observed worker ant formations, and optimized the patterns observed through mathematical analysis and approximation, in order to apply these behavior patterns to traffic management within packet communications networks. This basic research involved observation/recording of ant movements, image processing, behavior pattern analysis, mathematical analysis/approximation, the creation of a new communications algorithm, computer simulation, ROM prototyping, and proposal of a marine field test using free float buoys.

## 2. Metaheuristics

### 2.1 Assumption

We made the following assumption (A), shown in Fig.1: this indicates assumption that individual ants understand that the weight of the food is to be distributed among them, and calculate the distances and angles formed between neighboring ants.

### 2.2 Mathematical analysis and approximation

Since the ant behavior recorded represents a mere subset of all ant movements, not all ant behavior patterns can be analyzed. However, we believe certain mathematical analyses and approximations are feasible, based on the assumptions.

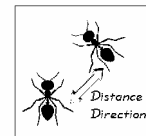


Figure 1. Mutual recognition (assumption)

### 2.3 Communication algorithm

Based on mathematical approximations, we developed an algorithm based on the above assumptions, substituting terminals for ants. We assumed wireless terminals using packet communication protocols. The important aspect here is the task of efficiently managing customer traffic to correspond to changes in the surrounding situation and topology.

## 3. Implementation

To gain an understanding of the network topology, terminals must obtain data such as signal intensity from the physical layer and GPS data from the application layer, and then seek to grasp the overall network topology by exchanging data such as information on administrative traffic. We checked our communications algorithm via computer modeling and calculated throughput for a few forms of network with a communication simulator (OPNET).

## 4. Considerations

METAHEURISTICS might be very useful under the environment of disaster where central control system will be broken and unable to conduct operations from the center. We intend to use this Metaheuristics for a new Search and Rescue System based on communication satellite; ETS-VIII, ROM prototyping, and proposal of a marine field test using free float buoys.

## 5. References

- [1] I. Nakajima, S. Hama: Wireless networks based on information exchange between ants, Proceeding of the 3rd International Symposium on Multi-Dimensional Mobile Communications, IEEE Vehicular Technology Society, pp.116-119, 1998

## **The Application and Realization Of XML Transform Middleware in Electronic Medical Records**

CHEN Yi<sup>1</sup>, HUANG Yong-feng<sup>2</sup>, ZHANG Si-fa<sup>1</sup>

(1 Department of Computer Science and Technology , China University of Geosciences , Wuhan 430074, China; 2. Department of Electronics and Information Engineering Tsinghua University, Beijing 100084, China )

([Eric\\_cyl126@126.com](mailto:Eric_cyl126@126.com))

### **Summary**

As one kind of descriptive language which is independent of the platform, XML is widely used for its high extension. When it comes to the data of a patient's medical record, they are dispersed at the hospital's different service database systems. Then how to integrate these data and set them in a document which is "patient centered" appears more and more essential. In view of this, we have realized an middleware for the transformation of XML. It can be used for the dynamical transformation between the data in medical database systems and XML. It also can be used to the uniform archive information integration of the electronic medical record.

Regardless of the interior or the outside data requested must hand in the form of XML to middleware. Realize the interactive with the database by middleware. The results of the middleware which operated the database still in form of the XML feedback to the request. We designed based on the XML database transform middleware, can transparently cause the application system to be highly effective, unified to visit the data which located in the different hardware platform, the operating system database, this middleware design may enhance the database efficiency, the system may act according to the client side the request, the tendency completes in the database the data and the XML document transformation, thus shields the SQL language completely to the procedure development personnel.

The XML transform middleware mainly realizes the function below :

First, completes the transformation between relation database and XML documents.

In order to correctly exchanges the data between the different databases should be supposed to guarantee both sides use in the exchange data table have the similar structure, and produced XML document which description database structure, for clearly describe the database structure to enable opposite part to be possible to produce the similar table to carry on the data exchange.

Second, Completes the Electronic Medical Record data unification pigeonhole with the medical record data integration.

The scope of the pigeonhole which according to the medical record document carries on the unification to the medical record documents to file away, realizes the static and the dynamic medical record data separation, disperses the patient in each service database system medical record data concentrates to together, forms take "the patient as the center" XML document.

It is proved that XML transform middleware in Electronic Medical Record design and the realization is completely feasible, moreover has the actual application values. Because of the patient medical record comes from various of service database system, these database system are develop by different merchant ,so the realization of the XML transform middleware it is very helpful for patient medical record information centralism gathering ,and make the pigeonhole of the Electronic Medical Record very convenience, will have the extremely vital significance on medical record data sharing and exchanging in future.





## Development of Teleconsultation System

Yuichi Ishibashi<sup>\*1</sup>, Atsuko Hara<sup>\*2</sup>, Muhammad A. Sadiq Sadiq<sup>\*1</sup>, Hiroshi Juzoji<sup>\*1</sup>, Isao Nakajima<sup>\*1</sup>  
<sup>\*1</sup> Tokai University School of Medicine, <sup>\*2</sup> Kitasato University School of Medicine  
 y-ishi@statlab.jp

### Abstract

*Tokai University has developed a standard telecardiology system and confirmed practical applicability in Bhutan. With this system information on patients, from ECG (Electrocardiogram), echocardiography and so on, is sent and medical consultation can be provided between distant places. In this case the data are transmitted through the internet and other information may be exchanged through E-mail, Fax and by telephone.*

*We are now introducing a new teleconsultation system in which almost all of the information transfer, whether for findings of ECG, laboratory examination and physical examination, or for diagnosis and other opinions, is by internet. This facilitates ease of communication and storage of all the information in databases for reference.*

### 1. Introduction

The standard telecardiology system in Bhutan is as follows. Patient's ECGs, echocardiograms and phonocardiograms are taken in primary care centers using multi-purpose ECG equipment made in Japan. To benefit from experience not available locally they are transferred with data from physical examinations through a simple phone, fax, internet connection and satellite between the primary care center and a secondary health care center. The 33.6 kbps connection speed by dial up in Bhutan imposes limitations of file size for image data, but carefully recorded still images of ECG and echocardiography can be transmitted for expert consultation.

This study has performed for the following purposes,

- To access the feasibility of transmitting ECGs and echocardiograms with an internet protocol.
- To enhance the diagnostic accuracy through telecardiology consultations diagnosing heart disease, and provide specialist care and consultation to places where no such facilities existed.
- To develop a standard telecardiology system for countries like Bhutan plagued with the shortage of human resources.

The following problems were encountered:

- The consultant receives many kinds of information through different media, such as internet, telephone, fax, E-mail, and then the received information needs to be sorted and filed.
- To record different kinds of information is not easy

without a computer base.

- There is no standardization regarding formats for information, with different levels and contents exchanged between clients and consultants.

We therefore developed a new consultation system to solve these problems.

### 2. Purpose and features of the system

The system was developed for medical teleconsultation through the internet, with the following functions. A medical worker as a client registers patient information and symptoms, and data from laboratory examination, ECG, echocardiography, X-ray, CT and histopathology into the system and then requests a consultation. A consultant indicates the necessary tests and provides a diagnosis and treatment recommendations. The client may ask different opinions from more than one consultant. The information, test data and image data are recorded in database, and then these are referenced as a case study.

### 3. The flow of consultation using teleconsultation system

The flow for consultation inside a hospital is as follows:

- (1) The patient exhibits some symptoms.
- (2) Doctor A in charge of the patient (client) immediately consults doctor B of the relevant department (consultant) with documents and oral reports after first aid is given.
- (3) The consultant examines and treats the patient directly, and indicates the necessity for additional checkups..
- (4) The consultant continues medical treatment where necessary
- (5) Consultation is finished at the time of recovery or improvement of the patient condition.

### 4. References

- [1]Sadiq, M., Nakajima, I., Juzoji, H., Nawaz, M., Tanabe, T., Standard Telecardiology System for Developing Countries, 2005, 3<sup>rd</sup> APT Telemedicine 2005, Kuala Lumpur, Malaysia.
- [2]Ahmad, S., Gilani, S., Malik, A. Z., Telemedicine software for a developing country, 2006, 4<sup>th</sup> APT Telemedicine, Rawalpindi, Pakistan.





## Correlation of EEG power ratio and severity of Alzheimer's disease

Chih-Chung Chen<sup>1,2</sup>, Hung-Wen Chiu<sup>2</sup>, Chien-Yeh Hsu<sup>2\*</sup>

<sup>1</sup>Department of Neurology, Bojen General Hospital, Taipei, Taiwan

<sup>2</sup>Graduate Institute of Medical Informatics, Taipei Medical University, Taipei, Taiwan

\*cyhsu@tmu.edu.tw

### Abstract

**Objectives:** To test the hypothesis that the electroencephalography (EEG) power ratio are closely related with the staging of Alzheimer's disease (AD).

**Methods:** Segments of resting, artifact-free, 16-electrode scalp EEG signals were obtained from 47 patients of different severity of probable Alzheimer's disease. Using fast Fourier transform (FFT), EEG spectral profile was obtained for each candidate and averaged spectrum was demonstrated. The ratio of slow waves (theta and delta bands) to fast waves (alpha and beta bands) power, defined as power ratio, was calculated. The severity of Alzheimer's disease was classified into 4 stages - very mild, mild, moderate and severe, according to Clinical Dementia Rating (CDR) and Global Deterioration Scale (GDS). Each candidate was classified into one of the four groups. The correlation of power ratio and groups was observed.

**Results:** Characteristic spectral profiles of each group were demonstrated. Group mean power ratios ( $\pm$  Standard deviation) were  $0.68 \pm 0.32$  ( $N=19$ ) in very mild group,  $1.12 \pm 0.57$  ( $N=12$ ) in mild group,  $1.84 \pm 0.85$  ( $N=9$ ) in moderate group and  $3.32 \pm 1.54$  ( $N=7$ ) in severe group. Very mild group can be differentiated from moderate group ( $P=0.019$ ) or severe group ( $P=0.023$ ) and mild group can be differentiated from severe group ( $P=0.049$ ). When comparing very mild plus mild and moderate plus severe groups, the discrimination was evident ( $P<0.01$ ). However, There's no statistically significant difference between any two adjacent groups.

**Conclusions:** EEG power ratios can reflect spectral profiles in AD patients and is a useful EEG parameter to represent the severity of the disease. Although the trend of increasing ratio in more advanced stage is evident, discriminating any two adjacent groups are still not applicable.

disability in daily activities and behavioral abnormalities. The incidence increases with age. The percentage of persons doubles with approximately every five years of age. However, the onset of the disease is insidious and the initial presentation is usually subtle. With application of the new drugs, cholinesterase inhibitor, the natural course of reluctantly progressive mental deterioration can be modified.

Two criteria are commonly used to diagnose AD, one from DSM-IV and another from National Institute of Neurologic and Communicative Disorders and Stroke Alzheimer's Disease and Related Disorders Association (NINCDS-ADRDA). To fulfill the criteria, typical history and presentations must be met and other alternative diagnosis must be excluded. Brain computerized tomography (CT) or magnetic resonance image (MRI) are commonly applied to evaluate the extent of cortical atrophy and to exclude other brain pathology. Traditional clinical assessments of severity of AD rely on psychometric scales, such as Mini-Mental Status Examination (MMSE), ADAS-cog, CASI, CIBIC-plus...etc. Clinical Dementia Rating (CDR) and Global Deteriorating Score (GDS) can be used to stage the illness course.

Electroencephalography (EEG) records electric activities from cerebral cortex and can reflect pathologic conditions in cortical and subcortical regions of cerebrum. Unlike CT or MRI, it is a simple, cheap, widely available and non-invasive procedure which may be a potential tool to monitor this disorder. However, it is not routinely applied during clinical practice either to diagnose or to monitor the illness course of AD. Typical EEG findings in Alzheimer's disease are decrease in fast (alpha and beta) activities and increase in slow (theta and delta) activities. Although these features have long been identified and described in the literatures, they are regarded as lack of sensitivity and specificity for AD. Textbooks and most clinicians believe that the EEG features won't appear until the late stage of the disease<sup>21</sup>. Recent studies<sup>1-5</sup> reveal that EEG can be helpful in many aspects to evaluate AD.

EEG can differentiate AD from depression<sup>6,14</sup>, VaD<sup>7-9,15</sup>, normal aging<sup>10-12,15</sup> and frontotemporal dementia<sup>13</sup>. EEG can also reveal abnormality at the very early stage of AD<sup>16-19</sup>, as known as mild cognitive impairment (MCI). In addition, EEG may predict the progress and outcome of individual patient<sup>13-18</sup>. EEG is also valuable

### 1. I. Introduction

Dementia is a common disorder among the elderly population. Major subtypes include Alzheimer's disease (AD), vascular dementia (VaD), Frontotemporal dementia (FTD), dementia of Lewy body (DLB)... etc. Among these, AD warrants special attention because of the popularity of victims and the new advance of effective drug treatment. AD is a progressive neurodegenerative of cerebrum, characterized by progressive cognitive decline, memory impairment,





In evaluating response to drug treatment. Several cholinesterase inhibitors have been shown to have positive effects on EEG. Among subjects receiving treatment, individuals with prominent EEG response may have better drug effect<sup>25,26</sup>.

Rodriguez et al<sup>28</sup> described typical EEG spectral profile of AD of different GDS and tried to differentiate them by means of relative band power at 1.5 Hz interval. However, the 7 quantitative EEG bands were far from the experience of conventional 4 frequency bands by most clinical electroencephalographers. Bennys K et al<sup>29</sup> used power ratio between slow activities and fast activities to differentiate AD from normal control. The discrimination was striking throughout different scalp regions (except frontal region) and cutoff values was proposed. The ratios were also found to be correlated with severity of the disease, but the cutoff values were not mentioned.

To verify the supposition that cutoff value of power ratios are useful in clinical practice to stage the disease, we conducted this study.

## II. Materials and methods

### Subjects

Candidates for this study are psychotropics-free probable AD and mild cognitive impairment (MCI). Two medical institutes in Taipei (Bojen General Hospital and Municipal JenAi Hospital) participated in this study. From separate registry system of the EEG laboratories, patients received EEG study from 2000 to 2005 with the impression of Alzheimer's disease, MCI, dementia or poor memory (ICD-9 332 and 290) were included. The medical chart of each patient were carefully reviewed by neurologists to confirm the diagnosis of probable Alzheimer's disease (NINCDS-ADRDA criteria<sup>24</sup>) and mild cognitive impairment (Peterson's criteria<sup>25</sup>). Brain CT or MRI, serum vitamin B<sub>12</sub> and folic acid level, thyroid function, VDRL and anti-HIV antibody were routinely performed to exclude vascular dementia, hydrocephalus, nutritional deficiency, hypothyroidism, neurosyphilis and AIDS dementia complex. Alternative diagnosis, such as frontotemporal dementia, dementia with Lewy body, depression, or other unclassified dementia were excluded from this study. AD patients whose EEG recordings were in the acute delirium state, receiving potent antipsychotics and sedatives or in the presence of major systemic disease were also excluded. Clinical informations of MMSE, CDR and GDS of the candidates within 2 weeks before or after EEG recording were collected. Candidates were further grouped as very mild (CDR 0.5, GDS 3), mild (CDR 1, GDS 4), moderate (CDR 2, GDS 5) and severe (CDR 3, GDS 6) according to illness severity.

### EEG recording

EEG studies were performed in quiet rooms. Sixteen

scalp electrodes (Fp1, Fp2, F3, F4, F7, F8, C3, C4, T3, T4, T5, T6, P3, P4, O1, O2) were referenced to electronically linked earlobes (A1 and A2) according to international 10/20 system. All electrode impedances were kept below 5 k $\Omega$ . Concurrent surface EMG and EKG were also recorded for reference, but were not included in the EEG analysis. EEG settings were identical in 2 laboratories with sampling rate at 200 Hz, bandpass at 1-70 Hz and a notch filter at 60 Hz. During recording sessions, subjects sat on a comfortable chair with eyes closed. Activation maneuvers, including eyes opening, hyperventilation or photic stimulation were performed under instruction by technicians and the events were marked on the EEG recordings. Average recording duration was 10 to 15 minutes.

### Quantitative EEG

All EEG data were reviewed by an experienced electroencephalographer by direct visual inspection. Artifact-free (free from muscle contraction, electro-oculographic deviations and body movements) epochs of 1-s duration were collected. The data were exported to Matlab for spectral power analysis. Power spectral profiles of 16 scalp electrodes as well as averaged spectrum were demonstrated for descriptive purposes. For averaged spectrum, absolute and relative power in four standard frequency band – delta (2-4 Hz), theta (4-8 Hz), alpha (8-13 Hz) and beta (13-32 Hz) – were calculated. We defined the "EEG power ratio" as the power ratio of slow to fast activities.

EEG power ratio=(theta band power+delta band power)/(alpha band power+beta bad power) (1)

### Statistics

Statistical analyses were performed using SPSS package. For each group, EEG power ratios were analyzed for the mean value, standard deviation, standard error and 90% CI. One-way ANOVA with Post Hoc multiple comparisons (Tamhane) was used to compared the power ratio between groups. Further analysis reclassified the candidates as early (very mild and mild groups) and advanced (moderate and severe groups) groups. Student's t test was used to compared these two groups. Statistical significance was defined as  $p < 0.05$ .

## III. Results

### Clinical characteristics of subjects

Forty-seven patients, 25 male and 22 female, were included in this study. Mean age is 77.9 (range 59-90). Nineteen patients were in the very mild group (MCI) with mean MMSE score at 22.5. Twelve were in the mild group with mean MMSE score at 19.3. Nine were in the moderate group with mean MMSE score at 12.0. Seven were in the severe group with mean MMSE score at 9.4. (Table 1).





## Development of the Home-Based Multimedia System for Mental Health Management in Adolescents via Information Superhighway

Jeongyee Bae

Department of Nursing, Inje University, Busan, Korea

jibai@inje.ac.kr

### Abstract

*The purpose of this study was to develop a web-based multimedia system for mental health management in adolescents via information superhighway. Research process includes needs assessment, needs analysis, design, development/testing, and application release. With this computerized system, adolescents can improved their mental health states. In addition, adolescents who have minor mental health problems can manage it by oneself.*

### keywords

Mental health, Information system, Nursing Informatics, Internet, Adolescents

### 1. Introduction

World Health Organization's warning last year that the fastest-growing mental health problem in the world, and particularly in the developed world, was among adolescents. According to the rapid socio environmental change In Korea, many adolescents suffer from their identity confusion. Most of Korean adolescents are under pressure to study hard for an examination. The failure to examine the possible causes of rising levels of mental ill health amongst adolescents provoked much critical commentary, even from those who welcomed efforts to draw attention to the problem. The incidence of depressive symptoms among Korean adolescents has increased to approximately 40%. Even 8% of them reported suicidal idea [1].

Worldwide, about 4.5% of all Internet searches are for health-related information [2]. Access to the rapidly increasing store of web-based information is essential for effective public health practice. In Korea, the Internet has become a most favored source to find health information. About 44.5% of all Internet searches are for health-related information [3]. Thus this project was conducted to develop a health information service system for mental health care in adolescents via internet.

### 2. Methods

This study used user-centered design model including needs assessment needs analysis, design, development/testing, and application release.

To identify users' information needs, the investigator will conduct a descriptive study by performing surveys, interviews, and workshops in major cities in Korea. 3487 Korean

adolescents participated in this study. A collaborative work between the health professions and information technical engineers began to produce a scrupulous program, which would accomplish meeting the established goals, contents, and methods.

The software used to produce the system includes Flash 4.0, Adobe Photoshop 5.0, Adobe illustrator, MS IIS 3.0, MS NT Serve 4.0, MS FrontPage, and MS SQL 7.0.

### 3. Results & Conclusion

This system will go through individual data and will be utilized to screen those who have mental health problems. In addition, this system includes intervention programs to relieve mental health problems such as education, exercise, relaxation, deep breathing, visualization, music therapy, family therapy and counseling with researcher or psychiatrist. Using this system, every adolescent who has minor mental health problems can manage it by oneself.

This research interested in the telemedicine and developing home-based multimedia system for management of the adolescent's mental health problems. Therefore, through these system adolescents can save time and cost. And it could be adequately applied to assessing mental health problems and as an intervention strategy for not only adolescents who have mental health problems but also normal adolescents. It is expected that this management system will contribute to the mental health promotion as well as community for the mental health related researchers, professionals and experts to share information on the adolescent's mental health.

This research interested in the telehealth, especially information service system for more efficient psychiatric mental health care. Therefore, this system will be a new way of nursing intervention and contribute to the psychiatric mental health promotion in adolescents.

### Acknowledgments

This work was supported by the Korea Research Foundation Grant funded by the Korean Government (R05-2004-000-12694-0)

### References

- [1] Kang K.M. (1994). Depression in Children and Adolescents, *Child Adolescent Psychiatry*, 5(1), 3-11.
- [2] Janet M., Morahan-Martin. How Internet Users Find, Evaluate and Use online health information; A Cross-Culture Review. *Cyber Psychology & Behavior*. 2004. 7(5), 497-510.





[3] Korean institute for health and social affairs, 2001.

**Address for correspondence**

Jeongjee Bae, RN, PhD,  
Department of Nursing, Inje Univeristy, 633-165 Kaekeum-dong,  
Busanjin-Ku. Busan 614-735, Korea.  
Tel: 82-51-890-6823. Fax: 82-51-896-9840  
E-mail: jibai@inje.ac.kr





## Validating the role of Mobile Telemedicine Package in Marine Time Teleconsultations

Muhammad Athar Sadiq<sup>1</sup>, Hiroshi Juzoji<sup>1</sup>, Isao Nakajima<sup>1</sup>, Kiyoshi Igarashi<sup>2</sup>, Kenji Tanaka<sup>2</sup>

<sup>1</sup>Nakajima Laboratory, Department of Emergency and Critical Care Medicine, Tokai University, Shimokasuya 143, Isehara, Kanagawa, 259-1193 Japan

<sup>2</sup>National Institute of Information and Communications Technology, 4-2-1 Nukui-Kitamachi, Koganei, Tokyo 184-8795 Japan

Email: [sadiq@is.icc.u-tokai.ac.jp](mailto:sadiq@is.icc.u-tokai.ac.jp)

### Abstract

A small and light mobile telemedicine package is quipped with mobile satellite communications and can be used anywhere in the world except in polar zones. The system aims to establish telemedicine linkages between remote areas and specialized centers and also between disaster-stricken district and non- devastated district using satellite communications. The purpose of this study was to verify the role of mobile telemedicine package in marine time telemedicine teleconsultations between Bosemaru, a cruiser in Pacific sea and lifesaving emergency medical center of Tokai University in Japan.

**Key Words:** Mobile telemedicine package, Satellite Communications, Bosemaru

### Background

Various telecommunications tools are currently investigated and utilized for remote telemedicine consultations [1]. Mobile telemedicine package is small, light and handy. It is designed for medical usage and connections to medical terminals are possible via wireless LAN. It is IP based and it can easily establish a flexible network. In this study we have studied the role of mobile telemedicine package in marine time telemedicine consultations via satellite connection between Bosemaru Cruiser at Pacific sea and Emergency Medical Center of Tokai University.

### Material and Methods

#### Equipments

Characteristically this telemedicine package is small, light and handy. The package is suit case sized kit (51cm x 40cm x 22 cm), weighing 10 kg, that contains PC, satellite antenna and modem, equipped with satellite communications and wireless LAN.

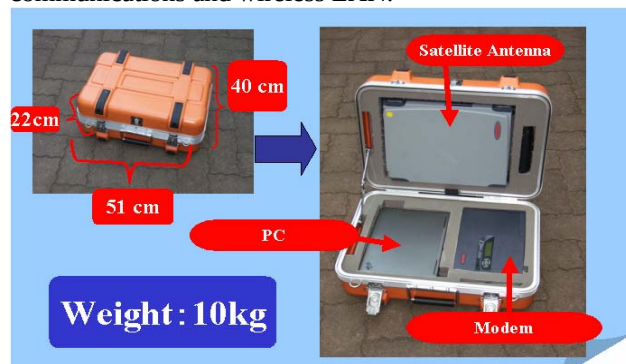


Fig. 1. Mobile Telemedicine Package

### Procedure

Telemedicine consultations were made via satellite communication (INMARSAT). The standard data transfer rate was at 64kbps ISDN connection.

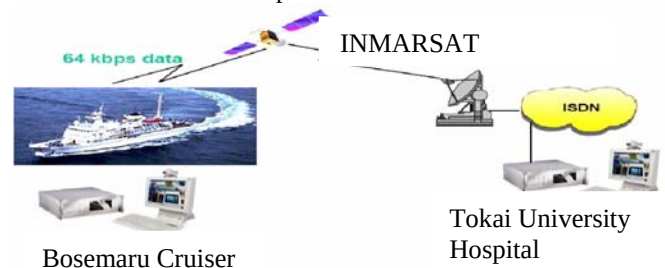


Fig. 2. Portable 64kbps satellite connection.

### Results and Discussion

Using mobile telemedicine package, six video conferences were carried out via satellite connection between Bosemaru medical staff and Tokai University's hospital at Isehara city, Japan. Both the audio and video communications were of satisfactory level at transfer rate of 64kbps. It was possible to get the opinion of emergency room doctor including one for a severe motion sickness patient. Although the communication via satellite connections was satisfactory but the cost of communication was relatively high about US\$ 250 per thirty minutes. One solution to reduce this high cost is Universal Service Fund (USF) for e-health services. In United States the role of USF has been extended to support cost communications services for schools, libraries and rural health care providers But in Japan USF only includes emergency call via voice, which is 110 and 119. There is strong need to include e-health applications by USF in Japan.

### Conclusion

Mobile telemedicine package, equipped with satellite communications and the wireless LAN to support telemedicine communications can be shipped easily to boats and ships for remote marine time telemedicine consultations from anywhere in the world. To reduce the cost of remote teleconsultations there is strong need to cover e-health services under Universal Access Fund.

### References

[1] Nagami k, Nakajima I et al. Telemedicine Package for Disasters. 4<sup>th</sup> APT Telemedicine, Pakistan, pp. 17-22





# A Rough Set-Based Prediction Approach to Medical Application

Jia-Yuarn Guo<sup>1</sup>, \*Chun Pei<sup>2</sup>, Yu-Tung Dai<sup>3</sup>, Han Hsin Chou<sup>4</sup>,

Chung-hsiu Ku<sup>5</sup>

<sup>1</sup>Dept. of Industrial Engineering and Management,

<sup>2</sup>The Graduate School of Gerontic Technology and Service Management,  
Nan Kai Institute of Technology

<sup>3</sup>Dept. of Occupational Safety and Health, Chung Haw College of Medical

<sup>4</sup>Department of Logistics and Shipping Management, Kainan University

<sup>5</sup>Ordnance Readiness and Development Center

Correspondence: Chun Pei, pantonpei@hotmail.com

## Introduction

It is noted that due to the typically huge size of today's information systems, real-world data tend to be incomplete due to missing some values. Hence, discover knowledge from incomplete information systems has received more and more attention in recent years.

The rough sets theory, proposed by Pawlak, provides a natural method to cope with incomplete or inconsistent information which has been the main impediment to the classification and rule induction of objects. Several techniques have been developed to extract decision rules from an incomplete information system. A key factor among them is using different methods to manage the missing data (unknown values). The simplest is removing the objects with unknown values. Other simple methods include replacing missing values with possible values calculated by statistics analysis. These methods all try to transform an incomplete system into a complete system by smoothing or extending the data. Other groups of techniques deal with the incomplete systems without changing the size of the data sets or making assumption of the missing values. These methods intend to induce every certain rule directly from the original data sets. Like the second group, our approach uses a rule generation algorithm to induce all certain rules and possible rules from the original incomplete data. No matter what the missing values will be, they won't affect the induction rules.

Compared with most of the classification and rule induction methods which induce knowledge from sets approximation concepts, the rule induction technique proposed in this paper applies a new modified rule-reduct generation algorithm (MRGI) and rule induction program (RIPI) to mine knowledge directly from the minimal set of decision rules (rule-reduct).

## Conclusions

In this article, a new rule-generation algorithm (MRGI) based on rough set theory has been proposed to generate the minimal set of rule-reduct which also represent the certain rules from incomplete information system. Some

examples have been illustrated to demonstrate the validation of the methods we presented.

Compared with other complex rough sets based rule induction approaches, MRGI presents not only the capacity to deal with uncertainty or inconsistent information but the high ability to generate concise and simple rules as well, and makes it easier to analyze and induce knowledge from large incomplete information systems. Without changing the size of the original system or adding possible values to the null values, the technique proposed in this paper generates knowledge rules directly from the original incomplete information systems.

We strongly believe that the rule induction technique presented in this paper will provide an efficient access for rule generation and knowledge induction, and will play a competent role to complement other existed rule induction Techniques for incomplete information systems.

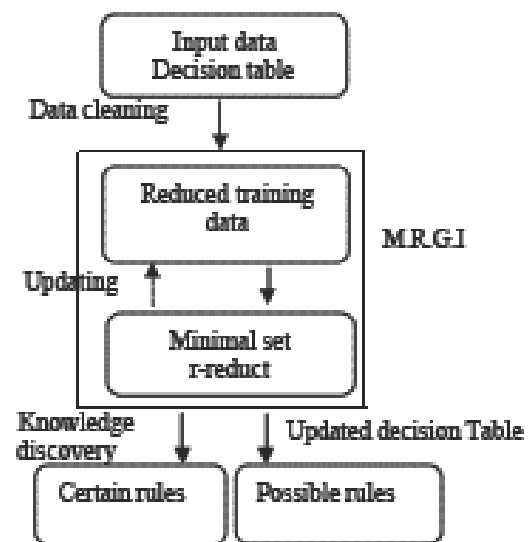


Fig.1 Mechanism of RIGPI including MRGI





## Practical operation of a document delivery system for medicine information, which used a hospital web-based system in risk management side

HIROHIKO MISHIMA, MUNEO KADOBAYASHI

Department of Pharmacy, Hospital of Hyogo College of Medicine, Hyogo, Japan  
ykmisi@hyo-med.ac.jp

### Abstract

In recent years, medical-supplies appending document information can be referred to from a hospital information system terminal. However, as for such medical supplies attachment document information search system, what is limited a practical use method to become a problem while there is it at a position becoming the pivot of medicine information because search feature is fragmentary. We can access it in web environment from an HIS terminal and it is easy and applies a synthesis medicine reporting system enabled by a high search of the degree of freedom. We report it about evaluation of a role of a medicine reporting system by web environment with expansion of a use terminal this time.

### 1. Introduction

When we carry out reasonable drug therapy, in late years, as for the medicine information, the system which medical supplies attachment document information can refer to is built in many institutions by a hospital information system terminal by being the required source of information. However, as for such medical supplies attachment document information search system, what is limited a practical use method to become a problem while there is it at a position becoming the pivot of medicine information because search feature is fragmentary. Therefore we can access it in Web environment from an HIS terminal and we build the synthesis medicine reporting system which enabled a high search of the degree of freedom and apply it easily. We go through a homepage of a medicine section from an Internet line from 2003 and start in particular an offer of medicine information same as an HIS terminal. With using a His terminal, the role of a medicine reporting system brought a big change. As for the performance of medical information, speedy qualitatively quantitatively is an element. We have to satisfy information, update frequency, an opinion of a worker to improve by a performance of medical information. If unified information is obtained, workers will take a benefit impartially, and, as for it, an immeasurable merit is provided. Therefore we planned

promotion of efficiency with a Web system. We summarize information to escape from the situation that workers cannot grasp professional medicine information. As for the worker, grasp of information becomes easy by converting information about a medicine into a unified document of expression. When we make Web data, it is necessary to review data structure of medicine information. It is offered text file with a tag to us by an outside organization. An order entry system is weak in indication of data with a tag. It is simple that the reporting that used a Web system lets a different kind document coexist. It loses the effectiveness to offer information one-sidedly. We should take a device of a system constantly. As a result, it becomes a custom that workers collect information actively. When I get rid of an offer of medicine information, I cannot secure security of a patient. It is necessary for us to build reporting of a high-risk medicine as a risk management system.

### 2. Method

- 1) A system summary: This system uses PRIMERGY - C600 of Fujitsu as development machinery of a search Web server for HIS terminals. Web server software uses Internet information service (IIS). A search Web server for HIS terminals connects with the LAN that HIS uses by TCP/IP. A data file uses ethical drug attachment document data in an HIS server. (Figure.1)
- 2) Construction of a general medicine reporting system: The data for searches were able to do communalization by letting a save folder name be common to. The data, which print needed, supported by a PDF file. The search menu constituted it by frame structure to shorten time to need it for a search of a user. (Figure.2)
- 3) A flow of data making: Medicine information consists of outside document and hospital document and drug information. Making of an outside document downloads data from Internet and uses it. A hospital document converts a distribution document into a PDF file and uses



it. Drug information uses a PDF file and image data as manipulation preventive measures. Probability to search medicine information when a medical worker uses an order terminal is high.

(Figure.3-4)

- 4) We conducted an investigation into the use situation of a system with access log.

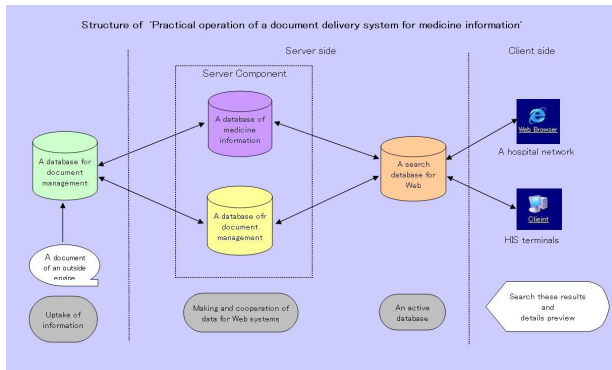


Figure.1

A figure of constitution of a Web system



Figure.2

A top page of a Web system

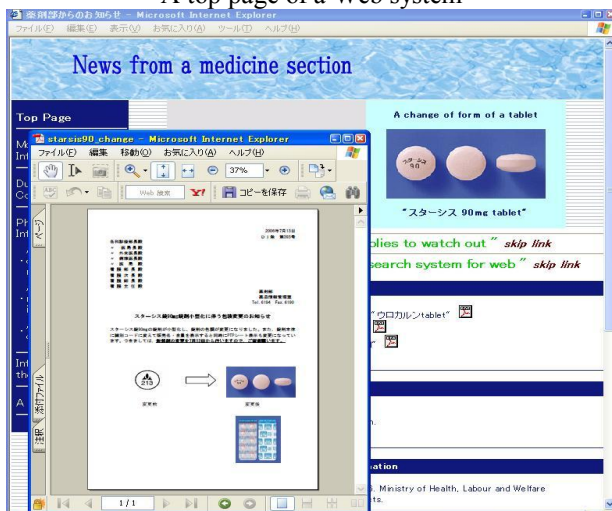


Figure.3

Medicine information of a Web system

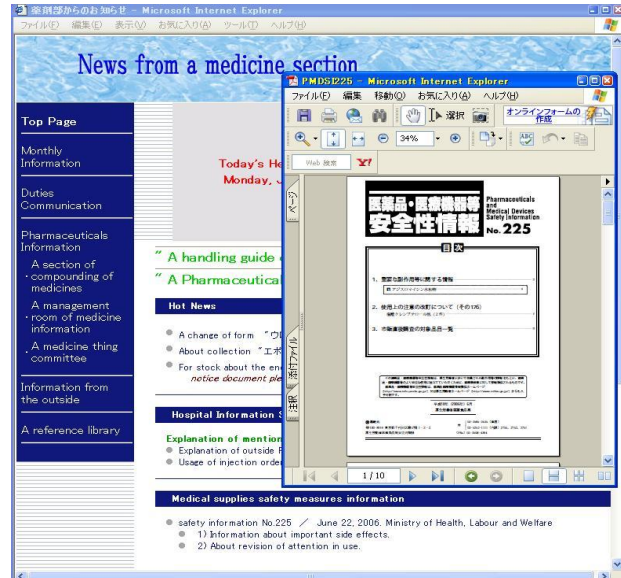


Figure.4

A PDF file of medical supplies safety information

### 3. Results

- 1) We analyzed a log file of three years.
- 2) Average access number of times 2003 was 95 times in annual for one day per one terminal. (Figure.5)
- 3) Average access number of times 2004 was 82 times in annual for one day per one terminal. (Figure.6)
- 4) Average access number of times 2005 was 79 times in annual for one day per one terminal. (Figure.7)
- 5) The average access number of times was 85 times for one day per one terminal of three years.
- 6) A lot of use of a medical worker of a new face is looked after in June from April. (Figure.8)
- 7) The access number of times is concentrated on Tuesday and Thursday among one week. There is it for reasons of being the next day of data update day. (Figure.9)
- 8) Access of a client terminal for medical staff is total 37%. The access number of times of a client terminal for consulting rooms is total 20%. The access number of times of a client terminal for medical technicians is total 15%. Emergency center has much access in Sunday and the night. (Figure.10)





## A study of formation and storage of laboratory results based on HL7 Clinical Document Architecture Standard

Hyung Hoi Kim<sup>1, 2</sup>, Shine Young Kim<sup>2</sup>, Hwa Sun Kim<sup>1</sup>, and Hune Cho<sup>1</sup>

<sup>1</sup>*Dept. of Medical Informatics Kyungpook National Univ., Republic of Korea*

<sup>2</sup>*Dept. of Laboratory Medicine and Biomedical informatics, Pusan National Univ., Republic of Korea*

### Abstract

**Background:** The Hospital Information System is used independently to other hospital information system, so Interoperability had been excluded. Clinical Document Architecture was applied in this study to maximize management of inter-hospital result information about bone marrow study and CBC test. For this, we suggested that template composition and schema design for test results of HL7CDA. **Methods:** After we defined template and item criteria of bone marrow and CBC test, we designed XML schema and created effective CDA documents. Schema was analyzed based on HL7 Reference Information Model and HL7 V2.4 was used for transporting protocol. **Results:** 1) It is the first trial to apply the definition of CDA for laboratory result in Korea as the basis. 2) It accomplished an expansion and purification of international standardization so that it could be applied to domestic standardization. 3) It represented the structure of future hospital systems based on the CDA. **Conclusion :** The study for CDA makes it possible that Clinical Data Repository which is knowledge based storage of evidence oriented medicine and Electronic Health Record in future. At this point of time, it makes possible to share information about lab finding which is very important evidence to clinical physicians for diagnosis between variable health medical institutions.



## **The development of additional automatic mapping protocols for local LIS system which do not make one-to-one correspondence with LOINC code set**

Hyung Hoi Kim<sup>1, 2</sup>, Shine Young Kim<sup>2</sup>, Hwa Sun Kim<sup>1</sup>, and Hune Cho<sup>1</sup>

<sup>1</sup>*Dept. of Medical Informatics Kyungpook National Univ., Republic of Korea*

<sup>2</sup>*Dept. of Laboratory Medicine and Biomedical informatics, Pusan National Univ., Republic of Korea*

### **Abstract**

The Logical Observation Identifier Names and Codes (LOINC) code set is the only publicly available universal standard for laboratory test names and essential for EMR system. Although Regenstrief Institute maintaining the LOINC database provide an automatic LOINC mapping tool (RELMA), it is not sufficient for applying LOINC code set to LIS system of local hospitals. LOINC code has up to six attributes and if one of six attributes has the different value, it means different LOINC code is defined. Because many local hospitals have used their own LIS system with arbitrary codes which is not followed by LOINC attributes and some codes contain several same attributes (especially specimen), one code of LIS system would match with several LOINC codes. So we cannot help developing protocols and tools to match codes of LIS system to codes of LOINC as one to one.

As the authors develop automatic tools to connect LIS system of Pusan National University Hospital with LOINC code set, we has also developed additional protocols converting own LIS data to correspond with LOINC. And it may be valuable as facilities for applying of LOINC code set essential for EMR system to the local hospitals which has their own LIS system.

# A u t h o r I n d e x

|                      |                           |                  |               |
|----------------------|---------------------------|------------------|---------------|
| Adi Utarini          | 68                        | Hsu-Hua Ho       | 94            |
| Akinori Abe          | 171                       | Hui Nar Quek     | S41           |
| Alvin B. Marcelo     | 110, S25                  | Hung-Pin Hou     | 94            |
| Anis Fuad            | 68                        | Hung-Wen Chiu    | 31, 270       |
| Athar Sadiq Muhammad | 49, 184, 220,<br>274, S24 | Hung-Wen Tung    | 128           |
| Atsuko Hara          | 269                       | Isao NAKAJIMA    | 257           |
| Augustine Tsai       | 128                       | Isao Nakajima    | 49, 184       |
| Austin H Chen        | 24                        | Isao NAKAJIMA    | 267, 269, S23 |
| Chen Gi              | S16                       | Jeongeun Kim     | 235           |
| Chen k. Feng         | 83                        | Jeongyee Bae     | 272           |
| Chia Ching Huang     | 146                       | Jia-Yuarn Guo    | 275           |
| Chien-Yeh Hsu        | 227, 236, 270             | Jin Mo Goo       | 39            |
| Chih-Cheng Lu        | 266                       | Jinsheng Roan    | 266           |
| Chih-Chung Chen      | 270                       | Johan Groothoff  | 68            |
| Chi-Pang Wen         | 78                        | John FULCHER     | 154           |
| CHIUNG HSU CHEN      | 132                       | Jong Hyo Kim     | 39, 45        |
| Cho Hune             | 118, 278, 279             | Jun Nakaya       | 1             |
| Chung Pao Lin        | 54                        | Ka-Lok Ng        | 15            |
| Chung-hsiu Kuo       | 275                       | Kaoru Sagara     | 171           |
| CP Wong              | S39                       | Kazuya Okamoto   | 177           |
| Datin Dr.S.Selvaraju | S31                       | Kei Kakehi       | 49            |
| David W Bates        | 235                       | Kenji TANAKA     | S26           |
| Da-Wei Wang          | 200                       | Khin Than Win    | 154           |
| Fei-Hung Hung        | 31                        | Kiyomu Ishikawa  | 249           |
| Fumio Taoka          | 102                       | Kiyoshi IGARASHI | S27           |
| Futoshi Naya         | 171                       | Kiyoshi Kogure   | 171           |
| Hao-Chung Hsu        | 19                        | Kuan Chih Lin    | 146           |
| Hari Kusnanto        | 68                        | Kumara Mendis    | 72            |
| Hidehiko Tsukuma     | 160                       | Kung Chen        | 200           |
| Hirohiko Mishima     | 276                       | Kuo Chin Tsai    | 146           |
| Hiromi itoh Ozaku    | 171                       | Kwang Gi Kim     | 39, 45        |
| Hiroshi Juzoji       | 49, 184                   | Li Baoluo        | S14           |
| Hiroyuki Yoshihara   | 177, 208                  | Li Xin Xu        | 241           |
| Horng-Shyang Liao    | 60                        | LungWen Tsai     | 188           |
|                      |                           | Makoto Fukui     | 142           |

|                   |          |                    |          |
|-------------------|----------|--------------------|----------|
| Mann-Li Chang     | 260      | Yasumitsu TOMIOKA  | S25      |
| Masaaki Teshima   | 102      | Yaw-Tang Shih      | 78       |
| Masatsugu - Tsuji | 102      | Yi Chen            | 268      |
| Meng Xiaoyang     | S17      | Yong Feng Huang    | 268      |
| Michio Kimura     | 1, S20   | Yoon Kim           | S9       |
| Naoki Aida        | 49       | Yoshie Ashida      | 224      |
| Naoki Nakashima   | 123      | Youko Ohashi       | 142      |
| Naoto Kume        | 208      | Young Moon Chae    | S1       |
| Nobuyuki Ashida   | 142, 224 | Yu Ju Chih         | 194      |
| Noriaki Kuwahara  | 171      | Yuan-Chii G Lee    | 19       |
| Norio KIMURA      | S22      | Yuan-Chun Chang    | 200      |
| PEI CHUN          | 275      | Yu-Chuan (Jack) Li | 227, S15 |
| Po-Huang Chiang   | 78       | Yuen Ho Chow       | S42      |
| Rasika Amarasiri  | 72       | Yuichi Ishibashi   | 269      |
| Ruey k. Chiu      | 83, 146  | Yu-jen Lee         | 266      |
| Sang Ho Lee       | 39, 45   | Yu-Lun Liu         | 227      |
| Sang-joon Park    | 39, 45   | Yung-Tai Yen       | 227      |
| Shen Tao          | S15      | Yung-Yu Su         | 154      |
| Shiou-Ping Lee    | 36       | Yu-Tung Dai        | 275      |
| Si Fa Zhang       | 268      | Zhi En Chen        | 24       |
| Sukwha Kim        | 235      |                    |          |
| Tadamasa Takemura | 177      |                    |          |
| Taemi Hirai       | 142      |                    |          |
| Tanaka Takeshi    | 138      |                    |          |
| Ten Fang Yang     | 188      |                    |          |
| Terumasa Higasi   | 142      |                    |          |
| Tomohiro Kuroda   | 177, 208 |                    |          |
| Toshihiko Kitano  | 49       |                    |          |
| Tsair Wei Chien   | 90       |                    |          |
| Tzu-Yi Yu         | 94       |                    |          |
| Walter T.L. Lim   | S16      |                    |          |
| Wansa Paoiin      | S29      |                    |          |
| Wei Yi Fan        | 236      |                    |          |
| Wenlung Shu       | 9        |                    |          |
| Wen-Yung Wu       | 94       |                    |          |
| Wetter Thomas     | 214      |                    |          |
| Woo Kyung Moon    | 45       |                    |          |
| Yasumitsu Tomioka | 49, 184  |                    |          |