

Measuring the Quality of Hospital Care

**Min Hua Jen
Imperial College London**

Contents

Background

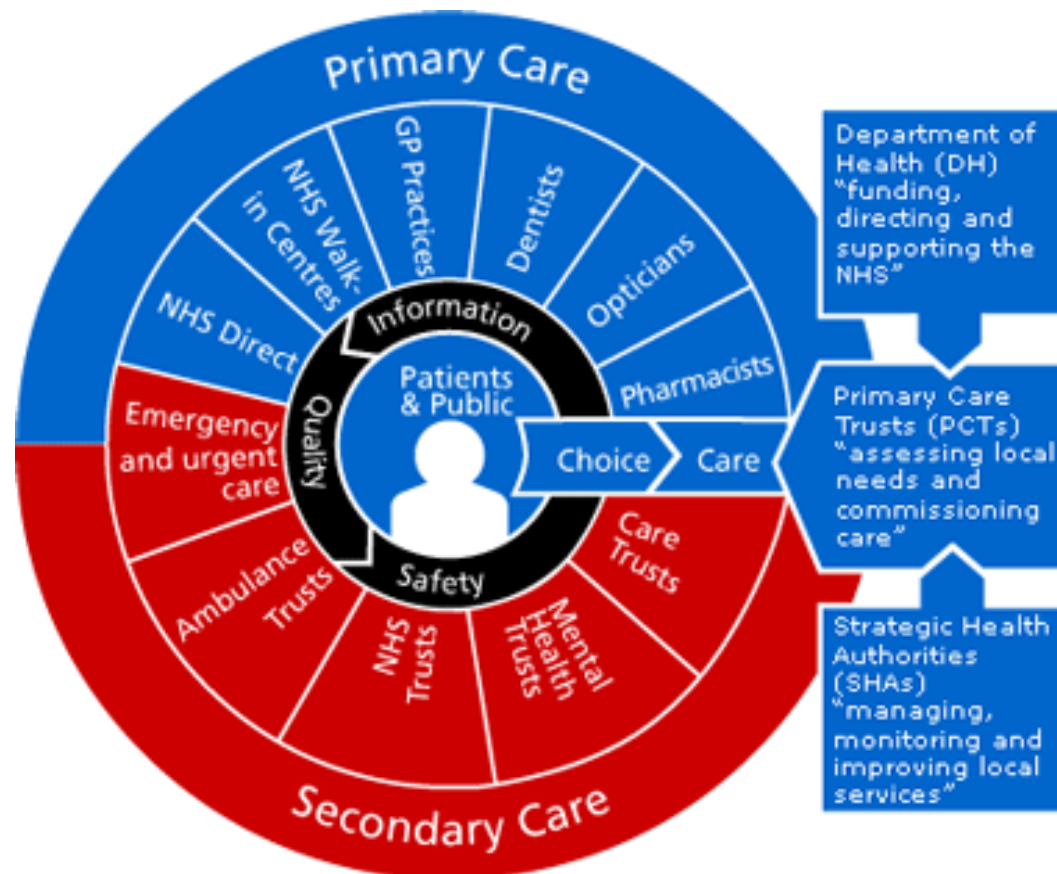
English Hospital Statistics

Case-mix adjustment

Presentation of performance data

- League tables
- Bayesian ranking
- Statistical process Control Charts

NHS structure



Key events

- Heart operations at the BRI

“Inadequate care for one third of children”



- Harold Shipman

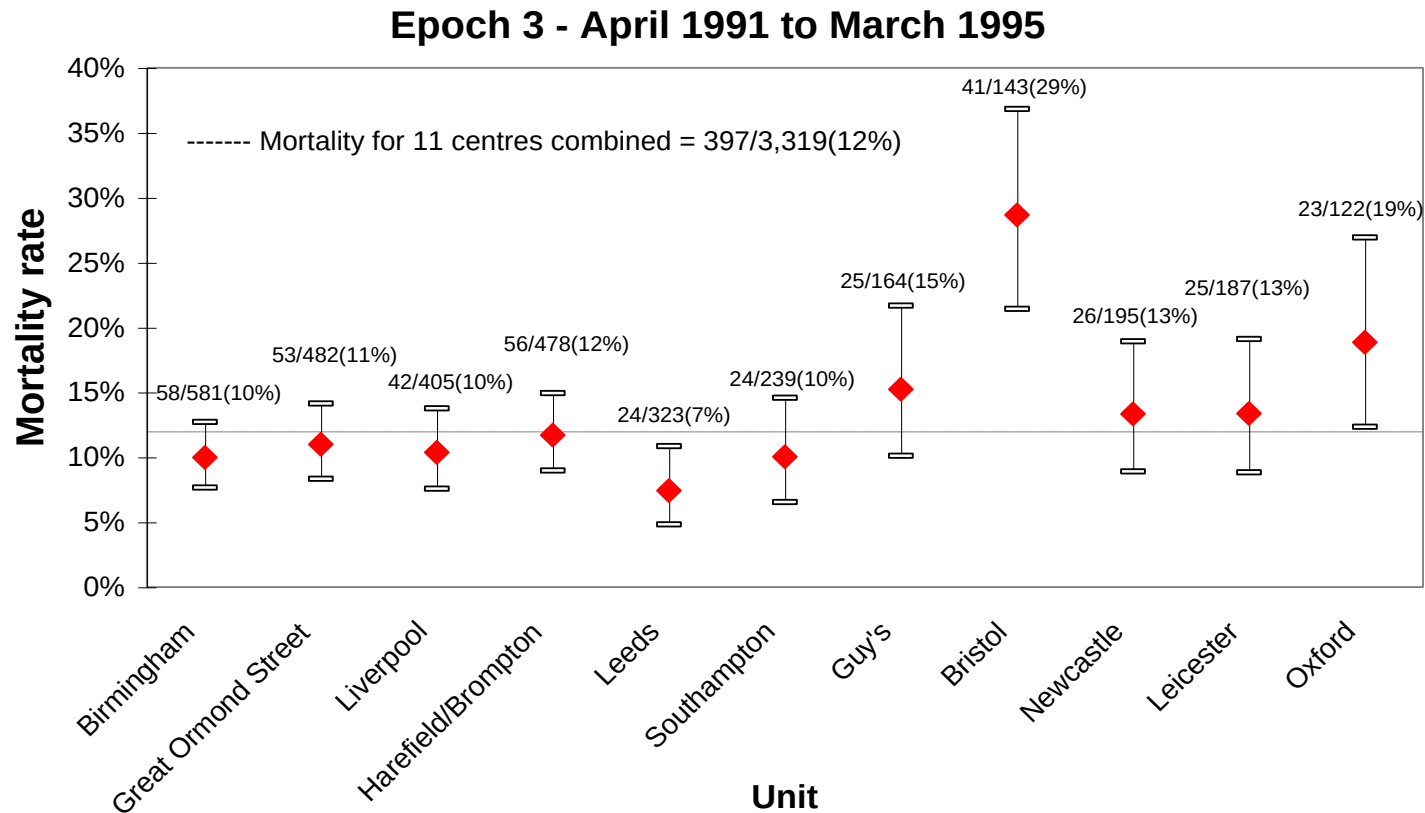
Murdered more than 200 patients



Bristol (Kennedy) Inquiry Report Data were available all the time

“From the start of the 1990s a national database existed at the Department of Health (the Hospital Episode Statistics database) which among other things held information about deaths in hospital. It was not recognised as a valuable tool for analysing the performance of hospitals. It is now, belatedly.”

Mortality from open procedures in children aged under one year for 11 centres in three epochs; data derived from Hospital Episode Statistics (HES)



Following the Bristol Royal Infirmary Inquiry

- Commission for Health Improvement (now Healthcare Commission) - regularly inspect Britain's hospitals and publish some limited performance figures.
- National Clinical Assessment Authority – investigates any brewing crisis.
- National Patient Safety Agency collates information on medical errors.
- Annual appraisals for hospital consultants
- Revalidation, a system in which doctors have to prove they are still fit to practice every five years

Hospital Episode Statistics

Electronic record of every inpatient or day case episode of patient care in every NHS (public) hospital

14 million records a year

300 fields of information including

- Patient details such as age, sex, address
- Diagnosis using ICD10
- Procedures using OPCS4
- Admission method
- Discharge method

Why use Hospital Episode Statistics

- Comprehensive – collected by all NHS trusts across country on all patients
- Coding of data separate from clinician
- Access
- Updated monthly from SUS (previously NHS Wide Clearing Service)

Case mix adjustment

Limited within HES?

- Age
- Sex
- Emergency/Elective

Use of administrative data or clinical databases as predictors of risk of death in hospital: comparison of models

Paul Aylin, clinical senior lecturer,¹ Alex Bottle, lecturer,¹ Azeem Majeed, professor of primary care and social medicine²

¹Dr Foster Unit, Imperial College London, London EC1A 9LA

²Department of Primary Care and Social Medicine, Imperial College London

Correspondence to: P Aylin
p.aylin@imperial.ac.uk

doi: 10.1136/bmj.39168.496366.55

ABSTRACT

Objective To compare risk prediction models for death in hospital based on an administrative database with published results based on data derived from three national clinical databases: the national cardiac surgical database, the national vascular database and the colorectal cancer study.

Design Analysis of inpatient hospital episode statistics. Predictive model developed using multiple logistic regression.

Setting NHS hospital trusts in England.

Patients All patients admitted to an NHS hospital within England for isolated coronary artery bypass graft (CABG), repair of abdominal aortic aneurysm, and colorectal excision for cancer from 1996-7 to 2003-4.

Main outcome measures Deaths in hospital. Performance of models assessed with receiver operating characteristic (ROC) curve scores measuring discrimination (<0.7 =poor, $0.7-0.8$ =reasonable, >0.8 =good) and both Hosmer-Lemeshow statistics and standardised residuals measuring goodness of fit.

Results During the study period 152 523 cases of isolated CABG with 3247 deaths in hospital (2.1%), 12 781 repairs of ruptured abdominal aortic aneurysm (5987 deaths, 46.8%), 31 705 repairs of unruptured abdominal aortic aneurysm (3246 deaths, 10.2%), and 144 370 colorectal resections for cancer (10 424 deaths, 7.2%) were recorded. The power of the complex predictive model was comparable with that of models based on clinical datasets with ROC curve scores of 0.77 (v 0.78 from clinical database) for isolated CABG, 0.66 (v 0.65) and

INTRODUCTION

Routine administrative databases are increasingly being used for performance monitoring in healthcare in the United Kingdom (such as www.healthcarecommission.org.uk and www.drfooster.co.uk), United States (such as www.ihl.org/IHL/Programs/Campaign/), and elsewhere.¹ In comparisons of performance between clinicians or organisations it is essential to adjust for several parameters including comorbidity and severity of disease (case mix). Routine data, however, might contain insufficient information for adequate adjustment. Clinical databases, run by various bodies including professional societies, could potentially record more detailed clinical information and might permit better adjustment for case mix. A survey of 105 multi-centre clinical databases (which included hospital episode statistics, the administrative database available within England) found that their distribution was uneven and that their scope and the quality of the data was variable.² The report from the public inquiry into deaths at a paediatric cardiac unit at Bristol criticised this "dual" system as "wasteful and anachronistic."³ It also suggested that hospital episode statistics should be supported as a major national resource and used to undertake monitoring of a range of healthcare outcomes.

We examined mortality for three index procedures (coronary artery bypass graft, abdominal aortic aneurysm repair, and colectomy for bowel cancer) used in three large clinical datasets (the national adult cardiac surgical database, the national vascular database, and a

CABG, the c statistic 0.83 from the national

is based on the most procedures giving is for tenths of risk t. Although there is nation of risk for an underestimation rhaps to be expected o be close agreement icted numbers and The Hosmer-Lemeshow suggest a highly significant difference between deaths predicted from r colorectal excision t model). Repair of eurysm also shows . The proportion of he range -1.96 to BG, 0.5% for repair neurysm, 7.3% for ortic aneurysm, and So influential points

tal models with good operative death as an with those derived

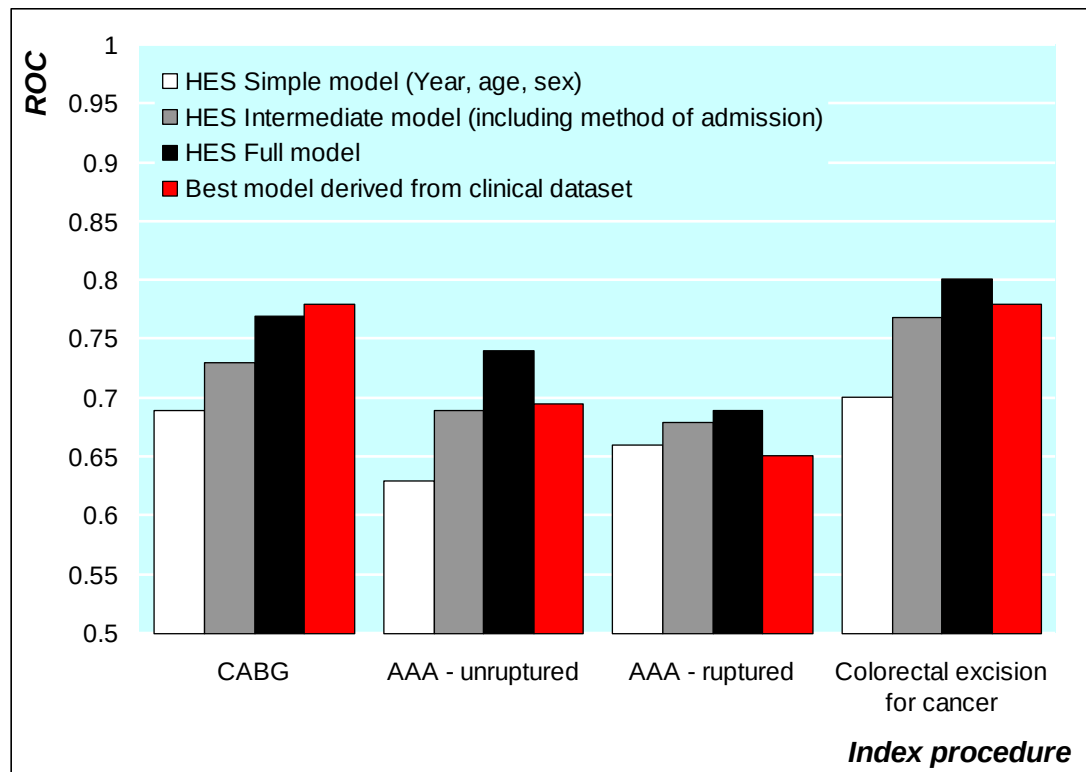
Risk adjustment models using HES on 3 index procedures

- CABG
- AAA
- Bowel resection for colorectal cancer

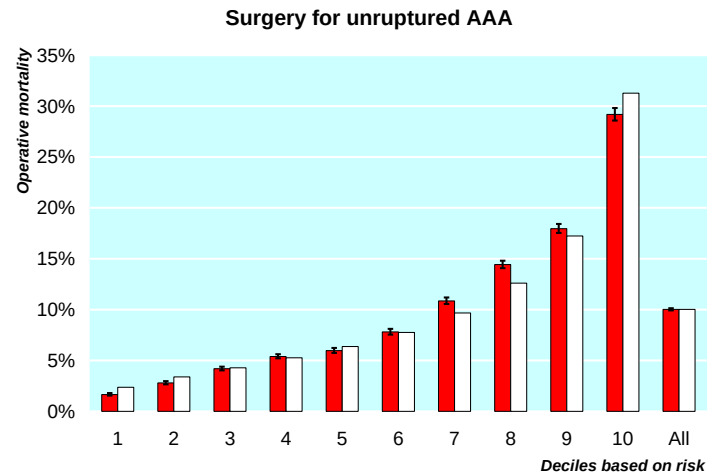
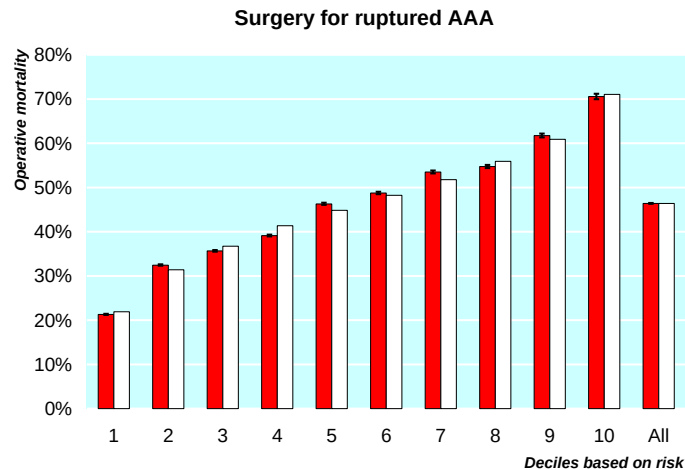
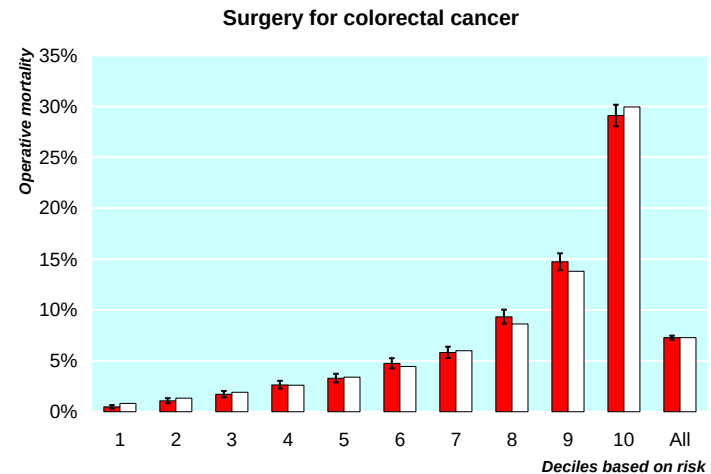
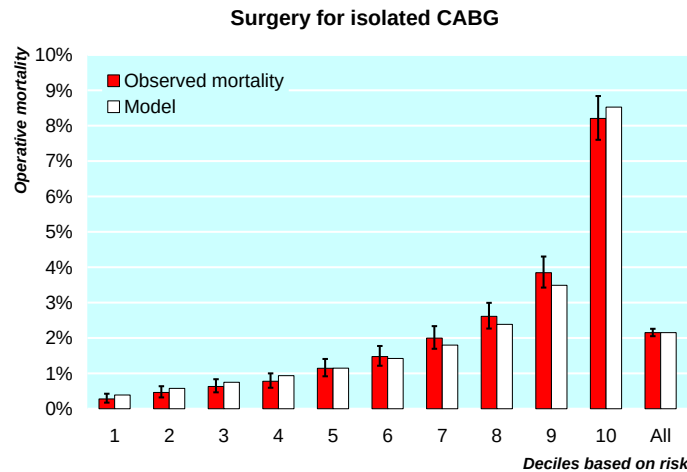
Risk factors

Age	Recent MI admission
Sex	Charlson comorbidity score (capped at 6)
Method of admission	Number of arteries replaced
Revision of CABG	Part of aorta repaired
Year	Part of colon/rectum removed
Deprivation quintile	Previous heart operation
Previous emergency admissions	Previous abdominal surgery
Previous IHD admissions	

ROC curve areas comparing 'simple', 'intermediate' and 'complex' models derived from HES with models derived from clinical databases for four index procedures



Calibration plots for 'complex' HES-based risk prediction models for four index procedures showing observed number of deaths against predicted based on validation set



Aylin P; Bottle A; Majeed A. Use of administrative data or clinical databases as predictors of risk of death in hospital: comparison of models. BMJ 2007;334: 1044

Current casemix adjustment model for each diagnosis and procedure group

Adjusts for

- age
- sex
- elective status
- socio-economic deprivation
- Diagnosis subgroups (3 digit ICD10) or procedure subgroups
- co-morbidity – Charlson index
- number of prior emergency admissions
- palliative care
- year
- month of admission

Current performance of risk models

ROC (based on 1996/7-2007/8 HES data) for in-hospital mortality

56 Clinical Classification System diagnostic groups leading to 80% of all in-hospital deaths

7 CCS groups 0.90 or above

- Includes cancer of breast (0.94) and biliary tract disease (0.91)

28 CCS groups 0.80 to 0.89

- Includes aortic, peripheral and visceral aneurysms (0.87) and cancer of colon (0.83)

18 CCS groups 0.7 to 0.79

- Includes septicaemia (0.77) and acute myocardial infarction (0.74)

3 CCS groups 0.60 to 0.69

- Includes COPD (0.69) and congestive heart failure (0.65)

Presentation of clinical outcomes

“Even if all surgeons are equally good, about half will have below average results, one will have the worst results, and the worst results will be a long way below average”

- *Poloniecki J. BMJ 1998;316:1734-1736*

SOUTHEAST HOSPITALS

SOUTHEAST NHS REGION

HOW THE NHS TRUSTS COMPARE

Ranked by mortality rate from low to high
Average index for Southeast region: 98

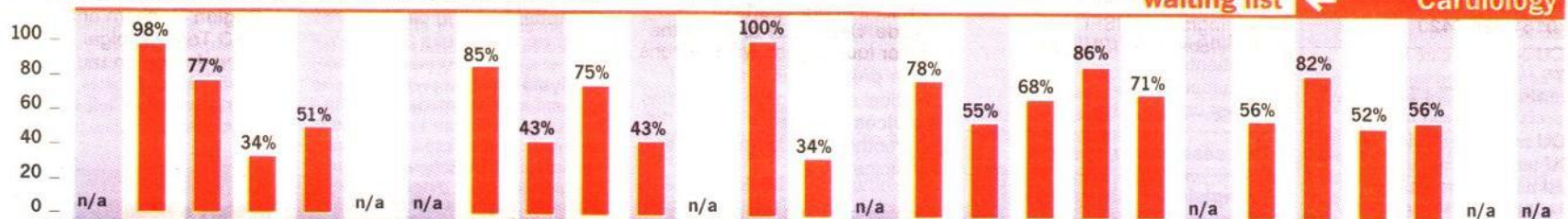
Low Average High

	Mortality index	Doctors per 100 beds	Nurses per 100 beds	Page
1 Royal West Sussex	81	37	123	19
2 North Hampshire Hospitals	88	25	167	-
3 Mid Sussex	90	32	132	-
4 Portsmouth Hospitals	90	41	135	13
5 Royal Surrey County Hospital	91	46	107	18
6 Stoke Mandeville Hospital	93	38	118	30
7 Winchester and Eastleigh Healthcare	94	32	133	40
8 Worthing and Southlands Hospitals	94	30	108	42
9 Isle of Wight Healthcare	95	29	130	-
10 Northamptonshire General Hospital	96	36	134	-
11 Eastbourne Hospitals	96	28	94	-
12 Ashford and St Peter's Hospital	96	43	132	-
13 Oxford Radcliffe Hospitals	96	70	176	10
14 Royal Berkshire and Battle Hospitals	97	39	140	15
15 Brighton Health Care	98	45	149	-
16 Frimley Park Hospital	98	34	105	-
17 Maidstone and Tunbridge Wells	99	37	116	-
18 Hastings and Rother	99	30	109	-
19 East Kent Hospitals	101	37	120	-
20 Southampton University Hospitals	104	56	120	26
21 Milton Keynes General	104	43	132	-
22 Kettering General Hospital	104	22	131	-
23 South Buckinghamshire	104	33	118	23
24 Surrey and Sussex Healthcare	105	32	96	30
25 The Medway	106	40	136	-
26 Dartford and Gravesham	111	29	106	-
27 Heatherwood and Wexham Park Hospitals	115	33	111	-

Ranking based on unrounded mortality figures

Where no page number is indicated
trust profile appeared in part one, last Sunday

OUTPATIENTS SEEN WITHIN 13 WEEKS



Criticisms of 'league tables'

- Spurious ranking – 'someone's got to be bottom'
- Encourages comparison when perhaps not justified
- 95% intervals arbitrary
- No consideration of multiple comparisons
- Single-year cross-section – what about change?

Bayesian ranking

Bayesian approach using Monte Carlo simulations can provide confidence intervals around ranks

Can also provide probability that a unit is in top 10%, 5% or even is at the top of the table

- *See Marshall et al. (1998). League tables of in vitro fertilisation clinics: how confident can we be about the rankings? British Medical Journal, 316, 1701-4.*

THE 15 ENGLISH TRUSTS WITH THE LOWEST MORTALITY RATES

Region	NHS Trust	Mortality Index (3yr)	Probability of being in top 10%	Page
1 London	Barts and The London	60.7	100%	(19)
2 London	University College London Hospitals	71.0	100%	(25)
3 Eastern	Bedford Hospital	73.5	100%	(12)

THE 15 ENGLISH TRUSTS WITH THE HIGHEST MORTALITY RATES

	Region	NHS Trust	Mortality Index (3yr)	Probability of being in top 10%	Page
1	West Midlands	Walsall Hospitals	126.0	100%	(70)
2	Northwest	Bolton Hospitals	121.0	100%	(35)
3	Southeast	Heatherwood and Wexham Park Hospitals	120.4	100%	(44)
4	Southeast	The Medway	120.2	100%	(46)
5	Southwest	The Royal Bournemouth and Christchurch Hospitals	119.1	100%	(56)
6	West Midlands	Sandwell Healthcare (Sandwell & West Birmingham)	118.2	100%	(69)
7=	Northwest	Tameside & Glossop Acute Services	116.2	79%	(40)
7=	West Midlands	George Eliot Hospital	116.2	91%	(66)
9	London	Barnet and Chase Farm Hospitals	116.1	98%	(18)
10	Eastern	The Princess Alexandra Hospital, Harlow	115.7	86%	(15)
11	London	Newham Healthcare	115.1	76%	(23)
12	Eastern	Essex Rivers Healthcare	115.0	93%	(12)
13	London	Queen Elizabeth Hospital	114.9	81%	(24)
14	London	Barking, Havering and Redbridge Hospitals	114.3	90%	(18)
15	Northwest	Blackpool Victoria Hospital (Blackpool, Fylde & Wyre)	114.1	87%	(35)

5	London	Barnet and Chase Farm Hospitals	125.6	12.1	(18)
6	Northern and Yorkshire	North Durham Health Care (Co Durham & Darlington Acute)	100.3	11.6	(29)
7	Southeast	Dartford and Gravesham	119.0	11.5	(43)
8	Southeast	Royal Berkshire and Battle Hospitals	110.2	10.7	(48)
9	Northern and Yorkshire	Northallerton Health Services (South Tees Hospitals)	88.5	9.1	(33)
10	Southeast	Portsmouth Hospitals	95.3	8.7	(48)

Source: Dr Foster

Statistical Process Control (SPC) charts

Shipman:

- Aylin et al, *Lancet* (2003)
- Mohammed et al, *Lancet* (2001)
- Spiegelhalter et al, *J Qual Health Care* (2003)

Surgical mortality:

- Poloniecki et al, *BMJ* (1998)
- Lovegrove et al, *CHI* report into St George's
- Steiner et al, *Biostatistics* (2000)

Public health:

- Terje et al, *Stats in Med* (1993)
- Vanbrackle & Williamson, *Stats in Med* (1999)
- Rossi et al, *Stats in Med* (1999)
- Williamson & Weatherby-Hudson, *Stats in Med* (1999)

Common features of SPC charts

Need to define:

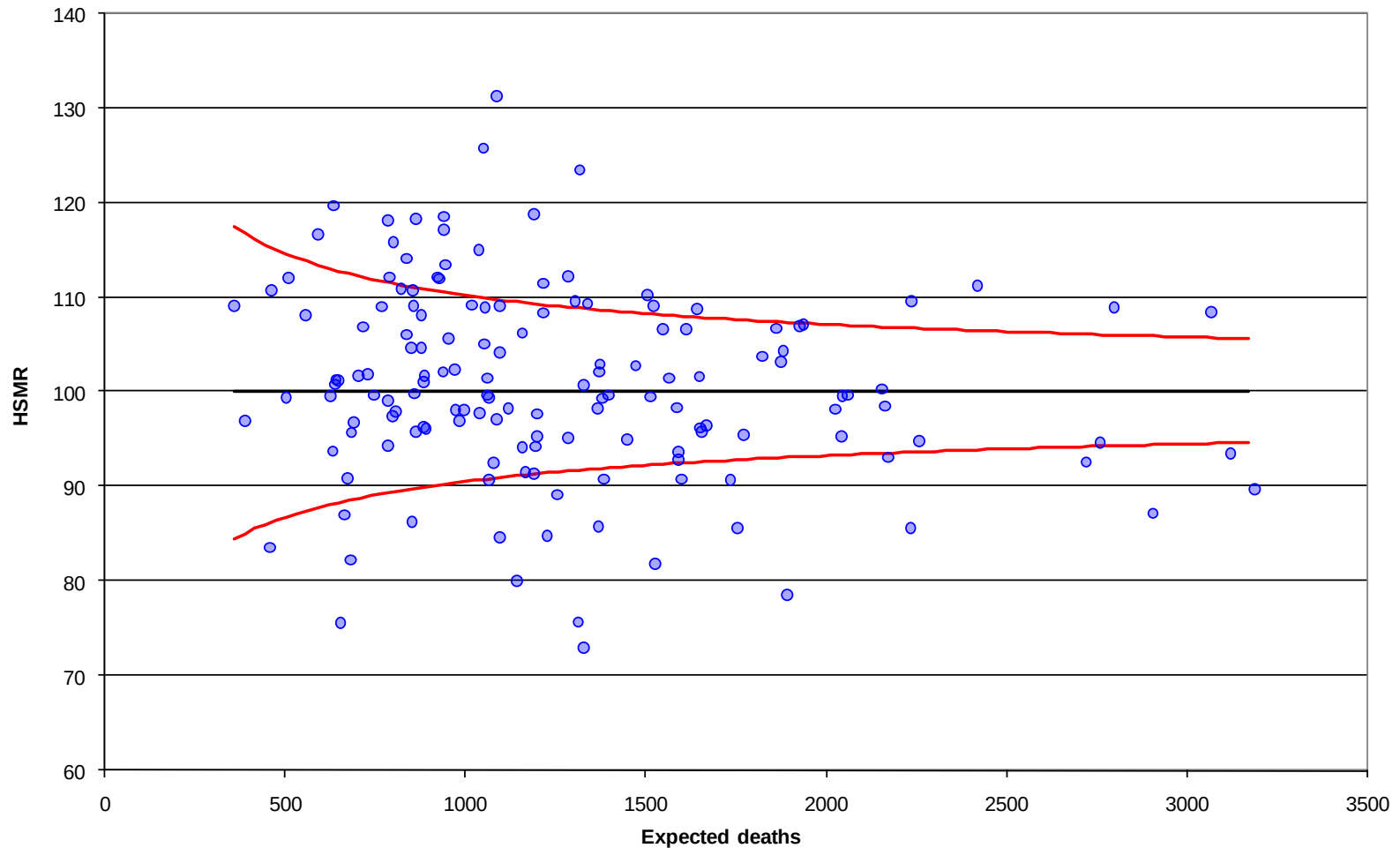
- in-control process (acceptable/benchmark performance)
- out-of-control process (that is cause for concern)

Test statistic

- Function of the difference between observed and benchmark performance
- calculated for each unit of analysis

HSMR 2007/8 with 99.8% control limits

HSMR with 99.8% control limits 2007/8



Funnel plots

No ranking

Visual relationship with volume

Takes account of increased variability of smaller centres

Risk-adjusted Log-likelihood CUSUM charts

- STEP 1: estimate pre-op risk for each patient, given their age, sex etc. This may be national average or other benchmark
- STEP 2: Order patients chronologically by date of operation
- STEP 3: Choose chart threshold(s) of acceptable “sensitivity” and “specificity” (via simulation)
- STEP 4: Plot function of patient’s actual outcome v pre-op risk for every patient, and see if – and why – threshold(s) is crossed

More details

- Based on log-likelihood CUSUM to detect a predetermined increase in risk of interest
- Taken from Steiner et al (2000); pre-op risks derived from logistic regression of national data
- The CUSUM statistic is the log-likelihood test statistic for binomial data based on the predicted risk of outcome and the actual outcome
- Model uses administrative data and adjusts for age, sex, emergency status, socio-economic deprivation etc.

Monitoring surgical performance using risk-adjusted cumulative sum charts

STEFAN H. STEINER*, RICHARD J. COOK

*Department of Statistics and Actuarial Sciences, University of Waterloo, Waterloo,
Ontario, Canada N2L 3G1
shsteine@uwaterloo.ca*

VERN T. FAREWELL

Department of Statistical Sciences, University College London, UK

TOM TREASURE

St. George's Hospital Medical School London, UK

$$W_t = \begin{cases} \log \left[\frac{(1 - p_t + R_0 p_t) R_A}{(1 - p_t + R_A p_t) R_0} \right] & \text{if } y_t = 1 \\ \log \left[\frac{1 - p_t + R_0 p_t}{1 - p_t + R_A p_t} \right] & \text{if } y_t = 0 \end{cases}$$

Currently monitoring

- 78 diagnoses
- 128 procedures
- 90% deaths
- Outcomes
 - Mortality
 - Emergency readmissions
 - Day case rates
 - Length of Stay



Performance Summary (Jun-08 to Aug-08)

Trust	Mortality	Length of Stay	Day Case Rate	Readmissions
Diagnoses	 	 	 	
Procedures		 		

Diagnoses	Mortality	Length of Stay	Day Case Rate	Readmissions
HSMR Basket of 56 Diagnoses Groups				
Abdominal pain				
Acute cerebrovascular disease				
Acute myocardial infarction				
Affective disorders				
Alcohol-related mental disorders				
Asthma				
Cancer of prostate				
Chronic obstructive pulmonary disease and bronchiectasis				
Leukaemias				
Nonspecific chest pain				
Normal pregnancy and/or delivery				
Other psychoses				
Pleurisy, pneumothorax, pulmonary collapse				
Residual codes, unclassified				
Respiratory failure, insufficiency, arrest (adult)				
Schizophrenia and related disorders				
Short gestation, low birth weight, and fetal growth retardation				
Syncope				

Procedures	Mortality	Length of Stay	Day Case Rate	Readmissions
Arthroscopy				
Bunion operations				
Carpal tunnel decompression				
Circumcision				
Correction of squint				
Diagnostic endoscopic procedures on upper GI tract				
Extraction of cataract with/without implant				
Extraction of tooth				
Haemorrhoidectomy				
Knee replacement				
Laparoscopy				
Myringotomy with/without grommets				
Reduction of fracture of bone				
Removal of metalware				
Submucous resection (septum or turbinate of nose)				
Termination of pregnancy (other)				

Reports

CUSUM

Relative Risk

Data Detail

Graph

Table

Description

Time Period

Level

Indicators to display

Performance Summary

3 months

Trust

Alerts (out of control)



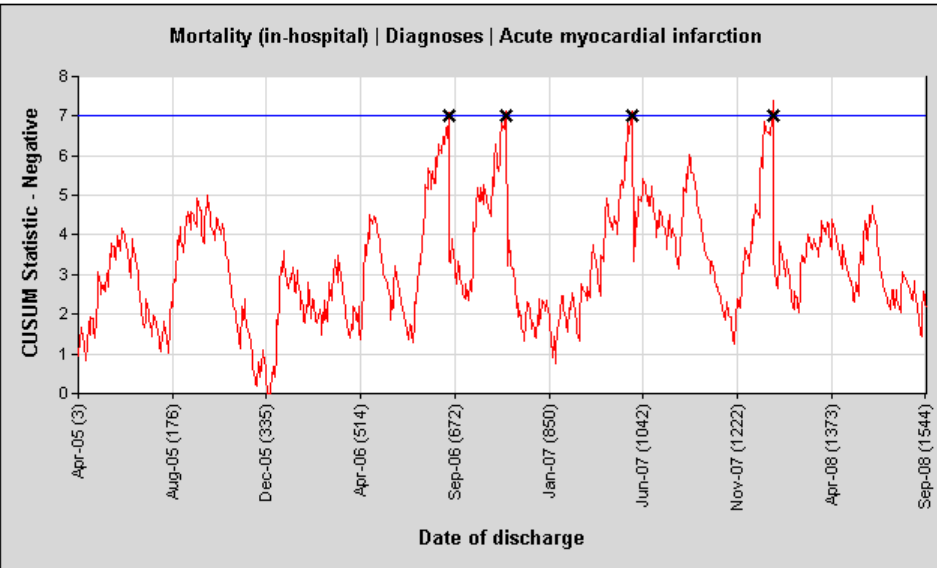
Provider: NHS Trust

Basket: Diagnoses - 56 HSMR groups
Outcome: Mortality (in-hospital)
Benchmarks: Data year

Chapter: Circulatory
Diagnosis: Acute myocardial infarction
Department: All
Team: All

Admission type: All
Sex: All
Age Range: All
Deprivation: All

Spells: 1546
Superspells: 1417 (1415/2)
First / Last: Apr-05 / Sep-08
Deaths: 240 (16.9%)
Expected: 163.5 (11.5%)
O-E: 76.5 (5.4%)
Relative Risk: 146.8 (128.8-166.6)
C-Statistic: 0.74 (Average)
LoS: 9.6 / 10.6
Alerts (X): 4 (Aug-06, Nov-06, May-07, Dec-07)
FAR / SDR: 0.8% / 97.3%



Reports

CUSUM

Relative Risk

Data Detail

Graph

Table

Basket: Diagnoses - 56 HSMR groups

Chapter: Circulatory

Diagnosis Group: Acute myocardial infarction

Department: -- All --

Team: -- All --

Admission Type: -- All --

Sex: -- All --

Deprivation: -- All --

Age Range: -- All --

View From: Apr 2005

View To: Most Recent

Benchmarks: Data year

View: Negative

Reset on alert: to Threshold+2

Start date: Apr 2005

Threshold: High (99%)

GENERATE



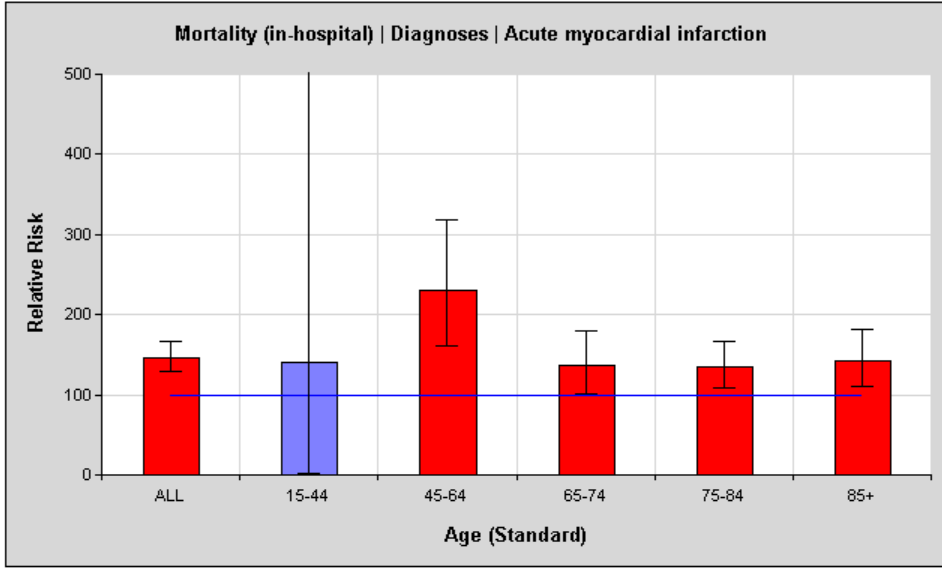


Basket: Diagnoses - 56 HSMR groups
 Outcome: Mortality (in-hospital)
 Benchmarks: Data year

Chapter: Circulatory
 Diagnosis: Acute myocardial infarction
 Department: All
 Team: All

Admission type: All
 Sex: All
 Age Range: All
 Deprivation: All

Spells: 1546
 Superspells: 1417 (1415/2)
 First / Last: Apr-05 / Sep-08
 Deaths: 240 (16.9%)
 Expected: 163.5 (11.5%)
 O-E: 76.5 (5.4%)
 Relative Risk: 146.8 (128.8-166.6)
 C-Statistic: 0.74 (Average)
 LoS: 9.6 / 10.6



Reports

CUSUM

Relative Risk

Data Detail

Graph

Table

Basket: Diagnoses - 56 HSMR groups

Chapter: Circulatory

Admission Type: -- All --

Sex: -- All --

Deprivation: -- All --

Age Range: -- All --

View From: Apr 2005

View To: Most Recent

Benchmarks: Data year

Outcome: Mortality (in-hospital)

Diagnosis Group: Acute myocardial infarction

Department: -- All --

Team: -- All --

Age (Standard): -- All --

Analyse by: Age (Standard)

GENERATE

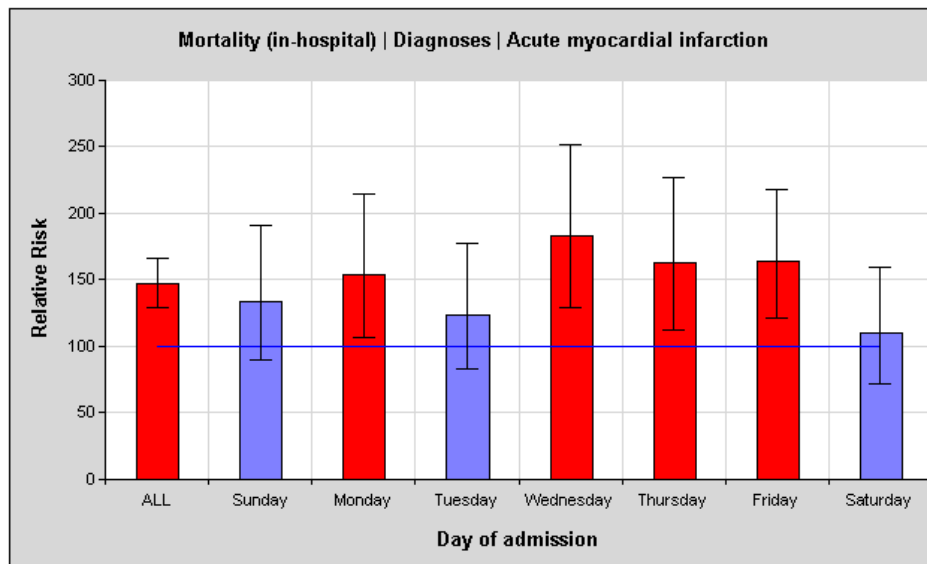


Basket: **Diagnoses - 56 HSMR groups**
 Outcome: **Mortality (in-hospital)**
 Benchmarks: **Data year**

Chapter: **Circulatory**
 Diagnosis: **Acute myocardial infarction**
 Department: **All**
 Team: **All**

Admission type: **All**
 Sex: **All**
 Age Range: **All**
 Deprivation: **All**

Spells: **1546**
 Superspells: **1417** (1415/2)
 First / Last: **Apr-05 / Sep-08**
 Deaths: **240** (16.9%)
 Expected: **163.5** (11.5%)
 O-E: **76.5** (5.4%)
 Relative Risk: **146.8** (128.8—166.6)
 C-Statistic: **0.74 (Average)**
 LoS: **9.6 / 10.6**



Reports	CUSUM	Relative Risk	Data Detail	Graph	Table
Basket Diagnoses - 56 HSMR groups			Outcome Mortality (in-hospital)		
Chapter Circulatory			Diagnosis Group Acute myocardial infarction		
Department -- All --			Team -- All --		
Admission Type -- All --	Sex -- All --	Deprivation -- All --	Age Range -- All --	View From Apr 2005	View To Most Recent
Benchmarks Data year				Day of admission -- All --	Analyse by Day of admission
GENERATE					

What to do with a signal

- Check the data
- Difference in casemix
- Examine organisational or procedural differences
- Only then consider quality of care

Future

- Patient Reported Outcomes (PROMs)
- Patient satisfaction/experience
- Safety/adverse events
- Pay for performance and quality