ORIGINAL ARTICLE

Effect of an acute medical admission unit on key quality indicators assessed by funnel plots

Edward D Moloney, Kathleen Bennett, Bernard Silke

Postgrad Med J 2007;**83**:659–663. doi: 10.1136/pgmj.2007.058511

See end of article for
authors' affiliations

Correspondence to:
Dr Bernard Silke,
Department of
Pharmacology &
Therapeutics, Trinity Centre
for Health Sciences, St
James' Hospital, James'
Street, Dublin 8, Ireland;
silkeb@tcd.ie

Received 9 February 2007
Accepted 26 May 2007

Objective: To determine the effect of the introduction of an acute medical admissions unit (AMAU) on key quality efficiency and outcome indicator comparisons between medical teams as assessed by funnel plots.

Methods: A retrospective analysis was performed of data relating to emergency medical patients admitted to St James' Hospital, Dublin between 1 January 2002 and 31 December 2004, using data on discharges from hospital recorded in the hospital in-patient enquiry system. The base year was 2002 during which patients were admitted to a variety of wards under the care of a named consultant physician. In 2003, two centrally located wards were reconfigured to function as an AMAU, and all emergency patients were admitted directly to this unit. The quality indicators examined between teams were length of stay (LOS) <30 days, LOS >30 days, and readmission rates.

Results: The impact of the AMAU reduced overall hospital LOS from 7 days in 2002 to 5 days in 2003/04 ($p<0.0001$). There was no change in readmission rates between teams over the 3 year period, with all teams displaying expected variability within control (95%) limits. Overall, the performance in LOS, both short term and long term, was significantly improved ($p<0.0001$), and was less varied between medical teams between 2002 and 2003/04.

Conclusions: Introduction of the AMAU improved performance among medical teams in LOS, both short term and long term, with no change in readmissions. Funnel plots are a powerful graphical technique for presenting quality performance indicator variation between teams over time.

A relentless rise in emergency admissions in recent years has coincided with a reduction in hospital beds, resulting in severe problems in most acute hospitals.^{1,2} Evidence suggests that with reorganisation it may be possible to improve patient care and reduce the pressure on beds within existing resources.³ Spare bed capacity is essential if an emergency admissions service is to operate efficiently. Moreover, it is recognised that an acute hospital can expect regular bed shortages for emergency department patients when hospital bed occupancy rates exceed 85%, and periodic bed crises if hospital bed occupancy rises to 90% or more.^{4,5} We have previously reported how the introduction of an acute medical admissions unit (AMAU) facilitated speedier access from the emergency department to the acute medical service, and reduced hospital costs by decreasing length of stay (LOS).^{6,7}

The efficient utilisation of public resources and demands for increased accountability has focused attention on institutional comparisons on the basis of quantitative outcome measures. Such comparisons commonly lead to the production of "league tables", in which institutions or groups are ranked according to a performance indicator.⁸ However, league tables polarise results, by creating a state of "sanctuary" in hospitals above the cut-off, and a state of "alarm" in those below it.

Funnel plots discourage inappropriate ranking, while performing strong visual indication of "divergent" performance or "special cause" variation, and are suggested as the display methods of choice for institutional comparisons.^{9,10} The aim of this study was to examine the effect of reorganisation of an acute admissions process on key quality efficiency and outcome variables in individual consultant groups within one hospital, as assessed by funnel plots.

PATIENTS AND METHODS

Data relating to emergency medical patients admitted to St James' Hospital (SJH), Dublin, between 1 January 2002 and 31

December 2004 were recorded. Data were not recorded before 2002. SJH, although a tertiary referral centre for various specialties, operates a daily sectorised acute general medical "take-in" serving as a secondary care centre for emergency medical admissions for its local Dublin catchment area. In 2002, emergencies in acute medicine were initially assessed by the staff of the emergency department and referred by them to the on-call medical team of the day. All such patients identified as requiring hospitalisation, apart from cases admitted directly to the coronary care or intensive care units, were admitted to a variety of wards, many of which were not affiliated with a medical specialty, under the care of a named consultant physician. Fourteen consultant physicians (all dual-accredited in internal medicine and a major subspecialty) were responsible for the management of these patients, of which 10 were full-time health service consultants and four held split service/academic appointments. The "on-call" roster is a 1:9 with two slots each operated by teams from respiratory medicine and gastroenterology, one slot each contributed by specialty teams from diabetes/endocrinology, clinical pharmacology, and rheumatology, and one slot each contributed by two teams of general internal medicine.

In 2003, two of the modern centrally located wards, with close proximity to the emergency department and diagnostic imaging department, were reconfigured to function as an AMAU. A detailed operational plan for the unit was devised following extensive discussions with all interested parties in the year before its inception. Emergencies in acute medicine were initially assessed by the staff of the emergency department and referred by them to the on-call AMAU team of the day. All such

Abbreviations: AMAU, acute medical admissions unit; CI, confidence interval; HIPE, Hospital In-Patient Enquiry; ICD-9-CM, International Classification of Diseases, ninth revision; Clinical Modification; IQR, interquartile range; LOS, length of stay; PAS, patient administration system; SJH, St James' Hospital

patients requiring hospitalisation were admitted directly to the AMAU from the emergency department. The 59 bed capacity of the AMAU is sufficient, so that with an average of 15 admissions each day, less than 70% of all admissions would be predicted to receive their entire hospital care within the unit (maximum permitted stay in AMAU is 5 days). Those patients requiring a longer stay were transferred from the AMAU to the appropriate specialty or general medical beds. In 2003 and 2004, the on-call roster remained at 1:9, with each physician on-call for 24 h, and a post-call ward round carried out each morning in the AMAU, with other fixed commitments cancelled to accommodate this. Radiology, endoscopy, laboratory services, physiotherapy, occupational therapy, and social services prioritised appropriate requests from the AMAU. All patients identified as suitable for fast track discharge had a provisional discharge date identified on the post-call ward round. Medical teams reviewed these patients early on the morning of discharge, so that discharge could be confirmed, and arrangements made to transfer the patient to the discharge lounge to free up beds for patients awaiting admission in the emergency department. The discharge manager's role was to help identify patients suitable for early discharge, and work with the multidisciplinary team to ensure timeliness of discharge.

A patient database was acquired by linking the patient administration system (PAS) to the Hospital In-Patient Enquiry (HIPE) Scheme. HIPE is a national database of coded discharge summaries from acute public hospitals in Ireland. Sixty hospitals participate in the system and it is an invaluable source of hospital activity level and accreditation. Ireland has used the International Classification of Diseases, ninth revision, Clinical Modification (ICD-9-CM) for both diagnosis and procedure coding since 1990, with updates every 5 years. Linking the HIPE dataset with the PAS dataset permits application of routinely collected data for the purposes of research, planning and quality control. Data collected include: hospital number; patient's name; dates of admission and discharge; date of birth; sex; area of residence by county; diagnosis—principal and up to nine additional secondary diagnoses; procedures—principal and up to nine additional secondary procedures; and consultant responsible for care. The HIPE dataset of all coded diseases at time of discharge/death, together with procedures and investigations undertaken during the in-hospital stay, was examined. Codes with <20 occurrences were not considered for analysis. Individual codes together with the combination of all related codes were evaluated.

Statistical methods

The patient and hospital stay factors by year of admission were compared for the three years 2002–2004, these being before the introduction of AMAU (2002) and the pooled two subsequent years (2003/04). The patient data (table 1) are presented as mean (SD), median (interquartile range (IQR)) or percentages with the appropriate respective between group comparison with *t* tests, non-parametric Wilcoxon rank sum or χ^2 tests. The Charlson comorbidity method was used to compute a weighted index for each patient.¹¹ A higher weighting score (based on 19 diagnostic categories) indicates more comorbid disease. Severity of illness was not recorded.

Three quality indicators were examined: (1) proportion of all admissions with LOS ≤ 30 days; (2) proportion of all admissions with LOS > 90 days (long term care); and (3) proportion of first admissions resulting in a hospital readmission within 28 days.

Funnel plots were used to evaluate how these quality indicators differ between the nine consultant slots. Each

indicator was plotted against the volume of cases, and “control” or prediction limits formed the funnels around the target outcome (in this case, the average proportion over all nine consultant groups). Points outside the control limits (here 95% and 99.8% limits), are 2 (or 3, respectively) standard deviations from the overall target value, and can be interpreted as significantly departing from this value (at $p < 0.05$ or $p < 0.02$, respectively). We have presented the funnel difference datasets for each quality indicator—that is, the change between the two periods for each quality indicator. Significance at $p < 0.05$ is assumed throughout. All analyses were performed using the JMPin (version 5.1, SAS Institute Inc), and SPSS (version 12).

RESULTS

A total of 17 211 episodes were recorded among 11 928 patients admitted acutely via the emergency department in the 36 month study period from 1 January 2002 to 31 December 2004. There was an overall increase of 4.4% (5476 to 5715 episodes) in acute medical episodes presenting to SJH requiring emergency admission from 2002 to 2004. The median age of admissions was 65.6 years (IQR 44.5–77.5); 10% of admissions were over 84 years. Less than half of patients (48.4%) were male. The median LOS was 6 days (IQR 2–13). There was no significant change in mortality rates between the years (12.6% in 2002, 11.7% in 2003, 10.8% in 2004; $p = 0.07$). We compared demographic characteristics for patients admitted between 2002 and 2003/04; while age ($p = 0.87$) did not differ between the years, gender differed with a fall in the proportion of males (–1.7%; $p = 0.03$). The Charlson case-mix index increased significantly over the time period ($p < 0.02$) (table 1).

Impact on short stay patients (≤ 30 days)

The introduction of the AMAU dramatically shortened the inpatient LOS from a median of 7 days in 2002 to 5 days in 2003/04 ($p < 0.0001$). Assuming the case-mix of emergency admissions was similar between teams over an extended period (with a sequential 1:9 rota, not an unreasonable assumption), then fig 1 would indicate a target percentage for short stay patients of just under 90%, with control limits of approximately 87–92%. Four of the teams were outside the 95% confidence interval (CI); the two lower points represent teams with significantly shorter LOS. At the upper limit, it is apparent that one team has an LOS outside the 99.8% CI.

Comparing the combined data for AMAU activity in its initial 2 years of 2003/04, the difference between the two periods is illustrated in the funnel plot (fig 2). The target percentage for short stay patients has increased to in excess of 92%, with control limits of approximately 90–94%. Only one team remained outside the 95% CI, with none outside the 99.8% CI.

Impact on 28 day readmission rate

There was no change in readmission rate over the 3 years 2002–2004, despite a substantially shortened in-hospital episode duration (fig 3). The target percentage for 28 day readmission rate from funnel analysis is just below 6%, with control limits of approximately 3–8%. All “on-call” teams had similar readmission rates with none outside control (that is, 95%) limits.

Impact on long stay patients (> 30 days)

A total of 1485 episodes (8.4%) had an LOS > 30 days. Funnel plots analysis for 2002 would indicate a target percentage for long stay patients of just over 10%, with control limits of approximately 8–12% (fig 4). However, by 2004 these target number were different at just under 8%, with control limits of approximately 6–10%. The difference from 2002 to 2003/04 is illustrated and shows little or slightly increased percentage in patients with LOS > 30 days; the rates became more homogeneous

Table 1 Comparison of patient and hospital stay factors by year of admission

Variable	2002 (n = 5476)	2003–2004 (n = 11984)	p Value
Age*	60.72 (20.5)	60.67 (20.3)	0.87
LOS†	7 (3–14.8)	5 (2–2)	<0.0001
LOS for long term care†	51 (37–81)	48 (37–79)	0.431
Charlson index†	0 (0–1)	0 (0–1)	<0.0001
Males‡ (%)	49.4%	47.7%	0.032
MDC‡			
4	1439 (34.9%)	3164 (36.8%)	0.02
5	922 (22.4)	1896 (22.1)	
1	842 (20.4)	1556 (18.1)	
6	607 (14.7)	1324 (15.4)	
7	314 (7.6)	655 (7.6)	
Long term care‡	10.3%	7.9%	<0.0001

LOS, length of stay ≤ 30 days. Long term care are those with LOS > 30 days.

MDC, major disease category: MDC 4 = respiratory system disorders; MDC 5 = circulatory system disorders; MDC 1 = nervous system disorders; MDC 6 = digestive system disorders; MDC 7 = hepatobiliary system and pancreatic disorders.

*Mean (SD) and t test for comparison.

†Median (interquartile range) and non-parametric Wilcoxon rank sum test for comparison.

‡Percentages and χ^2 test for comparison.

between consultants over the years (within the 95% and 99.8% confidence limits), suggesting that AMAU reduced variability between the “on-call” teams.

DISCUSSION

We describe a novel approach to determine the impact of an AMAU on performance indicator results, using funnel plots, between nine “on-call” medical teams. Despite a significant increase of acute medical episodes with more comorbidities requiring emergency hospitalisation between 2002 and 2004, following the introduction of the AMAU in 2003 the average hospital LOS decreased from 7 days in 2002 to 5 days in 2003/04. Readmission rates over the 3 year period did not change, with all teams displaying expected variability within control limits. Overall, the performance in LOS, both short term and long term, was significantly improved, and was less varied between medical teams between 2002 and 2003/04. There was a trend toward reduced mortality rates over the 3 year period that did not quite achieve significance.

We linked the hospital PAS and HIPE dataset to define a clinically useful database relating to emergency admissions for the years 2002 to 2004. Data were not recorded before 2002. Given the costs associated with such data collection, there is considerable literature supporting its use for research and monitoring purposes.^{12–13} However, the fact that the coding is done with a version of ICD-9-CM up to 5 years old means that advances in medical technology cannot always be captured in sufficient detail. Moreover, complex clinical documentation, inexperienced coding personnel, and illegible handwritten medical record entries all contribute to inaccurate classification.^{14–15} A capture rate of 56% and coding accuracy of 59% has previously been reported.¹⁶ The Department of Health uses HIPE data and hospital financial information from the specialty costing system to measure and compare hospitals' performance. The case-mix directly influences funding given to a hospital, with more efficient hospitals rewarded at the expense of the less efficient. Therefore, improvement in the quality of clinical coding is a desired goal to make such comparisons more meaningful. In our study we found that the HIPE database was very powerful in predicting differences in LOS, re-admissions, and consultant practice. We used a validated method in

Figure 1 Percentage of all patients with length of stay (LOS) ≤ 30 days for 2002 by volume for each “on-call” team ($n=9$). Black graph lines represent upper and lower 95% confidence intervals (CI). Grey graph lines represent upper and lower 99.8% CI.

adjusting for the differences in physicians' LOS and readmissions, with the frequency of comorbid diseases.¹¹ LOS adjusted for comorbidities provides a more accurate measure of LOS than unadjusted LOS, and the presence of comorbidity is significantly associated with longer LOS and hospital costs.^{17–18} Severity of illness was not recorded.

AMAUs provide a focus of clinical care for medical staff rather than have patients spread across several different wards, and facilitate an efficient high quality emergency admissions process with a view to a shorter LOS. We have previously shown that the introduction of an AMAU speeded access to an acute medical service, and reduced costs by reducing LOS.^{6–7} Variations between consultants' practice in the AMAU may reflect a genuine special interest bias, or absence of definitive guidelines for the management of common acute medical conditions.⁹ Differences between consultants' practice can be expected since there is often a lack of consensus or guideline approaching the investigation of a patient with a given

Figure 2 All patients with length of stay (LOS) ≤ 30 days; percentage change between 2002 and 2003/04 by volume. Black graph lines represent upper and lower 95% confidence intervals (CI). Grey graph lines represent upper and lower 99.8% CI.

Figure 3 Percentage change in readmissions within 28 days of first admission between 2002 and 2003/04 by volume. Black graph lines represent upper and lower 95% confidence intervals (CI). Grey graph lines represent upper and lower 99.8% CI.

constellation of symptoms and signs. For those instances in which guidelines do exist, a decision to investigate beyond these may be taken either in response to additional findings, increasing demand for care, fear of litigation, and the urge to make use of new technology.^{19 20} It would be of interest to examine the extent to which protocols and guidelines for the management of common acute medical conditions could reduce such variations.

Health authorities, health service managers, and clinicians involved in clinical governance are responsible for maintaining and improving quality of care. This has been prompted by incidents of failure of professional self-regulation, notably the Bristol and Shipman cases,^{21 22} and has resulted in the collection of comparative data at all levels of healthcare provision. Despite concerns about their usefulness, league tables have been the most commonly used graphical technique for presenting performance data.²³ The Department of Health's intention in publishing league tables is to ensure that "where there are large and unexplained variations in performance, every effort is made to find out why, and work is put in train to bring about an early improvement".²⁴ Such presentations have been criticised as leading to a spurious focus on rank ordering, when it is known that the rank of an institution is one of the most difficult quantities to estimate.^{25 26} Moreover, the league table mode of presentation spuriously identifies outliers when there are none—that is, false positives are generated.

It has been suggested that the use of funnel plots is a more appropriate means of presenting performance indicator results, without spurious ranking into league tables.^{9 10 27} Such plots are not novel, and they are a standard tool within meta-analysis as a graphical check of any relationship between effects estimates and their precision that might suggest publication bias²⁸; they have also been previously used for comparing mortality rates in the health service.^{29 30} Funnel plots plot the observed indicator against a measure of its precision, typically the sample size, and superimposes the target as a horizontal line, and indicates thresholds at which the observed indicator is significantly different from the target. The control limits, essentially corresponding to 2 and 3 standard deviations, form a "funnel" around the target outcome, and outliers stand out as having divergent performance, while the remainder display expectant variability.

When the performance of clinical units is compared using funnel plots, variation may be attributable to either "common

Figure 4 All patients with length of stay (LOS) >30 days; percentage change between 2002 and 2003/04 by volume. Black graph lines represent upper and lower 95% confidence intervals (CI). Grey graph lines represent upper and lower 99.8% CI.

cause variation" or "special cause variation". Units display "common cause variation" when their performance falls within the limit lines of the funnel plot, indicating that their performance varied only by an amount consistent with random chance. Common cause variation, because it is linked to chance, is greatest when numbers of patients are small and reduces as numbers of patients per unit increases. Units display "special cause variation" when their performance falls beyond the limit lines of the funnel plot. This type of variation is caused by external factors acting on a process, and is a signal that something unusual is affecting performance. Funnel plots permit units to appraise their local practices, and just as areas for improvement can be identified in some units, unit processes that have contributed to high attainment of national targets can be used as examples of "good practice" that might be reproduced in other hospitals. A limitation of funnel plots, like all performance measurement and analysis methods, is that it may produce both false negative and positive results. Performing within limits does not guarantee that a unit may not be under performing, though this may either be too slight to detect or masked by other factors. Moreover, being outside a control limit is not always abnormal either, as a few may be outside the limits by chance alone, but it may be useful to investigate reasons for the apparent discordance.

In our study we found that with the introduction of an AMAU, the performance in LOS, both short term and long term, was significantly improved and less varied between medical teams, as assessed by funnel plots, with no change in readmission rates. We found funnel plots are a powerful graphical technique for presenting quality performance indicator variation between teams over time. Credibility using data to raise local awareness among health care staff of the need for improvements in processes is vital when there is under achievement in clinical performance. The routine publication of funnel plots summarising performance attainments will encourage the understanding, reflection, processes analysis and subsequent improvements in healthcare at all levels of policy, planning and actual delivery.

ACKNOWLEDGEMENTS

We wish to recognise the contribution of our consultant medical colleagues and the non-consultant members of the "on-call" teams without which this initiative could not have progressed. In particular, their continued enthusiasm for the practice of acute medicine in

difficult circumstances within stringent resource constraints deserves recognition. The dedicated contribution of Sister S Donnelly, her clinical nurse managers and the ancillary professions related to medicine (SCOPE) is gratefully acknowledged.

Authors' affiliations

Edward D Moloney, Bernard Silke, Division of Internal Medicine, St James' Hospital, Dublin, Ireland

Kathleen Bennett, Department of Therapeutics and Pharmacology, Trinity Centre at St James' Hospital, Dublin, Ireland

Competing interests: None declared

REFERENCES

- 1 Capewell S. The continuing rise in emergency admissions. *BMJ* 1996;**312**:991–2.
- 2 Blatchford O, Capewell S. Emergency medical admissions: taking stock and planning for winter. *BMJ* 1997;**315**:1322–3.
- 3 Audit Commission. *Lying in wait: the use of medical beds in acute hospitals*, Honey Press Ltd, 1992.
- 4 Bagust A, place M, Posnett JW. Dynamics of bed use in accommodating emergency admissions: stochastic simulation model. *BMJ* 1999;**319**:155–8.
- 5 Cooke M. Whole system is responsible for solving overcrowding of departments. *BMJ* 2002;**325**:389.
- 6 Moloney ED, Smith D, Bennett K, et al. Impact of an acute medical admission unit on length of hospital stay, and emergency department wait times. *QJM* 2005;**98**:283–9.
- 7 Moloney ED, Bennett K, Silke B. Factors influencing the costs of emergency medical admissions to an Irish teaching hospital. *Eur J Health Econ* 2006;**7**:123–8.
- 8 Department of Health. *NHS performance indicators*. The Stationery Office: London, 2002, <http://www.performance.doh.gov.uk/nhsperformanceindicators/2002/trust.htm> (accessed February 2004).
- 9 Spiegelhalter D. Funnel plots for institutional comparison. *Qual Saf Health Care* 2002;**11**:390–1.
- 10 Spiegelhalter D. Funnel plots for comparing institutional performance. *Statist Med* 2005;**24**:1185–202.
- 11 Charlson ME, Pompei P, Ales KL, et al. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis* 1987;**40**:373–83.
- 12 Bain MR, Chambers JW, Brewster DH. Routinely collected data in national and regional databases – an underused resource. *J Public Health Med* 1997;**19**:413–8.
- 13 Berkman RL, Buehler JW. Public health surveillance of non-infectious chronic diseases: the potential to detect rapid changes in disease burden. *Int J Epidemiol* 1990;**19**:628–35.
- 14 MacIntyre CR, Ackland MJ, Chandraraj EJ, et al. Accuracy of ICD-9-CM codes in hospital morbidity data, Victoria: implications for public health research. *Aust N Z J Public Health* 1997;**21**:477–82.
- 15 O'Neill B. The hospital in-patient enquiry scheme – a study of data accuracy. *Ir Med J* 1982;**75**:238–9.
- 16 Mehanni M, Loughman E, Allwright SP, et al. The hospital in-patient enquiry scheme: a study of data accuracy and capture. *Ir Med J* 1995;**88**:24–6.
- 17 Kinnunen T, Saynajakangas O, Tuuponen T, et al. Impact of co-morbidities on the duration of COPD patients' hospital episodes. *Respir Med* 2003;**97**:143–6.
- 18 Matsui K, Goldman L, Johnson PA, et al. Co-morbidity as a correlate of length of stay for hospitalised patients with acute chest pain. *J Gen Intern Med* 1996;**11**:262–8.
- 19 Winkens R, Dinant GJ. Rational cost effective use of investigations in clinical practice. *BMJ* 2002;**324**:783–5.
- 20 Marik PE, Rakusin A, Sandhu SS. The impact of the accessibility of cranial CT scans on patient evaluation and management decisions. *J Intern Med* 1997;**241**:237–43.
- 21 Department of Health. *Harold Shipman's clinical practice, 1974–1998*. London: Stationary Office, 2001.
- 22 Anon. *Inquiry K. Report of the public enquiry into children's heart surgery at the Bristol Royal Infirmary*. London: The Stationery Office, 2001.
- 23 Adab P, Rouse AM, Mohammed MA, et al. Performance league tables: the NHS deserves better. *BMJ* 2002;**324**:95–8.
- 24 NHS Executive. *HSC 2000/023. Quality and performance in the NHS. Performance indicators: July 2000*. Leeds: Department of Health, 2001.
- 25 Goldstein H, Spiegelhalter DJ. Statistical aspects of institutional performance: league tables and their limitations. *Journal of the Royal Statistical Society, Series A* 1996;**159**:385–444.
- 26 Marshall EC, Spiegelhalter DJ. Reliability of league tables of in vitro fertilisation clinics: retrospective analysis of live birth rates. *BMJ* 1998;**317**:1701–4.
- 27 Marshall T, Mohammed A, Rouse A. A randomised controlled trial of league tables and control charts as aids to health service decision making. *Int J Qual Health Care* 2004;**16**:309–15.
- 28 Egger M, Davey Smith G, Schneider M, et al. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 1997;**315**:629–34.
- 29 Spiegelhalter DJ. An investigation into the relationship between mortality and volume of cases: an example in paediatric cardiac surgery between 1991 to 1995. *BMJ* 2002;**324**:261–3.
- 30 Tekkis PP, McCulloch P, Steger AC, et al. Mortality control charts for comparing performance of surgical units: validation study using hospital mortality data. *BMJ* 2003;**326**:786–8.