

Risk-adjusted continuous outcome monitoring with an EWMA chart: could it have detected excess mortality among intensive care patients at Bundaberg Base Hospital?

David V Pilcher, Toni Hoffman, Chris Thomas,
David Ernest and Graeme K Hart

Comparative monitoring of outcomes and performance is well established in intensive care medicine. Mortality prediction models such as Acute Physiology and Chronic Health Evaluation (APACHE) II and APACHE III were developed nearly 20 years ago.^{1,2} APACHE III-j (the 10th version of APACHE III) was released into the public domain in 2002. Models such as these are typically used to “risk adjust” observed mortality and to calculate an annual standardised mortality ratio (SMR). The SMR has formed the basis of reports on intensive care in Australia and New Zealand since the mid-1990s.

Continuous monitoring techniques were used as long ago as World War II³ to monitor quality control of military supplies. More recently, they have been applied to medical processes. Their advantage is that they can rapidly detect deteriorations in outcome, allowing the process of care to be inspected. In addition, they can assess the timing and impact of interventions. However, they are yet to be widely used for routine performance monitoring.

Examples of continuous monitoring techniques include risk-adjusted cumulative sum (CUSUM) and risk-adjusted sequential probability ratio tests; these have been used retrospectively to investigate excess deaths associated with cardiothoracic transplantation,⁴ paediatric cardiac surgery and the practice of British general practitioner Harold Shipman.³ Their uptake into routine clinical practice in Australia has been slow, in part because of the perceived difficulty of interpretation by clinicians. However, Queensland Health has introduced continuous outcome monitoring with variable life-adjusted displays, using hospital administrative data.⁵

Since 2007, the Australian and New Zealand Intensive Care Society (ANZICS) Centre for Outcomes and Resource Evaluation (CORE) has been assessing methods of continuous monitoring of intensive care practice in Australia and New Zealand. The exponentially weighted moving average (EWMA) chart, generally thought more “understandable” to a clinical audience than CUSUM charts, has been chosen for inclusion in standard reports to ICUs from ANZICS CORE. Here, we illustrate the application, advantages and limitations of an EWMA chart constructed with data submitted retrospectively from the ICU at Bundaberg Base Hospital, Queensland.

In 2005, Bundaberg Base Hospital received considerable publicity because of perceived excess mortality allegedly

ABSTRACT

Objective: To test whether applying a continuous risk-adjusted charting method, using an exponentially weighted moving average (EWMA) chart, would have been useful for monitoring outcomes of patients admitted to the intensive care unit at Bundaberg Base Hospital, Queensland, between November 2000 and December 2005.

Design, setting and participants: An EWMA chart was constructed to show the change in observed compared with predicted mortality over time, using data submitted to the Australian and New Zealand Intensive Care Society Adult Patient Database. Limitations and practical implications of this monitoring technique were evaluated and compared with a routine monitoring technique using the annual standardised mortality ratio.

Main outcome measure: In-hospital mortality.

Results: Data were submitted on three occasions (August 2002, November 2002 and February 2006). In each year before 2005, the EWMA chart showed periods when observed mortality appeared higher than predicted. These periods were not detectable by analysing the data with an annual standardised mortality ratio. Comparison of aggregated data from peer group hospitals suggested that the mortality prediction model (APACHE III-j) was an appropriate risk adjustment model for this analysis.

Conclusions: Continuous monitoring of outcomes using an EWMA chart may have advantages over other techniques. Had data been available, analysis with an EWMA chart might have prompted review of processes and outcomes among patients at Bundaberg Base Hospital ICU.

Crit Care Resusc 2010; 12: 36–41

related to a general surgeon who worked at the hospital between April 2003 and March 2005. Because of the interest in patient outcomes at this ICU, its data were an ideal example for testing retrospectively whether analysis with an EWMA chart would have provided useful, relevant information to clinicians at the hospital. However, it is important to note we had no intention of investigating the performance

Table 1. Annual numbers of patients admitted to the intensive care unit at Bundaberg Base Hospital, 2000–2005, numbers analysed and exclusions

	2000 (Nov–Dec)	2001	2002	2003	2004	2005	Total
No. of patients admitted	39	320	252	191	211	278	1291
No. of patients analysed	26	139	126	106	111	177	685
Mortality	3.8%	10.8%	15.9%	23.6%	14.4%	11.3%	14.1%
APACHE III-j risk of death (SD)	2.1% (10.9%)	9.1% (17.6%)	11.1% (19.5%)	19.7% (26.2%)	15.2% (22.2%)	12.3% (22.7%)	12.7% (21.3%)
Exclusions							
Transfer to another ICU*	1	42	42	35	28	52	200
Age < 16 years or ICU length of stay < 4 h [†]	5	51	30	24	30	31	171
No APACHE III-j risk of death or no hospital outcome recorded	7	88	54	26	42	18	235

* Excluded because eventual outcome was unknown. † Routine exclusions in keeping with the APACHE III-j algorithm.

of any individual at the hospital, nor would it be possible to do so with this analysis of overall ICU outcomes.

Methods

Data were extracted from the ANZICS Adult Patient Database, a high-quality binational database⁶ listing over 800 000 de-identified ICU admissions. This database has previously been used to develop and refine performance indicators for intensive care practice.⁷

All first admissions to the ICU between November 2000 (when data collection at the ICU began) and December 2005 inclusive were included. Hospital mortality and predicted risk of death derived from APACHE III-j were extracted for each admission. Patients with no recorded outcome or no available APACHE III-j risk of death were excluded. Patients transferred from Bundaberg Base Hospital ICU to another ICU were also excluded because they were considered to have an unknown mortality outcome. In accordance with the APACHE III-j algorithm, patients under 16 years of age and those with an ICU stay less than 4 hours were also excluded. Data were analysed both by calculation of annual SMRs and by construction of an EWMA chart.

The analytical principles used are described briefly below. The exact formulas for calculation and construction of these indices are detailed elsewhere.⁸

Standardised mortality ratio

The total number of observed deaths was divided by the total number of predicted deaths derived from APACHE III-j. The 95% confidence intervals were calculated (determined by the number of observed deaths, not predicted deaths). An SMR greater than 1.0 indicated

observed mortality was higher than predicted by APACHE III-j.

Continuous charting using EWMA analysis

Each admission was plotted sequentially along the horizontal axis of the EWMA. Observed mortality was plotted on the vertical axis as a running mean of the previous observations, where the most recent observations were given exponentially more weight than historically distant observations. A weighting of 0.005 was chosen in keeping with construction of standard EWMA charts by ANZICS CORE. Control limits were derived from a similarly weighted running mean of the APACHE III-j predicted risk of death (using 95% and 99% CIs). On the vertical axis, an arbitrary starting value of 16% was selected, as this was the overall mortality rate for the pooled cohort of five peer-group hospitals (see *Validation*). When the line for observed mortality rose above the control limits for predicted risk of death, observed mortality was interpreted to be significantly worse than predicted.

Validation of APACHE III-j as a risk adjustment model for this analysis

To validate APACHE III-j as an appropriate model for this analysis, aggregated data from five rural Queensland hospitals (including Bundaberg Base Hospital) were extracted for the same time period. The same exclusion criteria were applied to these data as to the data for the Bundaberg Base Hospital analysis. The identities of the other four hospitals were not known to the investigators. Model calibration was assessed using the SMR and Hosmer–Lemeshow C-statistic. An SMR close to 1.0 and a small C-statistic with a high *P* value indicated good calibration. Model discrimination was

assessed by estimating the area under the receiver operator curve, where an area greater than 0.8 indicated good discrimination. Data were analysed using Intercooled Stata version 9.2 for Windows, and figures were constructed using Microsoft Windows Excel 2007.

Input from staff at Bundaberg Base Hospital

The study was undertaken with the support of administrative and clinical staff at the Bundaberg Base Hospital ICU, who provided historical perspective and clinical context for the analysis.

Figure 1. Annual standardised mortality ratio for Bundaberg Base Hospital intensive care unit for 2000–2005 (with 95% CIs)

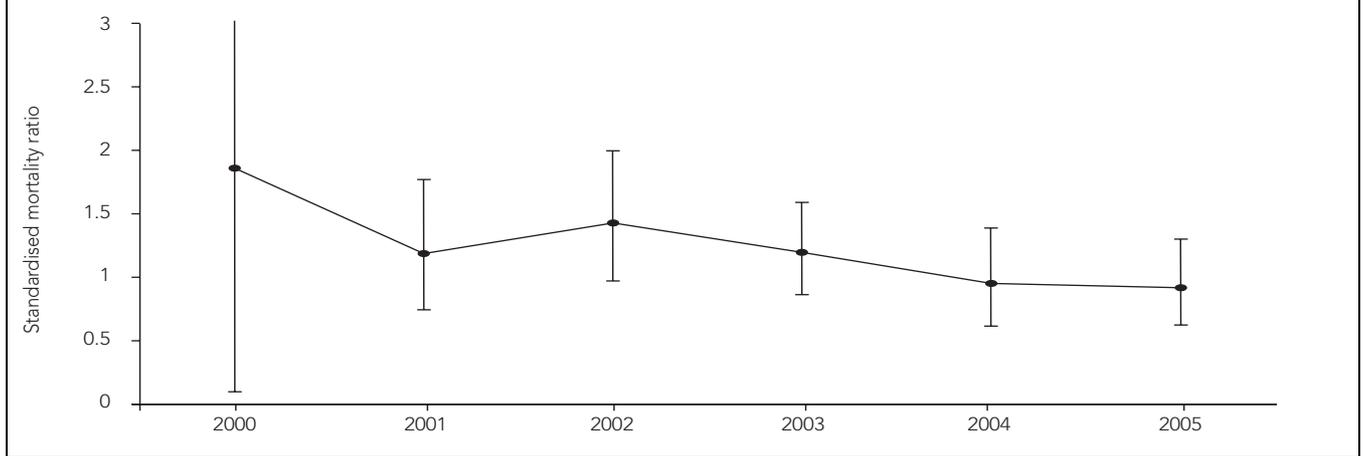


Figure 2. Exponentially weighted moving average (EWMA) chart for Bundaberg Base Hospital intensive care unit for 2000–2005

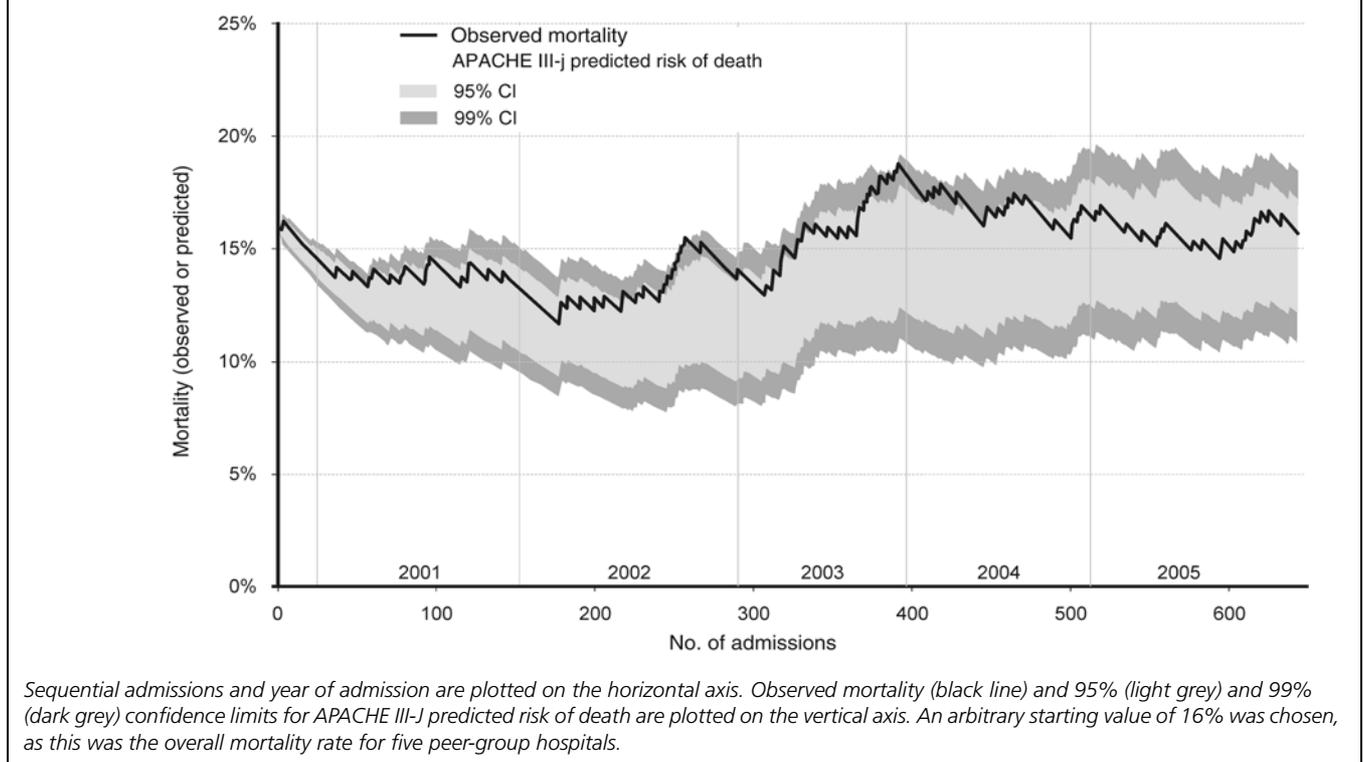


Table 2. Performance of APACHE III-J as a mortality prediction model for five rural Queensland ICUs*

Characteristic	Value
Validation data for APACHE III-j	
Time period	Jan 2000 – Dec 2005
No. of patients admitted to ICU	5889
Observed mortality	16.3%
APACHE III-J predicted risk of death (SD)	16.4% (24.3%)
Standardised mortality ratio (95% CI)	0.99 (0.95–1.04)
Discrimination and calibration statistics	
Area under receiver operator curve (95% CI)	0.90 (0.89–0.91)
Hosmer–Lemeshow C-statistic	13.1
<i>P</i>	0.22

* Including Bundaberg Base Hospital.

Results

There were 1291 patients admitted to the ICU at Bundaberg Base Hospital between November 2000 and December 2005. Data were submitted to ANZICS on three occasions during this period (August 2002, November 2002 and February 2006). Of the cohort, 235 admissions were excluded because of inadequate data quality (no APACHE III-j predicted risk of death or no recorded hospital outcome). A further 371 patients were excluded either in accordance with the APACHE III-j algorithm or because they were transferred to another ICU. This left 685 ICU admissions for analysis (Table 1).

The annual SMR for Bundaberg ICU is shown in Figure 1. In 2000, the confidence intervals were wide because of the small number of ICU admissions recorded (2 month's data only). In each year, the 95% CI included 1, suggesting there was no year when observed outcomes were significantly worse than predicted. There appeared to be a gradual reduction (improvement) in SMR from 2002 onwards.

The EWMA chart analysis is shown in Figure 2. Observed mortality appeared to reach the upper inner control limit on a number of occasions in 2001. Outcomes appeared to remain within acceptable limits between late 2001 and mid-2002. In the latter half of 2002, observed mortality surpassed the inner control limit and breached the outer control limit, again suggesting that outcomes were worse than predicted by APACHE III-j. A similar pattern was seen in late 2003, when observed mortality reached the outer control limit. From late 2004, one can be confident that observed outcomes were within acceptable limits (observed mortality lay within all control limits for APACHE III-j predicted risk of death).

Table 2 shows the analysis of aggregated data from five rural (peer group) Queensland hospitals for the years 2000–2005 (inclusive), used to validate APACHE III-j as a mortality prediction model. APACHE III-j provided an accurate estimation of mortality for this cohort of nearly 6000 admissions. It was a well calibrated and discriminatory prediction tool. The SMR for aggregated data from these five hospitals was very close to 1, indicating that observed mortality overall was almost identical to that predicted by APACHE III-j.

Discussion

Our analysis demonstrated the advantages of the EWMA chart over an SMR chart in detecting variation in mortality outcomes over time.

Initial inspection of annual SMRs appeared to show a consistent improvement in outcome from 2002 onwards. However, as confidence intervals were wide, the true SMR might lie anywhere within this range. The annual SMR also gave no indication of variation in mortality during the year. Furthermore, as it is a ratio, the SMR gave no indication of actual mortality or severity of illness.

In contrast to the SMR, the EWMA chart tracked fluctuations and allowed direct visual inspection of observed compared with predicted mortality. The EWMA chart continuously “built on” previous data, weighting recent observations more highly than distant ones. As seen in this case, an EWMA chart can be started at any time in the year, whenever recording of outcomes begins. Data from the first year analysed (in this case 2000) can be included and interpreted appropriately by the EWMA chart in the light of subsequent years. However, the SMR for 2000 provided little information because of the small numbers. The EWMA chart might thus be particularly useful for ICUs with a small number of admissions each year, which would have wide confidence intervals around an annual SMR.

Between 2001 and 2004, the EWMA chart suggested various instances when mortality appeared higher than predicted. There was limited continuity of senior staff in the ICU over this time, but whether this contributed to the observed outcomes is not known. From late 2004, the EWMA chart indicated that mortality was within acceptable limits. This marked a time when ICU staff decided to transfer more critically ill patients to larger hospitals with more specialised services. The outcomes of transferred patients were not known to the investigators, they were not included in this analysis, and it is not known whether their transfers contributed to the improved outcomes seen in 2005.

Monitoring is possible only when there are data to be analysed. There was a considerable gap in data submission

from Bundaberg ICU between November 2002 and February 2006. The Australian Council for Healthcare Standards has designated that data submission to ANZICS is a key performance indicator for intensive care practice. Inability to submit data may in itself be a signal of a problem. Smaller regional hospitals may have limited resources for data collection, data submission and audit of clinical outcomes. This was indeed the case at Bundaberg Base Hospital in the early 2000s. It is not known whether the perception or anticipation of poor performance was a disincentive to data submission.

On several occasions during the period investigated, ICU staff wished to assess whether outcomes were below expected, but did not have analysed data available. Had there been timely submission of data, with analysis and feedback of results through an appropriate governance structure (as now exists in Queensland), we speculate that EWMA analysis could have led to an inspection of processes at the hospital, and possibly early detection of deteriorating outcomes. At the minimum, such an analysis would have prompted review of data quality, highlighting the lack of resources for data collection and submission.

When monitoring clinical outcomes, it is vital to avoid over-interpretation of results and erroneous judgements about quality of care.⁹ There are significant caveats and limitations to our analysis. Firstly, this was an analysis of a cohort of ICU patients. EWMA charts constructed in this way are appropriate for monitoring the risk-adjusted outcomes of ICUs. No inference can be drawn as to whether the presence of any individual at the hospital directly influenced the outcomes of patients in the ICU. In addition, the APACHE III-j predicted risk of death was calculated using data collected in the first 24 hours after ICU admission. Thus, this analysis was not applicable to the assessment of practices before ICU admission.

Secondly, although all data underwent extensive checking and “cleaning” before analysis by ANZICS, the impact of data quality on this, and indeed any, analysis must be emphasised. There were many exclusions, with 47% of admissions to Bundaberg ICU not included in this analysis. Some of these were standard exclusions that are routinely applied to data when using APACHE III-j as the risk adjustor and would have been applied to data from any ICU. However, a considerable proportion were excluded because of “poor data quality” — no recorded hospital outcomes or no APACHE III-j predicted risk of death. The effect of these missing data on analysis and interpretation could not be determined. However, notably the proportion of ICU admissions without recorded hospital outcomes or APACHE III-j predicted risk of death was lowest in 2005, when outcomes were best. Whether this was due to general improvement in processes, which in turn led to both improved data collec-

tion and improved patient outcomes, or whether poor data quality led to an erroneous impression of poor outcomes in earlier years could not be determined. Despite the exclusions there appears to be a “signal” of periods when outcomes were worse than expected, highlighting a potential benefit of this type of analysis.

APACHE III-j appeared to be an appropriate risk adjustment tool. However, the performance of any risk adjustment model itself may be affected by many factors, including changes in observed mortality rate,¹⁰ patient casemix,¹¹ admission lead time,¹² discharge practices,¹³ data collection,^{14,15} rule interpretations and transcription errors.¹⁶ The contribution of each of these factors to the calculation of predicted risk of death for patients at the Bundaberg ICU was unknown.

Limitations to the introduction of EWMA charts into routine clinical practice throughout Australia require mention. In recent years, outcomes for patients admitted to Australian ICUs are consistently better than those predicted by APACHE III-j.¹⁷ Thus, deterioration in outcome may be more difficult to detect. A more accurate, appropriate and contemporary mortality prediction model may need to be developed.

Conclusions

With appropriate risk adjustment and interpretation, the EWMA chart can provide meaningful and important clinical results. It can detect fluctuations in outcome that are not apparent with an SMR. Had this analysis been available in 2000–2005, it might have prompted review of processes and outcomes among ICU patients at the Bundaberg Base Hospital. As part of future reports from the ANZICS CORE, techniques such as EWMA chart analysis may be beneficial for routine monitoring of intensive care practice in Australia and New Zealand. Hospitals and ICUs should be provided with adequate resources for the accurate and timely collection of these important quality assurance data.

Acknowledgements

Initial analysis of these data was presented, with the permission of administrative and clinical staff of Bundaberg Base Hospital ICU, at the ANZICS Annual Scientific Meeting in Rotorua, New Zealand, in 2007.

Author details

David V Pilcher, Director, ANZICS Adult Patient Database,¹ and Intensivist²

Toni Hoffman, Nurse Unit Manager³

Chris Thomas, Intensivist³

David Ernest, Intensivist⁴

Graeme K Hart, Chair,¹ and Intensivist⁵

- 1 Australian and New Zealand Intensive Care Society Centre for Outcome and Resource Evaluation.
 - 2 Department of Intensive Care, Alfred Hospital, Melbourne, VIC.
 - 3 Department of Intensive Care, Bundaberg Base Hospital, Bundaberg, QLD.
 - 4 Department of Intensive Care, Box Hill Hospital, Box Hill, VIC.
 - 5 Department of Intensive Care, Austin Hospital, Melbourne, VIC.
- Correspondence:** d.pilcher@alfred.org.au

References

- 1 Knaus W, Draper E, Wagner D, Zimmerman J. APACHE II: a severity of disease classification system. *Crit Care Med* 1985; 13: 818-29.
- 2 Knaus W, Wagner D. APACHE III study design: analytic plan for evaluation of severity and outcome in intensive care unit patients. Individual patient decisions. *Crit Care Med* 1989; 17 (12 Pt 2): S204-9.
- 3 Spiegelhalter D, Grigg O, Kinsman R, Treasure T. Risk-adjusted sequential probability ratio tests: applications to Bristol, Shipman and adult cardiac surgery. *Int J Qual Health Care* 2003; 15: 7-13.
- 4 Rogers CA, Ganesh JS, Banner NR, Bonser RS. Cumulative risk adjusted monitoring of 30-day mortality after cardiothoracic transplantation: UK experience. *Eur J Cardiothorac Surg* 2005; 27: 1022-9.
- 5 Duckett SJ, Coory M, Sketcher-Baker K. Identifying variations in quality of care in Queensland hospitals. *Med J Aust* 2007; 187: 571-5.
- 6 Stow PJ, Hart GK, Higlett T, et al. Development and implementation of a high-quality clinical database: the Australian and New Zealand Intensive Care Society Adult Patient Database. *J Crit Care* 2006; 21: 133-41.
- 7 Pilcher DV, Duke GJ, George C, et al. After-hours discharge from intensive care increases the risk of readmission and death. *Anaesth Intensive Care* 2007; 35: 477-85.
- 8 Cook DA, Duke G, Hart GK, et al. Review of the application of risk-adjusted charts to analyse mortality outcomes in critical care. *Crit Care Resusc* 2008; 10: 239-51.
- 9 Lilford R, Mohammed MA, Spiegelhalter D, Thomson R. Use and misuse of process and outcome data in managing performance of acute medical care: avoiding institutional stigma. *Lancet* 2004; 363: 1147-54.
- 10 Zhu BP, Lemeshow S, Hosmer DW, et al. Factors affecting the performance of the models in the Mortality Probability Model II system and strategies of customization: a simulation study. *Crit Care Med* 1996; 24: 57-63.
- 11 Murphy-Filkins R, Teres D, Lemeshow S, Hosmer DW. Effect of changing patient mix on the performance of an intensive care unit severity-of-illness model: how to distinguish a general from a specialty intensive care unit. *Crit Care Med* 1996; 24: 1968-73.
- 12 Dragsted L, Jorgensen J, Jensen NH, et al. Interhospital comparisons of patient outcome from intensive care: importance of lead-time bias. *Crit Care Med* 1989; 17: 418-22.
- 13 Cooper GS, Sirio CA, Rotondi AJ, et al. Are readmissions to the intensive care unit a useful measure of hospital performance? *Med Care* 1999; 37: 399-408.
- 14 Lawrance RA, Dorsch MF, Sapsford RJ, et al. Use of cumulative mortality data in patients with acute myocardial infarction for early detection of variation in clinical practice: observational study. *BMJ* 2001; 323: 324-7.
- 15 Chen LM, Martin CM, Morrison TL, Sibbald WJ. Interobserver variability in data collection of the APACHE II score in teaching and community hospitals. *Crit Care Med* 1999; 27: 1999-2004.
- 16 Fery-Lemonnier E, Landais P, Loirat P, et al. Evaluation of severity scoring systems in ICUs — translation, conversion and definition ambiguities as a source of inter-observer variability in APACHE II, SAPS and OSF. *Intensive Care Med* 1995; 21: 356-60.
- 17 Victorian Data Review Committee. Monitoring adult intensive care annual report 2004-5. Melbourne: Victorian Department of Human Services, 2007. □