Identification and assessment of potentially high-mortality intensive care units using the ANZICS Centre for Outcome and Resource Evaluation clinical registry

Kerry McClean, Daniel Mullany, Sue Huckson, Allison van Lint, Shaila Chavan, Peter Hicks, Graeme Hart, Eldho Paul and David Pilcher

ABSTRACT

Purpose: A hospital's highest-risk patients are managed in the intensive care unit. Outcomes are determined by patients' severity of illness, existing comorbidities and by processes of care delivered. The Australian and New Zealand Intensive Care Society (ANZICS) Centre for Outcome and Resource Evaluation (CORE) manages a binational clinical registry to benchmark performance, and report and assess ICUs which appear to have worse outcomes than others.

Methods: A descriptive retrospective cohort study was undertaken to detail processes, outcomes, limitations and practical lessons learnt from monitoring ICU performance throughout Australia and New Zealand. All ICUs contributing to the ANZICS Adult Patient Database between 2009 and 2014 were included. A potential outlier ICU was defined as one with a statistically significantly higher standardised mortality ratio (SMR) than its peer group.

Results: There were 757 188 admissions to 168 ICUs. Of these, 27 ICUs (16%) were identified as potential outlier ICUs at least once. Data quality problems led to inaccurate or artificially elevated SMRs at 16/27 ICUs. Variation in diagnostic casemix partly or completely explained the elevated SMR at 15/27 ICUs. At nine ICUs where data quality and casemix differences did not explain the elevated SMR, process-of-care problems were identified.

Conclusions: A combination of routine monitoring techniques, statistical analysis and contextual interpretation of findings is required to ensure potential outlier ICUs are appropriately identified. This ensures engagement and understanding from clinicians and jurisdictional health departments, while contributing to the improvement of ICU practices throughout Australia and New Zealand.

Crit Care Resusc 2017; 19: 230-238

Transparency of clinical performance has become increasingly important in the modern health care system.[1] Highly publicised adverse events could have been detected earlier if rigorous institutional or provider outcome monitoring had been in place.[2-4]

Intensive care units manage the hospital's most critically unwell patients, those with an increased risk of death and/or complications.[5] It is a high-cost service that admits patients across a range of specialties. Large volumes of clinical data are routinely collected and used in benchmarking mortality outcomes.[6] Outcomes after ICU admission are predominantly determined by the initial severity of illness and existing comorbidities,[7] but can also be affected by organisational factors and processes of care within each hospital.[8] They reflect the quality of care in surgical and other services which frequently admit patients to the ICU.[9] The combination of high-risk patients, high cost, broad casemix and the availability of clinical data for statistical risk adjustment make the ICU an ideal service on which to focus outcomes measurement.[10]

Variations in ICU outcomes that are not explained by severity of illness have previously been reported.[11-13] To improve outcomes, it is important to highlight and understand underlying causes within a framework of the quality improvement cycle,[14-15] as punitive approaches to performance measurement may lead to gaming, risk-averse practices or non-participation.[16-17]

The importance of national clinical quality registries is recognised globally.[18] The Australian and New Zealand Intensive Care Society (ANZICS) Centre for Outcome and Resource Evaluation (CORE) is a binational clinical quality registry organisation which benchmarks adult ICU performance. When a potential outlier ICU is identified, a structured program of notification and analysis is undertaken. Our aim is to describe this process, known as the ANZICS CORE outlier management program, and to highlight operational aspects, outcomes, limitations and lessons learnt from 2009 to 2014, using real examples.

Methods

Setting

We analysed de-identified data from adult ICUs between 2009 and 2014, submitted to the ANZICS Adult Patient Database, one of four registries run by ANZICS CORE. The Adult Patient

Database contains over 1.8 million patient records from 188 ICUs across Australia and New Zealand. ICUs were also surveyed annually to determine staffing, resources and processes of care within each ICU through the ANZICS CORE Critical Care Resources Registry.

The outlier management program

The ANZICS CORE outlier management program was established in 2009. The process undertaken for this analysis is current and ongoing, and is based on the Mohammed pyramid, adapted to the intensive care model.[19] It involves a stepwise process of analysis, where each subsequent step is undertaken only if the previous one does not adequately explain the findings.

1. Initial identification of a potential outlier

An ICU whose standardised mortality ratio (SMR) is above the 99% confidence intervals on a comparative funnel plot of peer group units is considered a potential outlier. The SMR is calculated as the ratio of observed to predicted number of in-hospital deaths. The risk of death is calculated using the ANZICS CORE adaptation of the Acute Physiology and Chronic Health Evaluation (APACHE) III-J model. An ICU with a lower SMR than their peer group would not undergo investigation. A summary of methods, limitations and controversies in the construction, interpretation and ongoing use of funnel plots and exponentially weighted moving average charts is available.[12,20-26]

Once an ICU is identified as a potential outlier, the relevant ICU director, jurisdictional governance body or health department is notified at the same time as confidential data analysis begins and is reported within 21 working days. The detailed report to the ICU director follows the analysis description here.

2. Assessment of data quality

Data accuracy and reliability are investigated by estimates of data completeness, analysis of inconsistencies in submitted variables which fail automatic validation rules, and by the findings of onsite data quality audits. Data quality reviews also involve examination of overall admission numbers submitted over longer periods, typically 3 to 5 years. A sample data quality audit is shown in Appendix eTable 1 (online at cicm.org.au/Resources/Publications/Journal).

3. Assessment of consistency of casemix and subgroups

Subgroups are examined on the basis of diagnostic casemix, severity of illness, chronic health comorbidities, age, elective surgical status, nursing dependency, ventilation status and temporal and seasonal variation. Casemix is defined by the APACHE III-J diagnostic codes (as used by ANZICS CORE and defined in the data dictionary). The first aim was to identify the presence and impact of subgroups where mortality predictions might be inaccurate (eg, APACHE III-J “under-predicted” mortality for cervical spine injuries, so large numbers might lead to a “false elevation” of the SMR). The second aim was to identify subgroups which appear to have worse mortality outcomes than others even after appropriate adjustment for severity of illness (ie, a “true elevation” in the SMR).

4. Assessment of processes and resources at the outlier hospital

When data quality and casemix cannot explain a unit’s outlier status, the reported processes of care within the unit are examined, such as provision of prophylaxis for venous thromboembolism, access to critical care services (eg, occupancy, refusals and after-hours discharge) and medical and nursing staff levels. Onsite visits to hospitals to assess processes of care are not performed by ANZICS CORE staff.

5. Action after the outlier review process

A detailed report is compiled and provided simultaneously to the ICU director, the hospital chief executive officer and the local jurisdictional governance review body. When no reason can be identified for elevation of the SMR, an onsite audit of data quality is offered. If data quality issues are found to be the predominant factor accounting for the potential outlier status, ANZICS CORE offers further training to data collectors, with resubmission of updated data. In all other cases, actions are determined by the local jurisdictional governance group and involve monitoring only, further analysis or onsite investigation of hospital processes. Jurisdictional liaison committees, ICU directors and jurisdictional health departments are responsible for confirmation and remediation of any identified performance issues.

Results

Overall number of reports produced

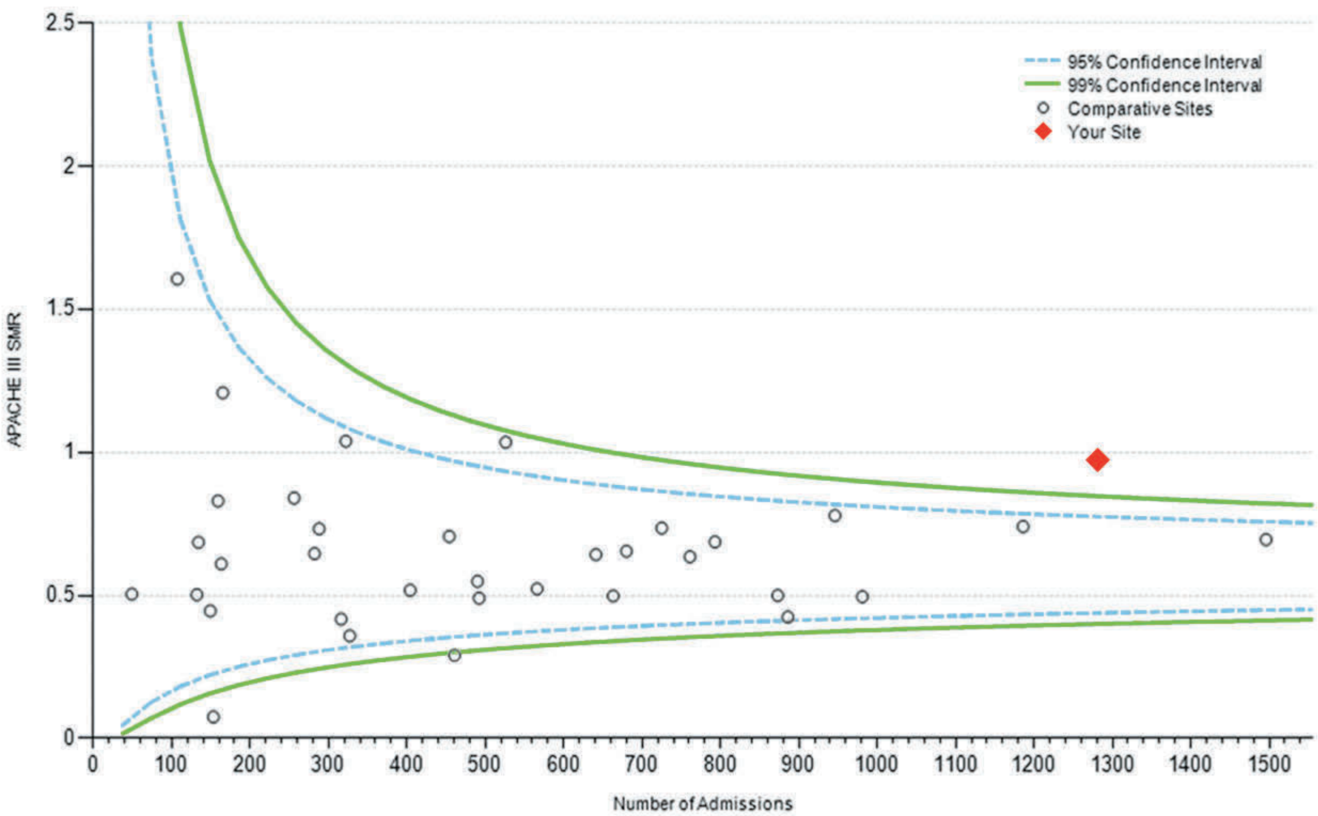
Between 2009 and 2014, 168 ICUs submitted 757 188 admissions to the Adult Patient Database (Table 1). The overall APACHE III-J SMR for these ICUs declined from 0.759 (95% CI, 0.745–0.774) in 2009 to 0.694 (95% CI, 0.682–0.706; $P < 0.001$) in 2014. Contributing units were provided with a benchmarking report every quarter. A total of 3446 individual benchmarking reports were produced and analysed, with SMR funnel plots used to identify potential outliers. Of these, just over 1% led to investigations of potential outliers, with 37 outlier analyses performed in 27 different ICUs. Two additional analyses were produced at the request of local jurisdictional bodies, in the absence of a high SMR, and are not included in this article. Figure 1 shows

Table 1. Number of adult ICUs and ICU admission episodes contributing to the ANZICS CORE registries,* 2009–2014

	Hospital type				
Measure	Tertiary	Metropolitan	Rural/regional	Private	Total
Number of adult ICUs in Australia and New Zealand	39	38	55	66	198
ICUs contributing to any ANZICS CORE registry	39	37	48	57	181
ICUs contributing to APD	37	35	40	56	168
Total ICU admission episodes submitted to the APD	331 120	129 210	104 508	192 350	757 188
Mean annual number of submitted ICU admission episodes per ICU (SD)	1584 (599)	673 (290)	471 (207)	704 (481)	845 (597)

ICU = intensive care unit. ANZICS = Australian and New Zealand Intensive Care Society. CORE = Centre for Outcome and Resource Evaluation. APD = Adult Patient Database. * Includes submission to one or more of the APD, Critical Care Resources Registry or the Central Line Associated Blood Stream Infection Surveillance reporting system; excludes the Australian and New Zealand Paediatric Intensive Care Registry.

Figure 1. Example of an SMR funnel plot used to identify a potential outlier ICU*



APACHE = Acute Physiology and Chronic Health Evaluation. SMR = standardised mortality ratio. * The funnels are drawn around the overall mean SMR of the group.

an example of a potential outlier ICU. Potential outliers in each calendar year since commencement are shown in Appendix eTable 2.

Explanations for potential outlier status are detailed below.

Data quality problems

Data quality problems were the most common reason for a unit appearing as a potential outlier. In 19/37 analyses (51%) (16/27 individual units), data quality problems appeared to be the predominant factor contributing to the high SMR.

These were due to submission of inaccurate data or due to missing data. The summary of data quality findings included the following:

- Sudden major changes in temporal outcome trends or patients' severity of illness typically indicated data quality issues. Figure 2 shows an ICU with a rapid change in observed mortality during the monitoring period. This hospital underwent a change of data collection software in March 2013, which resulted in the submission of data relating to survivors without hospital outcomes, who were excluded from further analysis. With survivors excluded, this unit's observed mortality appeared (incorrectly) to increase rapidly.

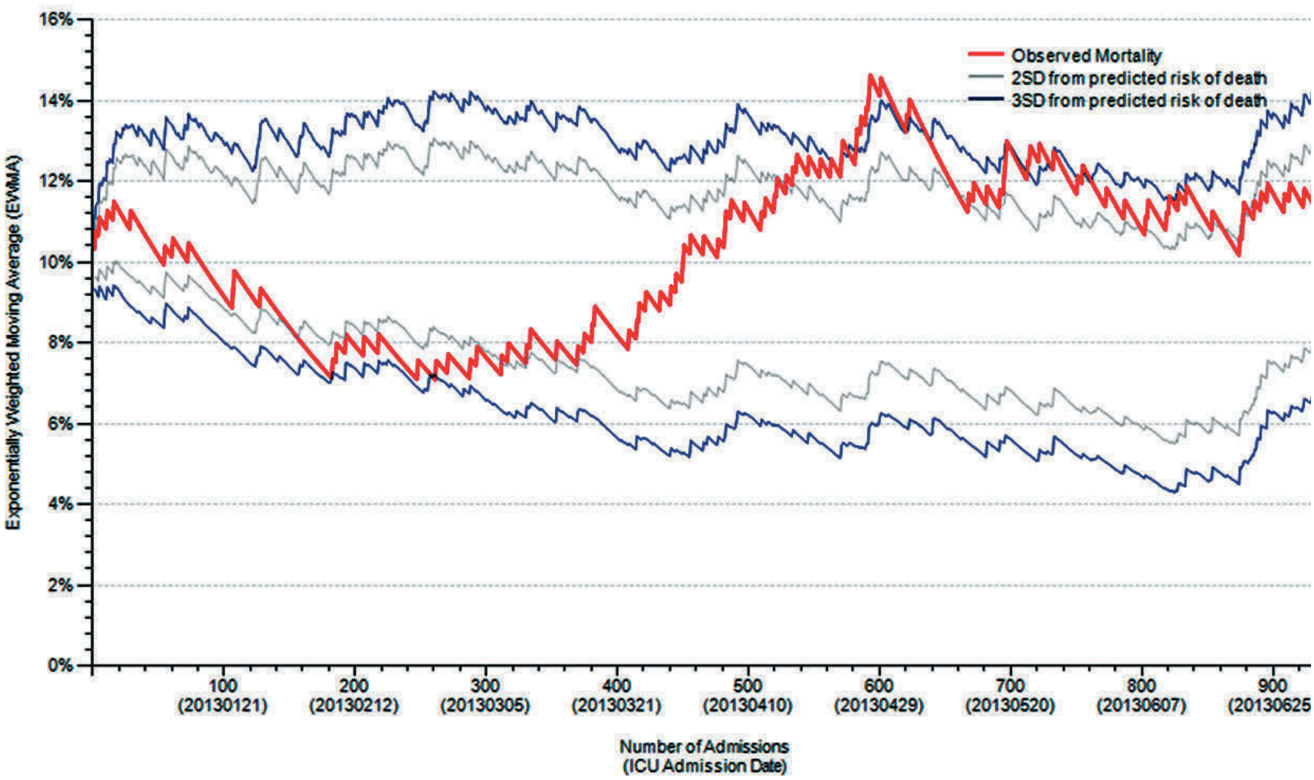
In most hospitals, it was easier to find the chart of a dead patient than a long-stay survivor. Thus, incomplete data sets were more likely to contain dead patients than long-stay survivors.

- Onsite audits suggested that overall data quality was good (Appendix eTable 1) but the effect of missing data was not uniform. Figure 3 shows that missing Glasgow coma score data could result in large variations in the calculated risk of death.
- Compared with data collected in ICUs without trained, dedicated data collectors, information collected in ICUs which did have trained, dedicated data collectors had higher levels of agreement with external auditors (eg, APACHE II score mean difference of 0.26 v 0.02; between-group difference, 0.24 [95% CI, –0.76 to 1.24]). Bland–Altman plots are shown in Appendix eFigure 1.
- Smaller rural and private hospitals and new contributors to the ANZICS CORE clinical registry tended to have a higher incidence of data quality problems.
- Missing physiological, biochemical and diagnostic data generally led to inaccurately low predicted mortality, which artificially elevated the SMR.
- When data quality issues were apparent, further training, feedback on data collection and clarification of variable definitions were offered. In all except two cases, data quality problems were rectified within 12 months.

Casemix variation

In 17/37 of the analyses (46%) (15/27 individual units), casemix differences either partly or completely explained

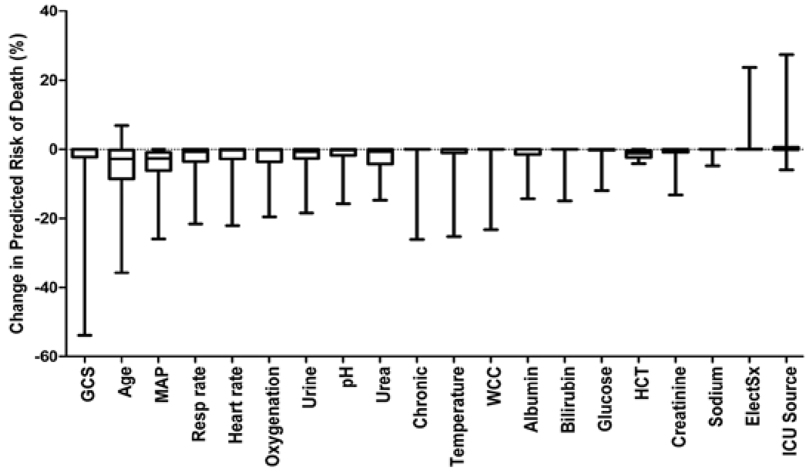
Figure 2. EWMA observed mortality (red) and upper and lower limits of APACHE III-J predicted mortality (grey and blue)



EWMA = exponentially weighted moving average. SD = standard deviation. ICU = intensive care unit. * The EWMA Chart shows the weighted moving average of observed in-hospital mortality for patients admitted to the ICU and the upper and lower limits of predicted mortality.

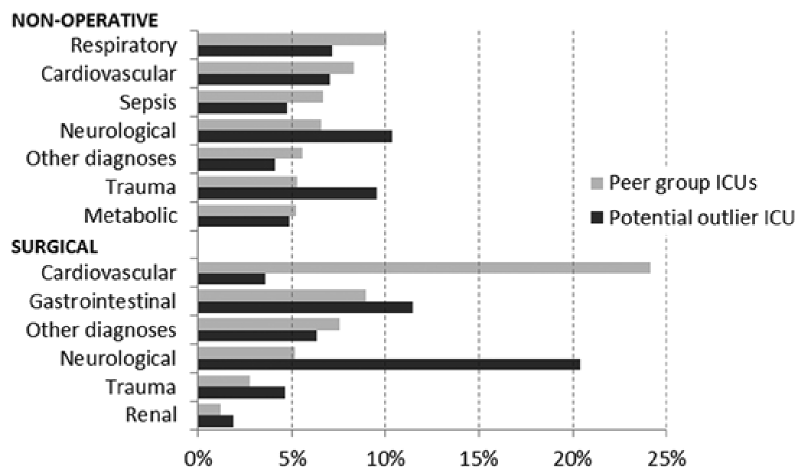
the elevated SMR. For example, in 2013 and 2014, patients with post-operative neurological diagnoses had higher observed and predicted mortalities (10.4% and 14.9%, respectively) than post-operative cardiovascular patients (1.6% and 4.4%, respectively). However, despite lower absolute values, "over-estimation" of predicted mortality was proportionately greater in post-operative cardiovascular patients. Figure 4 illustrates a case where a variation in casemix contributed to the finding of a high SMR, due to more post-operative neurological diagnoses and fewer post-operative cardiovascular diagnoses than similar hospitals. The potential effect of different hospital casemixes with APACHE III-J score is shown in Appendix eTable 3. This has been markedly reduced with the use of the Australian and New Zealand Risk of Death (ANZROD) mortality prediction model as the risk-adjustment model specific to Australia and New Zealand.

Figure 3. Audit review of 32 000 records, showing change in predicted risk of death (%) when a variable is missing and treated as normal*



GCS = Glasgow coma score. MAP = mean arterial pressure. Resp = respiratory. WCC = white cell count. HCT = haematocrit. ElectSx = elective surgery. ICU = intensive care unit. * Boxes represent median and interquartile estimated change in predicted risk of death as a result of a difference between audited and submitted values. 'Whiskers' represent maximum and minimum estimated change seen.

Figure 4. Casemix comparison for APACHE III major diagnostic groups*



APACHE = Acute Physiology and Chronic Health Evaluation. ICU = intensive care unit. * Peer group ICUs are shown for comparison with potential outlier ICUs. The potential outlier ICU has more postoperative neurological and fewer cardiovascular diagnoses than the peer-group hospitals.

Process and resources within the ICU

There were 14/37 analyses (38%) (9/27 individual units) for which data quality and casemix variation did not appear to explain the elevated SMR, and more extensive investigation of processes of care and resources was undertaken. The case study shown in Figure 5 summarises an example where ANZICS CORE identified high levels of staff vacancy, high occupancy and a high proportion of patients discharged after hours from an ICU in a potential outlier hospital.

Action after the outlier review process

The 37 reports from 27 ICUs were provided to the relevant ICU directors, hospital chief executive officers and the local jurisdictional liaison governance committees. In New South Wales, Queensland and Victoria these committees were comprised of intensive care specialists and representatives of the local department of health. In other regions, this governance liaison role was taken on solely by the health department.

In all except one case (where the ICU ceased further submission of data after receipt of the outlier report), ANZICS CORE continued to monitor the ICU outcomes. In eight cases, ICUs have appeared as potential outliers for more than 12 months, prompting repeat analyses. In one case, an ICU remained a potential outlier for 3 years.

Discussion

Our study showed that of 168 ICUs contributing to the ANZICS Adult Patient Database, 27 (16%) were flagged as potential outliers at least once during the study period.

Figure 5. Case study

CASE STUDY

Step 1 – Potential Outlier Identification
This ICU had a higher standardised mortality ratio (observed deaths / predicted deaths derived from the APACHE III-J scoring system) than other similar peer group ICUs. The ICU Director, hospital administration and local jurisidictional committee were informed.

Steps 2 and 3 – Data Quality and Subgroup Analysis
- An on-site audit demonstrated accurate data collection
- The data was of high quality with a high level of completeness
- Case-mix was similar to other peer group hospitals

Step 4 – Analysis of Processes and Resources
This ICU (compared to its peer group) had:
- Higher levels of after-hours discharge from the ICU
- Higher occupancy
- Higher nursing vacancy
- Lower medical and nursing staffing levels.

It was possible that these factors contributed to the poor mortality outcomes at this hospital.

Step 5 - A report was compiled by staff at ANZICS CORE and clinicians from the ANZICS Outlier Working Group. This was provided to the unit Director and to local jurisdictional governance body. The report lead to increased engagement from hospital administration to decrease after hours discharges from the ICU. The following year the local health department prioritised increased funding and resources for this ICU.

ICU = intensive care unit. APACHE = Acute Physiology and Chronic Health Evaluation. ANZICS CORE = Australian and New Zealand Intensive Care Society Centre for Outcome and Resource Evaluation.

Although data quality problems and variation in diagnostic casemix were frequent, there were nine ICUs where these did not explain the elevated SMR, and where potential process-of-care or resource problems were identified.

The outlier management program combined analytic processes and statistical methods within a structured governance framework and process which emphasised early notification, discussion and clinician involvement with jurisdictional oversight. This allowed not only the identification of "true outlier" ICUs but, in many cases, highlighted both resourcing and data quality problems. The outlier management program emphasised the pyramid of investigation[19] but did not formally assess the highest level (provider) which, if necessary, was undertaken by the ICU itself, the hospital administration and/or the local health department.

Comparison of the program with other benchmarking organisations

To the best of the authors' knowledge, the United Kingdom, Denmark, the Netherlands, Sweden, Finland, Norway, Australia, New Zealand and Malaysia manage national clinical ICU registries. The Nationwide Inpatient Sample[27] is the largest all-payer, publicly available, inpatient care database available in the US, containing over seven million hospital stays from about 1000 hospitals annually, and is used for large-sample ICU research. However, of about four million ICU admissions per year in the US, only 15% have clinical risk adjustment measures by organisations such as Project Impact.[6] Hospital benchmarking organisations such as Dr Foster (www.drfoster.com) and The Health Round Table (www.healthroundtable.org) take an approach where potential outliers do not undergo further investigation by the benchmarking organisation and the responsibility of investigation lies with the hospital. The use of registries to monitor performance and identify outliers has been highlighted recently by the Victorian Department of Health after excess neonatal deaths in a regional hospital.[4] Unlike adult and paediatric practice, neonatal ICUs in Australia and New Zealand do not participate in a registry-based outlier management program.

Benchmarking ICU performance as a marker of the wider healthcare system

With over 150 000 admissions per year, at $4000 per day,[28] and with an average length of stay of 3 days, ICU costs in Australia and New Zealand are likely to be nearly $2 billion per year. Thus, intensive care medicine represents a sector where monitoring of resource use and patient outcomes are both important. The ANZICS outlier management program, just one aspect of the ANZICS CORE benchmarking program, has been independently assessed as having resulted in a net saving of $26 million dollars to the Australian health care system (representing $4 dollars saved for every $1 spent).[29] ICUs are ideally placed to identify wider health care problems through their immediate exposure to the "sickest patients in the hospital", as was seen in 2009 with H1N1 influenza.[30] In addition, the use of ICU data as a focus for hospital performance measurement may offer advantages over the hospital SMR.[1] With widespread participation in the ANZICS CORE benchmarking program, surveillance of ICU outcomes may function as the "canary in the coal mine" for processes occurring within the rest of the hospital.[9]

Limitations and strengths

The creation of the ANZICS CORE registry in the late 1990s was driven by clinicians striving for a way to improve outcomes and performance. It is acknowledged, however, that the ANZICS CORE outlier management program is not a complete assessment of ICU quality, and multiple other performance measures are currently available. It is part of an overarching hospital quality-improvement process that has many facets (eg, accreditation, mortality and morbidity reviews and education).

There are many controversies in the area of detecting potential outlier providers with respect to the statistical methods[11,12,21,22] and overall effectiveness.[24,25] The predictive value of SMRs for detecting preventable mortality is highly sensitive to the proportion of deaths that are preventable and the extent to which the latter varies across hospitals.[24] There is ongoing debate that process measures, however defined, have an impact on the SMR.[25]

Potential limitations (common to all methods of performance measurement[31]) of the program include general limitations such as sample size and data accuracy, the requirement for accurate risk-adjustment models,[23] loss of model calibration,[32] sensitivity and specificity of funnel plots,[22] and limitations of short-term mortality as an endpoint. A statistical model cannot control for all potential confounding factors. The deteriorating calibration of APACHE III-J over time is a limitation of this study.[31] A risk model specific to ICUs in Australia and New Zealand has been introduced from 2014.[33] Comparisons made to previous methods of outlier detection suggest better adjustment for casemix variation.[6,34] The new benchmarking and monitoring process using ANZROD specifically excludes patients admitted for the purposes of palliative care or for organ donation, in an attempt to address this issue and limit the inclusion of “inevitable deaths”.

Processes of care before ICU admission and after ICU discharge are not captured in the database. Non-ICU care processes may affect outcomes of ICU patients, and need to be assessed when understanding ICU performance.

There is a potential trade-off between over detection and false positive enquiries versus obligations to public safety.[16] The outlier management program aimed to identify confounding factors as much as possible. ICUs used clinical datasets developed specifically for the purpose of risk adjustment; models were well developed and validated, event rates were sufficiently frequent and admissions for palliative care were rare. ANZICS is already recognised as a high-quality registry organisation.[35] Resubmission from sites with data quality problems led not only to more accurate benchmarking but also further improved the overall quality of information within the registry.

Sophisticated analyses of CORE data requiring advanced statistical knowledge have been published.[11-12] ANZICS CORE chose an approach similar to the Intensive Care National Audit and Research Centre[36] and supported by Bridgewater and colleagues.[17] This was a process which balanced appropriate statistical analysis with recognition of potential limitations of submitted data and the capacity to delve more deeply into processes of care. This ensured engagement and understanding from clinicians and policy makers.

Conclusion

We have presented a binational, outcome-based, ICU performance measurement process and a template for early detection and analysis of potential statistical outliers based on clinical data. The focus was a clinician-driven understanding of potential causes as a basis for further action. The spotlight on care was addressed every 3 months, keeping monitoring and action fresh and current. We believe that a system such as this is vital to monitor the health care system and contributes to the high-quality care and progressively improving outcomes of patients admitted to ICUs in Australia and New Zealand.[37]

Competing interests

Daniel Mullany, Peter Hicks and David Pilcher have been part of the ANZICS outlier working group. Eldho Paul is a PhD student funded by the ANZICS Centre for Outcome and Resource Evaluation and by the Australian and New Zealand Intensive Care Research Centre, Monash University. Other authors have no competing interests to declare.

Author details

Kerry McClean, Project Officer, Data Quality and Education[1]
Daniel Mullany, Intensivist[2]
Sue Huckson, Manager[1]
Allison van Lint, Project Lead[1]
Shaila Chavan, Senior Project Officer[1]
Peter Hicks, Director[1] and Intensivist[3]
Graeme K Hart, Intensivist[4]
Eldho Paul, PhD Candidate[5]
David Pilcher, Chair,[1] Adjunct Clinical Professor[5] and Intensivist[6]

1 Centre for Outcome and Resource Evaluation, Australian and New Zealand Intensive Care Society, Melbourne, Australia.
2 Critical Care Research Group, Adult Intensive Care Services, Prince Charles Hospital and University of Queensland, Brisbane, Australia.
3 Department of Intensive Care, Wellington Hospital, Wellington, New Zealand.
4 Department of Intensive Care, Austin Hospital, Melbourne, Australia.
5 Australian and New Zealand Intensive Care Research Centre, School of Public Health and Preventive Medicine, Monash University, Melbourne, Australia.
6 Department of Intensive Care, Alfred Hospital, Melbourne, Australia.

Correspondence: d.pilcher@alfred.org.au

References

1 Scott IA, Ward M. Public reporting of hospital outcomes based on administrative data: risks and opportunities. *Med J Aust* 2006; 184: 571-5.

2 Spiegelhalter D, Grigg O, Kinsman R, et al. Risk-adjusted sequential probability ratio tests: applications to Bristol, Shipman and adult cardiac surgery. *Int J Qual Health Care* 2003; 15: 7-13.

3 Mohammed MA, Cheng K, Rouse A, Marshall T. Bristol, Shipman, and clinical governance: Shewhart's forgotten lessons. *Lancet* 2001; 357: 463-7.

4 Victoria State Government. Targeting zero: supporting the Victorian Hospital system to eliminate avoidable harm and strengthen quality of care. Melbourne: Victorian Government, 2016. https://www2.health.vic.gov.au/hospitals-and-health-services/quality-safety-service/hospital-safety-and-quality-review (accessed May 2017).

5 Australian Institute of Health and Welfare. Meteor Metadata Online Registry. Intensive care unit. Canberra: AIHW, 2012. http://meteor.aihw.gov.au/content/index.phtml/itemId/327234 (accessed Jul 2014).

6 Breslow MJ, Badawi O. Severity scoring in the critically ill. Part 1: interpretation and accuracy of outcome prediction scoring systems. *Chest* 2012; 141: 245-52.

7 Afessa B, Gajic O, Keegan MT. Severity of illness and organ failure assessment in adult intensive care units. *Crit Care Clin* 2007; 23: 639-58.

8 Frankel SK, Moss M. The effect of organizational structure and processes of care on ICU mortality as revealed by the United States Critical Illness and Injury Trials group critical illness outcomes study. *Crit Care Med* 2014; 42: 463-4.

9 Pilcher DV, Hoffman T, Thomas C, et al. Risk-adjusted continuous outcome monitoring with an EWMA chart: could it have detected excess mortality among intensive care patients at Bundaberg Base Hospital? *Crit Care Resusc* 2010; 12: 36-41.

10 Cooke CR, Iwashyna TJ. Using existing data to address important clinical questions in critical care. *Crit Care Med* 2013; 41: 886-96.

11 Solomon PJ, Kasza J, Moran JL; Australian and New Zealand Intensive Care Society Centre for Outcome and Resource Evaluation. Identifying unusual performance in Australian and New Zealand intensive care units from 2000 to 2010. *BMC Med Res Methodol* 2014; 14: 53.

12 Moran JL, Solomon PJ; ANZICS Centre for Outcome and Resource Evaluation of the Australian and New Zealand Intensive Care Society. Statistical process control of mortality series in the Australian and New Zealand Intensive Care Society adult patient database: implications of the data generating process. *BMC Med Res Methodol* 2013; 13: 66.

13 Duckett SJ, Coory M, Sketcher-Baker K. Identifying variations in quality of care in Queensland hospitals. *Med J Aust* 2007; 187: 571-5.

14 Boots RJ, Singh S, Terblanche M, et al. Remote care by telemedicine in the ICU: many models of care can be effective. *Curr Opin Crit Care* 2011; 17: 634-40.

15 Mohammed MA, Stevens AJ. Performance monitoring in Australia and England: from scandals to action. *Med J Aust* 2007; 187: 549-50.

16 Shahian DM, Normand SL. Autonomy, beneficence, justice, and the limits of provider profiling. *J Am Coll Cardiol* 2012; 59: 2383-6.

17 Bridgewater B, Hickey GL, Cooper G, et al. Publishing cardiac surgery mortality rates: lessons for other specialties. *BMJ* 2013; 346: f1139.

18 Black N, Tan S. Use of national clinical databases for informing and for evaluating health care policies. *Health Policy* 2013; 109: 131-6.

19 Mohammed MA. Using statistical process control to improve the quality of health care. *Qual Saf Health Care* 2004; 13: 243-5.

20 Spiegelhalter DJ. Funnel plots for comparing institutional performance. *Stat Med* 2005; 24: 1185-202.

21 Spiegelhalter DJ. Handling over-dispersion of performance indicators. *Qual Saf Health Care* 2005; 14: 347-51.

22 Manktelow BN, Seaton SE, Evans TA. Funnel plot control limits to identify poorly performing healthcare providers when there is uncertainty in the value of the benchmark. *Stat Methods Med Res* 2016; 25: 2670-84.

23 Verburg, IW, Holman R, Peek N, et al. Guidelines on constructing funnel plots for quality indicators: a case study on mortality in intensive care unit patients. *Stat Methods Med Res* 2017; Jan 1. doi: 0962280217700169 [Epub ahead of print].

24 Girling A J, Hofer TP, Wu JH, et al. Case-mix adjusted hospital mortality is a poor proxy for preventable mortality: a modelling study. *BMJ Qual Saf* 2012: 21; 1052-6.

25 Pitches DW, Mohammed M, Lilford R. What is the empirical evidence that hospitals with higher-risk adjusted mortality rates provide poorer quality care? A systematic review of the literature. *BMC Health Serv Res* 2007: 7; 91.

26 Cook D A, 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.

27 Healthcare Cost and Utlization Project. National (nationwide) inpatient sample (NIS). Rockville, Md: HCUP. http://www.hcup-us.ahrq.gov/nisoverview.jsp (accessed Aug 2014).

28 New South Wales Government Ministry of Health. Costs of Care Standards 2009/10. Sydney: NSW Government, 2011. http://www1.health.nsw.gov.au/pds/ActivePDSDocuments/GL2011_007.pdf (accessed Sep 2016).

29 Australian Commission on Safety and Quality in Health Care. Economic evaluation of clinical quality registries: final report. Sydney: ACSQHC, 2016. https://www.safetyandquality.gov.au/wp-content/uploads/2016/12/Economic-evaluation-of-clinical-quality-registries-Final-report-Nov-2016.pdf (accessed Jan 2017).

30 Australia and New Zealand Extracorporeal Membrane Oxygenation Influenza Investigators, Davies A, Jones D, et al. Extracorporeal membrane oxygenation for 2009 influenza A (H1N1) acute respiratory distress syndrome. *JAMA* 2009: 302: 1888-95.

31 Kuzniewicz MW, Vasilevskis EE, Lane R, et al. Variation in ICU risk-adjusted mortality impact of methods of assessment and potential confounders. *Chest* 2008; 133: 1319-27.

32 Paul E, Bailey M, Van Lint A, et al. Performance of APACHE III over time in Australia and New Zealand: a retrospective cohort study. *Anaesth Intensive Care* 2012; 40: 980-94.

33 Paul E, Bailey M, Pilcher D. Risk prediction of hospital mortality for adult patients admitted to Australian and New Zealand intensive care units: development and validation of the Australian and New Zealand risk of death model. *J Crit Care* 2013; 28: 935-41.

34 Paul E, Bailey M, Kasza J, Pilcher D. The ANZROD model: better benchmarking of ICU outcomes and detection of outliers. *Crit Care Resusc* 2016; 18: 25-36.

35 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.

36 Intensive Care National Audit and Research Centre. Online reports. London: ICNARC. https://onlinereports.icnarc.org/Home (accessed Apr 2017).

37 Kaukonen KM, Bailey M, Suzuki S, et al. Mortality related to severe sepsis and septic shock among critically ill patients in Australia and New Zealand, 2000–2012. *JAMA* 2014; 311: 1308-16. ❏

CriticalCare and Resuscitation

Official journal of the College of Intensive Care Medicine of Australia and New Zealand

Chief Editor: Rinaldo Bellomo

Copyright transfer statement

The copyright to this article is transferred to the *College of Intensive Care Medicine of Australia and New Zealand* effective if and when the article is accepted for publication. The copyright transfer covers the exclusive right to reproduce and distribute the article, including reprints, translations, photographic reproductions, microform, electronic form (offline, online) or any other reproductions of similar nature.

We the undersigned declare that the material presented in the attached paper has not been published previously (nor is it currently being assessed for publication) by any journal other than *Critical Care and Resuscitation*.

MANUSCRIPT TITLE

....

....

....

....

....

AUTHORS' NAMES (Print)	**AUTHORS' SIGNATURES** (and Date)
....	....
....	....
....	....
....	....
....	....
....	....
....	....
....	....

Appendix

This appendix was part of the submitted manuscript and has been peer reviewed. It is posted as supplied by the authors.

Appendix eFigure 1: Bland Altman plots of the Effect of Different Forms of Data Collection on APACHE II scores.

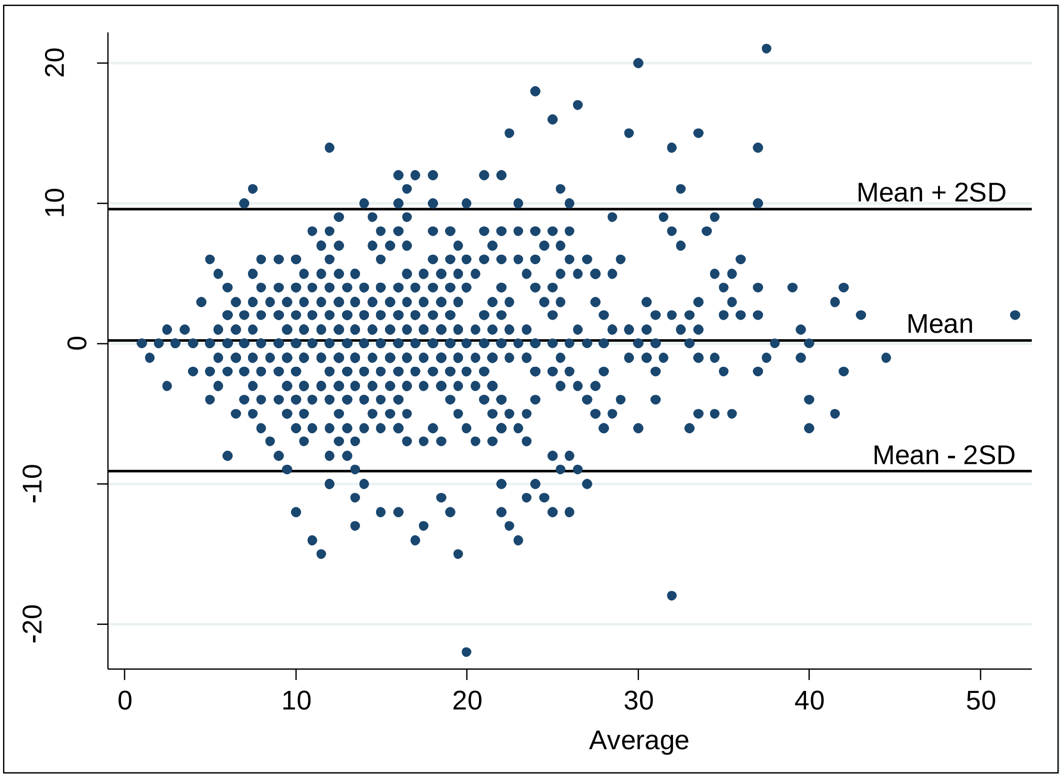


Figure 1a. Bland-Altman plot of difference against mean of audited and submitted APACHE II total scores (869 patient records from 43 ICUs)

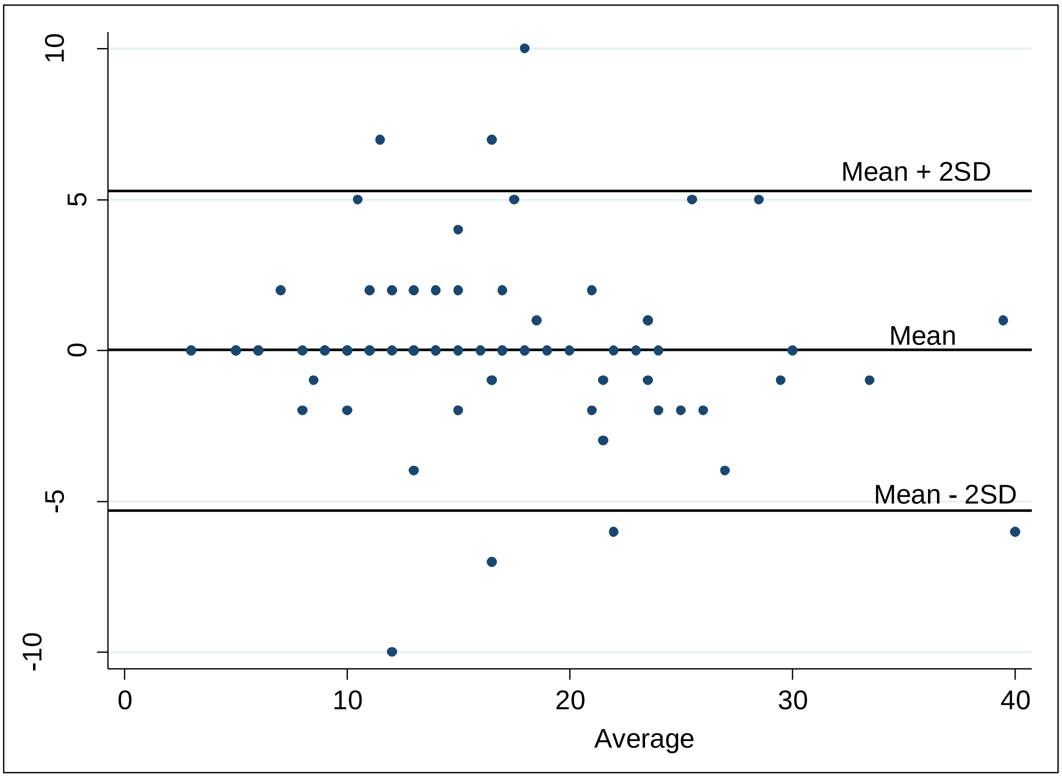


Figure 1b. Bland-Altman plot of difference against mean of audited and submitted APACHE II total scores in ICUs with trained dedicated data collectors (94 patient records)

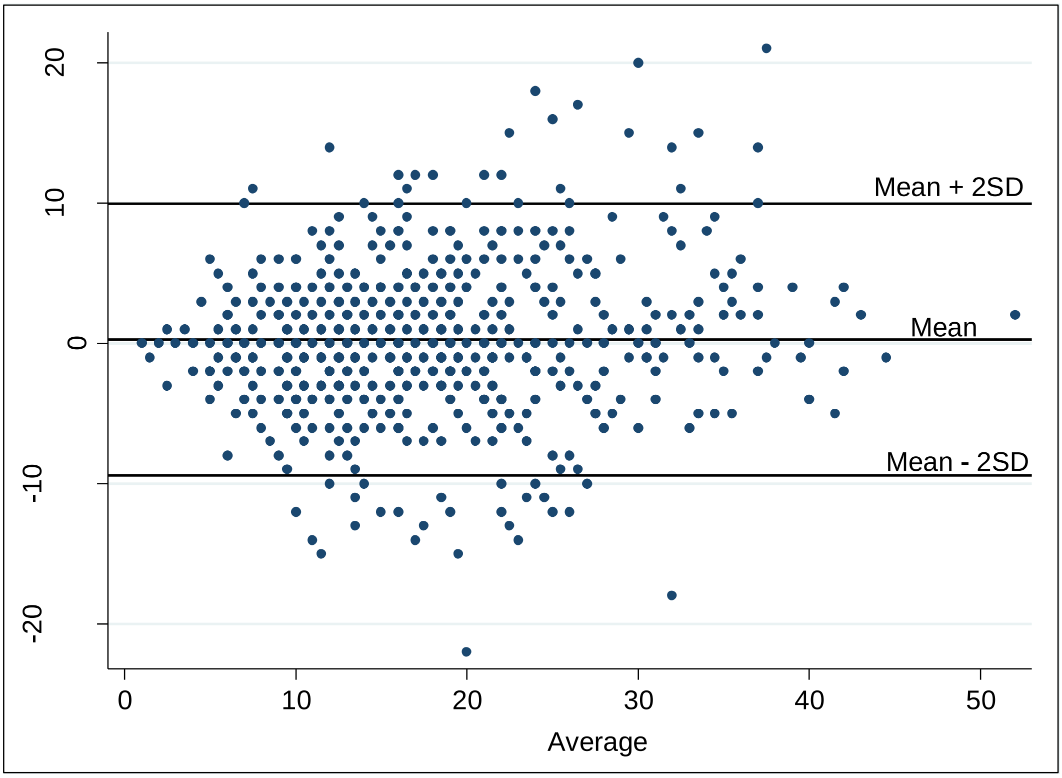


Figure 1c. Bland-Altman plot of difference against mean of audited and submitted APACHE II total scores in ICUs without trained dedicated data collectors (775 patient records)

Appendix eTable 1.

Levels of agreement between independently audited values and submitted APACHE II sub-score values obtained during audits of data quality (869 patient records from 43 ICUs)

Variable	Original score Mean (SD)	Audited score Mean (SD)	Mean difference (95% Confidence Interval)
Age score	3.424 (2.27)	3.422 (2.27)	0.002 (-0.002 to 0.007)
Temperature score	0.63 (0.82)	0.66 (0.78)	-0.03 (-0.06 to 0.01)
HR score	1.71 (1.10)	1.81 (1.07)	-0.10 (-0.16 to -0.05)
RR score	0.92 (1.01)	0.98 (0.99)	-0.06 (-0.11 to -0.01)
Sodium score	0.13 (0.55)	0.12 (0.51)	0.01 (-0.004 to 0.034)
Potassium score	0.375 (0.78)	0.361 (0.77)	0.014 (-0.02 to 0.048)
Creatinine score	1.22 (1.95)	1.11 (1.83)	0.11 (0.05 to 0.18)
Haematocrit score	0.85 (1.07)	0.83 (1.04)	0.02 (-0.03 to 0.07)
WCC score	0.556 (0.88)	0.558 (0.89)	-0.002 (-0.035 to 0.03)
MAP score	1.53 (1.13)	1.63 (1.10)	-0.10 (-0.17 to -0.04)
Oxygen score	1.02 (1.37)	1.25 (1.47)	-0.23 (-0.30 to -0.16)
pH score	1.08 (1.32)	0.99 (1.31)	0.09 (0.04 to 0.14)
GCS score	2.24 (4.01)	2.35 (4.10)	-0.11 (-0.31 to 0.09)
Chronic health score	1.38 (2.15)	0.77 (1.75)	0.61 (0.49 to 0.74)
Total score	17.05 (8.57)	16.81 (8.10)	0.24 (-0.07 to 0.55)

HR: Heart Rate; RR: Respiratory rate; WCC: White Blood Cell Count; MAP: Mean Arterial Blood Pressure; GCS: Glasgow Coma Score

Appendix eTable 2: Time trends in Outlier reports*

Year	Number of Eligible Admissions	Number of sites	Outlier Audit Reports	Potential New Outliers
2014	136338	151	4	2
2013	129153	151	7	5
2012	119687	143	11	8
2011	117831	145	5	2
2010	114146	146	9	9
2009	103946	139	1	1
Totals	721101		37	27

*This excludes admissions without a calculated APACHE III risk of death, repeated analyses and external requests for analysis

Appendix eTable 3: Differential risk of death with differing casemix for APACHE III. Based on Data from all ANZ ICUs in APD (2013 & 2014) with APACHE III exclusions applied.

	Neurological Diagnoses (includes traumatic brain injury)	Cardiac Surgery (CAGS & valves only)	All other Cases
Number of cases	37,249	34,761	192,607
Predicted mortality from APACHE III	14.9%	4.4%	15.1%
Observed mortality	10.4%	1.6%	9.4%
Ratio of Observed to Expected (SMR)	0.70	0.37	0.62