

Quality and Reliability of Data Collected in a Regional Hospital Intensive Care Unit

G. J. McHUGH

Intensive Care Unit, Palmerston North Hospital, Palmerston North, NEW ZEALAND

Objective: *To examine the accuracy of the existing data collection system in an intensive care unit (ICU).*

Methods: *Prospective audit of data collection for APACHE II in a six-bedded general ICU in a regional hospital for consecutive ICU admissions over a 12 month period. The existing resident medical officer (RMO) data collection system continued unaltered while concurrent collection of APACHE II data completed by one senior medical officer (SMO) was also performed. RMO-APACHE scores were compared to SMO-APACHE scores.*

Results: *The median (interquartile range) APACHE II scores were: 15 (9-21) for RMO data; 17 (12 - 22) for SMO data ($p = 0.0001$), with a mean difference (SD) of 4.2 (4.2) ($p < 0.0001$). A rapid survey of New Zealand ICU's indicated that only 4 of 11 APACHE II users employed substantial medical specialist involvement in APACHE II data collection and/or scrutiny prior to database entry.*

Conclusions: *The existing APACHE II data collection is often imperfect and requires further attention as it is not yet reliable enough to allow for accurate SMR calculations, or for external use for comparative purposes. (Critical Care and Resuscitation 2004; 6: 258-260)*

Key words: APACHE II, intensive care unit, data quality, interobserver variability

Planning of intensive care services, together with service evaluation and inter-regional comparisons, are heavily dependant upon data collected for, and collated within, computerised databases and/or registries. The accuracy and quality of the data upon which these databases and registries are based are generally assumed to be sound, often without acknowledgement of the potential uncertainties.^{1,2} Systematic verification of intensive care data quality has been infrequently reported.³⁻⁹ Once data have been entered into a database, processed and then subsequently reported in a collated format, the information is often accepted as fact upon which many important decisions are made.

The intensive care unit (ICU) in the Palmerston North hospital (PNH) has had a locally-developed computerised database maintained since 1986. Its current edition is a Microsoft Access[®] based system and data is collected by the resident medical officers (RMOs) working within the ICU. All are given an introductory tutorial on the collection of data, including

the use of the APACHE II system.¹⁰ Further instruction occurs on a day-to-day basis, particularly when it is perceived that the quality of data is poor or data collection is incomplete. Despite this on-going scrutiny, there have been recurrent concerns that the overall data quality may be less than acceptable. As part of the continuing quality assurance programme, it was decided that the quality of data collection would be audited prospectively, using APACHE II scores collected by a specialist medical officer (SMO) as an indicator of accuracy.

METHODS

The existing data collection and data entry continued unaltered throughout the audit period (January 1st, 2003 - December 31st, 2003 inclusive). The duties of the RMOs within the ICU included the completion of a data form for every patient admitted to the ICU, including the completion of an APACHE II score after the first 24 hours. The APACHE section of the data form excluded

Correspondence to: Dr. G. J. McHugh, Intensive Care Unit, Palmerston North Hospital, Palmerston North, New Zealand (e-mail: gerard.mchugh@midcentral.co.nz)

those patients who were aged less than 15 years and those patients whose length of stay (LOS) was less than 24 hours. The RMOs were not informed of the audit. One specialist medical officer with 10 years' experience in collecting data for the APACHE II system collected and collated APACHE II scores concurrently and separately for all admissions.

At the end of the 12 month period, all APACHE II scores generated by the RMOs (RMO-APACHE) were retrieved from the database to allow a comparison with the APACHE II scores generated by the specialist (SMO-APACHE).

Paired APACHE II scores were evaluated using Wilcoxon signed ranks test and mean absolute differences between scores were evaluated using a one sample Student's *t* test. A *p* value of 0.05 was considered significant. Statistical evaluation was completed using the Analyse-it™ General Statistics software package (version 1.71. Analyse-it Software Ltd., Leeds, UK). An informal 'real-time' email enquiry was sent to an SMO in each public hospital ICU in New Zealand to evaluate any SMO involvement in APACHE II data collection. Eleven current APACHE II-using ICUs self-identified and responded (matching the published data on use of APACHE II in New Zealand).¹¹

RESULTS

There were 290 admissions to the ICU during the twelve month audit period. Forty-one patients were transferred to other hospitals. Thirty-three patients died in the ICU (11%). Another 8 deaths occurred in hospital following ICU discharge (total deaths 41 or 14%). Median LOS was 44.5 hours (interquartile range 21 - 116 hours), with 194 (67%) receiving invasive mechanical ventilation and 114 (39%) admitted directly from the operating theatre. Eighty-eight patients had a LOS of less than 24 hours. Sixteen patients were aged < 15 years. Of the remaining 196 patients, APACHE II scores were used for subsequent evaluation. Raw data collected by RMOs were unavailable for scrutiny.

The median (range and interquartile range) APACHE II score generated by RMOs was 15 (0 - 47; 9-21), compared with 17 (4 - 42; 12 - 22) (*p* = 0.0001) for the SMO. When data were extracted from the database, it became apparent that the default APACHE II score in the database was zero. This meant that even when no data were supplied by the RMO an APACHE II score of 0 was recorded. Lest the mean difference be artificially exaggerated, RMO APACHE II scores of 0 were excluded (*n* = 21). For the remaining pairs of scores, the mean absolute difference (SD) was 4.2 (4.2) (*p* < 0.0001). The range of differences was -26 to 22. Only 22 scores agreed exactly.

Of the other 10 New Zealand ICUs using APACHE

II, only 4 reported substantial involvement of senior medical staff in the collection or scrutiny of APACHE II data prior to data entry.

DISCUSSION

The quality and reliability of data are fundamental to any process which incorporates data interpretation. The proliferation of computerised databases, be they held locally or centrally coordinated, has given rise to a 'culture' of data collection. Such data are used for a number of purposes, not all of which are necessarily appropriate to the method and rigour of original data collection.¹² A desire to pursue inter-ICU comparisons of activity and quality-of-care has led to widespread use of outcome prediction scoring (e.g. APACHE II,¹⁰ SAPS II,¹³ etc) and standardised mortality ratios (SMR).¹⁴ In so-doing, it is all too easy to assume that data reliability is assured, and to base important decisions and evaluations upon the information taken from the processed outputs of such databases. Multiple aspects of the process of data collection and recording can impact upon the reliability of the end-product.¹⁴ Those who are familiar with the practical and clinical limitations of the circumstances in which the data are collected are presumably less likely to misinterpret the data, or to treat processed outputs as irrefutable facts. Unfortunately important management decisions do not always pay due regard to the clinical setting of their basis.

This investigation has again demonstrated some of the potential deficiencies in the reliability of APACHE II scores, particularly the interobserver reliability.^{3-9,15} The significant differences between APACHE II scores collected by RMOs and by an experienced SMO have an immediate impact upon the apparent acuity of the population admitted to the PNH ICU. Further, the mean difference of 4.2 is also important, as each 5-point band of APACHE II scores is associated with a significant increase in predicted risk of death, right across the spectrum of severity.¹⁰ This difference would have been greater if RMO-scores of zero had not been excluded from the analysis. Even small differences in APACHE II scores have been shown to have important effects upon predicted mortality,⁴ as the impact of the discrepancy is compounded in any derivative, such as the standardised mortality ratio (SMR).

In response to these findings, the database at PNH ICU has been subsequently modified so that the default APACHE II score is null rather than zero. The relative likelihood of such faults being present in commercially developed software packages is perhaps less likely, but locally developed systems are more readily customised in response to day-to-day practicalities. In an effort to further address the observed shortcomings, additional

RMO training and closer scrutiny of data collection will also result. Consideration has been given to involvement of SMOs in data collection and/or checking.

The finding of interobserver differences in APACHE II score is not new.^{3,9,15} However if one assumes that an SMO score is a more accurate reflection of the patient's clinical features,¹⁵ it is interesting that only 4 of 11 APACHE II users in New Zealand use SMOs to collect or substantially review their APACHE II data. A recent governmental advisory document¹⁶ has signalled an expectation that inter-ICU comparison using mortality data will be developed. For this to be worthwhile, the PNH ICU, and presumably others, will need to address issues of data reliability and quality, including possible SMO involvement. Any comparison using existing local data would be inaccurate and potentially misleading. If these findings can be extrapolated nationally, outcome data and their derivatives should only be used for generalised comparisons with due attention having been paid to the reliability of base data.^{3,17,18}

In summary, shortcomings in the quality and reliability of data collection, particularly APACHE II scores, in the ICU of a regional hospital have been demonstrated. Data collected cannot be relied upon to provide an accurate external comparator and this may well reflect similar issues in other hospitals nationally.

Received 20 August 04

Accepted 7 September 04

REFERENCES

- West AF, West RR. Clinical decision-making: coping with uncertainty. *Postgrad Med J* 2002;78:319-321.
- Ridley SA. Uncertainty and scoring systems. *Anaesthesia* 2002;57:761-767.
- Holt AW, Bury LK, Bersten AD, Skowronski GA, Vedig AE. Prospective evaluation of residents and nurses as severity score collectors. *Crit Care Med* 1992;20:1688-1691.
- Goldhill DR, Sumner A. APACHE II, data accuracy and outcome prediction. *Anaesthesia* 1998;53:937-943.
- Polderman KH, Thijs LG, Girbes ARJ. Interobserver variability in the use of APACHE II scores. *Lancet* 1999;353:380.
- 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.
- Polderman KH, Girbes ARJ, Thijs LG, Strack van Schinjdell RJM. Accuracy and reliability of APACHE II scoring in two intensive care units. *Anaesthesia* 2001;56:47-50.
- Arts DGT, Bosman RJ, de Jonge E, Joore JCA, de Keizer NF. Training in data definitions improves quality of intensive care data. *Critical Care (London)* 2002;7:179-184.
- Arts D, de Keizer N, Scheffer GJ, de Jonge E. Quality of data collected for severity of illness scores in the Dutch National Intensive Care Evaluation (NICE) registry. *Intensive Care Med* 2002;28:656-659.
- Knaus WA, Draper EA, Wagner DP, Zimmerman JE. APACHE II: a severity of disease classification system. *Crit Care Med* 1985;13:818-829.
- Higlett T, Anderson T, Hart GK. Review of intensive care resources & activity 2001/2002. Melbourne: ANZICS, 2004.
- Wyatt J. Acquisition and use of clinical data for audit and research. *J Eval Clin Pract* 1995;1:15-27.
- Le Gall JR, Lemeshow S, Saulnier F. A new Simplified Acute Physiology Score (SAPS II) based on a European/North American multicenter study. *JAMA* 1993;270:2957-2963.
- Ridley S. Severity of illness scoring systems and performance appraisal. *Anaesthesia* 1998;53:1185-1194.
- Polderman KH, Jorna EMF, Girbes ARJ. Inter-observer variability in APACHE II scoring: effect of strict guidelines and training. *Intensive Care Med* 2001;27:1365-1369.
- Ministry of Health Intensive Care Clinical Advisory Group. Intensive Care Services in New Zealand. Wellington: Ministry of Health, 2004.
- Teres D, Pellow P. Assessing data elements in a severity scoring system. *Intensive Care Med* 2000;26:263-264.
- Rowan KM, Angus DC. Don't let perfection be the enemy of the good: it's time for optimism over the role of severity scoring systems in intensive care unit performance measurement. *Curr Opin Crit Care* 2000;6:153-154.