VALIDATION OF APACHE II SCORE IN A SURGICAL INTENSIVE CARE UNIT

F G Chen, K F Koh, M H Goh

ABSTRACT

The APACHE II scoring system was applied to 301 surgical intensive care admissions over a 9-month period. The mean age of patients admitted was 52.39 years (SD 19.3) and the mean duration of stay was 5.37 days (SD 8.93). The overall mortality was 17.27%. The mean APACHE II scores for survivors was 12.94 (SD 7.43) and non-survivors 28.19 (SD 10.43). There was good correlation between expected mortality predicted by the APACHE II system and observed mortality ($r=0.9732$). Using a predicted risk criterion of 0.5 to distingush between those predicted to survive and die, of the 45 patients predicted to die, only 30 actually did so. No patient survived with an APACHE II score of more than 40 and with a predicted risk of death greater than 0.87. We found the APACHE II system useful for evaluating ICU performance and risk stratification for the purpose of therapeutic trials but not as a triage tool.

Keywords: *surgical intensive care, APACHE II, mortality, validation.*

SINGAPORE MED J 1993; Vol 34: 322-324

INTRODUCTION

The intensive care unit (ICU) is an expensive resource for the management of patients requiring complicated medical and surgical care. Improving therapeutic capabilities and monitoring technology has escalated intensive care demand at a time of economic constraints. Accurate outcome prediction could optimise ICU bed usage by reducing unnecessary admissions of low risk patients and futile care of terminally ill patients.

The APACHE (acute physiology and chronic health evaluation) prognostic scoring system was developed in 1981 at the George Washington University Medical Centre as a method to measure disease severity(1,2). The APACHE score was found to correlate directly with hospital mortality. Because of the complicated method of obtaining physiological data, a simplified refinement of the original APACHE was introduced in 1985, allowing calculation of probability of death(3). Validation of the APACHE II scoring system has been carried out in the American and European general intensive care population but not to our knowledge in Asian and Singapore surgical intensive care patients. The aim of this study is to evaluate the APACHE II system in a clinical setting prospectively to validate its application in a study population totally different from the groups used to develop the model.

MATERIALS AND METHODS

The National University Hospital Surgical Intensive Care Unit is a 6-bed unit serving a 700-bed teaching hospital. Medical, paediatric, neonatal, cardiac surgical ICUs and coronary care unit are separate facilities. The Surgical Intensive Care Unit (SICU) is managed by a full-time staff composed of residents, registrars and consultants from the Department of Anaesthesia. Admission standards to the SICU are not strictly fixed but as a rule cover admission of patients after major or complicated surgery as well as patients with problems too ill to be managed in the surgical wards.

Over a 9-month study period data were collected prospectively on 301 consecutive admissions to intensive care. All patients were classified on admission by the primary reason of admission. The APACHE II score in the first 24 hours of admission and expected risks of death was calculated using the Microsoft Excel 3.0 computer program. The score and expected mortality was not used in the clinical management of the patient. A successful outcome was defined as a discharge from the intensive care and to our knowledge no patient died within 24 hours of discharge to the general wards.

The APACHE II scoring system has 3 parts. The first part, the Acute Physiological Score consists of a weighted sum of 12 physiologic measurements of 0 to 4 points being given to each measurement. The second part, the Chronic Health Evaluation is the chronic or pre-admission health status. Depending on whether the ICU admission was elective or emergency, postoperative or non-operative, 2 or 5 points are scored. Age points of 0 to 6 are assigned in the last part. The probability of death is derived from the formulae: $\mathrm{Ln}(R/1-R) = -3.517 + (\text{APACHE II} * 0.146) + D + S$ where Ln = natural logarithm, R = risk of hospital mortality, D = disease weight given to the primary reason for admission to ICU and S = additional weightage of 0.603 given for emergency surgery(3).

Validity was tested by analysing the accuracy of the predicted probabilities overall and within various APACHE score groups compared with observed mortalities using the chi square test and linear regression techniques. Comparisons were reported as significant at $p<0.05$. A Receiver Operating Characteristic (ROC) curve analysis was performed to determine the accuracy of the model in our population(4,5). The ROC curve is

Department of Anaesthesia
National University Hospital
5 Lower Kent Ridge Road
Singapore 0511

F G Chen, MBBS, M Med (Anaes), FFARACS
Senior Resident

K F Koh, MBBS, M Med (Anaes)
Senior Resident

M H Goh, MBBS
Resident

Correspondence to: Dr F G Chen

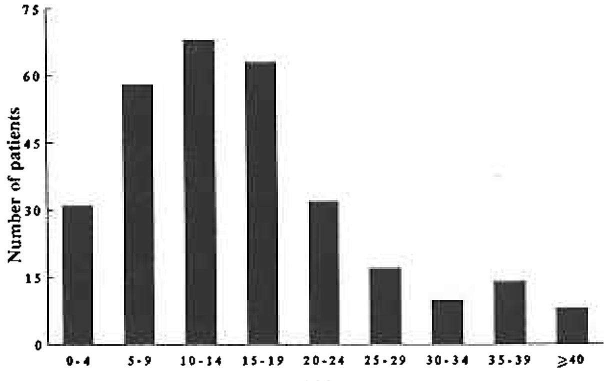
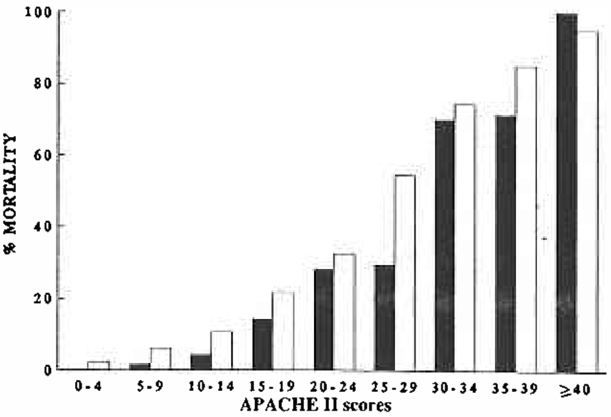


Fig 1 - The distribution of patients according to APACHE II score of 301 consecutive admissions at NUH-SICU.

Fig 2 - The relationship between observed and expected mortality according to APACHE II score on admission.
■ = observed mortality; □ = expected mortality



a plot of true positive ratio (expression of sensitivity) against the false positive ratio. The area under the ROC curve is an estimate of the correct classification rate of the prognostic scoring system. An accurate prognostic index has an area that approaches 1 that is, the largest area under the ROC curve.

RESULTS

There were 301 admissions from 1 July 1991 to 7 May 1992 inclusive, 174 males (57.8%) and 127 females (42.2%). The racial characteristics were: Chinese 73.4%, Indian 11.6%, Malay 10.7% and Others 4.0%. The mean age of patients admitted was 52.39 years (SD 19.3, range 13-93) and the mean duration of stay at the surgical intensive care was 5.37 days (SD 8.93, range 1-85). The primary systems affected which led to admission in SICU were: cardiovascular 24.6%, respiratory 25.6%, neurological 28.8%, gastrointestinal 13.6% metabolic and renal 4.32%, and other systems 3.65%. Although the surgical intensive care patients are generally from the surgical wards, 50 of the admissions (16.61%) were not postoperative patients and did not have surgery and anaesthesia during the first 24 hours of admission. Of all admissions, 91(30.23%) had emergency surgery.

In our series, there were 52 SICU deaths giving rise to a SICU mortality rate of 17.27%. Survivals were classified as those who were discharged from the SICU. It is the practice in our SICU not to discharge any terminally ill patients with hopeless prognosis and there were no deaths within 24 hours of discharge. We did not look into hospital discharge rates for patients admitted to our SICU as our purpose was to validate the APACHE II score in a surgical intensive care unit in Singapore and other factors beyond our control eg staff ratios, differences in training may influence the hospital outcome independent of ICU care. In patients who had surgery, 29 died (mortality rate 11.55%) whereas of the 50 who did not have any surgery, 23 died (mortality rate 46%). There were significant differences in mortality between those who had surgery and those who did not (chi square 34.62, d.f. 1, $p<0.001$). 19.78% of patients with emergency surgery died as opposed to 6.25% who had elective surgery (chi square 8.41 d.f. 1, $p>0.01$).

There were 12 readmissions in our series (3.98%), 5 of which were after elective surgery for postoperative monitoring and they were discharged 1 to 7 days later. A further 2 patients were readmitted after emergency laparotomy for acute surgical problems which developed in the ward, one of whom eventually succumbed to sepsis and died. The remaining 5 readmissions were for deterioration of clinical condition, 5 to 43 days after discharge from SICU all of whom eventually died.

The distribution of 301 admissions at various APACHE II groups is shown in Fig 1. The observed and predicted mortality rates at each APACHE II group is shown in Fig 2. At every APACHE II group except in those with APACHE scores more than 40, the observed mortality was less than the expected mortality. Correlation between the observed and expected mortality computed by linear regression test was $r = 0.9732$. The mean APACHE II score for survivals was 12.94 (SD 7.43) and non-survivals 28.19 (SD 10.43). Application of the unpaired Student's t test showed significant differences between the 2 means ($p<0.001$).

The receiver operating characteristic (ROC) curve for our data is shown in Fig 3. The area under the curve was computed to be 0.87. Using 0.5 as the decision criterion between survival and death, to look for applicability of APACHE II score for individual patients, 45 patients were predicted to die although the observed number of deaths in this group was 30. The best total correct classification was 89.4% at a decision criterion of 0.7 (sensitivity 50% and selectivity 97.6%).

DISCUSSION

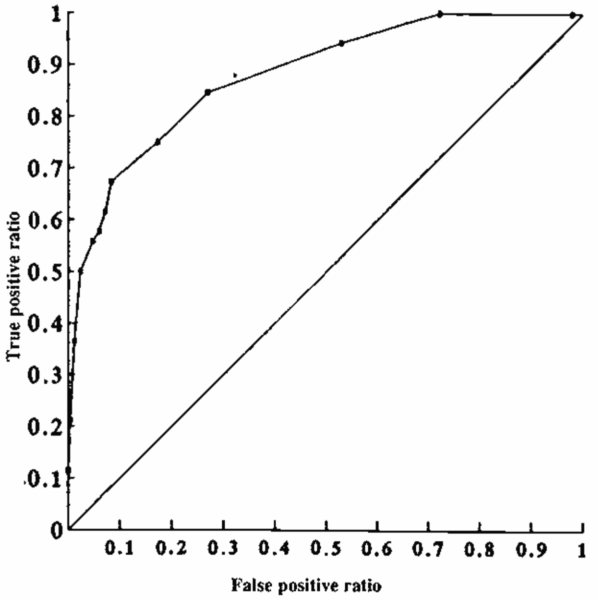
In developing the APACHE II model, Knaus and colleagues collected data from 5,815 patients from 13 medical centres in the US. Subsequent validation of the APACHE II system in different settings and populations have shown good correlation between the observed and expected mortality(6-8). The validation were mainly in general intensive care patients with mixed diagnoses rather than specifically for surgical patients in the American and European setting.

Applicability of the APACHE II system for surgical patients is controversial. In the population used to develop the APACHE II system, 785 patients who were post-coronary bypass graft patients were eliminated from the APACHE II study because "these patients represented a large group whose surgical and anaesthetic management resulted in high scores at ICU admission but very low hospital mortality rates"(9). This is typical of many postoperative surgical patients. The postoperative state is associated with an altered level of consciousness secondary to anaesthesia and residual paralysis is often present. This, together with the presence of an endotracheal tube, makes accurate determination of the Glasgow Coma Scale difficult, illustrating the difficulties encountered in using the APACHE II score for surgical patients(10). We adopted the practice proposed by Jacobs and colleagues, recording instead the "best Glasgow Coma Score value over 24 hours"(11).

Giangiuliani and colleagues used the APACHE II system on 598 consecutive surgical intensive care admissions and found good correlation between observed and expected outcome in their intensive care unit(12). Cerra and colleagues found that the admission APACHE II score did not predict the development of multiple organ failure and significantly underestimated the mortality rate for postoperative surgical patients(13).

Our findings suggest that there is good correlation between APACHE II scores and observed outcome in surgical intensive care patients on the whole at National University Hospital. Our mortality rates at various APACHE II scores were comparable with the results of intensive care units studied by Knaus and colleagues. We found, however, the predictive power for individual surgical patients limited. Using a predicted risk criterion of 0.5 as the cut-off between survival and death as suggested by Knaus and colleagues, the total correct prediction as 87.7%. Thirty out of 45 patients predicted to die died and 22 out of 256 patients expected to survive died. Our best correct prediction rate was 89.4% using a decision criterion of 0.7 to distinguish between survivors and non-survivors. This compares well with the results of Giangiuliani and colleagues who found the best total correct classification

Fig 3 - ROC curve analysis demonstrating the predictive ability of APACHE II system, based on 301 admissions at NUH-SICU. The diagonal line indicates an index that operates no better than chance.



of 84% using a decision criterion of 0.7 as risk of death. The trade off between false positive and false negative rates as the decision criterion varies is summarised by the ROC curve shown in Fig 3. The computed area under the ROC curve for our ICU data was 0.87 compared with 0.84 found by Giangiuliani in their surgical intensive care.

There were no survivals in our series when the APACHE II score was more than 40 and risk prediction more than 0.87. As there were only 8 admissions in this group, we feel that it would not be prudent to rely on this criterion to withdraw therapy. We found the mortality rates very low when the APACHE II score was less than 10. Only one death occurred in this group of 89 admissions.

Marks and colleagues in a prospective study of 568 patients admitted to a general intensive care unit compared the predictive abilities of APACHE II with subjective assessment by the ICU staff. He found the subjective assessments a more powerful predictor of individual outcome and concluded that although the APACHE II scores could be applied to the population as a whole, they were not suitable for predicting outcome in the individual patient(14).

The APACHE II system is easy to use. It has allowed us to compare our ICU with those in other institutions and to evaluate the performance of our surgical intensive care unit. It allowed stratification of patients into different degrees of severity of illness enabling therapeutic trials to be carried out. It does not , however, take into account modification of physiological variables induced by therapeutic interventions, which may not reflect organ dysfunction and could possibly alter the eventual score. Because the calculation of risk of death allows only one primary diagnostic reason for intensive care admission to be selected, we encountered difficulties in patients with multiple medical problems. We found the disease classification in APACHE II system imprecise and in a significant proportion of our patients who were in the intensive care for monitoring postoperatively without actual organ dysfunction eg patients for epidural narcotic pain relief, there were no appropriate specific diagnostic category to be selected and we had to choose a non-specific system dysfunction as the principal reason for admission. Although useful for retrospective assessment of intensive care performance and risk stratification, we found in agreement with other authors, the APACHE II system less satisfactory as a triage tool. Its prognostic power for the individual patients is limited and we would not base important clinical decisions on the APACHE II score.

The APACHE III system is now being developed to improve the risk prediction ability of the APACHE system by re-evaluating the selection and weightage of physiological variables, improving the precision of disease classification, identifying and quantifying factors in ICU care which contribute to variations in the final outcome, and increasing the database so that individual outcome prediction with narrow confidence limits could be obtained(15,16). Prospective data has been collected on 17,440 medical and surgical ICU admissions at 40 US hospitals randomly chosen to represent intensive care services. Initial validation of APACHE III by the authors have found that the overall predictive accuracy of the first day score was such that 95% of ICU admissions could be given a risk estimate of hospital death that was within 3% of observed mortality(17).

The exact details of the APACHE III system has yet to be published. Until the details are released to allow validation in our setting, caution has to be exercised in the application of any prognostic score to predict individual outcome. Triage decisions should continue to be based not on prognostic scores but rather on clinical judgement.

ACKNOWLEDGEMENTS

We would like to thank Ms Roslinda Yusoff for her help in data input and Ms Karen Ho for her help in the illustrations. We are indebted to the staff of the Department of Anaesthesia, National University Hospital for their co-operation and kind support.

REFERENCES

1. Knaus WA, Zimmerman JE, Wagner DP, Draper EA, Lawrence DE. APACHE - acute physiology and chronic health evaluation: a physiologically based classification system. Crit Care Med 1981; 9:591-7.
2. Wong DT, Knaus WA. Predicting outcome in critical care: the current status of the APACHE prognostic scoring system. Can J Anaes 1991; 38:374-83.
3. Knaus WA. Draper EA, Wagner DP, Zimmerman JE. APACHE II: a severity of disease classification system. Crit Care Med 1985; 13:818-29.
4. Hanley JA, Mcneil BJ. The meaning and use of the area under a receiver operating characteristic curve. Radiology 1982; 143:29-36.
5. McNeil BJ, Keeler E, Adelstein SJ. Primer on certain elements of medical decision making. N Engl J Med 1975; 293:211-5
6. Knaus WA, Draper EA, Wagner DP, Zimmerman JE. An evaluation of outcome from intensive care in major medical centres. Ann Intern Med 1986; 104:410-8.
7. Zimmerman JE, Knaus WA, Judson JA, Havill JH, Trubuhovich RV, Draper EA, et al. Patient selection for Intensive Care: A comparison of New Zealand and United States hospitals. Crit Care Med 1988; 16:318-26.
8. Kruse JA, Thill-Baharozian MC, Carlson RW. Comparison of clinical assessment with APACHE II for predicting mortality risk in patients admitted to a medical intensive care unit. JAMA 1988; 260:1739-42.
9. Dragsted L, Jorgensen J, Jensen N, Bansing E, Jacobsen E, Knaus WA, et al. Interhospital comparison of patient outcome from intensive care: Importance of lead time bias. Crit Care Med 1989; 17:418-22.
10. Rutledge R, Fakhry S, Rutherford E, Muakkasa F, Baker CC, Koruda M, et al. Acute Physiology and Chronic Health Evaluation (APACHE II) score and outcome in the surgical intensive care unit: An analysis of multiple intervention and outcome variables in 1238 patients. Crit Care Med 1991; 19:1048-53.
11. Jacobs S, Chang RWS, Lee B. One year experience with the APACHE II severity of disease of classification system in a general intensive care unit. Anaesthesia 1987; 42:738-44.
12. Giangiuliani G, Mancini A, Gui D. Validation of a severity of illness score (APACHE II) in surgical intensive care unit. Int Care Med 1989; 15:519-22.
13. Cerra FB, Negro F, Abrams J. APACHE II score does not predict multiple organ failure or mortality in postoperative patients. Arch Surg 1990; 212:216.
14. Marks RJ, Simons RS, Blizzard RA, Browne DRG. Predicting outcome in intensive care units - a comparison of APACHE II with subjective assessments. Intensive Care Med 1991; 17:159-63.
15. Knaus WA, Wagner DP. Individual patient decision. Crit Care Med 1989; 17:S204-9
16. Wagner DP, Draper EA, Knaus WA. Development of APACHE III. Crit Care Med 1989; 17:S199-203.
17. Knaus WA, Draper EA, Bergner M, Murphy DJ, Harrell FE. The APACHE III prognostic system. Risk prediction of hospital mortality for critically ill hospitalised adults. Chest 1991; 100:1619-36.