Intensive Care Med (2000) 26: 577–584

ORIGINAL

N. F. de Keizer
G. J. Bonsel
C. Goldfad
K. M. Rowan

The added value that increasing levels of diagnostic information provide in prognostic models to estimate hospital mortality for adult intensive care patients

Received: 24 September 1999
Final revision received: 24 January 2000
Accepted: 1 February 2000

N. F. de Keizer (✉)
Department of Medical Informatics,
J2–256, Academic Medical Center,
P. O. Box 22700, 1100 DE Amsterdam,
The Netherlands
e-mail: n.f.keizer@amc.uva.nl
Tel.: + 31-20-5 66 52 05
Fax: + 31-20-6 91 98 40

G. J. Bonsel
Department of Clinical Epidemiology
and Biostatistics, Academic Medical
Center, P. O. Box 22700,
1100 DE Amsterdam, The Netherlands

C. Goldfrad · K. M. Rowan
Intensive Care National Audit & Research
Centre, Tavistock House,
Tavistock Square,
London WC1H 9HR, UK

Abstract *Objective*: To investigate in a systematic, reproducible way the potential of adding increasing levels of diagnostic information to prognostic models for estimating hospital mortality.
Design: Prospective cohort study.
Setting: Thirty UK intensive care units (ICUs) participating in the ICNARC Case Mix Programme.
Patients: Eight thousand fifty-seven admissions to UK ICUs.
Measurements and results: Logistic regression analysis incorporating APACHE II score, admission type and increasing levels of diagnostic information was used to develop models to estimate hospital mortality for intensive care patients. The 53 UK APACHE II diagnostic categories were substituted with data from a hierarchical, five-tiered (type of condition required surgery or not, body system, anatomical site, physiological/pathological process, condition) coding method, the ICNARC Coding Method. The inter-rater reliability using the ICNARC Coding Method to code reasons for admission was good ($\kappa = 0.70$). All new models had good discrimination (AUC = 0.79–0.81) and similar or better calibration compared with the UK APACHE II model (Hosmer-Lemeshow goodness-of-fit H = 18.03 to H = 26.77 for new models versus H = 63.51 for UK APACHE II model).
Conclusion: The UK APACHE II model can be simplified by extending the admission type and substituting the 53 UK APACHE II diagnostic categories with nine body systems, without losing discriminative power or calibration.

Key words Intensive care · APACHE II · Prognostic models · Reason for admission · Mortality

Introduction

Patients admitted to intensive care units (ICU) form a heterogeneous group. Case mix (age, acute severity, comorbidity, surgical status and type of disease) varies widely. The development and application of models to estimate hospital mortality, defined as death before discharge from hospital following intensive care, is a growing field of research [1, 2, 3, 4, 5, 6, 7].

All prognostic models, such as the Acute Physiology And Chronic Health Evaluation (APACHE) I, II and III [8, 9, 10], the Simplified Acute Physiology Score (SAPS) I and II [11, 12] and the Mortality Prediction Models I and $II_{0, 24, 48, 72}$ [13, 14, 15], primarily rely on age, physiological variables and the existence of chronic, severe conditions as important explanatory variables. Some combine weightings for these variables into a score prior to inclusion into the model. In the APACHE II and III models, diagnosis or reason for admission is also included as an explanatory variable. A diagnostic category, defined as a simple dichotomous variable, is deemed to reflect information not explained by age, physiology and chronic illness.

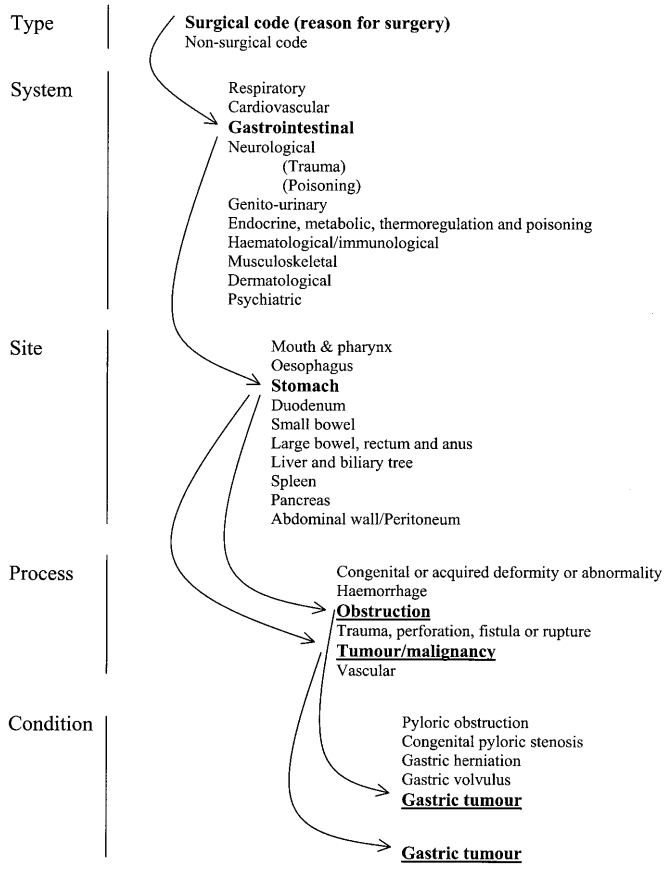

Fig. 1 Two "paths" to select the reason for admission for a patient whose ICU admission followed surgery for gastric tumour

The diagnostic categories included in the APACHE II and III models were derived from the literature and expert opinion, and are two-tiered. The first tier is based on the body system affected, for example, the respiratory system, the second tier on the specific process or condition, for example, infection. The second tier can be of varying level of detail from well defined conditions, for example diabetic ketoacidosis, to groups of conditions, for example, peripheral vascular surgery. Allocation of a patient to a diagnostic category is often troublesome due to the lack of unequivocal definitions. Where a patient cannot be coded to a specific process or condition, solely the first tier, the body system, is coded. This often results in a large group of heterogeneous admissions coded only to the first tier.

The purpose of this study was to investigate, in a systematic, reproducible way, the potential of adding increasing level of diagnostic information to a prognostic model. Systematic implies the use of a five-tiered, hierarchical coding method, the ICNARC Coding Method.

The ICNARC Coding Method was derived empirically from textual data describing the reason for admission for 10,806 patients from the Intensive Care Society's UK APACHE II study [1, 2]. It was developed and tested by a Working Group of seven senior, intensive care physicians and researchers involved in the Intensive Care National Audit & Research Centre (ICNARC) in the United Kingdom (UK), a centre established to undertake comparative audit and evaluative research of intensive care. The primary requirement for the method was to describe intensive care admissions better to enable future investigation of the explanatory power of diagnostic information in estimating hospital mortality. Due to its five-tiered hierarchy, the ICNARC Coding Method allows for stepwise analysis to investigate the potential value that each level of diagnostic information adds to a prognostic model. Each step allows inclusion of a new tier of information: tier one – type of condition required surgery or not; tier two – the body system; tier three – the anatomical site; tier four – the physiological or pathological process; and tier five – the condition or disease. The tiers are denoted by "type", "system", "site", "process" and "condition", respectively. At each of the five tiers, selection returns a unique code. The final code is the result of the five selections. The same conditions (on the fifth tier) can be the consequence of different paths (Fig. 1) and the 1140 final codes relate to 741 unique conditions. Each ICNARC Coding Method final code has been mapped to the original APACHE II diagnostic categories [9] and is currently being mapped to ICD-9-CM codes.

The first objective of this study was to investigate the inter-rater reliability of the ICNARC Coding Method and the second objective was to test the hypothesis that increasing levels of diagnostic information improve estimation of hospital mortality.

Methods

Data for investigating ICNARC Coding Method inter-rater reliability

To measure the inter-rater reliability of coding using the ICNARC Coding Method, a random sample of 43 admission records from two Dutch, mixed ICUs (one university hospital and one teaching hospital) were coded by two intensivists. They independently used the ICNARC Coding Method to code (multiple) reasons for admission of these 43 admissions, based on information available in the first 24 h of ICU admission.

Data for testing the added value of increasing level of diagnostic information

Data from the ICNARC Case Mix Programme in the UK was used for these analyses. The Case Mix Programme is the national, comparative audit of intensive care patient outcome [16]. The data made available for this study consisted of records, rendered anonymous, for 8057 consecutive admissions to 30 UK ICUs (from January 1997 to March 1998). Data included APACHE II score; APACHE II estimate of hospital mortality (based on the UK

APACHE II model); surgical status; ICNARC Coding Method codes for the reason for admission; and hospital outcome.

Data management

Before analysis of the added value of increasing level of diagnostic information, some rearrangement of data in the tiers of the ICNARC Coding Method was necessary. For convenience, tier 2 ("system") also contains three "systems" which do not correspond to conventional body systems. The psychiatric "system" is used to code for mental health conditions and has no anatomical tier. The trauma and poisoning "systems" are used to speed up the coding of those conditions given the frequency with which they occur. In the case where the "system" tier is trauma or poisoning, the hierarchy was changed from "type: process: system: site: condition" into the general order "type: system: site: process: condition". Paths with the psychiatric "system" were rearranged from "type: system: system: process: condition"into "type: system: *no site*: process: condition".

For each admission, a new category, "admission type", was derived. This categorised admissions into whether the admission was "elective surgical", "emergency surgical" or was "non-surgical". ICNARC Coding Method codes for the "site" and "process" tiers were aggregated to reduce the number of variables. Based on clinical judgement and the number of admissions per category, the 51 ICNARC Coding Method-alternatives for anatomical site were aggregated to 21 site-categories, for example the ICNARC Coding Method-sites "upper airway + trachea", "bronchi + airways" and "lungs" were aggregated into one site "airways". Similarly, the 56 physiological/pathological processes were aggregated into 15 process-categories, e.g. the ICNARC Coding Method-processes "infection", "inflammation" and "inflammation or intrauterine death" were aggregated into one process.

Statistical analysis for investigating the ICNARC Coding Method inter-rater reliability

Percentages of agreement and kappa statistics were calculated to measure the inter-rater reliability of the final code for the reason for admission using the ICNARC Coding Method. Different numerical final codes for the same condition were considered equal. Separate percentages of agreement and kappa statistics were calculated to measure inter-rater reliability at each tier of the code ("system", "site" and "process") of the selected condition.

Statistical analysis for testing the added value of diagnostic information

Following exclusion, either for missing data and/or for application of APACHE II inclusion criteria, the data were randomly split (50:50) into a training set and a test set for the purpose of unbiased comparison of accuracy of the resultant models. A number of prognostic models were developed by conventional stepwise logistic regression (inclusion 0.05, exclusion 0.10) on the training set. The dependent variable was hospital mortality and the independent variables were: APACHE II score, admission type (initially, post-emergency surgery or not, subsequently elective surgical, emergency surgical or non-surgical) and defined levels of diagnostic information using the ICNARC Coding Method. The UK APACHE II model served as the reference model. In the alternative models, the 53 UK APACHE II diagnostic categories were substituted with information from the ICNARC Coding Method tiers.

Models were first tested for robustness and stability by: (a) using forward and backward selection instead of stepwise selection; (b) inspecting the pattern of inclusion/exclusion of determinants and (c) using jackknife techniques.

Performance of the resultant models was judged by ROC curve techniques using the area under the curve as a measure of unweighted discriminative power of the prognostic model. The standard advantages and disadvantages of judgement by ROC curve comparison applied, that is, the absolute misclassification rate of a given sample of patients depended on the prevalence of the various prognostic categories and the threshold chosen to define a probability as predicting "alive" versus "dead". By computing the mortality ratio between the total number of observed and the total number of expected ultimate hospital deaths within a sample, the net misclassification rate within a sample can be estimated at the group level. The models' performance across admission type categories, severity categories, etc. was also tested, as uniform performance is necessary in practice. For this purpose, subgroup ROC curves and subgroup mortality ratios were calculated. To enable a fair comparison between the different models, we divided the mortality ratios for subgroups by 1.19, the overall mortality ratio for the UK APACHE II model on the UK data.

The Hosmer-Lemeshow goodness-of-fit H statistic [17] was used to evaluate the calibration of the models. All data analyses were performed with standard SPSS software version 8.0.

Results

The ICNARC Coding Method inter-rater reliability

The agreement between the two intensivists using the ICNARC Coding Method to code the reason for admission of 43 randomly selected Dutch ICU admissions was good on each level of diagnostic information (79 % agreement on the final code "condition", and 88 %, 81 %, 77 % on the levels "system", "site" and "process", respectively, $\kappa = 0.70$ on the final code "condition", and 0.77, 0.72, 0.66 on the levels "system", "site" and "process", respectively). Both observers noted that, in some cases, the ICNARC Coding Method codes and terms were either not as specific as they wished or too specific, which resulted in different choices of path and, occasionally, different resultant ICNARC Coding Method final codes. For example, a patient admitted after an aortic valve resection due to aorta stenosis was coded as "abnormality of aortic valve" by one physician (with process "congenital or acquired deformity") and as "chronic degeneration of aortic valve" by the other physician (with process "degeneration").

The added value of diagnostic information

The data were almost complete, only 20 admissions (0.02 %) missed ultimate hospital survival and 2 (0.002 %) admissions missed surgical status. Of the 8057 records, 1369 admissions (16.9 %) were not eligible for calculation of an APACHE II score and probability

Table 1 Variables available/selected for each model

	UK APACHE II model (reference)	Model A	Model B	Model C	Model D	Model E
Dependent variable	Hospital mortality	Hospital mortality	Hospital mortality	Hospital mortality	Hospital mortality	Hospital mortality
Explanatory variables fixed in the model	• APACHE II score • Post-emergency surgery or not • APACHE II diagnostic categories	• APACHE II score • Post-emergency surgery or not	• APACHE II score • Post-emergency surgery or not	• APACHE II score	• APACHE II score	• APACHE II score
Additional explanatory variables available for selection in stepwise logistic regression modelling	None	None	• 9 body systems	• Admission type • 9 body[1] systems	• Admission type[2] • 9 body systems[1] • 21 anatomical sites	• Admission type[2] • 9 body systems[1] • 21 anatomical sites[3] • 15 physiological/pathological processes[4]

[1] Body systems: respiratory/cardiovascular/gastrointestinal/neurological/genito-urinary/endocrine, matabolic, thermoregulation and poisoning/haematological, immunological/musculoskeletal/dermatological

[2] Admission type: elective surgical/emergency surgical/non-surgical

[3] Anatomical sites: blood vessels/heart/skin/endocrine organs/body composition, fluids and tissues/pancreas/abdominal wall or peritoneum/oesophagus, duodenum, stomach, large and small bowel/liver and biliary tree/mouth or pharynx/spleen/kidney, bladder or urethra/genitals/blood, marrow/muscles or connective tissue/pelvis, long bones or joints/vertebral column/brain, head and nerves/bronchi, trachea, lungs/chest wall/pleura or mediastinum

[4] Processes: intoxication/metabolic disturbance/collapse/coma/congenital or acquired deformity/dissection or aneurysm/organ failure/ haemorrhage/infection/obstruction/tachyarrhythmia/shock/trauma/tumour, malignancy/other

of hospital mortality as a result of application of the APACHE II exclusion criteria (age less than 16 years, ICU length of stay less than 8 h, readmission within the same hospital stay, admission for CABG or burns). After excluding admissions for missing data and application of APACHE II exclusion criteria, the data for 6,671 admissions remained, these were randomly split into equal-sized training set and test set.

The distribution of case mix (age, acute severity – APACHE II score and probability, co-morbidity – proportion with history of 16 serious, specified conditions, surgical status and reasons for admission – primary body system involved), outcome (ICU and hospital mortality) and activity (median ICU length of stay) were similar in the training and test sets.

Using the training set, five new models were developed to estimate hospital mortality. For all models, forward and backward selection resulted in the inclusion of the same variables. Jackknife estimation showed stability of the coefficients of selected variables. In model A, all variables from the original UK APACHE II model were used except for the 53 UK APACHE II diagnostic categories. In model B, the 53 UK APACHE II diagnostic categories were removed and nine body systems were available and were selected. In model C, post-emergency surgery or not was removed and the variable admission type "elective surgical", "emergency surgical" and "non-surgical" was available and was selected. In model D, 21 anatomical sites were available for selection in addition to the variables in model C. However, the anatomical sites were not selected during the stepwise logistic regression analysis. Hence, model D was equivalent to model C and was therefore excluded from any further analyses. In model E, 15 physiological/pathological processes were available and were selected as explanatory variables in addition to the variables available for selection in model D. Except for the anatomical sites available for model D, all available variables were selected during stepwise logistic regression analysis (Table 1).

Using the test set, the area under the ROC curve, the overall mortality ratio and the Hosmer-Lemeshow goodness-of-fit H statistic were calculated for each model (Table 2). All the new models had good discriminative power compared with the original UK APACHE II model. The area under the ROC curve varied between 0.79 and 0.81 (Fig. 2). Even model A, the simplest model, had only slightly less discriminative power compared with the original UK APACHE II model. The overall mortality ratio was similar for all models. Fig. 3 shows the calibration curves for each model.

Model C had the best calibration ($H = 18.05$, $df = 10$, $p = 0.05$) which was superior to the original UK APACHE II model ($H = 63.51$, $df = 10$, $p < 0.001$). Table 3 shows the number of patients in each diagnostic category used in model C.

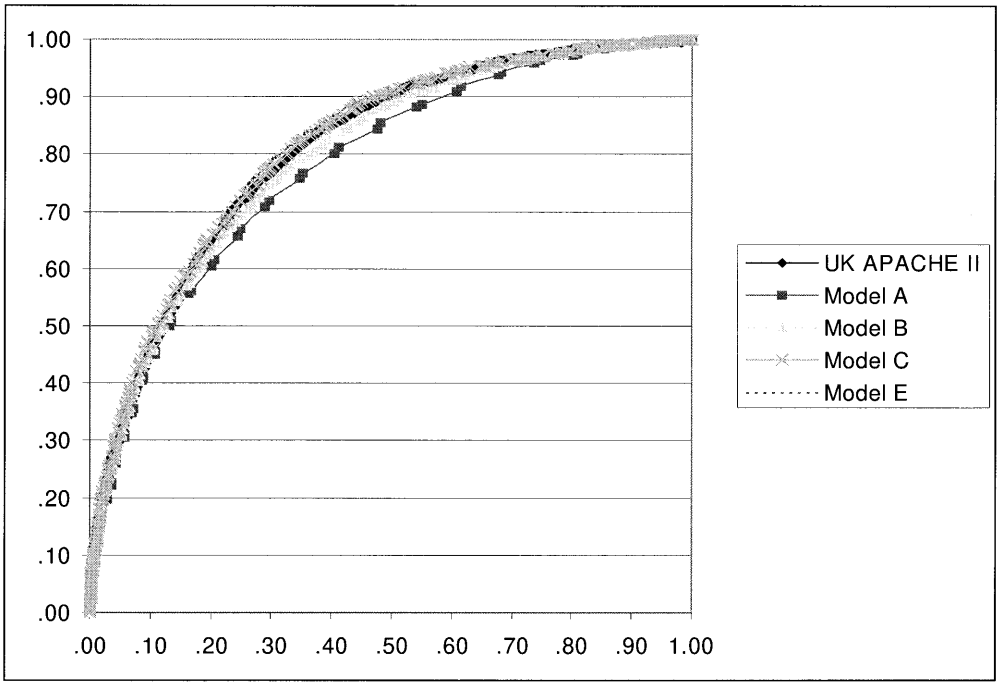

Fig. 2 ROC curves for the original UK APACHE II model and four new models

Using the test set, ROC curves and mortality ratios were calculated for subgroups for each model (Table 4). The discrimination (area under the ROC curve) for the three admission type subgroups was comparable among the models. Mortality ratio within the admission type subgroups were closer to 1.0 in model C and E compared with the original APACHE II model, model A and model B.

Discussion

The results of this research show that the inter-rater reliability for the empirically derived, five-tiered, hierarchical coding method, the ICNARC Coding Method, was good both overall and for each tier. Using the ICNARC Coding Method, the added value that increasing levels of diagnostic information provide in prognostic models to estimate hospital mortality for adult, intensive care patients was investigated. A widely accepted and well-described prognostic model, APACHE II (recalibrated for the UK in 1988–1990), served as the reference. Surprisingly, the additional value of increasing the level of detail of diagnostic information was low. However, a systematic disease grouping, based on body system (9 categories), was slightly superior to the 53 UK APACHE II categories, resulting in better calibration and discrimination. More detailed diagnostic information added little to this result. The Hosmer-Lemeshow H-statistic of the new models was lower (thus calibration was better) than for the original UK APACHE II model because the fit for low and middle risk patients, the largest group in the population, appeared to be improved.

Although these data did not encompass all ICUs in the UK, we believe these findings cannot be explained by uneven composition of cases as the ICUs were selected neither on the basis of the mix of admissions nor on

Table 2 Performance of models

	UK APACHE II model (reference)	Model A	Model B	Model C	Model E
Cox and Snell R^2	–	0.22	0.24	0.26	0.27
Area under ROC curve	0.81	0.78	0.79	0.81	0.81
Mortality ratio	1.19	1.03	0.96	0.97	0.96
Hosmer Lemeshow H-statistic	63.51	21.53	26.77	18.05	20.45

Model D was excluded because it was equivalent to model C

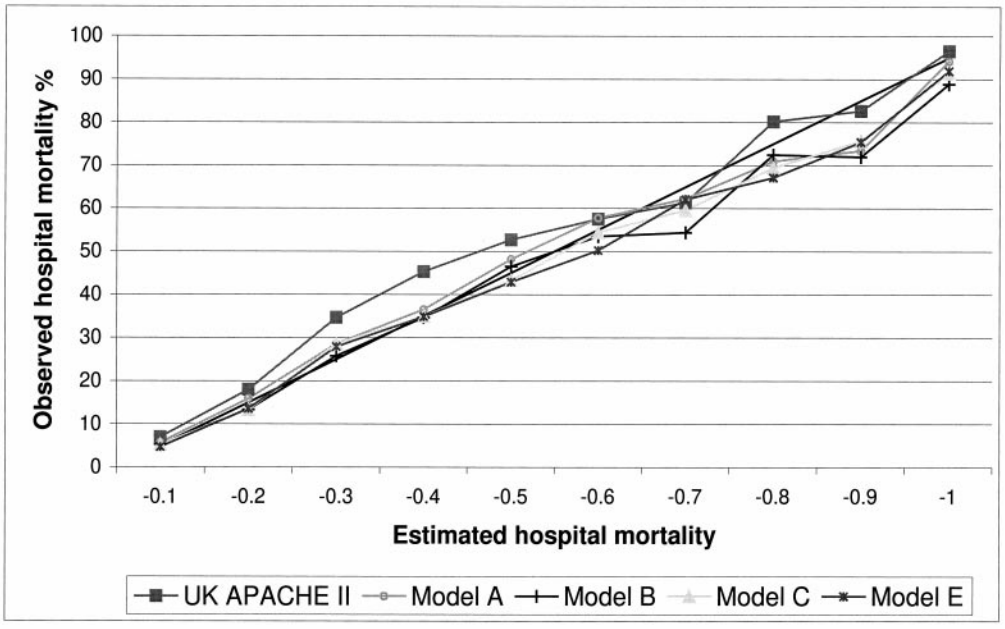

Fig. 3 Calibration curves for the original UK APACHE II model and four new models

ICU performance. Poor accuracy (validity and reliability) of diagnostic coding could explain the low additional value of more detailed diagnostic information, but misclassification at the higher levels of the ICNARC Coding Method seems unlikely, as the assignment to system, for example gastrointestinal or neurological system, was straightforward. Moreover, coding reliability with the ICNARC Coding Method was tested and appeared to be good, empirically. The inter-rater reliability was tested on Dutch patient records, which may not have the same characteristics as UK records. However, we do not believe this will lead to biased results, as reasons for admission are recorded in the same way in both countries and the use of the ICNARC Coding Method is unambiguous and well-structured to code the reason for admission. Furthermore, the Dutch intensivists performing this study followed the same data collection and coding training as all ICNARC Case Mix Programme participants.

The UK APACHE II model [3], was used as the reference model to compare with the new models so that a biased comparison of internal developed models by an external developed model was avoided. However, the UK APACHE II model was recalibrated on old data from a separate data collection period for 10,806 admissions to 26 ICUs during 1988–1990. Given the development of intensive care medicine, in terms of organisation and practice, it is probable that the performance of an old model is not optimal nowadays. Research is underway at ICNARC (a recently commenced

Table 3 Number of patients per diagnostic category used in model C

	Elective surgery	Emergency surgery	Non surgical	Total
Respiratory	77	81	1142	1300
Cardiovascular	639	367	1238	2244
Gastrointestinal	398	603	300	1301
Neurological	142	212	633	987
Genito-urinary	94	30	110	234
Endocrine, Metabolic, Thermoregulation and poisoning	21	3	120	144
Haematological/Immunological	0	3	58	61
Musculoskeletal	64	65	41	170
Dermatological	0	6	4	10
No system	0	0	220	220
Total	1435	1370	3866	6671

Table 4 Performance across subgroups for each model

	Number of admissions	UK APACHE II model (reference)	Model A	Model B	Model C	Model E
Area under the ROC curve						
Elective surgical	810	0.72	0.74	0.74	0.74	0.74
Emergency surgical	706	0.76	0.76	0.77	0.77	0.77
Non-surgical	1940	0.81	0.78	0.79	0.79	0.80
Mortality ratio[1]						
Elective surgical	810	0.73 (0.62–0.84)	0.52 (0.42–0.63)	0.52 (0.42–0.64)	0.94 (0.76–1.15)	0.93 (0.76–1.14)
Emergency surgical	706	1.05 (0.96–1.15)	1.37 (1.20–1.55)	1.00 (0.88–1.14)	0.96 (0.84–1.09)	0.95 (0.84–1.08)
Non-surgical	1940	1.03 (0.97–1.08)	1.09 (1.01–1.17)	1.08 (1.00–1.17)	0.98 (0.90–1.05)	0.97 (0.89–1.04)

Model D is excluded because it was equivalent to model C

[1] To achieve a fair comparison of the models, mortality ratios for subgroups were divided by 1.19 (based on UK APACHE II model)

2-year study) to compare all the current published methods leading to development and testing of the optimal method(s).

APACHE II was chosen as the reference model over other prognostic models such as SAPS II [12] or $MPM_{0/24}$II [14] because APACHE II was the only model that used reason for admission as an independent variable. The developers of SAPS II and $MPM_{0/24}$II concluded that it was not useful to include reason for admission as an independent variable because of the lack of unequivocal definitions for diagnostic categories and the complexity of selecting one reason for admission. This study suggested that a model incorporating nine body systems for the reason for admission, in which misclassification is very unlikely, performed well. We acknowledge, however, that the aggregation of sites and processes, to reduce the number of independent variables in models D and E, would not cause misclassification but may have led to loss of information.

Although prognostic models such as APACHE II cannot be used to predict individual hospital mortality risk, there is no obstruction to the aggregated use of these models, for example, for comparative audit and evaluative research to investigate the impact of the organisation and practice of intensive care on outcome. Based on our results, it appears that the current UK APACHE II model could be simplified to model C; the simplest model with the best performance. However further research is required before undertaking this change. Four approaches for further research may be suggested. The APACHE II score was left unchanged but the assumption that the explanatory power of age, physiology and chronic illness is optimal in the score needs further exploration. The impact of diagnostic information might increase after optimisation of the score.

The second approach involves the exploration of new, explanatory factors at the patient level, for example, new markers for physiological deterioration in the light of more recent diagnostic or therapeutical opportunities. A third approach is the investigation of dynamic (or trend) data instead of static data to estimate, for example, day-1 mortality separate from hospital mortality. The fourth approach explores the power of (alternative) disease classifications. Traditional classifications such as ICD-9-CM/ICD-10 [18, 19], Read [20] or SNOMED [21] are not appropriate for intensive care. Their underlying structure does not support the aggregation of unambiguous and complete diagnostic categories based on features of diagnostic concepts, for example, all *infectious* diseases which are *located in the gastrointestinal system* and which are *caused by a virus*. The structure of the ICNARC Coding Method supports the aggregation of diagnostic information at different levels of detail. However, the ability to select the same condition via different paths, for example, gastric tumour (Fig. 11), does not clarify whether a patient coded with a gastric tumour selected via "tumour/malignancy" had an obstruction or not. In other words, it is not clear whether a specific characteristic, such as obstruction, is an implicit or explicit characteristic of this diagnosis. Further investigation as to whether the physiological variables in these models provide sufficient explanatory power for such characteristics of a particular diagnosis is needed.

In addition, the level of detail of the ICNARC Coding Method needs scrutiny. For daily care, this level may be insufficient, for example it is impossible to compose complex diagnoses based on basic diagnoses such as "decompensatio cordis" due to "old anterior myocardial infarction". A semantic network, instead of a strict hierarchy, as the underlying structure might improve the utility of diagnostic information [22]. Modified to the intensive care domain, this network might support further exploration of the value of diagnostic information to estimate hospital mortality and, in time, to esti-

mate other outcome measures, such as functional health status [23] and costs [24]. Irrespective of the level of detail required for estimating prognosis, detailed diagnostic information will always be essential to describe and stratify populations of intensive care patients.

Acknowledgements All participants in the ICNARC Case Mix Programme. We thank Rob Bosman and Evert de Jonge for coding the Dutch patient records and we thank the reviewers for their helpful comments and suggestion on an earlier version of this paper.

References

1. Rowan K, Kerr J, Major E, McPherson K, Short A, Vessey M (1993) Intensive Care Society's APACHE II study in Britain and Ireland – I: Variation in case mix of adult admissions to general intensive care units and impact on outcome. BMJ 307: 972–976
2. Rowan K, Kerr J, Major E, McPherson K, Short A, Vessey M (1993) Intensive Care Society's APACHE II study in Britain and Ireland – II: Outcome comparisons of intensive care units after adjustment for case mix by the American APACHE II method. BMJ 307: 977–981
3. Rowan K, Kerr J, Major E, McPherson K, Short A, Vessey M (1994) Intensive Care Society's Acute Physiology and Chronic Health Evaluation (APACHE II) study in Britain and Ireland: a prospective, multicenter, cohort study comparing two methods for predicting outcome for adult intensive care patients. Crit Care Med 22: 1392–1401
4. Rivera-Fernandez R, Vazques-Mata G, Bravo M, et al. (1998) The APACHE III prognostic system: customized mortality prediction for Spanish ICU patients. Intensive Care Med 24: 574–581
5. Baston P, Sun X, Wagner D, Knaus W, Zimmerman J, Group TBAIS (1996) Application of the APACHE III prognostic system in Brazilian intensive care units: a prospective multicenter study. Intensive Care Med 22: 564–570
6. Beck D, Taylor B, Millar B, Smith G (1997) Prediction of outcome for intensive care: a prospective cohort study comparing Acute Physiology and Chronic Health Evaluation II and III prognostic scoring systems in a United Kingdom intensive care unit. Crit Care Med 25 (1): 9–15
7. Moreno R, Morais P (1997) Outcome prediction in intensive care: results of a prospective, multicentre, Portuguese study. Intensive Care Med 23: 177–186
8. Knaus W, Zimmerman J, Wagner D, Draper E, Lawrence D (1981) APACHE – Acute Physiology And Chronic Health Evaluation: a physiologically based classification system. Crit Care Med 9 (8): 591–597
9. Knaus W, Draper E, Wagner D, Zimmerman J 1985) APACHE II: a severity of disease classification system. Crit Care Med 13: 818–829
10. Knaus W, Wagner D, Draper E, et al. (1991) The APACHE III Prognostic System. Risk prediction of hospital mortality for critically ill hospitalized adults. Chest 100: 1619–1636
11. Le Gall J, Loirat P, Alperovitch A, et al. (1984) A Simplified Acute Physiology Score for ICU patients. Crit Care Med 12 (11): 975–977
12. Le Gall J, Lemeshow S, Saulnier F (1993) A new Simplified Acute Physiology Score (SAPS II) based on a European/North American multicenter study. JAMA 270: 2957–2963
13. Lemeshow S, Teres D, Pastides H, Spitz Avrunin J, Steingrub J (1985) A method for predicting survival and mortality of ICU patients using objectively derived weights. Crit Care Med 13 (7): 519–526
14. Lemeshow S, Teres D, Klar J, Avrunin J, Gehlbach S, Rapoport J (1993) Mortality Probability Models (MPM II) based on an international cohort of intensive care unit patients. JAMA 270: 2478–2486
15. Lemeshow S, Klar J, Teres D, et al. (1994) Mortality probability models for patients in the intensive care unit for 48 or 72 hours: A prospective, multicenter study. Crit Care Med 22 (9): 1351–1358
16. ICNARC (1995) ICNARC Case Mix Programme Dataset Specification. Intensive Care National Audit & Research Centre, London
17. Hosmer D, Lemeshow S (1989) Applied Logistic Regression. Wiley, New York
18. International Classification of Diseases, manual of the International Statistical Classification of diseases, injuries and causes of death: 9th revision. WHO, 1977
19. International Classification of Diseases, manual of the International Statistical Classification of diseases, injuries and causes of death: 10th revision. WHO, 1993
20. Robinson D, Comp D, Schulz E, Brown P, Price C (1997) Updating the Read Codes: user-interactive maintenance of a dynamic clinical vocabulary. J Am Med Inform Assoc 4 (6): 465–472
21. Rothwell D, Coté R (1996) Managing information with SNOMED: Understanding the model. Proc AMIA Annu Fall Symp, pp 80–83
22. De Keizer N, Abu-Hanna A, Cornet R, Zwetsloot-Schonk J, Stoutenbeek C (1999) Analysis and design of an ontology for intensive care diagnoses. Methods Inf Med 38: 102–112
23. De Keizer N, Bonsel G, Gemke R (1997) Health status prediction in critically ill children: A pilot study introducing standardized health ratios. Qual Life Res 6: 192–199
24. Edbrook D, Stevens V, Hibbert C, Mann A, Wilson A (1997) A new method of accurately identifying costs of individual patients in intensive care: the initial results. Intensive Care Med 23: 645–650