OUTCOME PREDICTION IN INTENSIVE CARE WITH SPECIAL REFERENCE TO CARDIAC SURGERY

John Scott Turner MBChB MMed (Cape Town) FCP (SA)

A thesis submitted to the University of Cape Town for the degree of Doctor of Medicine.

Declaration

I hereby empower the University of Cape Town to reproduce for the purpose of research either the whole or any portion of the contents of this thesis in any manner whatsoever.

I, John Scott Turner, hereby declare that the work on which this thesis is based is original (except where acknowledgements indicate otherwise) and that neither the whole work nor any part of it has been, is being, or is to be submitted for another degree in this or any other University.

Signed Signed by candidate
signature removed

Date 18.8.94

Dedication

To Roseanne, to my mother, and to my late father.

Acknowledgements

To Professor Peter Potgieter for his encouragement and to all the Respiratory ICU registrars (including me!) at Groote Schuur Hospital for the data they collected.

To Drs Cliff Morgan and Yugan Mudaliar for the data they collected for the second study in the Adult ICU of the Royal Brompton National Heart and Lung Hospital.

To Dr Rene Chang for his Riyadh Intensive Care Unit program and his advice on data analysis.

To Dr Cliff Morgan again for inviting me, and to the Doverdale Fellowship for generous funding during 1990 and 1993.

To the nurses and biomedical engineering people in the Adult ICU for their help with the subtleties of CareVue.

To Dr Tim Evans for his help with the process of writing it all up, and for the example he set in academic excellence.

To the late Vic Aber for his relaxed statistical advice.

To Dr Bharat Thakrar for agreeing to help a desperate man with the statistical analysis of the third study.

To Professor Rosemary Hickman for her encouragement and the final stamp of approval.

And finally (I admit I cribbed this from someone else's acknowledgements), while all these people have encouraged or advised me, any errors or heresies are entirely my own.

ABSTRACT

The development, use, and understanding of severity of illness scoring systems has advanced rapidly in the last decade; their weaknesses and limitations have also become apparent. This work follows some of this development and explores some of these aspects. It was undertaken in three stages and in two countries. The first study investigated three severity of illness scoring systems in a general Intensive Care Unit (ICU) in Cape Town, namely the Acute Physiology and Chronic Health Evaluation (APACHE II) score, the Therapeutic Intervention Scoring System (TISS), and a locally developed organ failure score. All of these showed a good relationship with mortality, with the organ failure score the best predictor of outcome. The TISS score was felt to be more likely to be representative of intensiveness of medical and nursing management than severity of illness. The APACHE II score was already becoming widely used world-wide and although it performed less well in some diagnostic categories (for example Adult Respiratory Distress Syndrome) than had been hoped, it clearly warranted further investigation. Some of the diagnosis-specific problems were eliminated in the next study which concentrated on the application of the APACHE II score in a cardiothoracic surgical ICU in London. Although group predictive ability was statistically impressive, the predictive ability of APACHE II in the individual patient was limited as only very high APACHE II scores confidently predicted death and then only in a small number of patients. However there were no deaths associated with an APACHE II score of less than 5 and the mortality was less than 1% when the APACHE II score was less than 10.

Finally, having recognised the inadequacies in mortality prediction of the APACHE II score in this scenario, a study was undertaken to evaluate a novel concept: a combination of preoperative, intraoperative, and postoperative (including APACHE II and III) variables in cardiac surgery patients admitted to the same ICU. The aim was to develop a more precise method of predicting length of stay, incidence of complications, and ICU and hospital outcome for these patients. There were 1008 patients entered into the study. There was a statistically significant relationship between increasing Parsonnet (a cardiac surgery risk prediction score), APACHE II, and APACHE III scores and mortality. By forward stepwise logistic regression a model was developed for the probability of hospital death. This model included bypass time, need for inotropes, mean arterial pressure, urea, and Glasgow Coma Scale. Predictive performance was evaluated by calculating the area under the receiver operating characteristic (ROC) curve. The derived model had an area under the ROC curve 0.87, while the Parsonnet score had an area of 0.82 and the APACHE II risk of dying 0.84. It was concluded that a combination of intraoperative and postoperative variables can improve predictive ability.

Table of contents

Title page i

Declaration ii

Dedication iii

Acknowledgements iv

Abstract vi

Table of contents viii

CHAPTER 1

Introduction and aims of the thesis 1

1.1 Introduction and general comments 1

1.2 Aims of the thesis 3

  1.2.1 Study 1 3

  1.2.2 Study 2 3

  1.2.3 Study 3 4

CHAPTER 2

Literature review. Severity scoring in general intensive care 5

2.1 Scope and mode of the literature review 5

  2.1.1 Scope of the literature review 5

  2.1.2 Mode of the literature review 6

2.2 History of scoring systems 6

  2.2.1 Basic principles of predictive indices 6

  2.2.2 Defining the outcome measure 8

  2.2.3 Early scoring systems 9

2.3 Rationale and ethics of scoring systems 10

2.3.1 Rationale of scoring systems 10

2.3.2 Prediction and probability 11

2.3.3 Ethics and the application of severity scoring systems 12

2.4 Principles of development of scoring systems 14

2.4.1 Dynamic vs static scores 15

2.4.2 Physiological vs binary variable scores 16

2.5 APACHE: development and refinements 17

2.5.1 APACHE 17

2.5.2 APACHE II 17

2.5.2.1 Other uses of APACHE II 18

2.5.2.1.1 Standardised mortality ratio (SMR) 18

2.5.2.1.2 Identification of low risk patients 19

2.5.2.1.3 Cost analysis and quality control 20

2.5.2.1.4 Allocation of funding 20

2.5.2.2 Criticisms and problems with APACHE II 21

2.5.2.2.1 Interpretation of the methodology 21

2.5.2.2.2 Interobserver variability 21

2.5.2.2.3 Timing of scores 21

2.5.2.2.4 Use of admission scores 22

2.5.2.2.5 Source of ICU admission 23

2.5.2.2.6 Response to therapy 24

2.5.2.2.7 Patient subgroups 25

2.5.3 Use of APACHE in general ICU settings 25

2.5.4 Use of APACHE II in specific areas 26

2.5.5 Derivatives of APACHE II 29

  2.5.5.1 Daily trend plus organ-system failure 29

  2.5.5.2 Sickness Score 30

  2.5.5.3 Rapid Acute Physiology Score 31

2.5.6 APACHE III 31

  2.5.6.1 Introduction 31

  2.5.6.2 Development of APACHE III 32

  2.5.6.3 Differences from APACHE II 33

  2.5.6.4 The APACHE III predictive equation 33

  2.5.6.5 Ongoing research 34

  2.5.6.6 Applications of APACHE III 34

2.5.7 Scoring systems and clinical judgement 36

  2.5.7.1 General ability of physicians 36

  2.5.7.2 APACHE and clinical judgement 38

2.6 Other severity of illness scoring systems 38

2.6.1 Simplified Acute Physiological Score (SAPS) 38

  2.6.1.1 SAPS II 40

2.6.2 The Mortality Prediction Model (MPM) 41

  2.6.2.1 MPM II 42

2.6.3 The organ-system failure (OSF) score 44

2.6.4 The ODIN (organ dysfunction and/or infection) model 45

2.6.5 Therapeutic Intervention Scoring System (TISS) 46

2.6.6 The system outcome score (SOS) and outcome index 48

2.6.7 Clinical sickness score 50

2.6.8 Simpler indicators of outcome 50

2.6.8.1 The Computerised Severity Index 50

2.6.8.2 The comorbidity score 51

2.6.8.3 The McCabe-Jackson scoring system 51

2.6.8.4 The ASA classification 51

2.6.8.5 The Visceral System Failure Index 52

2.6.8.6 The Hanover Intensive System 52

2.6.8.7 Critical Care Scoring System 53

2.6.8.8 POSSUM 53

2.7 New scoring systems 54

2.8 Specialised scoring systems for specific disorders 54

2.8.1 Trauma 54

2.8.2 Paediatrics 55

2.8.3 Pancreatitis 55

2.8.4 Burns 55

2.8.5 Septic shock 56

2.8.6 Non-traumatic coma 56

2.8.7 Comparison of general and specific scoring systems 56

2.9 Summary 56

CHAPTER 3

Literature review: severity scoring and predictive indices in cardiac surgery 58

3.1 General scoring systems in cardiac surgery 58

3.1.1 Introduction 58

3.1.2 APACHE II 58

3.1.3 APACHE III 59

3.1.4 SAPS II, MPM I and II 60

3.2 Cardiac surgery operative risk systems 60

3.2.1 Introduction 60

3.2.2 Interpretation of the literature 61

3.2.3 Independent validation 62

3.2.4 Collaborative Study in Coronary Artery Surgery (CASS) 63

3.2.5 Other large studies 64

3.2.6 Patients with an intraaortic balloon pump 65

3.2.7 Parsonnet score 65

3.2.7.1 Problems with the Parsonnet score 67

3.2.7.1.1 Methodology 67

3.2.7.1.2 Allocation of points 67

3.3 Individual risk factors in CABG surgery 68

3.3.1 Gender 68

3.3.2 Age 69

3.3.3 Impaired left ventricular function 70

3.3.4 Reoperation 70

3.3.5 Diabetes mellitus 71

3.3.6 Emergency surgery 72

3.3.7 Obesity 72

3.3.8 Confounding issues 72

3.4 Summary 73

CHAPTER 4

Methods 74

4.1 Introduction 74

4.2 Type of study 74

4.3 Scoring systems 74

    4.3.1 APACHE II 74

        4.3.1.1 Problems with the APACHE II methodology 76

        4.3.1.2 Management of methodological problems 76

    4.3.2 APACHE III 77

        4.3.2.1 Management of methodological problems 79

4.4 Computer hardware 80

    4.4.1 Apple (Apple Computer Inc.) 80

    4.4.2 IBM (International Business Machines Corporation) 80

    4.4.3 VAX mainframe computer (Digital Equipment Corporation) 80

4.5 Computer software 81

    4.5.1 dBASE II 81

    4.5.2 dBASE IV 81

    4.5.3 Riyadh ICU Program (RIP) 81

    4.5.4 Paradox 82

    4.5.5 ObjectVision 82

    4.5.6 Quattro Pro 82

    4.5.7 Epistat 83

    4.5.8 Instat 83

    4.5.9 StatGraphics 83

4.5.10 BMDP 83

4.6 Statistical methods 84

4.6.1 Statistical tests 84

4.6.2 Statistician 84

CHAPTER 5

Study 1 85

5.1 Introduction 85

5.1.1 Aims of this study 85

5.2 Protocol 86

5.2.1 Site 86

5.2.2 Entry criteria 87

5.2.3 Exclusion criteria 87

5.2.4 Data collection 87

5.2.5 Therapeutic Intervention Scoring System (TISS) 88

5.2.6 Organ failure score (Cape Town) 92

5.3 Results 94

5.3.1 Demographics 94

5.3.2 Scoring systems and mortality 94

5.4 Discussion 99

5.5 Conclusions 101

CHAPTER 6

Study 2 103

6.1 Introduction 103

6.1.1 Aims of this study 104

6.2 Protocol 104

6.2.1 Site 104

6.2.2 Entry criteria 105

6.2.3 Exclusion criteria 105

6.2.4 Data collection 105

6.3 Results 106

6.3.1 Demographics 106

6.3.2 APACHE II scores and mortality 108

6.3.3 Chronic health 111

6.3.4 Ventilation 111

6.3.5 Duration of admission 111

6.3.6 Emergency surgery 112

6.3.7 CABG patients only 112

6.4 Discussion 114

6.5 Conclusions 117

CHAPTER 7

Study 3 118

7.1 Introduction 118

7.1.1 Aims 120

7.2 Protocol 120

7.2.1 Site 120

7.2.2 Hewlett-Packard CareVue 9000 data management system 121

7.2.2.1 Introduction 121

7.2.2.2 Hardware 123

7.2.2.3 Software and facilities 124

7.2.2.4 Implementation 126

7.2.3 Entry criteria 126

7.2.4 Exclusion criteria 127

7.2.5 Ethical approval 127

7.2.6 Variable selection 127

7.2.7 Definitions for complications 128

7.2.8 Data entry 129

7.2.8.1 Computer software and program 133

7.2.9 Missing data 136

7.2.10 Outcome measures 136

7.2.11 Statistical analysis 136

7.2.11.1 Planning of the study 136

7.2.11.2 Statistical methodology and tests 137

7.3 Results 138

7.3.1 Demographics 138

7.3.1.1 All ICU admissions 138

7.3.1.2 Study patients 139

7.3.2 Selected single variables and mortality 142

7.3.3 Scoring systems and mortality 142

7.3.4 Bypass and ischaemic times and mortality 144

7.3.5 Simple combinations of variables and mortality 146

7.3.6 Univariate logistic regression 146

7.3.7 Forward stepwise logistic regression 148

7.3.8 ROC curve analysis 149

7.3.9 Subgroup of CABG patients 153

7.3.10 Complications 153

7.3.11 Length of stay in ICU 155

7.3.12 APACHE II vs APACHE III scores 156

7.3.13 Cardiac surgery patients not admitted to ICU 157

7.4 Discussion 157

7.5 Conclusions 162

CHAPTER 8

Conclusions, future directions, publications, presentations 163

8.1 Conclusions 163

8.2 Future directions 164

8.3 Publications related to this research 164

8.4 Abstracts related to this research 165

8.5 Presentations to learned societies related to this research 166

CHAPTER 9

References 168

APPENDIX 203

CHAPTER 1

Introduction and aims of the thesis

1.1 Introduction and general comments

Clinical prediction is an ancient medical art that has only recently developed a scientific aspect. This has been largely shaped by three elements, namely the construction of large patient databases, the use of multiple logistic regression and other statistical tests to analyse these databases, and the ready availability of microcomputers and sophisticated software to perform these analyses (Ephron & Tibshirani 1991). In the last decade predictive indices have proliferated, mainly in conjunction with or as an offshoot of severity of illness scoring.

Outcome prediction has become particularly important in critical care medicine with patients and their families expecting a high degree of accuracy in this respect, although accuracy has probably become more difficult in recent years as modern life-support technology becomes more successful in delaying death than restoring health. It has therefore become more important to stratify patients at an early stage into those with reversible conditions who can be saved and those who will inevitably die. In addition the cost of medical care is extremely high, with the mean daily cost of survivors in an Intensive Care Unit (ICU) in the United Kingdom (the country where two parts of this study were carried out) being £550 and that nonsurvivors £816 (Ridley et al 1993). Cost has increased dramatically in the last decade (Jacobs & Noseworthy 1990), with much of the expenditure occurring in the ICU setting and in the last few weeks or months of life. More recently cost has been highlighted by the development of new and expensive drugs and life support devices, such as monoclonal antibodies for sepsis, and Intravenacaval Oxygenation (IVOX) for ARDS. Many more will surely follow.

These developments have emphasised the need for accurate prediction of severity of disease, both for optimal patient care and appropriate resource utilisation in an era of shrinking medical funding, and also for acceptable scientific evaluation of these new therapeutic modalities. In addition, it has been postulated that we are in the midst of the "third revolution" in medical care, namely the era of assessment and accountability (Relman 1988). Severity scoring systems have been developed in this setting; they are now regarded by many as prognostic indicators and there has inevitably been much confusion and unrealistic expectation about their abilities which has led to inappropriate use in many instances. It is not generally realised that the technology is still evolving, and that the needs of physicians may have moved forward faster than the technology and its understanding. Indeed, although predictions for patient groups are good, the effectiveness of severity scores in predicting individual outcome has been clearly shown to be less useful (Schafer et al 1990).

Clinical judgement has long been the benchmark by which medical decision making has been measured. This has been based upon clinical experience, from which many valuable medical traditions and clinical aphorisms have arisen. It is therefore not surprising that predictive indices have met with some suspicion and there has been both reluctance to use them and fears that their use will negate clinical skills. This is not entirely unjustified as predictive indices will never be able to cater for all aspects of the clinical picture, will not always be able to assign the appropriate importance to certain variables, may not precisely fit the clinical scenario, and may not take into account newer therapies. They also cannot take account of common sense or "gut feeling". Although it is reassuring to have numerical data and accurate statistics to back a decision, it is important to remember that (as commented in an editorial on TPN and APACHE II scoring in *The Lancet*), "statistics should be used as the drunken man uses the lamppost - for support rather than illumination" (Anonymous 1986).

1.2 Aims of the thesis

1.2.1 Study 1

The first study aimed to evaluate the application and usefulness of three severity of illness scoring systems in a general ICU. These included the APACHE II score (which was newly developed at the time), TISS (recently updated at the time), and an organ failure score which had been empirically derived locally.

1.2.2 Study 2

Recognising some of the diagnosis-specific problems of severity scoring systems, the second study aimed to take the theme of APACHE II scoring further by evaluating its predictive value in more detail and in a more homogeneous setting, namely a cardiothoracic ICU.

1.2.3 Study 3

Having found difficulties in predicting outcome in individual cardiothoracic surgery patients, the third study aimed to use a novel combination of preoperative, intraoperative, and postoperative predictive indices (including APACHE II and III scores) to develop a better prognostic index for cardiac surgical patients, and to develop a predictive index for length of ICU stay and for the occurrence of complications.

CHAPTER 2

Literature review. Severity scoring in general intensive care

2.1 Scope and mode of the literature review

2.1.1 Scope of the literature review

This literature review concentrates on the more frequently cited references in major medical journals in two main areas, namely general scoring systems in Intensive Care (discussed in Chapter 2) and predictive indices for cardiac surgery (discussed in Chapter 3). It examines their history, rationale, ethical applications, principles of development, and clinical and research utilisation. The emphasis on the Acute Physiology And Chronic Health Evaluation (APACHE) scoring system and its derivatives is deliberate and is due to the extensive research that has gone into each of its developments, and to the number of publications about its use.

This review mentions only briefly specialised scoring systems for specific conditions, while abstracts of papers or posters presented at scientific meetings are not included, unless they are particularly important or are known to have been submitted for publication and are in press. In addition, only some of the more recent references were felt to be relevant in the field of cardiac surgery, as the indications, techniques, and results have changed so significantly in the past two to three decades.

2.1.2 Mode of the literature review

A search was made of Index Medicus by the CD Plus software on CD ROM in December 1992 and again in June 1994. The following key words were found in various combinations to produce the broadest yet most specific search:

- Intensive care units
- Critical care
- Coronary artery bypass (adverse effects, mortality, statistics, trends, utilisation)
- Severity of illness index
- APACHE
- Medical audit

The search was limited to human studies, those published in the English language, and adult patients (including middle age and aged). It included the years 1976-1993. Further references were found by review of the bibliographies in the selected papers, by careful scanning of the current literature, and by communication with experts in the field of severity scoring systems. More than 600 references were found in this way and the more important ones were selected as mentioned above.

2.2 History of scoring systems

2.2.1 Basic principles of predictive indices

Although relatively new in critical care (which is itself fairly new), predictive indices have been in use for many years, with conditions for fair comparison between clinical judgement and predictive indices having been specified at an early stage (Meehl 1954). Their evaluation and application have been fully reviewed and contrasted in a wide variety of clinical scenarios (Shapiro 1975, Wasson et al 1985, Dawes et al 1989). The conclusions varied, and may well be related to the timing of the papers. The first author assessed the skill of rheumatologists in predicting certain clinical outcomes regarding creatinine clearance, biopsy results, and serum complement levels (but not death) in patients. He presented a formula for measuring predictive accuracy and by this method found that statistical algorithms were not as accurate as the best physicians tested (Shapiro 1975). The second paper reviewed published reports of clinical prediction rules and found that the prediction rules, patients, and clinical setting were often inadequately defined. They warned that this inattention to basic principles might cause the objectives of these indices to not be achieved (Wasson et al 1985). Three potential deficiencies in study design might affect the validity of a prediction rule. These were poor definition of the outcome event, imprecise definition of the predictive findings, and failure to blind the investigator who assigns the diagnosis.

The last study was a broad overview comparing the predictive index (actuarial method) with clinical judgement in a variety of circumstances (Dawes et al 1989). In the actuarial method human judgement is eliminated and conclusions are dictated by the relationship between established data and the condition investigated. Examples given included the distinction between neurosis and psychosis, the diagnosis of progressive brain dysfunction based on intellectual testing, and the prediction of survival time following the diagnosis of Hodgkin's disease. The actuarial method was shown to be superior in these and other scenarios. Resistance to the actuarial method certainly exists; the authors describe it as irrational and abusing the basic principles of probability. Subsequent studies comparing clinical assessment with APACHE II and III scoring are reviewed below.

2.2.2 Defining the outcome measure

This thesis concentrates on mortality which is a binary variable and therefore much simpler to define than quality of life. However it may be argued that quality of life (rather than simple survival) measured at some point after ICU discharge is the most useful measure of the effectiveness of ICU therapy, as it would be counterproductive to put severely handicapped individuals back into the community. No scoring system has been designed to predict quality of life, although there is no shortage of instruments to measure this entity. Some of the more recently developed and general indices include the Sickness Impact Profile (Bergner et al 1981), and the Uniscale (Spitzer et al 1981).

Quality of life has been measured following intensive care, but the studies are difficult to compare as they differ in methods, patients, and follow-up time. Some of them are briefly reviewed. One year after ICU discharge 60% of those previously employed had returned to work (Goldstein et al 1986), while in another study performed at the same time interval 44% of patients had returned to their normal pre-hospital activity (Dragsted & Qvist 1989c). Also at one year, there was little difference in activities of daily living between elderly and young patients (Rockwood et al 1993). In a further study, 62% of patients not retired or home-makers were working between 16 and 20 months after ICU discharge (Sage et al 1986). Two years after ICU discharge 66% of patients were working (Parno et al 1984).

2.2.3 Early scoring systems

Early accounts of predictive indices in intensive care are scanty. The Therapeutic Intervention Scoring System (TISS), an important landmark which appeared in its first form in 1974 (Cullen et al 1974), is fully reviewed below. The same author identified physiological abnormalities as being significantly related to mortality; it was also recognised that it was becoming possible to document results of intensive care in comparable groups of patients (Cullen 1977). Thibault et al identified the aged and chronically ill as the major consumers of intensive care, noted that mortality was high in the aged (more so in the long-term than in the short-term), and concluded that there was a need to identify further predictors of mortality (Thibault et al 1980). An early predictive model for ICU patients (the Condition Index Score) was developed in 1981 (Snyder et al 1981). This used regression analysis to weight a variety of "conditions" that might occur in ICU patients, and to formulate a model to calculate probability of survival. The patient numbers were small (455 patients in the estimation group and 43 in the verification group) and the number of conditions was large (225 potential conditions of which 195 were actually observed). It is not an easy or practical score, but the most important part of the study was probably the recognition that such scores could be used to establish objective criteria for admission and discharge to ICU, compare the quality of care between ICUs, be used as a basis for multicentre studies, and to establish appropriate numbers of ICU beds for a hospital or region. These principles are equally important today. APACHE was also published in 1981 (Knaus et al 1981); the development of modern scoring systems began here and will be followed in detail below.

2.3 Rationale and ethics of scoring systems

2.3.1 Rationale of scoring systems

The original aim of APACHE was to classify groups of patients according to severity of illness (Knaus et al 1981). The authors felt this could be useful in comparing outcome, evaluating new technology, and planning for ICU needs. The ability of APACHE II to calculate a disease-specific risk of dying for the individual patient (Knaus et al 1985a) by a logistic regression equation led the way for the goal of severity of illness scoring to become mortality prediction. This unfortunately has led to much misunderstanding and disappointment: the ability of a scoring system to predict mortality for a group of patients cannot be extrapolated to individual patient predictions. Nevertheless, most severity of illness scoring systems can be appropriately used for comparison of ICU results, adjusted for diagnosis and severity of illness. They can also be used for stratification of patients into randomised trials.

For mortality prediction models to improve medical care, they ideally need the following characteristics (Selker 1993).

1. They should be accurate and easy to use in the clinical setting. In addition they should have the same mortality prediction when used prospectively as when used retrospectively.
2. Mortality risk should be predictable from data available in the first minutes of hospital admission.
3. Hospitalisation should not affect the mortality predictor.
4. They should require only data that are collected in the usual care of the patient.

5. They should use only objective and universally collected information.
6. They should not be susceptible to differences in ICU admission policies and practice.
7. They should be accurate at all levels of mortality risk i.e. be well calibrated.
8. They should not add to the burden of hospital record-keeping.
9. They should be open to inspection and testing.

These characteristics are presently met by no scoring system currently available although the more recently developed APACHE III, Simplified Acute Physiology Score (SAPS) II, and Mortality Probability Model (MPM) II go some way towards meeting them. Nevertheless they give excellent (if ambitious) guidelines and direction for further development.

2.3.2 Prediction and probability

It is important to separate the concept of prediction from that of probability (all current ICU severity systems provide estimates of the probability of hospital mortality). Prediction is a binary variable: the patient is predicted either to live or to die. On the other hand, probability expresses a risk for a group of patients. For the individual patient this may not be particularly useful as even when the probability of death is for instance 0.99, one cannot say whether any specific individual patient will be one of 99 patients expected to die or the single patient expected to live. That predictive indices work at all is in fact unexpected, as general chaos theory shows that "predictability is the exception rather than the rule" (Firth 1991).

The conversion of a probability to a prediction may be based on a decision rule to divide the range of probabilities into dichotomous groups of "predicted to live" and "predicted to die". The usual cutpoint for this is a probability of 0.5, but any other cutpoint could be used instead.

2.3.3 Ethics and the application of severity scoring systems

A certain proportion of ICU deaths will follow some restriction of therapy; 39% of ICU deaths in 7265 admissions were preceded by DNR (Do Not Resuscitate) orders (Zimmerman et al 1986), while 45% of deaths in another ICU study were preceded by withdrawal of therapy (Smedira et al 1991). The principal reason for the decisions in the latter study was the clinical impression of a poor prognosis; there was no mention of predictive indices being used.

Policy statements on withholding and withdrawing mechanical ventilation issued by the National Institutes of Health (NIH Workshop summary 1986) and on life-sustaining therapy by the American Thoracic Society (American Thoracic Society 1991) and Stanford University (Ruark et 1988) have clarified many of the most controversial issues in these areas, but although they make some powerful statements, they do not go into specific details of when and how to withhold or withdraw therapy. Current practices worldwide may be guided by these statements but the patient selection and actual withdrawal practices probably differ. The principal medical reasons or justification for withdrawal of life support include brain death (Black 1978, Pallis 1983), severe neurological dysfunction after 2-7 days of intensive care (Levy et al 1981, Levy et al 1985, Jennett et al 1979), and 3 or more organ failure after 3 or 4 days of organ failure (Knaus et al 1985b). More sophisticated computer-generated prognostic indices exist (Chang et al 1988, Knaus et al 1991a) but are presently not widely available. Occasionally a severe but not in itself fatal illness in a patient with chronic ill-health will result in withdrawal of life support although the more usual circumstance is that such therapy would not be initiated. In addition there are many subsets of patients in whom mortality is known to be very high from the evidence in published studies; this information may be used to help make withdrawal decisions (Osborne 1992). The definitions, legal opinions, and policy statements that guide current practice are constantly changing (Luce & Raffin 1988, Teres 1993) and presumably will continue to evolve.

There is always the possibility that clinical application of predictive indices without due caution will simply cause the predictions they make to be self-fulfilling and thereby perpetuate the impression that they are perfectly calibrated and mistakes are not being made. An editorial following the paper of Levy et al on prognosis in hypoxic-ischaemic coma (Levy et al 1985) strongly warned against using hypoxic-ischaemic coma outcome data to make clinical decisions (Black 1985).

It would appear that availability of predictive indices has already been shown to alter patient management in patients with head trauma (Murray et al 1993), and in general ICUs (Knaus et al 1990). In the former study, provision of a predictive index caused less aggressive therapy to be used in patients with a poor prognosis and more active therapy to be used in those with a good prognosis. On a similar note, in the latter study, when feedback was provided on patients with 3 or more organ-systems in failure (Knaus et al 1985b), there was a small but significant increase in decisions to stop active treatment, while ICUs that did not participate in the feedback showed no change in decision making. The authors claimed, however, that it was difficult to eliminate other influences. The ethical appropriateness of using prognostic scoring systems in clinical management has been recently reviewed (Luce & Wachter 1994).

Already physiology based severity scores can be mathematically calculated by direct input of variables from both the patient monitoring system (data from which need to be verified by a nurse to eliminate artefacts) and laboratory computers. A score is thereby automatically generated. In the case of APACHE III, a prognostic estimate can be produced on line and in real time if one has the right hardware and software (APACHE Medical Systems Inc.), as has been pictorially demonstrated (Knaus et al 1991b). This facility is not widely available but it is of great concern that the prognostic estimate so given may influence both management decisions and the attitude of medical and nursing staff towards patients.

2.4 Principles of development of scoring systems

The steps needed for the development of a scoring system have now been clearly defined (Kollef & Schuster 1994). They can be listed as follows:

1. Selection of the outcome variable.
2. Selection of the patient population.
3. Selection of variables as risk factors.
4. Collection and analysis of the database so acquired.
5. Statistical development of the prediction model.
6. Validation of the model.
7. Evaluation of the model's impact and utility.

8. Updating of the system.

The "new generation" (APACHE III, MPM II, SAPS II) of severity scores are based on these sound principles.

2.4.1 Dynamic vs static scores

Dynamic scores are scored on consecutive days (and may use some comparison between the daily scores) while static scores are based on a single day's data. There are several reasons why day one scoring alone may be inadequate for outcome prediction (Chang et al 1988):

1. The pathophysiology of disease is a dynamic process and cannot be fully assessed by a single day's measurements.
2. The development of complications and organ-system failure have important prognostic implications.
3. Major clinical decisions should not be based on a single assessment.

The vast majority of ICUs do not have the resources to collect scores on consecutive days of a patient's admission, making the debate somewhat academic. However, automated data collection by data management systems is already achievable and, if the high price falls, may become widely available in the not too far distant future. Automated scoring and outcome prediction could be simultaneously developed (as has already happened with APACHE III), making dynamic scoring systems easily accessible.

The development of a dynamic scoring system based on APACHE II and organ-system failure is discussed below. APACHE III is also meant to be scored daily and risk estimates are thereby updated. No other scoring systems are dynamic in nature although MPM has a separate model to be applied at 24 hours and there are further models being developed (Lemeshow et al 1993).

2.4.2 Physiological vs binary variable scores

APACHE and SAPS are based on abnormal physiology, while MPM is based on the presence or absence of binary variables. Physiological variables are routinely collected and are familiar to most observers with their abnormalities (and the degree and implications of those abnormalities) being easily recognised. Binary variables are simply either present or absent, and they stand or fall by the ease of application of their definitions. Early examples of scores based on binary variables are the Goldman index to predict cardiac complications in patients undergoing noncardiac surgery (Goldman et al 1977) and the asthma severity score (Fischl et al 1981). Many of the variables comprising both these scores were subjective clinical variables, subject to interobserver bias. In addition, misclassification of a single variable resulted in a large error in prediction.

The variables in the more recent MPM are much more objective, with well explained definitions in the text (Lemeshow et al 1993). In addition, there are 15 and 13 variables in the admission and 24 hour models respectively, which would lead to a smaller error if any single variable were misclassified.

2.5 APACHE: development and refinements

2.5.1 APACHE

The forerunner of current severity of illness scoring systems was APACHE (acronym for Acute Physiology And Chronic Health Evaluation), which was introduced in 1981 (Knaus et al 1981). Seven experienced physicians combined their wisdom to agree on 34 physiological variables likely to reflect severity of illness, assigning each a weight from 0 to 4 depending on the degree of derangement from normal. These variables included simple and routine measurements such as temperature, heart rate, and blood pressure, as well as less routine ones like serum lactate, ECG abnormalities, and serum osmolarity. To this was added a preadmission health evaluation to derive the APACHE score. The score was well correlated with mortality in 582 admissions to a university hospital ICU and 223 admissions to a community hospital ICU. Full statistical validation for APACHE has been presented (Wagner et al 1983). Although APACHE was both novel and logical in its concept, it was impractical for routine use because of the large number of physiological measurements (many of which were not routinely performed) that had to be collected.

2.5.2 APACHE II

It was thus the much slimmed-down APACHE II score published in 1985 (Knaus et al 1985a) that rapidly gained widespread acceptance and that has become extremely widely used in both general ICUs and those managing specific clinical problems. APACHE II evaluates the deviation from normal of 11 physiological variables (similar to the original APACHE, each variable scoring 0 points if within the normal range and from 1 to 4 points as the deviation from normal increases); it also includes the Glasgow Coma Scale, and allocates points for age and chronic health. The selected variables were the smallest number that reflected physiological derangement of all organs and still maintained statistical precision. Full details of these variables and the method of application of APACHE II appear in Chapter 4. APACHE II has now been validated worldwide. However in no study has APACHE II been accurate enough to confidently predict outcome in the individual patient, even though it has the ability to calculate a disease-specific risk of dying for the individual patient. APACHE II is easy to use, needing only simple physiological measurements and routinely performed blood tests, and all the information needed to do so is available in the original paper, although there are parts of the methodology that are problematic (see below). APACHE in all its developments has been thoroughly reviewed (Wong & Knaus 1991, Wisner 1992).

2.5.2.1 Other uses of APACHE II

2.5.2.1.1 Standardised mortality ratio (SMR)

The calculation of a disease-specific risk of dying for the individual patient allows the comparison between predicted deaths and actual deaths (standardised mortality ratio or SMR) and thereby can contrast the performance of different hospitals while taking into account severity of illness (Knaus et al 1986). The performance of this function in 5030 patients from 13 USA hospitals showed a wide variation in SMRs (from 0.59 to 1.58, $P<0.01$) and allowed for the identification of positive and negative features in the participating ICUs. The variations in performance appeared to be principally related to the structure, interaction, and coordination of the medical and nursing staff. In particular, a full-time unit director, controlling policy for therapy as well as for admission and discharge, and 24 hour in-unit physician coverage, seemed to be important. From the nursing side, important factors were adequate staffing, continuity of care, and a consistent senior charge nurse.

The SMR may be the only way to adjust results of intensive care according to severity of illness. However, just like the equation for risk of dying (from which it is of course derived) it relies heavily on the correct reason for admission being chosen, and this is not always easy. Lead-time bias (discussed below) may also alter mortality rates and hence the SMR.

2.5.2.1.2 Identification of low risk patients

Another use of APACHE II scoring has been identifying patients at low risk of requiring any ICU intervention other than monitoring (Wagner et al 1987). A risk threshold of the Acute Physiology Score component of the APACHE II score was derived for three primary categories of ICU admission (elective surgery, emergency surgery, and nonoperative). This application was remarkably successful with only 4.4% of the validation sample patients identified as low-risk actually receiving active treatment. APACHE II could therefore be used to guide more appropriate utilisation of ICU facilities.

2.5.2.1.3 Cost analysis and quality control

The cost of the first day of ICU management was significantly related to the APACHE II score in a United Kingdom study of 90 general ICU patients (Ridley et al 1993); a simple model was developed to predict the cost of the first day in ICU from the APACHE II score. In contrast to this finding, APACHE II was found to be unsuitable for cost containment and quality assurance in a much larger USA study of 372 surgical ICU patients, as grouped APACHE II scores did not correlate with total hospital charges or length of ICU stay (Civetta et al 1990). These differences may be explained by the size of the studies, the types of patient, or by different medical practices and costs in the respective countries (Osborne & Evans 1994).

2.5.2.1.4 Allocation of funding

In order to confirm anecdotal evidence that patients admitted as emergencies for acute chest problems to a United Kingdom hospital had become more ill between 1985 and 1990, APACHE II scoring was rather ingeniously used. The 1990 APACHE II scores were significantly higher than those of 1985 and enabled the medical services to obtain an extra £100000 of funding from the authorities (Irwin & Jessop 1993).

2.5.2.2 Criticisms and problems with APACHE II

2.5.2.2.1 Interpretation of the methodology

There are many pitfalls in the precise interpretation of the method of scoring; a recent paper outlines and clarifies some of these (Palazzo & Patel 1993). They include difficulties with definitions of variables (see Chapter 4) but most importantly, difficulty in defining a single reason for ICU admission (an important factor in the equation for risk of dying). These differences in interpretation may lead to very different scores being obtained for the same patient by different data collectors.

2.5.2.2.2 Interobserver variability

Although interobserver variability between APACHE II data collected by residents and nurses is statistically small in groups of patients, there may be a significant variability in the individual patient (Holt et al 1992). In their study evaluating APACHE II scoring in 120 consecutive patients, the scores of residents and nurses differed in 79 patients. In most, but not all cases, the differences were small. The most common causes of error were incorrect choice of highest or lowest value as the worst, and error in Glasgow Coma Scale score.

2.5.2.2.3 Timing of scores

A criticism that has been made of the APACHE II method is that early and efficient treatment (e.g. in an operating theatre or Accident and Emergency department) correct physiologic abnormalities and make the score lower than it would have been and therefore invalid for prognostic purposes. The data on this issue is contradictory. One study using data collected on 756 patients either in the Accident and Emergency unit or immediately on admission to ICU showed that APACHE II scored this way significantly underestimated mortality in drug overdose and trauma (Waters et al 1990). In other words, the scores were lower than would have been expected with the observed mortality (this observation may also be interpreted as showing that the APACHE II coefficients are not validated for admission scoring). Another study compared emergency room and conventional (worst over the first 24 hours) APACHE II scores in trauma patients (McAnena et al 1992). In this case the conventional scores were significantly lower than those obtained in the emergency room.

2.5.2.2.4 Use of admission scores

In a subset of patients from the original APACHE II database, the worst physiological value over the first 24 hours was also the admission value (Knaus et al 1985a). Although the full admission APACHE II scores were close to the worst scores, they were not identical and the authors recommended that worst values should continue to be used.

The use of admission scores throughout the patient's ICU admission may not reflect events that occur subsequent to admission. This has been highlighted in a study in which multivariate analysis was carried out to determine patient variables associated with outcome (Ferraris & Propp 1992). The occurrence of iatrogenic complications, renal failure, or sepsis were all significantly associated with mortality, more so than the admission APACHE II score.

The predictive accuracy of APACHE II appears to decrease with the length of time a patient stays in the ICU. Predicted and actual outcomes were compared in patients grouped according to their length of ICU stay. Using a risk of death of 0.4 as a cutoff point (patients with a risk of less than 0.4 were predicted to live, and those with a risk greater than 0.4 were predicted to die), the misclassification rate was 11.6%-15.7% for patients staying between 1 and 3 days, but rose to 38.7% for patients staying for 6 or 7 days (Sleigh et al 1992).

2.5.2.2.5 Source of ICU admission

Another criticism (that has subsequently been addressed in the APACHE III methodology) is that APACHE II does not take the source of ICU admission into account. The importance of this was demonstrated in a study of 235 Medical ICU patients where the predicted and actual mortality rates of those admitted from the emergency department were similar, but there were differences in those patients admitted from hospital wards, the intermediate care unit, and other hospitals (Escarce & Kelley 1990). In addition, logistic regression analysis showed an independent association between admission source and risk of death.

Lead-time bias is another aspect of this problem. Differences in ICU admission policies and practice have been addressed in a study in which lead-time bias was felt to influence the differences in mortality between two Danish ICUs (Dragsted et al 1989a). Measured severity of illness was similar, but 35% of patients in the hospital with the higher mortality rate were transfers from other units. This potentially confounding bias needs to be identified if comparisons of hospitals' performance are to be made.

2.5.2.2.6 Response to therapy

APACHE II scores were originally thought (but not proven) to be independent of therapy. However total interventions scored in a point system according to their intensity were inversely related to APACHE II scores in patients who died and directly related to APACHE II scores in patients who survived (Civetta et al 1992). This strongly suggests that APACHE II is in fact dependent on therapeutic manoeuvres.

Along the same lines, it was suggested that as physiological data can be strongly influenced by medical and nursing intervention, scoring systems based on them should not be used for audit (Boyd & Grounds 1993). They argue that ICUs that are performing well may have the same SMR as ICUs that are performing badly. As an example, one ICU could be promptly diagnosing and treating physiological abnormalities. The APACHE II scores might be low because the physiology has not been allowed to become too abnormal. This would lead to a low prediction of mortality, but the mortality rate would also be low because of efficient treatment. Another ICU might be slower to respond. APACHE II scores would be higher, as would predicted risk of dying. Mortality might also be higher. The ratio between predicted and actual mortality (i.e. the SMR) might however be similar for these two hypothetical ICUs.

This paper stimulated much discussion, with no fewer than five letters appearing in *The Lancet* in response. One writer felt the authors' argument to probably be correct, but that it should not obscure the value of scoring systems in directing the audit of patient care in ICUs (Rothwell 1993), while others went so far as to say that physiological data alone were inappropriate for comparison of ICU performance (Hasibeder et al 1993). In contrast, their argument was felt to be improbable (Chang & Bihari 1993), and their advice misguided (Palazzo et al 1993). Finally, it was felt that potential biases should be recognised and understood, but that they should not invalidate the audit process (Holt et al 1993).

2.5.2.2.7 Patient subgroups

APACHE II has been perceived to be less successful in postoperative surgical patients (Cerra et al 1990). those receiving total parenteral nutrition (Hopefl et al 1989), oncology patients (Abbott et al 1991, Dart et al 1991), those with cardiogenic pulmonary oedema (Fedullo et al 1988), AIDS patients with respiratory failure (Chu 1993), and one study of patients with acute renal failure (Schaefer et al 1991). Reasons may include the small numbers of such patients in the original APACHE II database, or to unusual characteristics of the subgroups themselves.

2.5.3 Use of APACHE in general ICU settings

APACHE II has been successfully used in a number of general ICUs (including general surgical ICUs) throughout the world. The larger studies include those from the USA (Knaus et al 1985a, Marsh et al 1990, Lockrem et al 1991, Rutledge et al 1991), the Middle East (Jacobs et al 1988), Africa (Turner et al 1988, Joshua et al 1989), Europe (Dragsted & Qvist 1989b, Giangiuliani et al 1989, Berger et al 1992), the Far East (Chen et al 1993, Lee et al 1993, Oh et al 1993), and the UK (Chisakuta & Alexander 1990, Rowan et al 1993). The use of APACHE II scoring has also been reported in a comparative study between New Zealand and the USA of patient selection for intensive care (Zimmerman et al 1988). All of these studies showed that in general terms APACHE II was applicable worldwide (Kruse & Carlson 1993); the large UK study was in some respects an exception. This study evaluated 8796 admissions to 26 general ICUs in Britain and Ireland. The predicted and actual mortality rates were generally close but were significantly different in certain disease categories; it was felt that the APACHE II predictive equation (USA derived) did not fit all of the local data and might need modification for use in the UK. This is the largest study of APACHE II outside the USA (in fact it is larger than the original study used to derive APACHE II), and it creates concern that larger studies in other centres might have reached the same conclusion.

2.5.4 Use of APACHE II in specific areas

APACHE II has also been used in a surprisingly wide variety of specific disorders. Those in which the prime purpose has been the evaluation of APACHE II scoring include the following (Figure 2.1).

Figure 2.1

Specific disorders (with references) in which APACHE II has been used.

1. Outcome in haematological malignancy (Johnson et al 1986).
2. Usefulness of total parenteral nutrition (Chang et al 1986).
3. Outcome in abdominal sepsis (Bohnen et al 1988, Sleigh et al 1989).
4. Outcome in haemodialysis (Dobkin & Cutler 1988).
5. Outcome in cardiogenic pulmonary oedema (Fedullo et al 1988).
6. Outcome in haematological malignancy (Lloyd-Thomas et al 1988).
7. Outcome in patients receiving total parenteral nutrition (Hopefl et al 1989).
8. Outcome in mechanically ventilated patients (Knaus 1989).
9. Outcome in pancreatitis (Larvin & McMahon 1989, Wilson et al 1990, Roumen et al 1992).
10. Outcome in acute renal failure (Maher et al 1989, Schaefer et al 1991).
11. Outcome in acute myocardial infarction (Moreau et al 1989).
12. Outcome in upper gastrointestinal haemorrhage (Schein & Gecelter 1989).
13. Organ failure and mortality in postoperative surgical patients (Cerra et al 1990).
14. Outcome in trauma (Rhee et al 1990, Vassar et al 1992).
15. Outcome in oncology patients (Abbott et al 1991, Dart et al 1991).
16. Prediction of nosocomial infection risk (Bueno-Cavanillas et al 1991).
17. Outcome after cardiac arrest (Nikanen et al 1991, Ebell & Preston 1993).
18. Outcome in coronary care (Teskey et al 1991).
19. Outcome in cardiothoracic surgery (Turner et al 1991).

(Figure 2.1 continued...)

20. Prediction of morbidity and mortality after liver resection (Gagner et al 1991).
21. Outcome in breast cancer (Headley et al 1992).
22. Outcome in rheumatologic disease (Kollef & Enzenauer 1992).
23. Prediction of septic complications after cardiac surgery (Kreuzer et al 1992).
24. Outcome in asthma (Day et al 1993).
24. Outcome in AIDS patients with respiratory failure (Chu 1993).

Most of the above predictive applications of APACHE II have been perceived to be successful although some of this may be related to differences in the expectations of the users and to observer bias. Less successful applications include the following: postoperative surgical patients, those receiving total parenteral nutrition, oncology patients, those with cardiogenic pulmonary oedema, AIDS patients with respiratory failure, and one of the studies of patients with acute renal failure (Schaefer et al 1991). In two studies not specifically evaluating the performance of APACHE II, the scoring system underestimated mortality in adult respiratory distress syndrome (ARDS), with predicted mortality being 39% and actual mortality 45% (Kraus et al 1993), while in the other study it significantly overestimated mortality: predicted mortality was 39.6% and actual mortality 16% (Hickling et al 1990).

The reasons for these "failures" of APACHE II are perhaps less than complex, and involve the inaccuracy of individual patient predictions. APACHE II was often misunderstood (Knaus et al 1991a); the above is perhaps one of the misunderstandings.

The application of APACHE II scoring in mechanically ventilated patients (Knaus 1989) illustrated two important points about the system which are in fact pertinent to most of its applications. The first was that when the APACHE II score was more than 25, mortality varied from 28% to 100% according to the disease category. The second point was that APACHE II scores on days 4 and 7 improved the ability to estimate a fatal outcome; very few studies have performed sequential APACHE II scoring.

2.5.5 Derivatives of APACHE II

2.5.5.1 Daily trend plus organ-system failure

A derivative of APACHE II, using the trend of daily scores and integrating organ-system failure (Chang et al 1988), has shown promise in prediction of outcome but has not gained as wide acceptance as the original APACHE II. It analyses the rate of change relative to the previous day's APACHE II score and organ failure coefficient. An algorithm predicts those patients who will eventually die (Chang et al 1988). It has been compared with clinical judgement (Chang et al 1989). The computer model had no false positive predictions of death while doctors and nurses had a false positive predictive rate for death of between 7.7% and 16.7% (Chang et al 1989). The authors stress that computer predictions were not acted upon. The software is called (rather unfortunately, and soon to be changed) RIP (Riyadh ICU Program, Copyright RWS Chang, Medical and Associated Software House, London, UK), now in version 4, and is commercially available. The authors of a small study from Ireland found incorrect predictions in two patients scored in this way; although they did not apply the methodology strictly or correct for organ failure (Chisakuta & Alexander 1990). When correctly used on 1155 patients in a teaching hospital ICU in Wales, the RIP predictions had a sensitivity of 14.8% and specificity of 99.8%, incorrectly predicting three deaths (Jacobs et al 1992). RIP has also been used to classify ICU patients into four groups: those who benefitted from intensive care, those who might have benefitted, those who would never or would no longer have benefitted, and those who did not need intensive care (Jacobs et al 1989). The authors conclude that more appropriate utilisation of limited facilities could result from such use.

2.5.5.2 Sickness Score

Another scoring system based on APACHE II is the Sickness Score (Bion et al 1988), derived from the analysis of 128 ICU patients. This score modifies the physiology components of the APACHE II score by conversion to SI units, changing from haematocrit to haemoglobin, and assessing oxygenation using the ratio of inspired oxygen concentration and the arterial oxygen tension. In addition the cardiovascular variables are scored two-hourly, the scores summed, and then rescored according to an averaging protocol (Bion et al 1988). The definitions for chronic health have also been broadened. The proportion of change in score from day 1 to day 4 is calculated. Estimation of probability of death is calculated by logistic regression analysis. In the above study, the predictive power of the Sickness Score on admission was 74.4%, while that of the staff nurses was 73.9%, the sisters 70.3%, and the resident SHO and consultant both 65.8% (Bion et al 1988).

2.5.5.3 Rapid Acute Physiology Score

Also based on APACHE II is the Rapid Acute Physiology Score (RAPS). This was developed and tested on 283 patients as a severity indicator for transport of critically ill patients (Rhee et al 1987). It incorporates the blood pressure, heart rate, and respiratory rate components of the APACHE II score (with identical point allocation) and allocates from 0 to 4 points for the Glasgow Coma Scale. It correlated well with worst APACHE II variables; the best predictor of outcome was worst APACHE II score but pretransport RAPS also had significant predictive power for mortality. The advantage of RAPS is that it is easy to calculate, relying only on easily available and simple clinical variables (without the need for laboratory tests); the authors conclude that it complements APACHE II and may have limited utility when used alone.

2.5.6 APACHE III

2.5.6.1 Introduction

The most recent version of APACHE is the APACHE III prognostic system (Knaus et al 1991a), designed to be capable of delivering objective probability estimates. The predictions from the original database are most impressive: within 24 hours of ICU admission, 95% of patients could be given a risk estimate within 3% of that observed with the area under the receiver operating characteristic (ROC) curve of 0.90; this is in fact only a modest improvement over APACHE II which had a ROC area of 0.86 (Knaus et al 1991a). However APACHE III requires a greater amount of data than APACHE II and uses complicated score algorithms which remain copyright material; it is marketed as a hardware and software package by APACHE Medical Systems Inc. (Washington, USA).

2.5.6.2 Development of APACHE III

The description of the development of APACHE III took up a whole supplement (funded by the George Washington University research group) of Critical Care Medicine (APACHE III study design: analytic plan for evaluation of severity and outcome, Volume 17, Number 12, December 1989), leading to an editorial justification for this amount of attention (Shoemaker 1989), and a well argued perspective and appraisal demanding careful attention, scrutiny, and peer review of severity of illness scoring (Teres 1989).

For the APACHE III data collection, twenty-six hospitals and ICUs were selected randomly through the USA and 14 volunteered to participate. Variable selection included 212 disease categories and 20 physiological parameters (selected on the basis of past experience and clinical judgement). Data collection ran from May 1988 to November 1989, and all patients were followed up to hospital discharge. The methodology excluded patients with burns, patients younger than 16 years of age, and patients admitted to exclude myocardial infarction. In addition, patients had to remain in the ICU for a minimum of 4 hours. Data on coronary artery bypass patients was collected for separate analysis. The integrity of data collection was ensured by a training course for participants followed by quality assurance and reliability studies. Statistical analysis involved multivariate logistic regression, with categoric and continuous weighting to determine component variables and weights.

There were 17 440 patients in total. Ninety percent of the patients were randomly divided into estimation and validation halves.

2.5.6.3 Differences from APACHE II

In the final version, APACHE III differed from APACHE II as follows. Potassium and bicarbonate measurements were abandoned, while the pH score incorporated measurement of $PaCO_2$, and the neurology score was refined. Both of the latter are scored by referring to tables. Urea, urine output, albumin, bilirubin, and glucose were added. The age and chronic health weightings were changed, as was the weighting for each individual variable. The APACHE III score comprises the sum of the physiology points, the age points, and the chronic health points. More precise details of its application appear in Chapter 4.

2.5.6.4 The APACHE III predictive equation

The APACHE III score (published, "shareware") is only one part of the APACHE III predictive equation (copyright) which includes weighting for 78 different disease categories, the treatment location prior to ICU admission, and the occurrence of emergency surgery. The equation then employs natural antilogarithm mathematics to calculate the probability of dying for the individual patient. The natural antilogarithm of the equation equals the risk of hospital mortality divided by 1 minus the risk .

2.5.6.5 Ongoing research

Data collection and database accumulation is ongoing, under the auspices of APACHE Medical Systems Inc. The database had over 120000 patients by December 1993 (WA Knaus, personal communication), and is expanding internationally. In the United Kingdom, 19 ICUs (17 from the South West Thames region) are taking part in APACHE III data collection; when there is enough data the predictive equations will be refined for local use. APACHE III may also be used to calculate standardised mortality ratio and predict length of ICU stay, allowing comparison between the performance of ICUs (Knaus et al 1993b).

2.5.6.6 Applications of APACHE III

APACHE III has not yet been applied in a wide variety of scenarios, although this will almost certainly change.

The role of the Glasgow Coma Scale score in the APACHE III predictive ability was examined in 15973 non-trauma patients from the APACHE III database (Bastos et al 1993). It was found that the Scale itself had a significant but non-linear relationship with outcome, although discrimination in intermediate score levels was reduced. In addition, the overall predictive ability of APACHE III was improved by the incorporation of the Glasgow Coma Scale score.

APACHE III was used to quantify the mortality risk of 519 ICU patients (drawn from the APACHE III database of 17440 patients) admitted with a primary diagnosis of sepsis (Knaus et al 1992). While the mortality rates of patient groups with "sepsis syndrome" and "septic shock" differed significantly, the range of individual patient risks in these groups and in the patients not meeting these criteria merged into one another.

A further and more innovative use of APACHE III scoring was to develop a model to predict survival for patients entering clinical trials. Such a model was created from the analysis of APACHE III predictions for 1195 patients with sepsis syndrome, based on 58737 patients admitted to 107 hospitals in the USA and Europe (Knaus et al 1993a). This model was used in the retrospective analysis of a phase III clinical trial using an Interleukin 1 (IL-1) receptor antagonist in sepsis syndrome (to date only in abstract form). This study included 893 patients from the USA, Canada, and Europe. Although overall there was no difference in 28 day mortality between placebo and treatment with high or low dose IL-1 receptor antagonist, there was a similar benefit in both dosages when the predicted mortality was more than 24%. In all clinical trials there are low-risk patients who will survive without a new therapy, and high risk patients who will die whatever therapy they are given. Risk stratification such as that delivered by APACHE III, provided it is available at the time the therapy needs to be given, may distinguish patients in whom new (and expensive) therapy is likely to be of benefit.

APACHE III has also been used to compare case-mix, structure, resource utilisation, and outcome performance between ICUs in teaching and non-teaching hospitals (Zimmerman et al 1993a). The teaching hospital ICUs dealt with more complex patients, had twice the number of physicians providing services for these patients, used more resources (with increased risk of diagnostic tests and invasive procedures), and mostly had better risk-adjusted survival rates. In a further study, nine ICUs were selected from the APACHE III database for on-site analysis

(Zimmerman et al 1993b). Although there was a wide variation in risk-adjusted survival rates, this could not be correlated with differences in ICU structure and organisation, or by on-site clinical judgement. However the authors felt that there were good and bad practices observed; the better practices included a patient-oriented culture, strong leadership, good and effective communication and coordination, and an open collaborative approach to problem solving and conflict management.

2.5.7 Scoring systems and clinical judgement

2.5.7.1 General ability of physicians

Before comparing scoring systems with clinical judgement, the predictive skills of physicians need to be examined. In intensive care, the accuracy of physicians' outcome predictions is poorly defined. Firstly, individual predictive ability may be poor. When 20 ICU physicians were asked to predict survival on each patient's fourth ICU day, they predicted death in only 41% of patients who died and survival in 87% of those who survived (Perkins et al 1986). In another study, many patients thought certain to die (100% certainty) lived, and patients thought certain to live, died (Poses et al 1989). Secondly, physicians may disagree amongst themselves. When physicians' quantitative prognostic judgements for 269 ICU admissions were assessed, there was good overall discriminating ability but more than 40% of patients provoked disagreements of at least 20 percentage points between physicians (Poses et al 1989).

Reasons for these discrepancies are complex but may include the following arguments. People have been shown to have an optimistic bias about personal risks. This has been reviewed (Weinstein 1989). Optimism is greatest for risks where the person has little personal experience of that risk, for risks of low probability, and for risks that may be controllable by personal action. Thus, physicians may extrapolate their own optimism about personal risks to the condition of their patients. Alternatively, they may be misguided by their personal overoptimistic view of overall ICU mortality. Poses et al evaluated the effects of "ego bias" on the prognostic judgements of physicians. Ego bias includes the optimistic bias about personal risks described above; it also involves an optimistic bias about personal good fortune. It may be strongest in people who have little control over their environment. In their study of 201 medical and surgical ICU patients, predictions for individual patients were compared with predictions for the overall ICU mortality rate. House officers made predictions for individual patients that were significantly more optimistic than their judgements of the overall survival rate, thus displaying ego bias. However, the attending physicians showed the reverse, with more optimistic predictions for the overall survival rate than for individuals (Poses et al 1991). This latter characteristic is known as "reverse ego bias". The authors felt that the difficult conditions, beyond their control, experienced by the house officers may have caused their ego bias, while the more secure attending physicians had no such need and may have realised that down-playing their chances of success would be more productive.

2.5.7.2 APACHE and clinical judgement

Both APACHE II and APACHE III have been subjected to extensive comparison with clinical judgement. Mortality predictions by clinicians were better than those of APACHE II in two studies in general ICU patients (Brannen et al 1989, Marks et al 1991) and one in surgical patients (Meyer et al 1992), while there was no significant difference in a third study (Kruse et al 1988). In a further study physicians again performed better than APACHE II in estimating mortality, but APACHE II was better calibrated in the central probability ranges (McClish & Powell 1989). APACHE III, on the other hand, was slightly better at mortality prediction than clinicians and in addition was more consistently accurate at all levels of mortality (Knaus et al 1991b). However. the differences determined in the above studies are small and are probably of minimal or no clinical significance.

2.6 Other severity of illness scoring systems

2.6.1 Simplified Acute Physiological Score (SAPS)

The original APACHE score (Knaus et al 1981) had too many physiological variables for routine use; this problem was addressed by the much shorter Simplified Acute Physiological Score (SAPS), using only 14 easily obtained and readily available physiological variables (Le Gall et al 1984). These include age, heart rate, systolic blood pressure, temperature, respiratory rate, urine output, blood urea, haematocrit, white cell count, serum glucose, potassium, sodium and bicarbonate, and the Glasgow Coma Scale. For reasons that are not entirely clear, the APACHE II score published shortly afterwards (Knaus et al 1985a) caught on more readily and widely, although SAPS remained popular in Europe. The SAPS principles are similar to those of APACHE II, with component weights ranging from 0 to 4 depending on the degree of derangement from normal; the variables and the component weights are only slightly different from those used in APACHE II.

In 679 unselected patients in 8 French ICUs, SAPS correlated well with hospital mortality (Le Gall et al 1984). SAPS has also been used in clinical studies of ARDS. Although SAPS on admission was shown to be not significantly different between survivors and nonsurvivors, it improved in survivors and got worse in nonsurvivors (Mancebo et al 1987). This is stated as having been confirmed in the European Collaborative Study of ARDS (Artigas et al 1992), although the actual figures are not presented.

SAPS was compared with an oxygen consumption index in critically ill ventilated surgical ICU patients (van Lanschot et al 1988). SAPS was significantly lower in survivors than nonsurvivors but failed to be useful in providing a prognosis in the individual patient. However there was no significant difference in oxygen consumption index between survivors and nonsurvivors.

A simplified version of SAPS has been proposed by using a multiple logistic regression analysis of the 14 original variables; only five variables needed to be retained (Viviand et al 1991). This model was compared with SAPS in a further set of 446 patients and there was no difference in discriminating power.

2.6.1.1 SAPS II

An update of SAPS, named SAPS II, is based on analysis of 13152 patients (randomly divided into developmental and validation models) in 137 ICUs in 12 countries (Le Gall et al 1993). It includes 12 physiology variables (Figure 2.2, ranges of scores and points allocated not presented here), age, type of admission (scheduled surgical, medical, and unscheduled surgical), and underlying disease (metastatic cancer, haematologic malignancy, and AIDS). SAPS II excludes burn patients, coronary care patients, and all cardiac surgery patients.

No primary diagnosis or reason for admission is needed to perform SAPS II scoring. From the SAPS II score a probability of hospital mortality may be calculated. The equation for this calculation includes both the actual SAPS II score and a shrinking power transformation of the score using the natural logarithm

Figure 2.2

Physiological variable making up the SAPS II score

Heart rate
Systolic blood pressure
Temperature
Arterial oxygenation (PaO_2/FiO_2 ratio)
Urine output
Urea
White cell count
Potassium

(Figure 2.2 continued...)

Sodium

Bicarbonate

Bilirubin

Glasgow Coma Scale

When predicted and actual mortalities were compared, the area under the receiver operating characteristic curve was 0.86 in the validation sample. Future directions for SAPS II include scoring on a daily basis and enlarging the database with periodic quality checking of the predictive model.

2.6.2 The Mortality Prediction Model (MPM)

The Mortality Prediction Model (MPM) (Lemeshow et al 1985) differs significantly from other scoring systems. It employs 7 binary (i.e. the variable is either present or absent) clinical variables selected by multiple logistic regression and weighted by a statistical technique known as maximum likelihood. In addition, different variables are scored at admission (none of which are treatment-dependent) and at 24 hours (these reflect both treatment and clinical condition). From these variables is calculated a probability of hospital mortality, rather than a score. The predictions were closely correlated with actual outcome, and although MPM performed better than the Acute Physiology Score component of APACHE II and SAPS in the hands of the authors (Lemeshow et al 1987), in another relatively small study it did not show significant predictive ability in patients with septic shock (Arregui et al 1991). In a further study goodness-of-fit was better for

APACHE II than MPM when comparing observed and predicted outcome; MPM overestimated deaths (Castella et al 1991).

The MPM was validated in a separate group of 1997 patients (Teres et al 1987) where the admission but not the 24 hour model accurately predicted mortality. The authors subsequently described an improvement, deriving mortality models from data gathered at admission, at 24 hours, and at 48 hours (Lemeshow et al 1988). The variables for each of these models differ slightly. The serial observations substantially enhanced predictive ability. The MPM has also been effectively used to assess the effect of timing of ICU admission on outcome. Patients admitted to ICU within 1 day of hospital admission had a lower actual than predicted mortality, while patients admitted later had the opposite (Rapoport et al 1990a). In a large study of 2749 patients, MPM was used to study resource use. The relationship between severity of illness (as measured by MPM) and resource use was not linear, but within each diagnosis-related group, MPM explained a statistically significant percentage of the variability in resource use (Rapoport et al 1990b).

2.6.2.1 MPM II

MPM II (this time called Mortality Probability Model) was published late in 1993 (Lemeshow et al 1993). The model was developed on 12610 ICU patients in 12 countries and validated on a further 6514 patients. Patients younger than 18 years, burn patients, and coronary care and cardiac surgery patients were excluded. The result was an admission (MPM_0) and 24 hour (MPM_{24}) model; both models were well calibrated and discriminated well. The MPM_0 contains 15 variables (Figure 2.3) readily obtained on admission and well defined in the text; it is the only scoring system available for use on admission to ICU. The MPM_{24} contains five of the admission variables as well as 8 variables easily obtained at 24 hours. Both MPM_0 and MPM_{24} were well calibrated (by Hosmer-Lemeshow goodness-of-fit testing) and both discriminated well (by ROC curve analysis).

Figure 2.3

Variables in the MPM_0

Physiology	Coma or deep stupor
	Heart rate >= 150 beats/minute
	Systolic Blood Pressure <= 90 mm Hg
Chronic diagnosis	Chronic renal insufficiency
	Cirrhosis
	Metastatic neoplasm
Acute diagnosis	Acute renal failure
	Cardiac dysrhythmia
	Cerebrovascular incident
	Gastrointestinal bleeding
	Intracranial mass effect
Other	Age (10-year odds ratio)
	Cardiopulmonary resuscitation prior to admission
	Mechanical ventilation
	Nonelective surgery

2.6.3 The organ-system failure (OSF) score

The organ-system failure (OSF) score (Knaus et al 1985b) was developed in parallel with APACHE II. Strict definitions were developed for 5 organ-systems, namely cardiovascular, respiratory, renal, haematologic, and neurologic (see Figure 2.4). The numbers and duration of OSF were evaluated against hospital outcome for 5677 ICU admissions (2719 of whom developed OSF) from 13 hospitals. A very accurate predictive value for death resulted from the analysis. Thus the failure of three or more organ-systems on the fourth day of OSF carries a 96% mortality, whilst two OSF carries a 56% to 68% mortality from the third day of OSF. The predictive abilities of this system, together with the tight 95% confidence intervals presented, would potentially make it a useful tool for helping to make decisions about withdrawal of therapy.

Figure 2.4

Definitions of organ-system failure (Knaus et al 1985b) converted to SI units where appropriate.

1. Cardiovascular failure (presence of one or more of the following):

A. Heart rate ≤ 54/min.

B. Mean arterial pressure ≤ 49 mmHg.

C. Occurrence of ventricular tachycardia and/or fibrillation.

D. Serum pH ≤ 7.24 with a $PaCO_2 \leq 6.53$ kPa.

(Figure 2.4 continued...)

2. Respiratory failure (presence of one or more of the following):

A. Respiratory rate ≤ 5/min or ≥ 49/min.

B. $PaCO_2 \geq 6.67$ kPa.

C. $AaDO_2 \geq 46.7$ kPa.

D. Dependent on ventilator on fourth day of organ-system failure.

3. Renal failure (presence of one or more of the following):

A. Urine output ≤ 479 ml/24 hours or ≤159 ml/8 hours.

B. Serum BUN ≥ 100 mg/100 ml.

C. Serum creatinine ≥ 308 µmol/l.

4. Haematologic failure (presence of one or more of the following):

A. White cell count ≤ $1000/mm^3$.

B. Platelet count ≤ $20\ 000/mm^3$.

C. Haematocrit ≤ 20%.

5. Neurologic failure:

Glasgow Coma Scale ≤ 6 (in the absence of sedation).

2.6.4 The ODIN (organ dysfunction and/or infection) model

The ODIN system evaluates the presence or absence of dysfunction in six organ systems (respiratory, cardiovascular, renal, haematologic, hepatic, and neurologic) and the presence or absence of infection (Fagon et al 1993). The definitions of organ dysfunction are slightly different from the organ-system failure definitions above (Knaus et al 1985b); they are said to have been derived from a review of the literature. There was a direct and statistically significant increase in mortality with increase in ODIN. Multiple logistic regression analysis was then used to demonstrate the relationship between each ODIN and calculate a probability of death. When predicted and actual death rates for ODIN, SAPS, and APACHE II were compared, there was no statistical difference in area under the receiver operating characteristic curve. The ODIN score is computed for the first 24 hours of ICU admission only and has not been used in a dynamic fashion.

2.6.5 Therapeutic Intervention Scoring System (TISS)

The Therapeutic Intervention Scoring System (TISS) first appeared in 1974 (Cullen et al 1974). Fifty-seven monitoring or therapeutic interventions were scored from 1 to 4 points depending on their intensity. The sum of the points acquired in a 24 hour period make up the TISS score. TISS was found by the authors to be useful in determining appropriate utilisation of ICU facilities, providing information on nurse staffing ratios, validating a classification of patients into categories of severity, and analysing cost relative to extent of ICU care.
TISS was updated in 1983 (Keene & Cullen 1983) to include 76 variables. This update took into account newer therapeutic modalities; some items were added, some deleted, and the points for some adjusted. A comparison of the updated score with the original score in 100 consecutive patients revealed no significant difference (Keene & Cullen 1983), although the method of comparison might not

be acceptable today (Bland & Altman 1986). There has been no update since this time, although various hospitals have introduced their own minor variations to incorporate new and expensive treatment modalities, and an “official” update, which will be orientated towards fairer evaluation of medical patients, is expected in the near future.

It was suggested (perhaps rather optimistically) that TISS could provide the following information (Keene & Cullen 1983): percentage of bed occupancy, classification of patients into classes of severity, identification of inappropriate ICU admissions, daily intensity of care delivered, identification of patients discharged too soon or not soon enough, follow up of care after discharge, nurse to patient ratio, and number of ICU beds needed.

The TISS has proved to be a strong predictor of total ICU admission costs (Slatyer et al 1986). In addition, it has been used to suggest ICU admission and discharge criteria (with a cut-off of 10 TISS points), as well as effective utilisation of nursing staff (Adam et al 1989). In this study the authors suggest that a single nurse could manage two patients with a total of 24 TISS points between them. These points seem rather low; in the Adult ICU at the Royal Brompton National Heart and Lung Hospital a single nurse manages a patient with up to 80 TISS points (the mean daily TISS score is around 50 points) and patients are usually discharged with 30 to 40 TISS points (unpublished data, collected at the time of the third study).

The TISS has also been used to predict the risk of nosocomial infection (Bueno-Cavanillas et al 1991). In this study a day one TISS score of more than 20 points and an APACHE II score of more than 12 points were both positively associated with nosocomial infection; on multivariate analysis there was no such association with APACHE II but each TISS point suggested an infection risk increase of 6%. On a theoretical basis however, (as it would be impossible to prove scientifically) TISS may not accurately reflect severity of illness as much as it reflects the level of medical and nursing intervention. For the same patient, the score may vary according to the aggressiveness or invasiveness of different treatment philosophies, as well as to the availability of resources and financial constraints. Details of the interventions, their scores, and their practical application appear in Chapter 5.

2.6.6 The system outcome score (SOS) and outcome index

Another scoring system developed from stepwise logistic regression of a variety of variables is the system outcome score (Gilbert & Schoolfield 1991). At the end of two stages of analysis of 2777 consecutive patients, five clinical variables were identified; stepwise discriminate analysis was then applied and points were allocated to the variables by expressing each variable's discriminate index as a percentage of the sum of all indices and then normalising to a total summation maximum of ten. The final SOS components and values are detailed in Figure 2.5.

Figure 2.5

SOS components and values.

Component	**Value**
Glasgow Coma Scale < 5	3.75
$FiO_2 > 0.5$	0.75
Administration of sympathomimetic amine > 12 hours*	1.75
Oliguria < 0.5 ml/kg/hour for 8 hours	2.50
Coagulopathy after first admission day	1.25
TOTAL	10.00

* Excludes renal dose dopamine

Cluster analysis was then used to provide mortality predictions and the system was then validated in a further 2860 patients. Three patterns of unpredicted mortality were identified. These included sudden deterioration, delayed clinical deterioration, and patients dying with inappropriately low maximum SOS. Review of the care of these patients could then take place to see if there were any identifiable deficiencies in their medical care. The authors feel that their system is especially useful in assessing ongoing quality of patient care and medical surveillance, the latter use being especially relevant after an episode of malicious interference in their own institute (Istre et al 1985).

2.6.7 Clinical sickness score

The clinical sickness score (Watters et al 1989) was developed in Zambia, an area where laboratory tests are not easily available. It was derived from 624 consecutive admissions to a Surgical ICU in Lusaka and allocates scores from 0 to 4 for pulse, systolic blood pressure, respiratory rate, urine output, temperature, Glasgow Coma Scale, and age. This score is simple to perform using the above clinical variables, and was significantly associated with outcome in this study. The clinical sickness score was compared with APACHE II in 97 admissions to a District General Hospital ICU in England (Sinclair et al 1991). This is a small study, and the age weighting was excluded from the methodology. Nevertheless there was significant correlation between the two scoring systems both for all patients and for hospital survivors; correlation for nonsurvivors was not significant.

2.6.8 Simpler indicators of outcome

The systems described here are noteworthy for their simplicity, ease of use, and wide area of application. They will be briefly described here.

2.6.8.1 The Computerised Severity Index

The Computerised Severity Index calculates a severity score that incorporates the ICD-9-CM code for patient diagnosis and a severity score between 1 and 4 (Horn & Horn 1986). Severity is based on a set of objective signs. Its uses include adjustment of cost to severity of illness, quality of care assessment, prediction of posthospital needs, the incentive to efficiency, and physician practice profiles.

2.6.8.2 The comorbidity score

The comorbidity score defines a comorbidity as a preexisting condition (in addition to the active primary diagnosis) which has to either require treatment during the hospital admission or has permanently altered organ function (Gross et al 1988). The authors found that the number of comorbidities varied directly with the development of nosocomial infection, appearance of new complications, and length of ICU stay.

2.6.8.3 The McCabe-Jackson scoring system

The McCabe-Jackson scoring system uses clinical judgement to assign one of three risk groups to an individual patient: rapidly fatal (e.g. acute leukaemia), ultimately fatal (e.g. chronic leukaemia, cirrhosis with bleeding varices), and nonfatal (McCabe & Jackson 1962).

2.6.8.4 The American Society of Anesthesiologists (ASA) classification

The American Society of Anesthesiologists (ASA) classification uses clinical data to assign a patient to a Physical Status class between I and V. Class I describes a healthy patient, Class II patients have mild systemic disease but no functional limitation, Class III have severe systemic disease and definite functional limitation, Class IV patients have severe systemic disease that is a threat to life, and Class V patients are moribund and unlikely to survive 24 hours with or without operative intervention (Dripps et al 1961).

The above four systems have been compared with each other and with APACHE II in 105 patients (Gross et al 1991). All systems correlated well with the authors' own comorbidity score. In addition, APACHE II, the Computerised Severity Index, and the McCabe-Jackson scoring system were comparable predictors of comorbidity. The authors concluded that the selection of one or other system would depend on the resources available and the intended uses.

2.6.8.5 The Visceral System Failure Index

The Visceral System Failure Index incorporating 11 empirically derived visceral "systems" (including conventional organ-systems and severe sepsis, major surgery, disseminated malignancy, major blood transfusion, and major trauma) was used in a small study of 79 patients in Australia (Kelly et al 1986). It was significantly correlated with mortality only when there were 4 system failures (and not at less than 4 failures, or at 5 or more failures). The numbers are too small to draw any conclusions.

2.6.8.6 The Hanover Intensive System

The Hanover Intensive System (HIS) seems to have been empirically derived and includes simple assessments of cerebral, cardiovascular, respiratory, gastrointestinal, renal, and immunological function (Lehmkuhl et al 1989). In a study of 215 surgical ICU patients it was found to be superior to APACHE II and TISS in predicting death; it was also useful in clinical decisions about reoperation and for administrative decisions predicting intensity of nursing care. The authors seem hostile towards APACHE II (describing the variables as a "potpourri") and TISS (describing the measurements as "bewildering"). The HIS looks simple and easy to use; it remains to be seen if it becomes popular and can be widely validated.

2.6.8.7 Critical Care Scoring System

The Critical Care Scoring System (Yeung et al 1990) is applicable only in patients managed with pulmonary artery catheters. This score is largely based on haemodynamic data; the rationale for its development was that in patients with circulatory or respiratory failure and needing pulmonary artery catheters the authors found a surprisingly high mortality (45.7%) with relatively low APACHE II scores (less than 14). Their scoring system managed to significantly improve mortality prediction in such patients.

2.6.8.8 POSSUM

The POSSUM (acronym for Physiological and Operative Severity Score for the enUmeration of Mortality and morbidity) was developed as a simple scoring system for general surgical patients, whose main use would be for audit (Copeland et al 1991). When compared with APACHE II in 117 admissions to a post-surgical high dependency unit, POSSUM had a superior predictive value for 30 day mortality (Jones et al 1992).

2.7 New scoring systems

The need for further scoring systems is perhaps summed up by the titles of two recent editorials: *"New and improved" scoring systems* (Civetta 1990) and *Do we need a new severity score?* (Le Gall & Lemeshow 1991). The former was prompted by the appearance of the Critical Care Scoring System (Yeung et al 1990), which met with guarded approval, and the latter by a new score based on logistic regression analysis of all the variables making up the currently used prognostic indices (Sarmiento et al 1991). This score was felt to be seriously flawed on the basis of the methodology employed. Le Gall and Lemeshow felt that it would be better that current severity scores were further studied and periodically modified to cater for differences in ICU patient mix and the development of new technologies.

2.8 Specialised scoring systems for specific disorders

There is a vast number of scoring systems applicable only to specific disorders. Only the better known ones will be described here and then not in any depth. In their own circumscribed diagnostic categories they may be extremely effective; this will not be critically reviewed here.

2.8.1 Trauma

For head injury the main scoring systems are the Glasgow Coma Scale (Teasdale & Jennett 1974, Jennett et al 1979), and for general trauma the Injury Severity Score (Baker et al 1974, Copes et al 1988), the Trauma Score (Champion et al 1981), the

Abbreviated Injury Scale (Civil & Schwab 1988), the TRISS Method (Boyd et al 1987) and the ASCOT score (Champion et al 1990).

2.8.2 Paediatrics

The best known score for paediatrics is the neonatal Apgar score (Apgar 1953). More recent scores include the Pediatric Risk of Mortality (PRISM) score (Pollack et al 1988) and the neonatal CRIB score (The International Neonatal Network 1993).

2.8.3 Pancreatitis

The Ranson score (Ranson et al 1974) is probably the best known and most widely used score for patients with pancreatitis. A more recent addition is the Glasgow Score (Blamey et al 1984).

2.8.4 Burns

Burn patients are commonly scored by the Burn Index (Feller et al 1980). Other scores include those of Baux (Stern & Waisbren 1978), Bull (Bull 1971), and Roi (Roi et al 1983). More recently, logistic regression indices have been used and compared with the established indices (Zoch et al 1992).

2.8.5 Septic shock

Patients with septic shock may be scored by the Complete and Simplified Septic Shock Scores (Baumgartner et al 1992).

2.8.6 Non-traumatic coma

Prognostic estimates for patients with non-traumatic coma may be made from the detailed descriptions of the clinical course of patient groups (Levy et al 1981, Levy et al 1985, Edgren et al 1994).

2.8.7 Comparison of general and specific scoring systems

The performance of specific scoring systems is sometimes not as good as general scoring systems, even in their perceived speciality. In pancreatitis for instance, APACHE II performs as least as well as the Ranson score (Larvin & McMahon 1989, Wilson et al 1990, Roumen et al 1992). However, in head trauma the Glasgow Coma Scale was superior to the Acute Physiology Score component of APACHE II, SAPS, and TISS in predicting outcome, as measured by both correct prediction of outcome and area under the ROC curve (Rocca et al 1989).

2.9 Summary

Severity of illness scores have developed substantially over the past ten to fifteen years and current scoring systems are based on mathematically and statistically sound principles. Databases are continuing to expand and predictive equations to be refined. The prognostic power of scores is increasing and there are few ICUs that do not use some form of severity scoring as part of their basic data collection. Economics may demand the expansion of medical audit and severity scoring is at its best here. Economics may also demand more appropriate use of ICU facilities and it is here, at "the intersection of economics and ethics in the intensive care unit" (Lanken 1994) that the most interesting developments in severity scoring may yet be seen.

CHAPTER 3

Literature review. Severity scoring and predictive indices in cardiac surgery

3.1 General scoring systems in cardiac surgery

3.1.1 Introduction

Cardiac surgery patients (and in particular Coronary Artery Bypass Graft [CABG] patients) have been almost entirely avoided or ignored by evaluations of general severity of illness scoring systems, especially the more recent APACHE III score (Knaus et al 1991), MPM II (Lemeshow et al 1993) and SAPS II (LeGall et al 1993).. This may well be a reflection of the difficulties both expected and encountered in predicting outcome in such patients, as well as to the availability of specific risk predictors for cardiac surgery. This theme is more fully explored in the introductions to Chapters 6 and 7. The risk predictors are not ideal either, and it has been recently stated that "an objective scoring system pertinent to the cardiothoracic ICU would be extremely timely and useful" (Rafkin & Hoyt 1994).

3.1.2 APACHE II

In the APACHE II database, there were 90 patients classified as "Postoperative, admission due to chronic cardiovascular disease" and 225 as "Postoperative, heart valve surgery", out of a total of 5030 patients. Our study evaluating the use of

APACHE II scoring in 811 cardiothoracic surgery patients, including 527 CABG patients out of 765 cardiac surgery patients (Turner et al 1991), is thus the largest study of a general severity of illness scoring system to be published in this field. It forms part of this thesis and is fully described below. The conclusions were that low APACHE II scores and risk of dying predicted survival with some certainty (the mortality was 0.93% when the APACHE II score was less than 10) but that only very high scores accurately predicted death, and then only in a small number of patients. The relationship between the APACHE II score and mortality was linear and statistically significant but APACHE II did not have the power to definitively predict death and thereby influence decisions to withdraw therapy. APACHE II has however been used in the prediction of septic complications after cardiac surgery. It was found to be superior to single variables such as fever, white cell count, and cardiac output in differentiating between patients who developed sepsis and those who did not. An APACHE II score of 19 had a positive predictive value of 86% and a negative predictive value of 96% in predicting septic complications (Kreuzer et al 1992).

3.1.3 APACHE III

Patients admitted to ICU after CABG were excluded from the APACHE III data collection and prognostic system (Draper et al 1989, Knaus et al 1991) and although they are to be analysed separately (Knaus WA, personal communication) this data is not yet available. Patients undergoing heart valve surgery have however been used in the original APACHE III database although in small numbers: there were only 211 such patients (1.2%) out of a total of 17440 patients

(Knaus et al 1991). APACHE III has not been independently validated in this (or any other) setting.

3.1.4 SAPS II, MPM I and II

No cardiac surgery patients at all were evaluated in SAPS II (LeGall et al 1993) or the Mortality Prediction Models I and II (Lemeshow et al 1988, Lemeshow et al 1993).

3.2 Cardiac surgery operative risk systems

3.2.1 Introduction

A variety of cardiac surgery operative risk systems have been developed from analyses of large numbers of patients, the largest and most important of which are detailed in Table 3.1 together with their methodology and end points; they will not be individually described further.

Table 3.1

Cardiac surgery risk analyses with their descriptions and end-points.

First author (date)	Description	End-point
Oldham (1972)	High/low risk groups	Mortality
Loop (1975)	Risk factors	Mortality
Kennedy (1980)	Risk factors, calculation	Mortality (CASS)
Pelletier (1980)	Risk factors	Mortality
Kennedy (1981)	Risk factors	Mortality (CASS)
Junod (1987)	CASS calculation	Mortality
Edwards (1989)	Bayesian analysis	Mortality
Parsonnet (1989)	Simple score	Mortality
Bolsin (1990)	Simple score	Mortality, duration
Hammermeister (1990)	Regression model	Major complications
Hannan (1990)	Regression model	Mortality
Higgins (1992)	Simple score	Morbidity, mortality
Tuman (1992)	Simple score	Morbidity, mortality
Geraci (1993)	Risk factors	Mortality, morbidity
Daly (1993)	Simple score	Mortality (males only)
Grover (1993)	Regression model	Mortality

3.2.2 Interpretation of the literature

Several of the above risk systems derive a numerical score, some of which are simple sums of points allocated for the presence or absence of risk factors, while others involve calculation of exponential functions. The former are easy to apply at the bedside (which is where they are needed if they are to be used clinically and just for research purposes) but perhaps less accurate than those that employ logistic regression analyses and may need scientific calculators or microcomputer programs to perform. Some sort of balance between these extremes would be useful.

Unfortunately interpretation of the literature is very confusing, as all the risk systems evaluate different patient populations and types of operation (some include CABG only), as well as different combinations of predictive factors and end-points (with different definitions for each of these), and although some more recent ones have used earlier studies as their basis, no single system has gained widespread acceptance. This is a pity, as the studies involve a total of tens of thousands of patients.

A further problem may be that the above studies span many years; patient profile and mortality have changed during that time (Naunheim et al 1988) and cardiac surgery continues to change. Results from the earlier studies may no longer be applicable.

Finally, apart from the Bayesian model (Edwards et al 1989), there is no risk stratification in any of the studies. Raw mortality data may not be the most appropriate outcome measure.

3.2.3 Independent validation

It would appear that only the Parsonnet system has been independently validated in any study other than its instigative one (Nashef et al 1992). This brief study of

1071 patients in a British hospital, with minimal description of the methods used, has been criticised for both its methodology and its conclusions (Spiegelhalter 1992). In the same letter this statistician criticises the unclear and unconventional derivation of the Parsonnet score and the tests used to correlate its predictions and outcomes (Spiegelhalter 1992).

Some uniformity in risk determination may emerge in the UK, where the Association of Cardiothoracic Anaesthetists (ACTA) has determined that the score of Bolsin et al (Bolsin et al 1990) should be used to assess operative risk. A large multicentre database has been established with substantial funding and there is a strong statistical input.

3.2.4 Collaborative Study in Coronary Artery Surgery (CASS)

The largest study of coronary artery bypass surgery patients, with more than 25000 patients when fully reviewed in the literature (Anderson 1986), is almost certainly the Collaborative Study in Coronary Artery Surgery (CASS), which has been reported at various times and with regard to various subgroups. These subgroups include left main-stem disease (Chaitman et al 1980), left ventricular aneurysm (Faxon et al 1982), unstable angina (McCormick et al 1985), patients older than 65 years (Gersh et al 1983), the effects of sex and physical size (Fisher et al 1982), and repeat surgery (Foster et al 1984).

CASS has been comprehensively reviewed (Anderson 1986). It is impressive in size and scope and consists of two elements: a randomised trial of surgical versus medical therapy, and a registry which includes a risk equation to estimate hospital mortality. The risk equation incorporates (in descending order of importance) age, left main coronary artery stenosis greater or equal to 90%, female sex, left ventricular score (a combination of ejection fraction and wall motion abnormality), and left ventricular end-diastolic pressure (Kennedy et al 1980); it involves exponential functions and needs a computer or programmable calculator to perform. A simplified method of determining left ventricular function (as a substitute for the complex left ventricular score) has been described (Pierpont et al 1985).

3.2.5 Other large studies

Another impressively large study of over 12000 coronary artery surgery patients comes from Australia (Iyer et al 1993). The study analyses operative mortality and myocardial infarction in patients undergoing CABG for the first time between 1978 and 1990. The overall operative mortality was impressively low at 0.99%, and showed what superficially appeared to be a decreasing trend between 1978 and 1990, with a mortality of 0.4% for the final year of the study, although univariate analysis showed epoch of operation not to influence the outcome.

By stepwise logistic regression, the authors found that age, female sex, poor ventricular function, the presence of unstable angina, and bypass time were related to higher operative mortality, and they derived a general logistic model risk equation from this data. By similar statistical methods, the epoch of the operation, bypass time and unstable angina were associated with postoperative myocardial infarction.

3.2.6 Patients with an intraaortic balloon pump

In a more limited or defined setting, variables to predict mortality in patients who need an intraaortic balloon pump to separate from cardiopulmonary bypass have been identified from an analysis of 322 such patients. These include need for temporary pacing, advanced age, preoperative blood urea nitrogen concentration, and female sex (Baldwin et al 1993).

3.2.7 Parsonnet score

The Parsonnet score is described in some detail, as it was selected by the Royal Brompton National Heart and Lung Hospital to assess preoperative risk for patients undergoing cardiac surgery. The authors retrospectively analysed 17 variables (all of which had to be available for every patient, and to be objective and not derived measurements) in 3500 adult patients undergoing open-heart surgical procedures (Parsonnet et al 1989). The outcome measure was death within 30 days. Using univariate and multivariate logistic regression a model was derived, using 15 statistically significant variables. This model was then tested prospectively in 1332 patients, and the predicted vs observed mortality correlation coefficient was 0.99. The following variables were excluded from analysis as they were either too subjective, indefinable, or unavailable: the presence of chronic obstructive airways disease, bypass time, the number of bypass grafts and the use of the internal mammary artery, and operative priority. The variables making up the score appear in Figure 3.1.

Figure 3.1

Variables making up the Parsonnet score, with their respective points.

Female sex		1
Morbid obesity (>1.5X ideal weight)		3
Diabetes		3
Hypertension (Systolic BP>140 mm Hg)		3
Ejection fraction	Good (≥ 50%)	0
	Fair (30-49%)	2
	Poor (<30%)	4
Age	70-74 years	7
	75-79 years	12
	≥ 80 years	20
Reoperation	First	5
	Second	10
Preoperative Intraaortic Balloon Pump		2
Left ventricular aneurysm surgery planned		5
Emergency surgery after PTCA*/catheter		10
Dialysis dependency		10
Catastrophic states		10-50
Other rare events		2-10
Valve surgery	Mitral	5
	Pulmonary artery pressure ≥ 60 mm Hg	8
	Aortic	5
	Peak Systolic Gradient >120 mm Hg	7
CABG and valve surgery		2

* PTCA Percutaneous Transluminal Angioplasty

3.2.7.1 Problems with the Parsonnet score

3.2.7.1.1 Methodology

The methodology of the score has been criticised on two points (Spiegelhalter 1992). Firstly it is not clear in the text as to where univariate and multivariate analyses were performed. Secondly, the comparison between the anticipated and the observed mortality, and between two different models, was tested by simple regression. The quoted measure of correlation was a correlation coefficient. The correct method of comparing two methods of clinical measurement, by plotting the difference against the mean, has been described (Bland & Altman 1986).

3.2.7.1.2 Allocation of points

In general, the definitions and criteria for the variables that make up the score are objective and simple to apply. The most glaring weakness in the score is the arbitrary allocation of points in the "catastrophic states" and "other rare events" categories (up to 50 and 10 points respectively at the discretion of the surgeon). This is subject to wide variations between scorers, even if a single scorer always allocates the same number of points for a certain condition. In addition, the exclusion of potentially important variables such as operative priority and chronic obstructive airways disease may weaken the predictive ability. Finally, there are differences in interpretation as to whether patients undergoing valve surgery and with pulmonary artery pressure ≥ 60 mm Hg or peak systolic gradient > 120 mm Hg should score points for both that valve surgery and the pressure abnormality.

3.3 Individual risk factors in CABG surgery

In a recent review, the effects of impaired left ventricular function, older age, sex, and reoperation were singled out as consistent predictors of high operative risk (Rosenfeldt & Wong 1993). Another review again singles out the risk factors analysed in the above papers, but adds that anaesthetic strategy and the prevention, detection, and treatment of perioperative ischaemia are important (Koch & Estafanous 1993).

A wide variety of risk factors have been identified in cardiac surgery; most of these appear in the risk systems evaluated in the section above. Even the anaesthetist has been implicated as an independent risk factor. In a recent study, both anaesthetist and bypass time were identified as being significantly associated with hospital death or raised cardiac enzymes (Merry et al 1992).

The individual factors described below have been well established in CABG surgery where patient numbers are large; they are generally accepted but less well established for other cardiac surgery procedures. Each factor is well represented in the literature and only the most important or recent studies will be reviewed.

3.3.1 Gender

Operative mortality in women is higher than in men. This was clearly shown in the CASS study, where most of this risk was attributed to their smaller vessels and stature (Fisher et al 1982). Subsequent studies have also shown female sex to be an independent risk factor (Loop et al 1983, Hannan et al 1992). The effect continues in the long term; ten and fifteen year survival is significantly worse in females than in males (Rahimtoola et al 1993). However concerns have been raised that this increased risk may have become a self-fulfilling prophecy, as women may be referred for surgery later than men (Tobin et al 1987, Khan et al 1990, Bickell et al 1992).

3.3.2 Age

Older age is an independent predictor of operative mortality (Weintraub et al 1991, O'Connor et al 1992). Patients over 65 years were shown to have significantly higher mortality in the CASS study (Gersh et al 1983). The effect of age has been recently reviewed (Mohan et al 1992). Twenty-one studies of at least 100 patients from 1973 to 1991 were included in the review, which examined both mortality and morbidity. In general terms, both these end-points were higher in the elderly. Not all studies reached statistical significance; nevertheless there were no studies that showed significantly worse mortality or morbidity in the younger age ranges. However female gender ceases to be a risk factor in the elderly; in octogenarians there was a trend towards lower operative mortality than men (Mohan et al 1992), and older women (70 years or older) were at no greater risk of mortality or postoperative complications when compared with younger women or older men (King et al 1992). In addition, in patients aged 80 to 93 years, female sex was not associated with a worse functional outcome (Glower et al 1992).

3.3.3 Impaired left ventricular function

Poor left ventricular function consistently emerges as an important predictor of operative risk. In a recently published study of 12471 patients undergoing CABG, operative mortality was 2.3% for left ventricular ejection fraction greater than 40%, 4.8% for left ventricular ejection fraction of 20% to 40%, and 9.8% for left ventricular ejection fraction of less than 20% (Christakis et al 1992). In the CASS study, operative mortality was 1.9% when the left ventricular ejection fraction was more than 49% and 6.7% when it was less than 19%; in addition mortality increased as the left ventricular wall motion score became more abnormal (Kennedy et al 1981). More recently, both left ventricular ejection fraction and end-diastolic pressure were found to be significantly associated with mortality (O'Connor et al 1992). Long term results are also associated with left ventricular function; ten and fifteen year survival was significantly worse in patients with abnormal left ventricular function (Rahimtoola et al 1993).

3.3.4 Reoperation

Reoperation is technically more difficult than first-time surgery; in addition as surgical results improve and patients outlive their grafts the incidence of reoperation will increase from the current 8.3 % (Akl et al 1992). Reoperation rates increase with time. For the first five postoperative years the rate of reoperation was 0.6% per year but it rose to 2% per year from years 6 to 15 (Rahimtoola et al 1993). Not unexpectedly, reoperation carries higher risks. This is confirmed in the CASS data (Foster et al 1984) and in a large Australasian study

(Osaka et al 1988) as well as a study from the Netherlands where the 30 day mortality was 7.5% (Verheul et al 1991) and one from the UK where the 30 day mortality was 5.2% (Akl et al 1992). In addition, postoperative morbidity is high and perioperative myocardial infarction rates may be more than 9% (Salomon et al 1990). The risks are not generally thought to be unacceptable when compared with the benefits: actuarial survival rates of 90% at 5 years and 88% at 10 years have been achieved (Akl et al 1992).

Even third-time operations are being performed. Numbers are obviously small but a recent study of 13 patients observed a hospital mortality of 7.7% and impressive symptomatic improvement (Merrill et al 1993).

3.3.5 Diabetes mellitus

The data on diabetes mellitus is not as clear. In addition, diabetes may be difficult to separate from associated factors such as atherosclerosis, hypertension, and renal failure. Diabetics with autonomic dysfunction who underwent ophthalmic surgery had a greater incidence of cardiovascular lability than nondiabetic controls (Burgos et al 1989). Although hospital stay was increased in diabetic patients undergoing CABG, the overall morbidity was not increased (Clement et al 1988). In another study, there was an increased incidence of postoperative arrhythmias, wound infections, and respiratory failure (Fietsam et al 1991). In insulin dependent diabetics with chronic renal failure and silent coronary artery disease, revascularisation decreased cardiac morbidity and mortality (Manske et al 1992).

3.3.6 Emergency surgery

The CASS data for surgery in patients with unstable angina demonstrated an operative mortality of 3.9% and 7 year cumulative survival rate of 79% (McCormick et al 1985). Emergency operation after failed angioplasty led to significantly higher mortality (12%), morbidity, and length of hospital stay (Parsonnet et al 1988). "True" emergency surgery was analysed in 117 patients over a five year period; hospital mortality was 14.5% and 36% of patients developed major complications (Edwards et al 1990).

3.3.7 Obesity

Obesity has been shown to be an independent risk factor for perioperative morbidity in patients undergoing CABG (Prasad et al 1991). In addition obese patients had a greater incidence of recurrent angina at a mean follow up time of 36.9 months.

3.3.8 Confounding issues

Potential confounding issues when examining risk factors over a period of time include differences in operative technique and in patient population. With regard to operative technique, the use of arterial grafts has now become widespread and long term graft patency is impressive. They include the internal mammary artery (Acinapura et al 1992, Naunheim et al 1992), gastroepiploic artery (Suma et al 1993), and more recently a re-emergence of the use of the radial artery (Acar et al 1992).

Patient populations also change over time. Higher risk patients are now likely to be considered for surgery (Naunheim et al 1988, Jones et al 1991). Populations will continue to change as more operations are performed. In addition, with patients outliving their grafts or valve replacements, the number of reoperations is certain to increase.

3.4 Summary

Dissatisfaction with general severity scoring systems may be related to unrealistic expectations but areas such as cardiac surgery remain problematic. In this group, the changes in both the operative techniques and the patient profiles have led to difficulties in interpreting the literature and the large number of preoperative predictive systems available may bear witness to this problem. With general severity of illness scoring systems avoiding cardiac surgery patients, there certainly may be a need to combine preoperative data with intraoperative and postoperative physiology, in order to obtain better predictive ability.

CHAPTER 4

Methods

4.1 Introduction

Methods applicable to the thesis as a whole are discussed in this chapter. More specific methods used only in, or relevant only to the separate studies are discussed in the Protocol sections of the relevant chapters.

4.2 Type of study

All three studies undertaken were prospective and descriptive, with no interventions performed upon the patients.

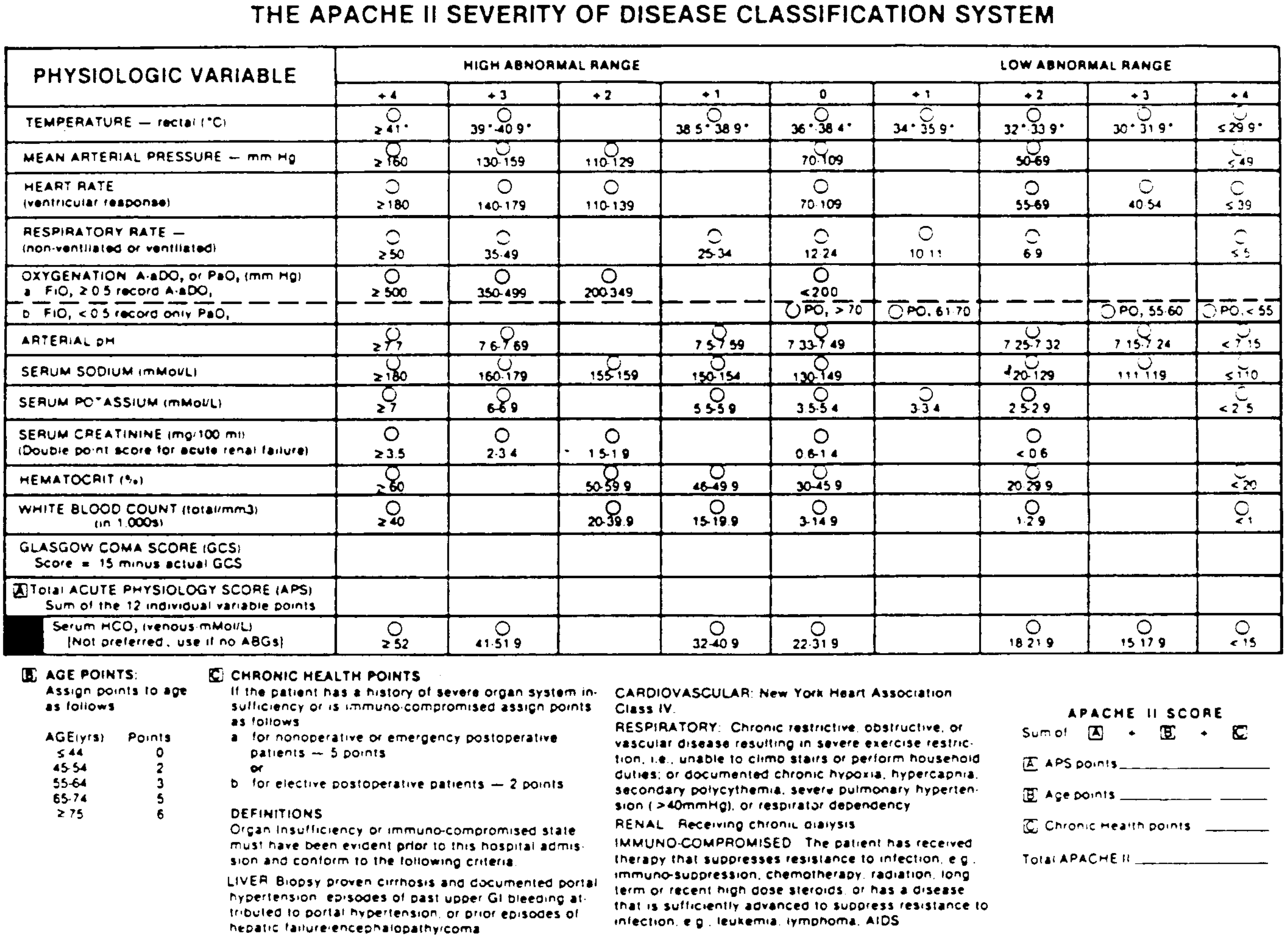
4.3 Scoring systems

4.3.1 APACHE II

The APACHE II score (Knaus et al 1985) is based on 11 physiological measurements, the Glasgow Coma Scale, age, and previous health status. Each physiological measurement is scored from 0-4 points depending on its deviation from normal and the total score is determined from the most deranged value of each measurement in a designated 24 hour period (generally the first 24 hours in the ICU). An example of an APACHE II scoring sheet with definitions for age and chronic health points is shown in Figure 4.1.

Figure 4.1

An example of an APACHE II scoring sheet with definitions for age and chronic health points (reproduced from Critical Care Medicine 1985;13:820).



THE APACHE II SEVERITY OF DISEASE CLASSIFICATION SYSTEM

PHYSIOLOGIC VARIABLE	HIGH ABNORMAL RANGE					LOW ABNORMAL RANGE			
	+4	+3	+2	+1	0	+1	+2	+3	+4
TEMPERATURE — rectal (°C)	○ ≥41°	○ 39°-40 9°		○ 38 5°-38 9°	○ 36°-38 4°	○ 34°-35 9°	○ 32°-33 9°	○ 30°-31 9°	○ ≤29 9°
MEAN ARTERIAL PRESSURE — mm Hg	○ ≥160	○ 130-159	○ 110-129		○ 70-109		○ 50-69		○ ≤49
HEART RATE (ventricular response)	○ ≥180	○ 140-179	○ 110-139		○ 70-109		○ 55-69	○ 40-54	○ ≤39
RESPIRATORY RATE — (non-ventilated or ventilated)	○ ≥50	○ 35-49		○ 25-34	○ 12-24	○ 10-11	○ 6-9		○ ≤5
OXYGENATION A-aDO, or PaO, (mm Hg) a FiO, ≥0 5 record A-aDO,	○ ≥500	○ 350-499	○ 200-349		○ <200				
b FiO, <0 5 record only PaO,					○ PO, >70	○ PO, 61-70		○ PO, 55-60	○ PO, <55
ARTERIAL pH	○ ≥7 7	○ 7 6-7 69		○ 7 5-7 59	○ 7 33-7 49		○ 7 25-7 32	○ 7 15-7 24	○ <7 15
SERUM SODIUM (mMol/L)	○ ≥180	○ 160-179	○ 155-159	○ 150-154	○ 130-149		○ 120-129	○ 111-119	○ ≤110
SERUM POTASSIUM (mMol/L)	○ ≥7	○ 6-6 9		○ 5 5-5 9	○ 3 5-5 4	○ 3-3 4	○ 2 5-2 9		○ <2 5
SERUM CREATININE (mg/100 ml) (Double point score for acute renal failure)	○ ≥3.5	○ 2-3 4	○ 1 5-1 9		○ 0 6-1 4		○ <0 6		
HEMATOCRIT (%)	○ ≥60		○ 50-59 9	○ 46-49 9	○ 30-45 9		○ 20-29 9		○ <20
WHITE BLOOD COUNT (total/mm3) (in 1,000s)	○ ≥40		○ 20-39.9	○ 15-19.9	○ 3-14 9		○ 1-2 9		○ <1
GLASGOW COMA SCORE (GCS) Score = 15 minus actual GCS									
[A] Total ACUTE PHYSIOLOGY SCORE (APS) Sum of the 12 individual variable points									
Serum HCO, (venous-mMol/L) [Not preferred, use if no ABGs]	○ ≥52	○ 41-51 9		○ 32-40 9	○ 22-31 9		○ 18-21 9	○ 15-17 9	○ <15

[B] AGE POINTS:
Assign points to age as follows

AGE(yrs)	Points
≤44	0
45-54	2
55-64	3
65-74	5
≥75	6

[C] CHRONIC HEALTH POINTS
If the patient has a history of severe organ system insufficiency or is immuno-compromised assign points as follows

a for nonoperative or emergency postoperative patients — 5 points
or
b for elective postoperative patients — 2 points

DEFINITIONS
Organ Insufficiency or immuno-compromised state must have been evident prior to this hospital admission and conform to the following criteria.

LIVER Biopsy proven cirrhosis and documented portal hypertension episodes of past upper GI bleeding attributed to portal hypertension, or prior episodes of hepatic failure/encephalopathy/coma

CARDIOVASCULAR: New York Heart Association Class IV.

RESPIRATORY: Chronic restrictive, obstructive, or vascular disease resulting in severe exercise restriction, i.e., unable to climb stairs or perform household duties; or documented chronic hypoxia, hypercapnia, secondary polycythemia, severe pulmonary hypertension (>40mmHg), or respirator dependency

RENAL Receiving chronic dialysis

IMMUNO-COMPROMISED The patient has received therapy that suppresses resistance to infection, e g., immuno-suppression, chemotherapy, radiation, long term or recent high dose steroids, or has a disease that is sufficiently advanced to suppress resistance to infection, e g., leukemia, lymphoma, AIDS

APACHE II SCORE
Sum of [A] + [B] + [C]

[A] APS points ______________
[B] Age points ______________
[C] Chronic Health points ______________
Total APACHE II ______________

4.3.1.1 Problems with the APACHE II methodology

There have been many criticisms levelled at the objectiveness of the APACHE II method. Areas of particular concern which were not discussed in Chapter 1 include the following:

1. Arterial blood gas analysis. This may be uninterpretable or unrepresentative of the acute physiological status of the patient in congenital heart disease, chronic pulmonary disease, and in ventilatory manoeuvres such as permissive hypercapnia. In addition, if this test happens to be performed more frequently, the likelihood of an abnormal result and higher APACHE II points might increase.
2. Glasgow Coma Scale. Extreme caution may be needed in sedated patients, those with polyneuritis, or those who are simply deaf.
3. Emergency surgery. The precise definition of what constitutes emergency surgery is open to interpretation.
4. Reason for admission. For the same patient there may be several reasons for admission, which may occasionally give very different risks of death for the same APACHE II score.

4.3.1.2 Management of APACHE II methodological problems

Specific controversial points (Palazzo & Patel 1993) were dealt with in the following way in these studies:

1. Timing of scores. All APACHE II data were collected from measurements performed in the first 24 hours after ICU admission (Knaus et al 1985). In the second and third studies, if a patient was admitted to ICU just prior to surgery, the 24 hour period began on return from the operating theatre.

2. Missing data. There were no missing data for APACHE II scoring.

3. Age. Although APACHE II should not strictly be applied to patients under 18 years of age, we have used it in a few such patients as they have been considered "adult" enough to be in an adult ICU.

4. Chronic health points. These were rigorously applied according to the criteria of Knaus et al (Knaus et al 1985).

5. Blood gases. The pH was taken from the blood gas with the worst value for oxygenation (the worst pH during the first 24 hours was independently and separately documented in the third study).

6. Acute renal failure. This was regarded a rise in serum creatinine (to more than 135 μmol/l within the last 24 to 48 hours) and oliguria (less than 135 ml urine over a consecutive 8 hour period).

7. Glasgow Coma Scale. This was always scored as 15 (best possible value) unless there was clear evidence of neurological dysfunction over and above that caused by anaesthetic agents, sedatives, opioid analgesics, and muscle relaxants. It was retrospectively scored if abnormalities presumed to have been present all along were noted once sedation had worn off.

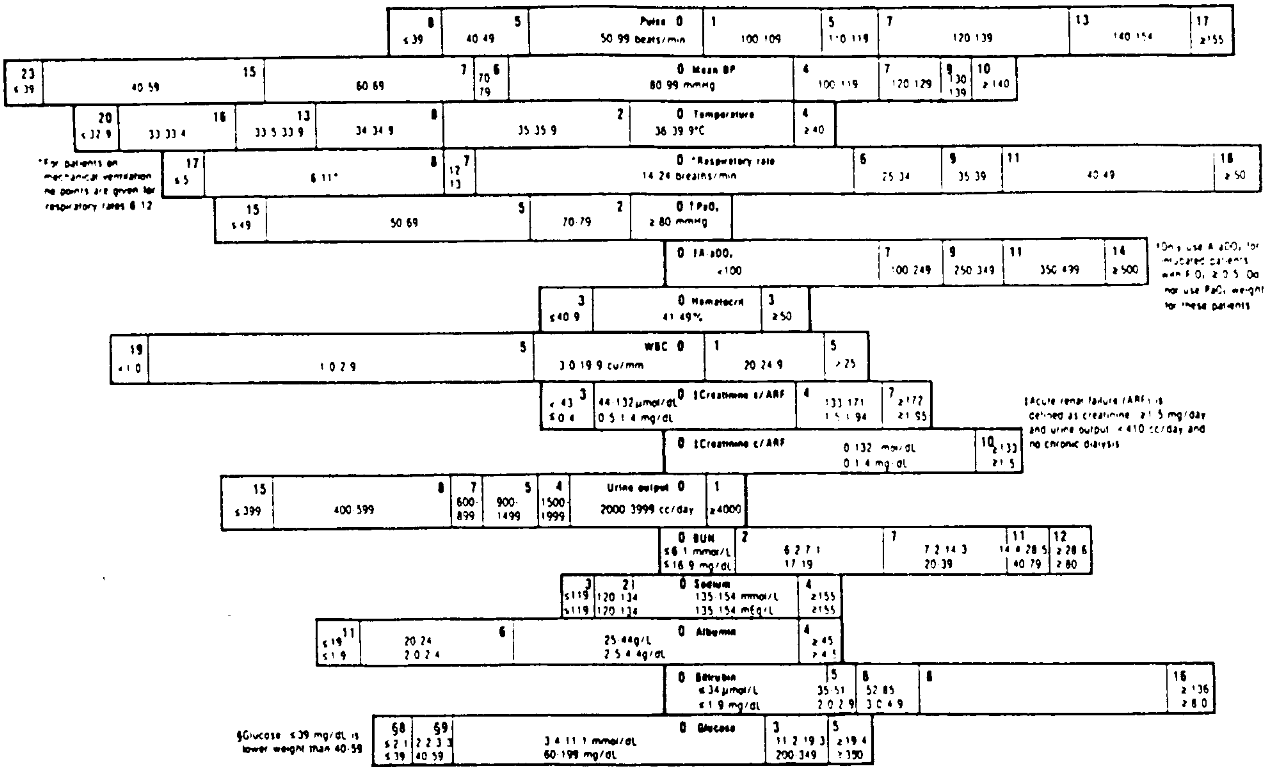
4.3.2 APACHE III

The APACHE III Prognostic system (Knaus et al 1991) is described in Chapter 1. The component variables and their weightings that make up the APACHE III score are outlined in Figure 4.2.

Figure 4.2

Component variables and their weightings that make up the APACHE III score (reproduced from Chest 1991;100:1622-1624).

APACHE III PHYSIOLOGIC SCORING FOR VITAL SIGNS AND LABORATORY TESTS



Pulse	8	5	0	1	5	7	13	17
	≤39	40-49	50-99 beats/min	100-109	110-119	120-139	140-154	≥155

Mean BP	23	15	7	6	0	4	7	9	10
	≤39	40-59	60-69	70-79	80-99 mmHg	100-119	120-129	130-139	≥140

Temperature	20	16	13	8	2	0	4
	≤32.9	33-33.4	33.5-33.9	34-34.9	35-35.9	36-39.9°C	≥40

*Respiratory rate	17	8	7	0	6	9	11	18
	≤5	6-11*	12-13	14-24 breaths/min	25-34	35-39	40-49	≥50

*For patients on mechanical ventilation no points are given for respiratory rates 6-12

†PaO$_2$	15	5	2	0
	≤49	50-69	70-79	≥80 mmHg

†A-aDO$_2$	0	7	9	11	14
	<100	100-249	250-349	350-499	≥500

†Only use A-aDO$_2$ for intubated patients with FiO$_2$ ≥ 0.5. Do not use PaO$_2$ weights for these patients

Hematocrit	3	0	3
	≤40.9	41-49%	≥50

WBC	19	5	0	1	5
	<1.0	1.0-2.9	3.0-19.9 cu/mm	20-24.9	≥25

‡Creatinine s/ARF	3	0	4	7
	≤43 ≤0.4	44-132 μmol/dL 0.5-1.4 mg/dL	133-171 1.5-1.94	≥172 ≥1.95

‡Creatinine c/ARF	0	10
	0-132 mol/dL 0-1.4 mg/dL	≥133 ≥1.5

‡Acute renal failure (ARF) is defined as creatinine ≥1.5 mg/day and urine output <410 cc/day and no chronic dialysis

Urine output	15	8	7	5	4	0	1
	≤399	400-599	600-899	900-1499	1500-1999	2000-3999 cc/day	≥4000

BUN	0	2	7	11	12
	≤6.1 mmol/L ≤16.9 mg/dL	6.2-7.1 17-19	7.2-14.3 20-39	14.4-28.5 40-79	≥28.6 ≥80

Sodium	3	2	0	4
	≤119 ≤119	120-134 120-134	135-154 mmol/L 135-154 mEq/L	≥155 ≥155

Albumin	11	6	0	4
	≤19 ≤1.9	20-24 2.0-2.4	25-44 g/L 2.5-4.4 g/dL	≥45 ≥4.5

Bilirubin	0	5	6	8	16
	≤34 μmol/L ≤1.9 mg/dL	35-51 2.0-2.9	52-85 3.0-4.9		≥136 ≥8.0

Glucose	§8	§9	0	3	5
	≤2.1 ≤39	2.2-3.3 40-59	3.4-11.1 mmol/dL 60-199 mg/dL	11.2-19.3 200-349	≥19.4 ≥350

§Glucose ≤39 mg/dL is lower weight than 40-59

APACHE III PHYSIOLOGIC SCORING FOR NEUROLOGIC ABNORMALITIES

Eyes open spontaneously or to painful/verbal stimulation

verbal / motor	oriented converses	confused conversation	inappropriate words and incomprehensible sounds	no response
obeys verbal command	0	3	10	15
localizes pain	3	8	13	15
flexion withdrawal/ decorticate rigidity	3	13	24	24
decerebrate rigidity/ no response	3	13	29	29

Eyes do not open spontaneously or to painful/verbal stimulation

verbal / motor	oriented converses	confused conversation	inappropriate words and incomprehensible sounds	no response
obeys verbal command				16
localizes pain				16
flexion withdrawal/ decorticate rigidity			24	33
decerebrate rigidity/ no response			29	48

Table 2—*APACHE III Points for Age and Chronic Health Evaluation*

	Points
Age, yr	
≤44	0
45-59	5
60-64	11
65-69	13
70-74	16
75-84	17
≥85	24
Comorbid condition*	
AIDS	23
Hepatic failure	16
Lymphoma	13
Metastatic cancer	11
Leukemia/multiple myeloma	10
Immunosuppression	10
Cirrhosis	4

*Excluded for elective surgery patients.

(Figure 4.2 continued...)

APACHE III PHYSIOLOGIC SCORING FOR ACID BASE ABNORMALITIES

<table>
<tr><th>pCO^2
pH</th><th><25</th><th>25-<30</th><th>30-<35</th><th>35-<40</th><th>40-<45</th><th>45-<50</th><th>50-<55</th><th>55-<60</th><th>≥60</th></tr>
<tr><td><7.15</td><td colspan="6" rowspan="2">12</td><td colspan="3" rowspan="2">4</td></tr>
<tr><td>7.15- <7.2</td></tr>
<tr><td>7.20- <7.25</td><td colspan="2" rowspan="3">9</td><td colspan="2" rowspan="2">6</td><td colspan="2" rowspan="2">[illegible]</td><td colspan="3" rowspan="2">2</td></tr>
<tr><td>7.25- <7.30</td></tr>
<tr><td>7.30- <7.35</td><td colspan="3" rowspan="3">0</td><td colspan="4">1</td></tr>
<tr><td>7.35- <7.40</td><td colspan="2" rowspan="3">5</td><td colspan="4" rowspan="2">1</td></tr>
<tr><td>7.40- <7.45</td></tr>
<tr><td>7.45- <7.50</td><td>0</td><td colspan="2">2</td><td colspan="4"></td></tr>
<tr><td>7.50- <7.55</td><td colspan="4" rowspan="2">3</td><td colspan="5" rowspan="4">12</td></tr>
<tr><td>7.55- <7.60</td></tr>
<tr><td>7.60- <7.65</td><td rowspan="2">0</td><td colspan="3" rowspan="2"></td></tr>
<tr><td>≥7.65</td></tr>
</table>

4.3.2.1 Management of APACHE III methodological problems

Much of the APACHE III methodology is very similar to APACHE II and basic principles will not be repeated here. Specific points were dealt with in the following way in the third study:

1. Blood gases. The pH/PCO$_2$ score was taken from the blood gas with the worst value for oxygenation.

2. Urine output. If a patient who remained in the ICU for less than 24 hours had a 24 hour urine output of less than 2000 ml, that patient was automatically given a value of 2000 ml. If the output was greater than 2000 ml, the value was rounded off to the nearest 100 ml.

3. Missing data. Missing data points were tagged as such and the patients were given the mean values for the whole study and scored no APACHE III points for those variables.

4.4 Computer hardware

4.4.1 Apple (Apple Computer Inc.)

The data collection for the first study was performed using an Apple II personal computer.

4.4.2 IBM (International Business Machines Corporation)

Data collection and all analyses (database, spreadsheet, graphics, and statistics) for the second and third studies were performed using IBM-compatible personal computers. They ranged from an IBM XT personal computer with 4 MHz 8086 microprocessor, through those with 80286 and 80386 microprocessors, to a more recent machine with 66 MHz i486 DX2 microprocessor and 8 Mb of RAM.

4.4.3 VAX mainframe computer (Digital Equipment Corporation)

A VAX mainframe computer at the University of Cape Town was used to run the BMDP statistical software (see below).

4.5 Computer software

4.5.1 dBASE II

dBASE II (Ashton-Tate, Torrance, California, USA) is an early version of a programmable relational database (used on an Apple personal computer).

4.5.2 dBASE IV

dBASE IV (Ashton-Tate, Torrance, California, USA) is a further development of the above relational database with its own programming language and structured query language.

4.5.3 Riyadh ICU Program (RIP)

The Riyadh ICU Program (RIP) 3.0 (Copyright R Chang 1988, Medical and Associated Software House, London, UK) is a DOS based data management program consisting of a data entry interface, a database, a number of simple statistical functions, and a structured query option which can provide a wide variety of reports. It is designed by clinicians specifically for use in an ICU and its analyses and predictions are based on a refinement of the APACHE II scoring system (Chang et al 1987). It was produced to be used for audit, for resource management, and as an aid to clinical decision making.

4.5.4 Paradox

Paradox 2.0 (Ansa Software, Belmont, California, USA) and Paradox for Windows 1.0 (Borland International Inc, Scotts Valley, California, USA) are two developments of a powerful relational database. The latter program has the advantage of an ability to form direct links with ObjectVision and to be read without the need for conversion in a Quattro Pro spreadsheet.

4.5.5 ObjectVision

ObjectVision 2.1 (Borland International Inc, Scotts Valley, California, USA) is a program designed to visually create Windows applications without experience in computer programming. The interface is a form which is customised by the user to create an unlimited variety of configurations. Procedures and calculations, as well as mathematical functions, can be embedded by means of a value tree into an application to calculate further field values or carry out commands. Data entered into such a program can be automatically linked to a database (in this case a Paradox database). The final program can be extremely sophisticated and can be used by someone with minimal knowledge of computers. Examples of an ObjectVision form and value tree are shown in Chapter 7.

4.5.6 Quattro Pro

Quattro Pro for Windows 5.0 (Borland International Inc, Scotts Valley, California, USA) is a Windows based spreadsheet with numerous graphical, mathematical, and statistical functions. It has the advantage of being able to directly read files of several formats, including (importantly in this study) Paradox files.

4.5.7 Epistat

Epistat (Copyright Tracy L Gustafson, Round Rock, Texas, USA) is a simple shareware statistical package with basic tests such as 2x2 table analysis.

4.5.8 Instat

Instat 2.04 (GraphPad Software, San Diego, California, USA) is a slightly more sophisticated statistical package which is extremely user-friendly and has a number of help-screens and explanations of the tests and the results.

4.5.9 StatGraphics

StatGraphics 4.0 and 5.0 (Manugistics Inc, Rockville, Maryland, USA) are developments of a powerful personal computer based statistical analysis package. They contain more than 250 statistical and system procedures as well as more than 50 types of graphical function.

4.5.10 BMDP

BMDP 7.0 (BMDP Statistical Software Inc., Los Angeles, California, USA) is a powerful statistical package capable of performing logistic regression analyses and receiver operating characteristic (ROC) curve analysis. It was run on a VAX mainframe computer at the University of Cape Town.

4.6 Statistical methods

4.6.1 Statistical tests

There was no statistical analysis performed in the first study, apart from the calculation of means.

In the second study, data was statistically analysed using the chi-square test (for two-by-two tables), chi-square test for trend and departure from trend (for relation between scores and mortality), and the Mann-Whitney test (for two-sample analysis of non-parametric data).

The statistical tests used in the third study are fully described in Chapter 7.

Statistical analysis was performed using the statistical packages described above.

4.6.2 Statistician

A qualified statistician (Vic Aber from the National Heart and Lung Institute, London) assisted with the analysis of the second study and with the planning of the third study. He advised on the way data should be entered and stored for the latter study, the best computer hardware and software to use, and on what statistical tests would be necessary for the analysis of the data. He tragically died during the study and Dr Bharat Thakrar BSc MSc PhD kindly stepped in and helped analyse the data.

CHAPTER 5

Study 1

SEVERITY OF ILLNESS SCORING IN A MULTIDISCIPLINARY INTENSIVE CARE UNIT

5.1 Introduction

At the beginning of this study, severity of illness scoring systems were a relatively novel phenomenon with exciting potential applications. None of those then available had been extensively validated or gained widespread acceptance, and experience with them in South Africa was limited and had not been published. These scoring systems included the TISS (Keene & Cullen 1983), SAPS (Le Gall et al 1984), and APACHE II score (Knaus et al 1985a). Full descriptions of these systems appear in Chapter 2. In addition there were no established definitions of organ failure, an area thought by many clinicians as likely to be of major prognostic significance, and possibly more important for this purpose than severity of illness scores.

5.1.1 Aims of this study

This study therefore aimed to evaluate the TISS and APACHE II systems in a general ICU, with respect to general utility and outcome predictive ability, and to prospectively evaluate a set of locally developed definitions of organ failure (see below for these definitions). ICU mortality was the proposed endpoint.

5.2 Protocol

5.2.1 Site

Groote Schuur Hospital in Cape Town is a general hospital functioning as a primary, secondary, and tertiary referral centre. It is an undergraduate and postgraduate teaching hospital affiliated to the University of Cape Town, as well as being a basic and post-basic training centre for nurses and paramedical personnel. It has a potential total of approximately 1700 beds though due to staffing constraints this is not realised and only about 1450 beds are utilised. The hospital is state-funded and admits all patients; the fee structure is based on income and patients unable to pay receive free treatment. It is important to state that even at the time of the study the hospital was fully racially integrated in terms of both patients and staff.

The Respiratory ICU is a multidisciplinary ICU which admits all categories of patient. It has 10 beds in two open-plan units of 6 and 4 beds; there is a separate 3-bedded isolation unit. It is staffed by a full-time director (who controls policy for admission, discharge, and therapy), two full-time consultant anaesthetists, and four part-time consultant pulmonologists who take part in the consultant call roster. There are two registrars from the Department of Medicine and two from the Department of Anaesthetics who provide 24 hour in-house physician cover.

5.2.2 Entry criteria

All patients admitted to the Respiratory ICU at Groote Schuur Hospital during 1985 and 1986 were included in this study. The patients were admitted from all disciplines including medicine, surgery, trauma, obstetrics and gynaecology, and there were also referrals from outside hospitals, both private and state-funded. Criteria for ICU admission were stringent: the majority of patients admitted required ventilatory support and/or intensive haemodynamic monitoring. Patients with primary cardiac disease were usually admitted to a specialised coronary care unit while neurosurgery and cardiac surgery patients were also usually admitted to specialised units unless they had multisystem dysfunction or developed multiple organ failure.

In addition, all patients admitted to the Respiratory ICU in 1989 had their risk of dying calculated from the APACHE II risk equation in order to assess the performance of the ICU (Knaus et al 1985a).

5.2.3 Exclusion criteria

There were no exclusion criteria.

5.2.4 Data collection

All data were collected prospectively onto a specially designed data sheet. This sheet was initially filled in by the registrar looking after the patient, and on discharge of the patient a consultant checked it for accuracy and data integrity. Data collected included demographic information, primary and secondary diagnoses, and detailed information on the clinical course of the patient, procedures performed, drugs used, complications, and outcome. Patients were scored on TISS, APACHE II and the Cape Town organ failure score (see section below for definitions) during the first 24 hours of admission. From the data sheet, data were entered into a dedicated database for analysis (dBase II, Ashton-Tate, Torrance, California, USA) on an Apple (Apple Computer Inc.) microcomputer. When new computer hardware and software became available, the database program was rewritten for dBase IV (Ashton-Tate, Torrance, California, USA) on an IBM (International Business Machines Corporation) microcomputer. This program included the calculation of the risk of dying for each patient (Knaus et al 1985a).

5.2.5 Therapeutic Intervention Scoring System (TISS)

The Therapeutic Intervention Scoring System (TISS) was introduced in 1974 (Cullen et al 1974) and updated in 1983 (Keene & Cullen 1983) to address innovations in the practice of critical care. Some items were deleted, some added, and others adjusted. The update evaluates 76 therapeutic tasks in 3 categories: active management, monitoring and ward care. Each applicable task is evaluated on a score of 1-4 depending on the intensity of intervention of medical or nursing care. TISS tasks and scores appear in Figure 5.1. The detailed explanations appear in the original text (Keene & Cullen 1983). On a daily basis, points from all tasks applicable to the patient are summed; the total is the TISS score.

Figure 5.1

TISS tasks with point allocation.

4 points

1. Cardiac arrest and/or countershock within past 48 hours.
2. Controlled ventilation with or without PEEP.
3. Controlled ventilation with intermittent or continuous muscle relaxants.
4. Balloon tamponade of varices.
5. Continuous arterial infusion.
6. Pulmonary artery catheter.
7. Atrial and/or ventricular pacing.
8. Haemodialysis in unstable patient.
9. Peritoneal dialysis.
10. Induced hypothermia.
11. Pressure-activated blood infusion.
12. G-suit.
13. Intracranial pressure monitoring.
14. Platelet infusion.
15. Intra-aortic balloon assist.
16. Emergency operative procedures (within past 24 hours).
17. Lavage of acute GI bleeding.
18. Emergency endoscopy or bronchoscopy.
19. Vasoactive drug infusion (> 1 drug).

(Figure 5.1 continued...)

3 points

1. Central IV hyperalimentation (includes renal, cardiac, hepatic failure fluid).
2. Pacemaker on standby.
3. Chest tubes.
4. Intermittent mandatory ventilation or assisted ventilation.
5. Continuous positive airway pressure.
6. Concentrated potassium infusion via central catheter.
7. Nasotracheal or orotracheal intubation.
8. Blind intratracheal suctioning.
9. Complex metabolic (frequent intake and output).
10. Multiple ABG, bleeding, and/or STAT studies (> 4 per shift).
11. Frequent infusions of blood products (> 5 units per 24 hours).
12. Bolus IV medication (nonscheduled).
13. Vasoactive drug infusion (1 drug).
14. Continuous antiarrhythmia infusions.
15. Cardioversion for arrhythmia (not defibrillation).
16. Hypothermia blanket.
17. Arterial line.
18. Acute digitilisation-within 48 hours.
19. Measurement of cardiac output by any method.
20. Active diuresis for fluid overload or cerebral oedema.
21. Active treatment for metabolic alkalosis.
22. Active treatment for metabolic acidosis.

(Figure 5.1 continued...)

23. Emergency thoracentesis, paracentesis, or pericardiocentesis.
24. Active anticoagulation (first 48 hours).
25. Phlebotomy for volume overload.
26. Coverage with more than 2 IV antibiotics.
27. Treatment of seizures or metabolic encephalopathy (within 48 hours of onset).
28. Complicated orthopaedic traction.

2 points

1. CVP monitoring.
2. 2 peripheral IV catheters.
3. Haemodialysis (stable patient).
4. Fresh tracheostomy (less than 48 hours).
5. Spontaneous respiration via endotracheal tube or tracheostomy.
6. GI feedings.
7. Replacement of excess fluid loss.
8. Parenteral chemotherapy.
9. Hourly neuro vital signs.
10. Multiple dressing changes.
11. Pitressin infusion IV.

1 point

1. ECG monitoring.
2. Hourly vital signs.

(Figure 5.1 continued...)

3. 1 peripheral IV catheter.
4. Chronic anticoagulation.
5. Standard intake and output.
6. STAT blood tests.
7. Intermittent scheduled IV medication.
8. Routine dressing changes.
9. Standard orthopaedic traction.
10. Tracheostomy care.
11. Decubitus ulcer.
12. Urinary catheter.
13. Supplemental oxygen (nasal or mask).
14. Antibiotics IV (2 or less).
15. Chest physiotherapy.
16. Extensive irrigations, packings or debridement of wound, fistula, or colostomy.
17. GI decompression.
18. Peripheral hyperalimentation/Intralipid therapy.

5.2.6 Organ failure score (Cape Town)

This score was conceptualised and developed at about the same time as definitions for organ-system failure were published by Knaus et al (Knaus et al 1985). The definitions were derived from local clinical experience and are therefore somewhat different from the above well-used and now familiar ones. Each organ failure is more simply but less tightly defined and hepatic failure is included. The complete definitions are tabulated below (Figure 5.2).

Figure 5.2

Definitions of organ failure (Cape Town)

Renal	Serum creatinine > 150 μmol/l.
Cardiovascular	Systolic blood pressure < 80 mmHg (or requiring inotropes to maintain blood pressure) with adequate volume replacement.
Respiratory	Requiring mechanical ventilation or $PaO_2 < 15$ kPa on $FiO_2 \geq 0.5$.
Neurological	Unconscious (no response to verbal command in absence of sedation).
Hepatic	Bilirubin or liver enzymes > twice normal.
Haematological	Platelet count $< 50 \times 10^9$ or white cell count $< 2 \times 10^9$.

5.3 Results

5.3.1 Demographics

Seven hundred and twenty-eight patients (375 males and 353 females) were admitted to the Respiratory ICU during 1985 and 1986. The mean age was 43 years (range 12-88 years). Of these patients 522 (71.7%) received intermittent positive-pressure ventilation (IPPV), 72 (9.9%) received continuous positive airway pressure (CPAP) by facemask, and 101 (14.9%) received oxygen by facemask. There were 130 deaths, giving an ICU mortality rate of 17.9%.

In 1989 there were 374 admissions, with a mean age of 41.2 years (range 12-85 years), with 300 patients (80.2%) receiving intermittent positive-pressure ventilation.

5.3.2 Scoring systems and mortality

The major diagnostic categories, mean TISS and APACHE II scores and mortality rates for 1985 and 1986 are shown in Table 5.1.

Table 5.1

Major diagnostic categories, mean TISS and APACHE II scores and mortality rates for the years 1985 and 1986.

	Number		**TISS**		**APACHE II**		**Mortality (%)**	
Year	1985	1986	1985	1986	1985	1986	1985	1986
Diagnosis								
Pneumonia	43	54	24.3	26.1	20.4	18.2	42	35
Asthma	60	36	13.7	19.3	15	16	3.3	2.7
Flail chest	44	45	20.9	19.6	9.9	9.5	6.8	6.6
ARDS	41	32	23.5	30.4	15.5	17.5	24.4	31
Postop*	27	23	18.7	23.4	7.6	10.1	3.7	0
Overdose	25	16	18	21.8	15.8	15.3	16	0
Other	131	151						
TOTAL		728		19.9		14.1		17.9

* Elective postoperative admission

The relationship between severity of illness scores (TISS, APACHE II and organ failure) and mortality for the total number of patients in 1985 and 1986 are shown in Figures 5.3-5.5 (reproduced from South African Medical Journal 1989;76:18-19).

Figure 5.3

TISS score vs ICU mortality (all patients)

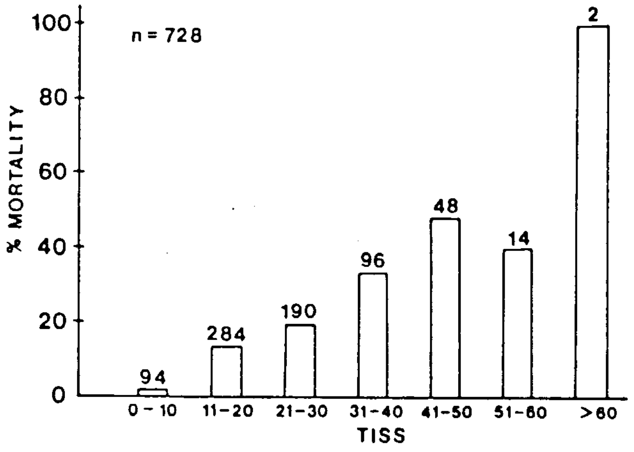


Figure 5.4

APACHE II score vs ICU mortality (all patients)

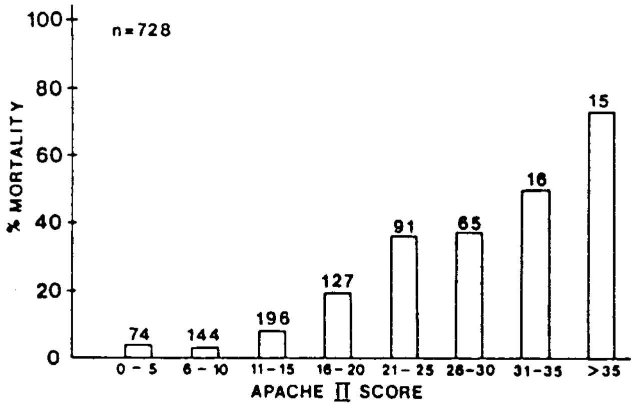
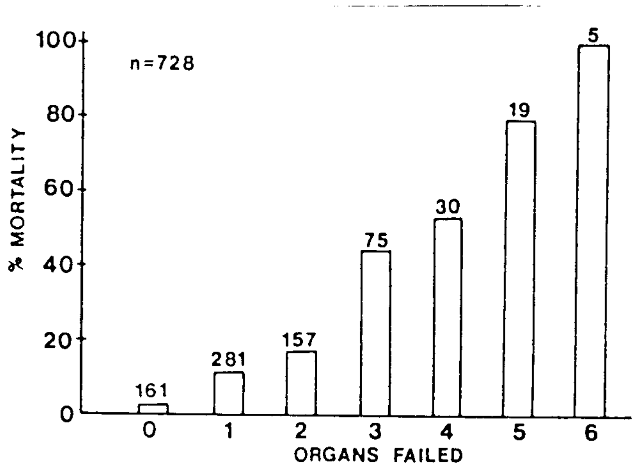


Figure 5.5

Organ failure vs mortality (all patients)



Selected disease categories (pneumonia and adult respiratory distress syndrome (ARDS)) with the relation between APACHE II score and mortality in 1985 and 1986 are shown in Figures 5.6 and 5.7 (reproduced from South African Medical Journal 1989;76:19).

Figure 5.6

APACHE II score vs mortality in patients with ARDS

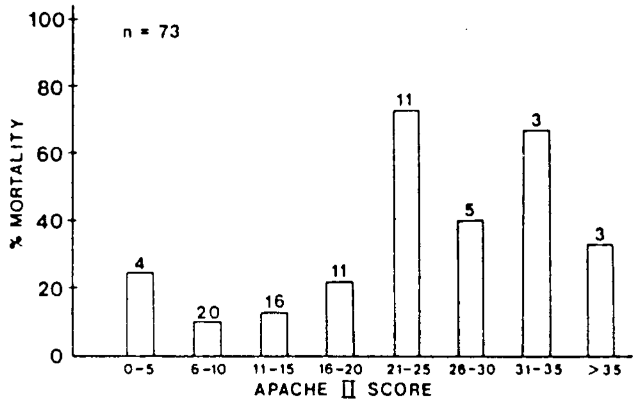
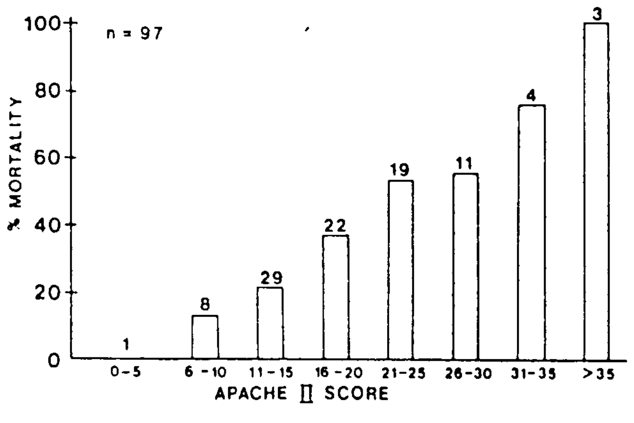


Figure 5.7

APACHE II score vs mortality in patients with pneumonia



In 1989 the average APACHE II score was 14.1 (range 1-44). The average risk of dying was 17.6% (range 12% - 94.3%). From this figure 65.7 deaths were predicted. There were 51 actual deaths (mortality rate 13.6%). The standardised mortality ratio (actual deaths divided by predicted deaths) was therefore 0.776.

5.4 Discussion

These three systems were selected for evaluation because they cover a broad spectrum of ways to score severity of illness. The APACHE II score has been shown to predict severity of illness efficiently if mortality is used as the end-point (Knaus et al 1985a). It has also been found to be useful in comparing the quality of care provided by different ICUs (Knaus et al 1986) and may be of value in assessing therapeutic protocols and evaluating efficacy of new treatments. In our study we have confirmed that TISS, APACHE II and the organ failure score all show a good relation with mortality in all patients admitted to the Respiratory ICU when they were scored over the first 24 hours after admission. Mortality in relation to severity of illness in individual disease groups (e.g. pneumonia) indicates a similar outcome in our ICU for 1985 and 1986, which suggests that the quality of care for the 2 years remained the same. The organ failure score appeared to show the best relation between increase in score and mortality. APACHE II and the organ failure score proved to be easy and quick to perform. Both score the severity of illness by measuring physiological deviation (directly in the case of APACHE II, indirectly in the organ failure score) from normal and thus would appear more valuable than TISS, which measures the degree of invasiveness of management and investigation and degree of monitoring of the patients. Although

TISS may determine the severity of illness in an individual ICU, it may be of less value when comparing different ICUs where different levels of invasive management are used, either because of availability of resources or treatment philosophy. Although the predictive value of these scoring systems was high there were notable exceptions, since some patients with low scores died. These deaths were usually due to sudden unexpected events (intracranial haemorrhage, unexplained cardiac arrest), late complications (progressive multiple organ failure secondary to sepsis) or delayed effects (paraquat poisoning).

In certain individual diseases, notably pneumonia, the relation between severity of illness scores and mortality is excellent. However, in the group of patients with ARDS the relation between APACHE II and mortality is very poor. This is probably due to the heterogeneous causation of ARDS and subsequent complications which may develop. If ARDS were to be defined by specific causes the relation with these scoring systems might be improved. However, numbers were too small in our study to separate aetiological groups. Even in individual diseases, however, the correlations between mortality and score were insufficiently accurate to determine individual patient outcome. Scoring patients on subsequent days after initiating therapy, which would evaluate response to treatment, might improve the prognostic value in individual patients.

The organ failure score showed the best overall correlation with mortality and is the simplest system to use. Knaus et al have used and validated a system which examines only 5 organ systems, namely cardiovascular, respiratory, renal, haematological and neurological failure (Knaus et al 1985b). They have shown that 3-organ failure on day 1 predicts about 80% mortality and a 3-organ failure score on day 5 predicts 100% mortality, as does 3-organ failure on subsequent days, thus allowing prognostic estimates accurate enough to support clinical decisions and thereby providing the most appropriate care for the patients. Scoring patients later in their illness allows better predictability and decision-making for individual patients, and to achieve maximum benefit from severity of illness scores these should be repeated on successive days following admission if predictive decisions are to be made in individuals.

The ICU mortality ratio of 0.776 calculated for 1989 admissions compares favourably with those reported by Knaus et al (Knaus et al 1986) when evaluating outcome in 13 major medical centres in the USA. In this study, mortality ratios ranged between 0.59 and 1.58. Knaus et al found that a full-time director, 24-hour in-unit physician coverage, and a well coordinated staffing structure were most important in determining the performance of ICUs in the USA. In the Respiratory ICU, all these positive features are present.

5.5 Conclusions

Scoring of severity illness is valuable for measuring the standard of care in intensive care units. The APACHE II scoring system is particularly well suited for this purpose, since it is easy to perform and, done on admission, will allow results of intensive care management to be compared nationally and internationally. In addition, the results of different therapies can be compared, because APACHE II accurately predicts severity of illness in a large group of patients. It may also be valuable in identifying problem areas and allowing changes in therapy to be carefully monitored. For making a decision regarding an individual patient, however, the organ failure system is more valuable, using the definitions of Knaus et al (Knaus et al 1985b) and particularly if it is extended up to and possibly even beyond day 7.

Having evaluated APACHE II in a general ICU setting and found the above limitations, especially the disease-specificity, it seemed pertinent to examine its performance in a narrower patient profile. An opportunity presented itself to study this in cardiothoracic surgery patients, and this is described in the next chapter.

CHAPTER 6

Study 2

APACHE II SCORING IN A CARDIOTHORACIC SURGERY INTENSIVE CARE UNIT

6.1 Introduction

The APACHE II score has not previously been evaluated in a large number of patients undergoing cardiothoracic surgery, apart from those patients used in the original APACHE II database (where the numbers were given in the details for calculations for risk of dying (Knaus et al 1985a). In addition, patients admitted to ICU after coronary artery bypass grafts (CABG) were later to be excluded from the APACHE III data collection and prognostic system (Draper et al 1989, Knaus et al 1991a)).

Cardiothoracic surgical patients are unique in several ways. Firstly, they may have severely abnormal physiology after surgery and cardiopulmonary bypass, most of which will revert to normal spontaneously and over a short period of time, and which may have no bearing on outcome. Secondly and conversely, many physiological abnormalities may be masked by multisystem support devices such as pacemakers, inotropic drugs, intra-aortic balloon pumps, mechanical ventilation, and haemofiltration. Thirdly, morbidity and mortality may be related to the presence of chronic health status (which scores only 2 APACHE II points in elective postoperative patients) and to unpredictable and unexpected events occurring in the immediate postoperative period. It would therefore not be unexpected if the APACHE II scoring system, essentially only looking at a short period of physiological measurements, were a poor predictor of outcome in this group of patients.

6.1.1 Aims of this study

This study was set up to evaluate the applicability of the APACHE II scoring system in cardiothoracic surgical patients by correlating APACHE II scores and predicted risk of dying with mortality.

6.2 Protocol

6.2.1 Site

The Royal Brompton National Heart and Lung Hospital in London is a tertiary cardiothoracic referral centre and as such deals with a large number of high-risk patients. Both NHS and private patients are admitted, and there are also contracts with other health authorities and foreign governments.

The Adult ICU has a potential of 14 beds and deals principally with immediate post-surgical cardiac and thoracic patients. However not all such patients are admitted to the ICU; low risk patients go from the Operating Theatre to a Recovery area for a few hours and from there to a High-Dependency Unit in the ward. Other less frequently admitted patients include those with acute lung injury, major vascular surgery, and cardiology patients needing mechanical support (mechanical ventilation, haemofiltration, or intra-aortic balloon pump).

The ICU is staffed by a director and a team of intensivists (consultants from anaesthetics and thoracic medicine), as well as a rotating surgical senior house officer and an anaesthetic registrar. The patients are managed jointly by their surgeons and the intensivists. There is 24-hour in-house physician cover provided by a surgical senior house officer and an anaesthetic registrar; there is always a consultant intensivist on call.

6.2.2 Entry criteria

All patients admitted to the Adult ICU of the Royal Brompton National Heart and Lung Hospital between 25/09/1987 and 24/11/1988 were consecutively entered into the study. A subset of patients was subsequently excluded for purposes of analysis (see Section 6.3.1 below).

6.2.3 Exclusion criteria

There were no exclusions.

6.2.4 Data collection

Physiological data appropriate to the APACHE II scoring system (Knaus et al 1985a) were collected daily, verified by checking patient charts, and together with demographic data were entered into a dedicated computer program for analysis (Riyadh ICU Program 3.0, Copyright R Chang 1988, Medical and Associated

Software House, London, UK). Risk of dying was calculated for each patient using the APACHE II score and a logistic regression formula (Knaus et al 1985a). Some data fields (including demographics, reason for admission, details specific to each operation performed, duration of ventilation and ICU stay, chronic health, emergency status, and outcome) were later entered from the original data sheets into another database (Paradox 2.0, Ansa Software, Belmont, California, USA) for ease and flexibility of analysis and to facilitate transfer of data to a statistical analysis program (Statgraphics 4.0, Manugistics Inc, Rockville, Maryland, USA). The original data collection and data entry was performed by Drs Cliff Morgan and Yugan Mudaliar. Data entry to the subsequent database and all analyses were performed by myself.

The physiological values used for the APACHE II score were based on the worst values for that 24 hour period except for the Glasgow Coma Scale which was scored as 15 (best possible value) unless there was clear evidence of neurological dysfunction. This was retrospectively scored if abnormalities were noted once sedation had worn off.

6.3 Results

6.3.1 Demographics

A total of 869 consecutive patients was entered into the study. The data on 12 patients were incomplete, 43 patients had non-surgical diagnoses, and 3 patients had non-cardiothoracic operations. This left a total of 811 cardiothoracic surgical patients for analysis, and all the results below pertain to this group of patients alone.

The mean age was 57 years (range 15-87 years), 78% of patients were male and 22% were female, and the mean duration of admission was 2.3 days (median 1 day, range 1-39 days). Elective surgery was performed in 90.3% of patients and emergency surgery in 9.7%.

Surgical procedures performed on the study patients are listed in Table 6.1 together with mean APACHE II scores, and predicted and actual mortality rates.

Table 6.1

Surgical procedures with mean APACHE II scores and predicted (by APACHE II) and actual ICU mortality rates.

Procedure	Number	APACHE II	Predicted (%)	Actual (%)
CABG*	527	8.9	3.4	3.2
Valve surgery	186	10.4	5.7	4.3
Thoracic surgery	37	10.6	9.1	13.5
Congenital disorders	26	8.6	4.0	11.5
Other cardiac	14	11.3	10.7	14.3
Transplants	12	13.9	13.4	8.3
Aortic aneurysm	9	12.3	12.2	11.1

* CABG = Coronary Artery Bypass Graft

6.3.2 APACHE II scores and mortality

The mean APACHE II score was 9.5 (median 9, range 0-33). The ICU mortality rate was 4.56% and the hospital mortality rate was 5.43%. The relationship between APACHE II score and ICU mortality is shown in Table 6.2 and Figure 6.1 (chi-square for trend 97.8, $p<0.001$, chi-square for departure from trend 53.4, $p<0.001$). The overall predicted risk of dying was 4.59%. The relationship between predicted risk of dying and mortality is shown in Table 6.3 and Figure 6.2

(chi-square for trend 139.0, p<0.001, chi-square for departure from trend 2.1, p>0.1).

Table 6.2

APACHE II scores and ICU mortality.

APACHE II	Number	Mortality (%)
0 - 5	128	0.79
6 - 10	410	0.98
11 - 15	208	5.29
16 - 20	42	23.81
>20	23	47.82

Figure 6.1

The relationship between APACHE II score and ICU mortality.

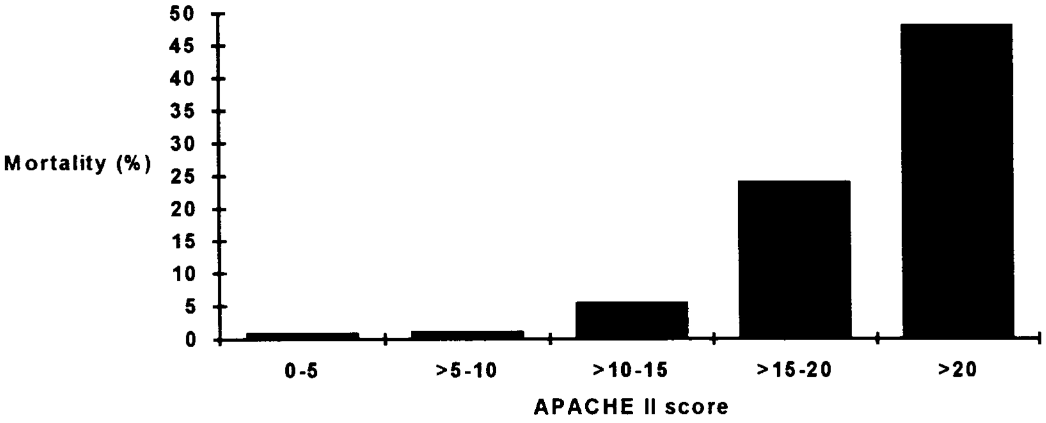


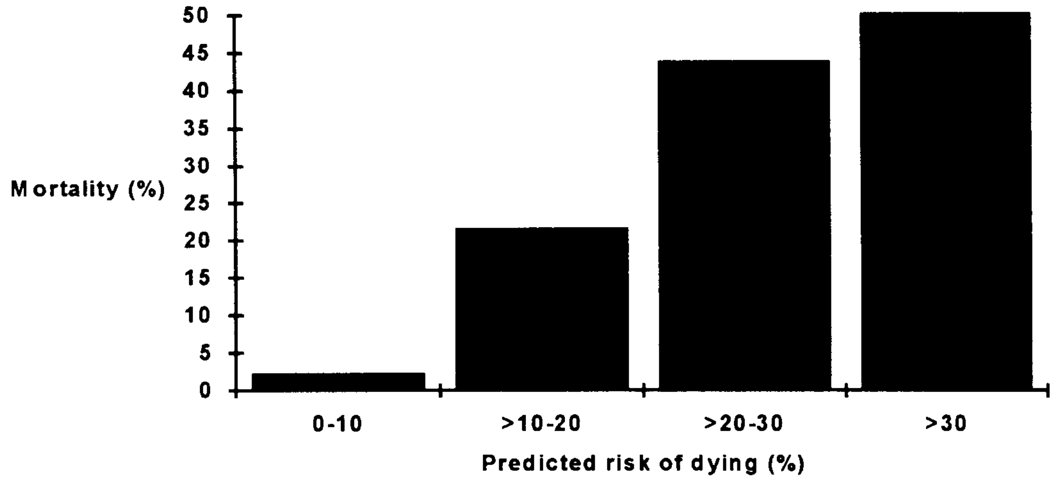
Table 6.3

Predicted risk of dying and actual ICU mortality.

Predicted risk (%)	Number	Mortality (%)
0 - 10	743	2.15
>10 - 20	42	21.43
>20 - 30	16	43.75
>30	10	50

Figure 6.2

The relationship between predicted risk of dying and ICU mortality.



The mean APACHE II score of survivors was 9.1 (median 9, range 0-28), while that of nonsurvivors was 17.0 (median 16, range 5-33), p<0.001, Mann-Whitney test.

There was a significant correlation between APACHE II score and duration of ventilation (p<0.001) and between APACHE II score and duration of ICU admission (p<0.001).

6.3.3 Chronic health

There was a significant difference in mortality between patients with a chronic health history as defined by the APACHE II scoring system (n=145, mortality 11.0%, mean APACHE II score 12.2) and patients without (n=666, mortality 3.15%, mean APACHE II score 8.9), p<0.001, chi-square test.

6.3.4 Ventilation

ICU survivors were ventilated for a mean duration of 1.6 days (median 1 day, range 0-29 days) while nonsurvivors were ventilated for a mean of 9.3 days (median 9 days, range 1-37 days), p<0.001, Mann-Whitney test.

6.3.5 Duration of admission

The mean duration of admission of patients who were discharged alive from ICU was 2.0 days (median 1 day, range 1-30 days) while that of patients who died in ICU was 9.8 days (median 5 days, range 1-39 days), p<0.001, Mann-Whitney test.

6.3.6 Emergency surgery

Patients who had elective surgery had a mortality rate of 4.0%, while those who had emergency surgery had a mortality rate of 11.4%, p<0.01, chi-square test.

6.3.7 CABG patients only

The subgroup of patients undergoing CABG (n=527) had a mean age of 57.5 years (range 33-79 years) and a mean duration of admission of 1.86 days (range 1-27 days). Elective surgery was performed in 91.8% of patients, and emergency surgery in 8.2%. There were only two patients who survived the ICU to die in hospital, so hospital mortality will not be separately analysed.

The relationship between APACHE II score and ICU mortality is shown in Table 6.4 (chi-square for trend 69.6, p<0.001) and Figure 6.3, while the relationship between predicted risk of dying and mortality is shown in Table 6.5 (chi-square for trend 131.4, p<0.001), with no bar chart to accompany it.

Table 6.4

APACHE II scores and ICU mortality in CABG patients.

APACHE II	Number	Mortality (%)
0-5	90	1.11
6-10	287	0.69
11-15	122	2.46
16-20	14	28.57
>20	12	58.33

Figure 6.3

The relationship between APACHE II score and ICU mortality in CABG patients.

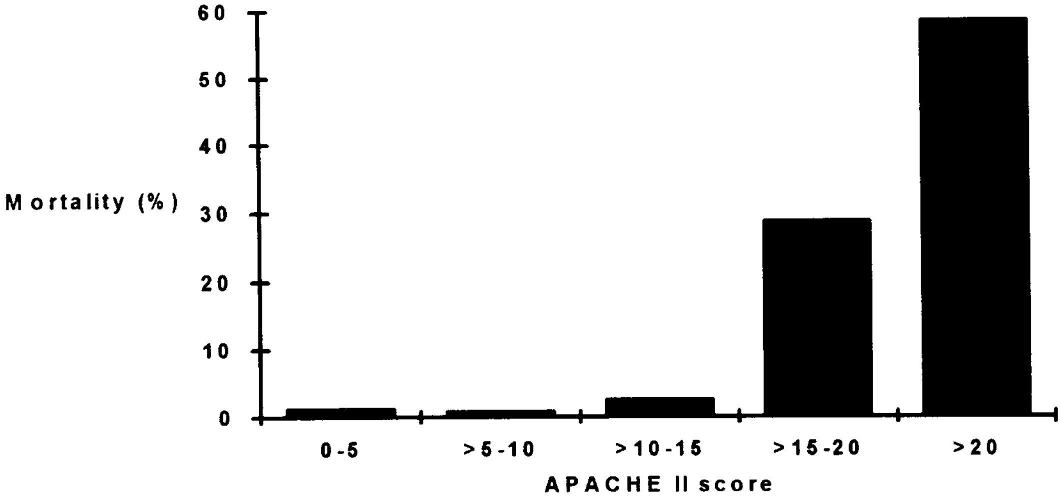


Table 6.5

Predicted risk of dying and actual ICU mortality in CABG patients.

Predicted risk (%)	Number	Mortality (%)
0-10	510	1.76
>10-20	12	33.33
>20	5	80

6.4 Discussion

Although the APACHE II scoring system has been evaluated in large studies in a wide number of general and specific applications, this is the first large study of its use in cardiothoracic surgical patients. It is also the first independent study (outside the APACHE II database). Due to the unique physiological disturbances in these patients (not necessarily indicative of a poor outcome) it was necessary to assess its validity and predictive powers as well as its shortcomings.

The patient profile reflects the type of patient undergoing cardiothoracic surgery in First World countries today: generally older (mean age 57 years), males (78% males in this study), and often with a positive chronic health history (18% in this study). The majority (65%) of patients had coronary artery bypass grafts, a result of the high incidence of ischaemic heart disease in such countries.

The APACHE II scores in this study were generally lower than those found in general ICUs worldwide (Marsh et al 1990, Lockrem et al 1991, Rutledge et al 1991, Jacobs et al 1988, Turner et al 1989, Joshua et al 1989, Dragsted & Qvist 1989b, Giangiuliani et al 1989, Berger et al 1992, Oh et al 1993, Chisakuta &

Alexander 1990, Rowan et al 1993). This is not unexpected as the cardiovascular system is often the only organ-system involved in these patients; even when patients score APACHE II points for temperature and blood gases, biochemical abnormalities are rare. Other reasons are that the Glasgow Coma Scale was scored as 15 out of 15 in the APACHE II score (leading to no neurological points being scored) unless there was obvious neurological impairment, and that patients undergoing elective surgery score only 2 APACHE II points for chronic health.

No patient with an APACHE II score of less than 5 died, and the mortality rate for the 538 patients with scores less than 10 was 0.93%. This is useful prognostic data for survival. Unfortunately, prognostic data for death is less helpful. Mortality was only 47.8% when the APACHE II score was greater than 20, with only 23 patients in this group. Death was certain only when the APACHE II score was greater than 30, but there was only one patient in this category. Statistical analysis suggested a significant relationship between increasing APACHE II score and mortality which may be exponential.

The mortality rate for patients with a predicted risk of dying of less than 10% was 2.15%. Statistical analysis suggested a significant linear relationship between increasing risk of dying and mortality. Only a predicted risk of dying of more than 50% was universally associated with death, but there were only two patients in this group.

The APACHE II score correlated well with duration of ventilation and duration of ICU admission. This suggests that patients with high scores might benefit from more aggressive early therapy in an attempt to shorten these times. However the abnormal physiology would in itself probably lead to the use of such therapy.

The predicted mortality rate (4.59%) was very close to the actual ICU mortality rate (4.56%). Firstly this would suggest the accuracy of the weighting for risk of dying based on the APACHE II score in cardiothoracic surgical patients. This study includes a larger group of cardiothoracic surgical patients than those used in the original calculations (Knaus et al 1985a) in which 225 patients had heart valve surgery and 90 had surgery as a result of chronic cardiovascular disease. Secondly it reflects the performance capability of the ICU. Knaus et al found that a full-time director, 24-hour in-unit physician coverage, and a well coordinated staffing structure were most important in determining the performance of ICUs in the USA (Knaus et al 1986). In the ICU in which this study was performed, there is a full-time director and 24 hour in-unit physician coverage although care of the patients is sometimes loosely defined. However, in the various subgroups there were differences between predicted and actual mortality, although the patient numbers are obviously smaller. APACHE II underestimated mortality for thoracic surgery and congenital heart disease surgery. This may be due to referral patterns, as the thoracic surgical patients often had advanced neoplastic disease and the patients with congenital heart disease had particularly complex problems and had frequently had multiple previous operations. On the other hand, APACHE II overestimated mortality for heart transplants, though the numbers are too small (n=12) to draw any major conclusions from this.

The markers for mortality were presence of a chronic health history ($p<0.001$), length of ICU stay ($p<0.001$), a high APACHE II score ($p<0.001$), and the need for emergency surgery ($p<0.01$). None of these markers on its own was able to adequately predict mortality.

Patients undergoing CABG (65% of patients in the study) had lower mean APACHE II scores and mortality than the group as a whole. The predicted and actual mortality rates were very close (3.4% and 3.2% respectively), and the relationship between increasing APACHE II score and mortality was good. This is the largest group of CABG patients studied to date, and would suggest that APACHE II is useful in these patients as a group. However, once again the prognostic data for survival, with a mortality rate of 0.79% for patients with APACHE II scores of less than 10, is more useful than the data for mortality.

6.5 Conclusions

Low APACHE II scores and predicted risk of dying predicted survival with some certainty but only very high scores accurately predicted death (in a small number of patients). In the final analysis the APACHE II scoring system is useful in generally predicting outcome (and specifically predicting survival in patients with low scores) but does not have the power to predict death and thereby potentially influence decisions to withdraw therapy. This power might, however, be increased by inclusion of preoperative, intraoperative, and postoperative variables.

CHAPTER 7

Study 3

A NOVEL COMBINATION OF PREOPERATIVE, INTRAOPERATIVE, AND POSTOPERATIVE FACTORS TO PREDICT OUTCOME IN CARDIAC SURGERY

7.1 Introduction

Although the major severity scoring systems for general Intensive Care Unit (ICU) purposes have matured through two or three generations, they still do not adequately address cardiac surgery patients. Patients undergoing coronary artery bypass grafts (CABG) were excluded from the APACHE III data collection (Knaus et al 1991a), and all cardiac surgery patients from the Mortality Prediction Model II (Lemeshow et al 1993) and Simplified Acute Physiology Score (SAPS) II (LeGall et al 1993). Although mortality for cardiac surgery patients in our unit is acceptably low at 3.2% for CABG and 4.3% for heart valve surgery, non-survivors stay in the ICU for longer (mean of 9.8 days vs 2 days for survivors), using more resources (Turner et al 1991).

Responsible use of ICU facilities and cost containment are becoming more important and are significant political issues in the UK, with sweeping changes proposed for the National Health System. In addition the cost of intensive care is high, with the mean daily cost of survivors in an ICU in the UK being £550 and nonsurvivors £816 (Ridley et al 1993). Medical costs have increased dramatically in the last decade (Jacobs & Noseworthy 1990), with much of the expenditure occurring in the ICU setting.

A variety of cardiac operative risk systems have been developed from analyses of large numbers of patients (see Chapter 3), several of which derive a numerical score. Some of these scores are sums of points allocated for the presence or absence of risk factors, while others involve calculation of exponential functions. Interpretation of the literature is confusing, as all risk systems evaluate different patient populations and types of operation (some include coronary artery surgery only), as well as different combinations of predictive factors and end-points (with different definitions for each of these), and although more recent ones have used earlier studies as their basis, no single system has gained widespread acceptance. In a recent review, the effects of impaired left ventricular function, older age, sex, and reoperation were singled out as consistent predictors of high operative risk (Rosenfeldt & Wong 1993). The Parsonnet system has been independently validated (Nashef et al 1993), although this relatively small study has been criticised for both its methodology and its conclusions (Spiegelhalter 1992). Part of the problem may be that the above studies span many years; patient profile and mortality has changed during that time (Naunheim et al 1988) and continue to change. The largest study of coronary artery bypass surgery patients is almost certainly the Collaborative Study in Coronary Artery Surgery (CASS), which has been reported at various times and with regard to various subgroups and has been comprehensively reviewed (Anderson 1986). Impressive in size and scope, it consists of two elements: a randomised trial of surgical versus medical therapy, and a registry which includes a risk equation to estimate hospital mortality. The risk equation involves exponential functions and needs a computer or programmable calculator to perform.

More and more cardiac surgery is being performed on higher-risk patients (including more redo operations as more patients outlive their grafts or valve replacements), many of whom will do poorly. There was thus a need to develop a better prognostic system for these patients, in order to assist in the prediction of mortality, complications, and length of ICU stay.

7.1.1 Aims

The aims of this study were as follows:

1. To develop a prognostic index for cardiothoracic surgical patients which may ultimately be used to help make decisions to withdraw therapy in hopeless cases.
2. To develop a predictive index for length of ICU stay and for the occurrence of complications.

7.2 Protocol

7.2.1 Site

(See Section 6.2.1 for basic description). There were some important changes between the second and third studies. Firstly the Royal Brompton National Heart and Lung Hospital moved about 200 metres to a newly constructed site with new equipment in Sydney Street in December 1990. The new ICU has a physical potential of 19 beds but is only routinely staffed for 11 beds; additional patients are looked after by staff from nursing agencies. Secondly the junior staffing changed to there being three senior house officers directly responsible for the intensivists' patients and one responsible for the surgeons' patients. Thirdly the responsibility for patient management became more formalised: cardiac surgery patients were looked after by the surgeons initially (with exceptions made in the case of the more critically ill) and then handed over to the care of the intensivists if they stayed longer than 48 hours in ICU.

7.2.2 The Hewlett-Packard CareVue 9000 data management system

7.2.2.1 Introduction

The Royal Brompton National Heart and Lung Hospital was the first site in the United Kingdom to install the Hewlett Packard CareVue 9000 data management system (Hewlett Packard, Andover, Massachusetts, California, USA). Its current and potential ability for data collection and storage makes it an integral part of this study. A full description therefore follows.

The Hewlett Packard CareVue 9000 data management system consists of a computer hardware system upon which runs a specialised software package enabling the rapid recording, storage, display and reporting of a wide range of clinical data. It addresses current and future projected needs for patient monitoring and the increasingly obvious need for sophisticated online data management. It includes a new range of bedside monitoring (the Merlin system) and a comprehensive and high powered computer solution to the management of the enormous quantity of data that is produced by the monitor and from other sources such as laboratory measurements and patient record data base. Hewlett Packard believe they have achieved future proofing by the use of high quality engineering and a component based design. That is, the core of the system is a computer type system unit with an expansion bus which can contain a range of hardware and software expansion cards which can control various types of video display unit and provide interfaces to other equipment. Monitoring is broken down into individual physiological parameters such as ECG, invasive blood pressure, oxygen saturation, etc., and each of these functions is subserved by a separate compact robust module which can be plugged in to a purpose built rack system. Currently available and soon to be available options cover the whole range of monitoring requirements for use in operating theatres, intensive care and coronary care units.

Actual configurations are variable; currently there is one work station per patient bed. The all important software and the appearance of the high resolution multicolour display units of the work stations has been designed to maximise user friendliness and ease of use whilst the redundancy and obsessional backup of the central file server ensures reliability. The screen display can be custom configured to almost mimic pre-existing paper records and in so doing to make those paper records redundant. The obvious difference is that data entry into the computer version is automatically channelled effortlessly and accurately without transcription error from the digital output from the network of individual patient monitors. This data can then be displayed on the computer screen in either tabular or graphical forms depending on preference and individual users can manipulate the display for their own personal preference, time scales, combinations of parameters, etc.

This system totally replaces nursing charts and is an attractive option in terms of labour saving, accuracy, convenience and immediacy.

7.2.2.2 Hardware

In normal operation the system runs on two Hewlett Packard 9000 S/300 computers. In the event of one processor failing the system automatically falls back to running on the remaining functioning processor, although there will be some loss of speed. The data is stored on two 300 Megabyte hard disks. The disks are arranged to back each other up. Thus there is no loss of information in the event of a disk failure. A quarter inch magnetic tape drive is provided for archive and long term backup. At each bed a workstation is provided. This consists of a 17 inch high-resolution colour display, a keyboard, and a trackerball. The workstation can either be on a table mounting or fitted to a specially designed trolley. The workstations are connected to the processors by an Ethernet bridge. A maximum of 20 workstations can be supported by a pair of processors.

There are a number of routes for data input to the system. Information from patient monitors is transmitted to a Hewlett Packard Careport. Manually entered data can be entered from the bedside workstation using either the keyboard, trackerball or screen. In addition facilities are provided to link to other computer systems or non Hewlett Packard monitoring instruments via an Ethernet bridge. Two laser printers are provided for producing hard copy reports.

The computers can be sited remote from the clinical area and require no special environmental or power supply requirements. A back up power supply is provided with the system. The bedside unit is fairly large (on a trolley 800mm x 800mm x 1400mm) and requires a mains supply and signal link cable.

Facilities are provided for multi user access. This will enable remote review or entry of data e.g. from the consultant's office or central station. One workstation can support more than one patient, enabling economies to be made in the less critical areas. Data can be transferred from one workstation to another. This is useful when a patient is moved from one area to another.

7.2.2.3 Software and facilities

The application software runs on the Hewlett Packard UNIX operating system. It comprises an ALLBASE relational data base on top of which lies object orientated code (C ++) to provide a high degree of system configurability. X windows is used for the human interface. Facilities are provided via a remote procedure call to a programmatic interface enabling links to other data bases for example Oracle and DBase using simple C programs.

Manual data entry can be in the form of structured text by selection with a trackerball or free text may be entered via the keyboard. Facilities may in future be provided for data input via a barcode reader.

Data can be presented in graphical, tabular or text form. The amount of data displayed at any instant is determined by the user. Facilities exist for scrolling both through data lists and time. The maximum time to present new data on the screen is specified as 2 seconds. The screens can be individually configured to suit a hospital, unit, patient type or individual. Some configuration can be performed by the bedside e.g. adding drugs whilst others will be performed by the local system support team.

Facilities exist to change any data entered. This action is recorded by the system and a comprehensive audit trail provided. The previous data, together with the identification of the person changing the data and the time at which this took place, is recorded. Password protection is provided and this could be implemented via the barcode reader. There are checks on the input of data and if "limits of unreasonableness" are violated an alarm is shown. All data is initially stored as non validated. This can be changed to validated by the action of an authorised person. The software is non hardware specific. This will enable it to be transferred to more powerful computers as they become available and thus enable advanced features to be developed e.g. artificial intelligence and expert systems. Custom databasing will be possible.

Interfaces: Hewlett Packard undertake to provide software or hardware interfaces to any existing hospital data bases or laboratory systems. The potential for automatic transfer of any existing computerised data into the Intensive Care data base is an extremely attractive option in terms of labour saving and accuracy, convenience and immediacy.

Severity and Acuity: Most of the components of the APACHE II severity of illness scoring system are available from the physiological monitor data and the typically available laboratory data. The latest phase of development of the software is to automatically compute an APACHE II score daily for each patient. The Therapeutic Intervention Scoring System (TISS) has been chosen as the most representative acuity score and again can be computed mostly from already available data. The automatic or semi-automatic availability of these measurements offer an enormously valuable insight into intensive care practice.

Medical Records: The system provides a configurable template for the construction of accurate, prospective comprehensive and legible records. The core data is shared by all users but separate sub-units serve the specific purposes of nurses, doctors, physiotherapists, pharmacists etc.

7.2.2.4 Implementation

CareVue came on line in the Adult ICU in July 1991 after a 6 week training period; a senior nurse was employed specifically for the purpose of training and to solve problems during the initial 18 months of implementation. There was no crossover period with the simultaneous use of conventional nursing charts and the CareVue system.

*7.2.3 **Entry criteria***

All patients undergoing cardiac surgery and subsequently admitted to the ICU were consecutively entered into this study. Specifically this did not include surgery for coarctation of the aorta, thoracic aortic aneurysm repair (unless the aortic valve was involved), or repair or rewiring of sternal wounds. However patients who had cardiac procedures which did not involve cardiopulmonary bypass (for example closed mitral valvotomy) or aortic cross-clamping (for example pericardiectomy) were included.

The study period ran from 25/1/93 to 15/1/94 and all ICU admissions during this time were eligible for entry into the study.

7.2.4 Exclusion criteria

Patients excluded from this study included those having other surgical procedures, medical patients, and patients who had cardiac surgery but went straight from Theatre Recovery to a High Dependency Unit in a surgical ward (locally known as "fast track" patients). The reason for not including the latter patients is that their APACHE scores are not comparable with those of the ICU patients, as both physiological measurements and blood tests are carried out at different intervals in the High Dependency Unit; their demographic details are however presented. Failed "fast-track" patients are however included in the study group.

7.2.5 Ethical approval

Ethical approval for this study was obtained from the Ethics and Research Committee of the Royal Brompton National Heart and Lung Hospital (Reference 93/43). The full application is shown in Appendix 1

7.2.6 Variable selection

A very large number of demographic, clinical, physiological, haematological, microbiological, and biochemical variables can be measured but obviously not all can be selected for collection and analysis. The precise choice of variables reflects the hope and the unstated hypothesis that that combination will be significantly related to an outcome measure. The process involved in defining the variable selection began with a review of the literature (see Chapters 2 and 3). Then followed identification of routinely performed measurements and tests, together with data availability and ease of access. From this a data sheet was produced (see below).

7.2.7 ***Definitions for complications***

Postoperative myocardial infarct: Myocardial infarct confirmed by sequential ECG changes and cardiac enzymes. Chest pain not a prerequisite as patients may have been anaesthetized/sedated.

Inotropes: Postoperative introduction of inotropic drugs. Dopamine at a dose of less than 5 ug/kg/min not included (such a dose is often used as "renal support").

IABP: Postoperative insertion of intra-aortic balloon pump.

VAD: Postoperative insertion of ventricular assist device (left ventricular assist device, right ventricular assist device, or biventricular assist device).

Ventilated >48 hours: Mechanical ventilation (not including Continuous Positive Airway Pressure) for more than 48 hours.

Reintubation: Does not include accidental extubation and subsequent reintubation, or insertion of mini-tracheostomy.

Pneumonia: Diagnosed by clinical criteria of pyrexia, raised White Cell Count, purulent secretions, new infiltrates on chest radiograph, deterioration in gas exchange. Protected specimen brush or lavage confirmation not necessary.

Creatinine x 2: Based on immediate preoperative value.

Dialysis: Any form of renal support, usually involving Continuous Veno-Venous Haemofiltration (CVVH), or Continuous Veno-Venous Haemofiltration with Dialysis (CVVHD).

Neurological deficit: The presence for more than 6 hours of any neurological deficit from confusion to focal deficit. The effects of anaesthetic drugs or sedation must be excluded.

Cardiac arrest: The cessation of effective circulation leading to the initiation of Cardiopulmonary Resuscitation. May be caused by arrhythmia, bradycardia, asystole, electromechanical dissociation.

Reoperation: Reopening of the chest for any indication, either in the ICU or in the operating theatre.

Readmission: Readmission to the Adult ICU following ICU discharge but not hospital discharge.

Other: Includes any other complications thought by the investigator or attendant surgeons to be of significance.

7.2.8 Data entry

A data sheet (Figure 7.1) was developed for use in this study after due consideration had been taken of the variables chosen to be collected. Demographic data, medical histories, physiological data (starting at the induction of general anaesthesia), and laboratory investigations were collected from the clinical records, the CareVue 9000 system, and the McDonnell-Douglas laboratory computer and entered on to the above data sheets for subsequent entry into the computer program described below. Patients were followed up to hospital discharge.

All data were checked for errors once the data sheet was completed. Data were then entered into the microcomputer program, and checked again.

If the patient was unable to be discharged from the ICU simply because there was no ward bed available, the discharge time was taken as the time he would have been discharged and the duration of admission calculated from this figure.

Figure 7.1 Data sheet

OUTCOME STUDY DATA SHEET

Study number

Hospital number

Name

Age Sex M F

Admission date

Diagnosis

Operation

Knaus reason for admission

Emergency surgery Y N

Intraoperative details

Time on bypass mins

Ischaemic time mins

Inotropes

IABP

VAD

Arrhythmias

Haemofiltration

Pacing

Time in ICU hours

Cause (if stay > 24 hours)

ICU outcome A D

Hospital outcome A D

Complications

Postop MI

Inotropes (new)

IABP (new)

VAD (new)

Ventilated > 48 hours

Reintubation

Pneumonia

Creatinine x 2

Dialysis

Neurological deficit

Cardiac arrest

Reoperation

Other

Date of death (if applicable)

PARSONNET score

Female		1
Morbid obesity (>1.5X ideal weight)		3
Diabetes		3
Hypertension (SBP>140 mm Hg)		3
Ejection fraction	Good (>=50%)	0
	Fair (30-49%)	2
	Poor (<30%)	4
Age	70-74	7
	75-79	12
	>=80	20
Reoperation	First	5
	Second	10
Preoperative IABP		2
LV aneurysm		5
Emergency surgery after PTCA/cath		10
Dialysis dependency		10
Catastrophic states		10-50
Other rare events		2-10
Valve surgery	Mitral	5
	PA pressure >=60 mm Hg	8
	Aortic	5
	PSG >120 mm Hg	7
CABG and valve surgery		2

APACHE II and III scores

Temperature
Mean arterial pressure
Heart rate
Respiratory rate Ventilated Y N
FiO_2
PaO_2
$PaCO_2$
pH Worst pH
APACHE III pH/$PaCO_2$
Haematocrit
White cell count
Sodium
Potassium
Urea
Creatinine Acute renal failure Y N
Glucose
Bilirubin
Albumin
Urine output
Glasgow coma scale
APACHE III neurology

APACHE II chronic health APACHE III chronic health

7.2.8.1 Computer software and program developed for the study

An ObjectVision (Borland International Inc, Scotts Valley, California, USA) program was developed to handle the data collected. The requirements were for a simple and aesthetic screen interface (preferably Windows based) for data entry, the ability to allocate points to demographic and physiological information and moreover add up these points to calculate scores, and an automatic link to a database.

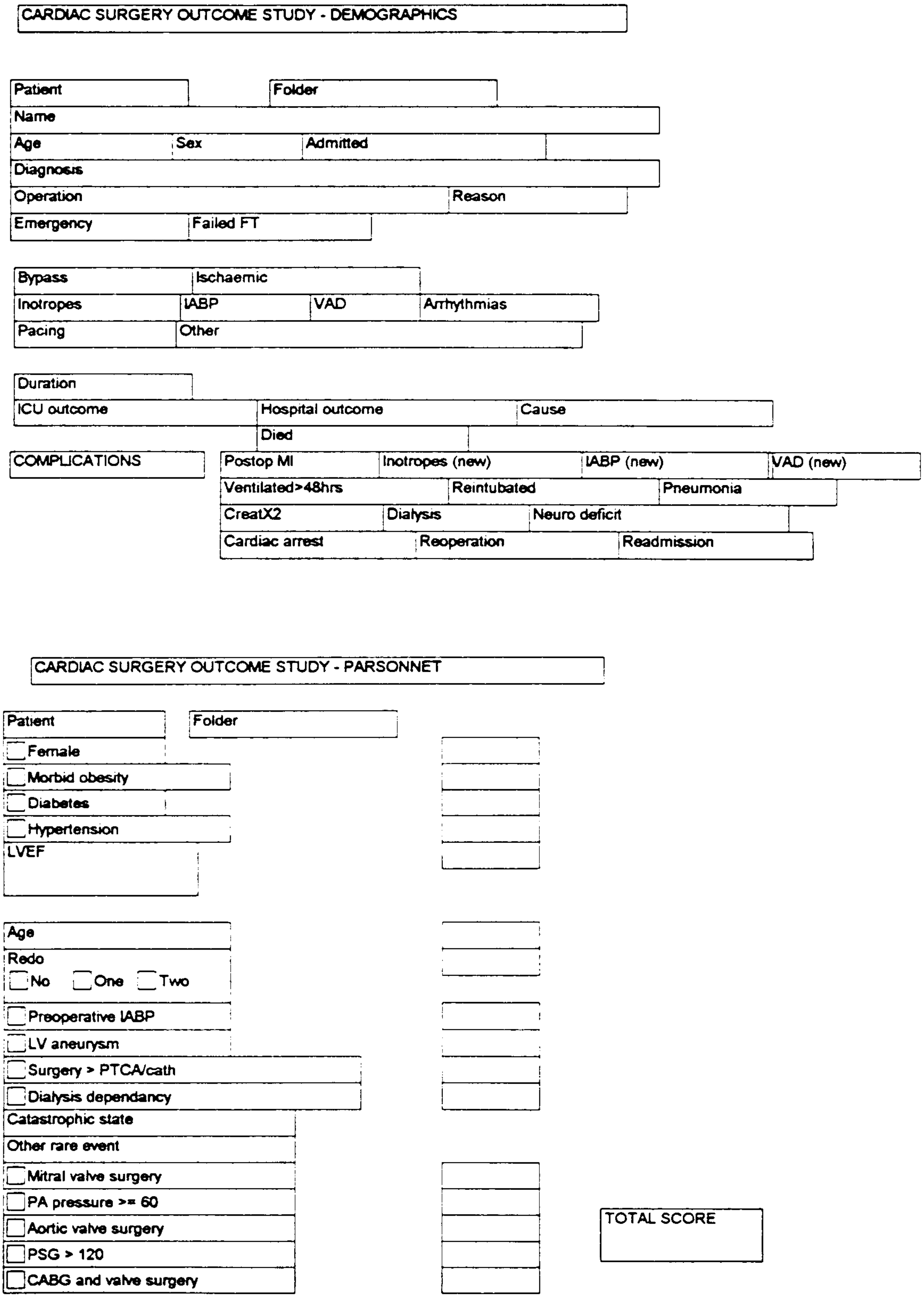
The final program consisted of three separate computer screen forms: demographics (including intraoperative data, complications, and outcome), details relevant to the Parsonnet score, and measurements to perform APACHE II and III scores. An example of two of the forms appears in Figure 7.2.

Raw data were entered into each form, and a value tree for each field allocated points relevant to the above scoring systems. An example of a value tree appears in Figure 7.3. From these allocated points a further value tree performed the calculations to give the Parsonnet, APACHE II, and APACHE III scores.

Each form was linked to a separate Paradox (Borland International Inc, Scotts Valley, California, USA) database, with the study number being the index field and the common link. Both the raw data and the calculated scores were stored in the databases.

Figure 7.2

Examples of ObjectVision forms (in this case the forms for demographic data and Parsonnet score).



CARDIAC SURGERY OUTCOME STUDY - DEMOGRAPHICS

Patient
Folder
Name
Age
Sex
Admitted
Diagnosis
Operation
Reason
Emergency
Failed FT

Bypass
Ischaemic
Inotropes
IABP
VAD
Arrhythmias
Pacing
Other

Duration
ICU outcome
Hospital outcome
Cause
Died

COMPLICATIONS
Postop MI
Inotropes (new)
IABP (new)
VAD (new)
Ventilated>48hrs
Reintubated
Pneumonia
CreatX2
Dialysis
Neuro deficit
Cardiac arrest
Reoperation
Readmission

CARDIAC SURGERY OUTCOME STUDY - PARSONNET

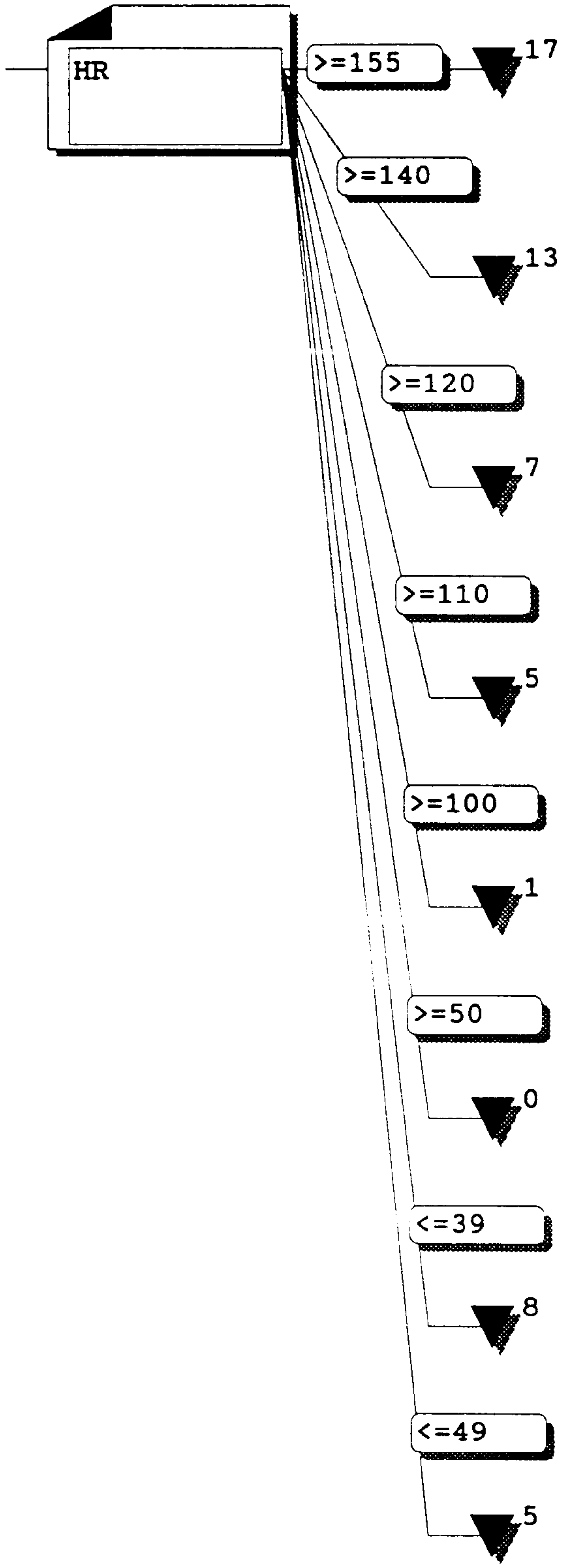
Patient
Folder
☐ Female
☐ Morbid obesity
☐ Diabetes
☐ Hypertension
LVEF

Age
Redo
☐ No ☐ One ☐ Two
☐ Preoperative IABP
☐ LV aneurysm
☐ Surgery > PTCA/cath
☐ Dialysis dependancy
Catastrophic state
Other rare event
☐ Mitral valve surgery
☐ PA pressure >= 60
☐ Aortic valve surgery
☐ PSG > 120
☐ CABG and valve surgery

TOTAL SCORE

Figure 7.3

Example of an ObjectVision value tree (allocating APACHE III points to the heart rate value).



7.2.9 Missing data

All data were available apart from bilirubin and albumin measurements (only relevant to the APACHE III score) in only 12 of 1008 (1.19%) patients. By default, no APACHE III points were scored for these variables in these patients.

7.2.10 Outcome measures

Hospital mortality (or in some studies, 30 day mortality) is a better representation of outcome and is much more widely used than ICU mortality. In this study variables were thus assessed against hospital mortality.

7.2.11 Statistical analysis

The basic principles of statistical analysis, were discussed in Chapter 4. Statistical analyses were performed using Instat 2.04 (GraphPad Software, San Diego, California, USA) on a microcomputer, and BMDP 7.0 (BMDP Statistical Software Inc., Los Angeles, California, USA) on a VAX mainframe computer at the University of Cape Town.

7.2.11.1 Planning of the study

A statistician was consulted during the planning of the study and directed the analysis. In planning the study, it was estimated that there would be approximately 1000 patients enrolled in the space of one year; with a hospital mortality rate of approximately 4-5 % this would mean 40-50 deaths. The possibility of a type II error emerges when death is used as the outcome measure; this is unlikely to be a problem when assessing complications or length of stay. Ideally data collection would have continued for longer and involved more patients, but this was not possible due to constraints on time and funding. However the integrity of the data collection (collected by a single person dedicated to the project) may partially make up for this potential weakness.

7.2.11.2 Statistical methodology and tests

Data were statistically analysed as follows.

Firstly the association between selected simple binary variables and outcome was examined by Fisher's exact test. These variables were selected from previous experience and from review of the literature.

Then scores (APACHE II, APACHE III, Parsonnet) and times (ischaemic, bypass) were put into groups and tests of association with outcome were carried out using the chi-squared for trend test.

Then comparisons of samples of non-parametric data were performed using the Mann-Whitney test.

Then univariate logistic regression of all preoperative, intraoperative, and postoperative variables with hospital outcome was performed. Complications were excluded from this model as they occurred at varying times after surgery and their presence in the model would have been of little practical use.

Finally forward stepwise logistic regression was used to find the best predictive variables for hospital outcome, important and evaluable complications, and ICU stay of more than 24 hours. The BMDP LR program was used for these analyses.

All variables relevant to the nature of the models were entered into them. The entry and removal P values were 0.1 and 0.15 respectively, and categorical variables had to be present in at least 2% of the study population for entry.

Probabilities were calculated in two steps.

Firstly the logit was computed, using the following formula:

Logit = $\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k$, where β_0 is the constant and β_n is the coefficient for the nth variable and x_n the value of that variable, with n taking the values from 1 to k, and k being the total number of variables in the model.

Secondly the probability of death was calculated using the formula:

Probability = $e^{\text{logit}} / 1 + e^{\text{logit}}$

Finally the fit of the model was assessed by the Hosmer-Lemeshow goodness-of-fit statistic (Hosmer & Lemeshow 1989), while the performance of the derived model, the Parsonnet score, and the APACHE II risk of dying were evaluated by calculating the area under the receiver operating characteristic (ROC) curve (Hanley & McNeil 1982). Both of these tests are available on the BMDP LR program.

7.3 Results

7.3.1 Demographics

7.3.1.1 All ICU admissions

There were 1298 patients admitted to the ICU during the study period. The main diagnostic categories are listed in Table 7.1. The categorisation is observer dependent and some categories may be subject to different interpretation. For example, a patient admitted with left ventricular failure may be found to have mitral stenosis and then go on to have cardiac surgery. Such a patient might be classified as medical or surgical (in this study they are classified as surgical). The cardiac surgery grouping here includes surgery for coarctation of the aorta, thoracic aortic aneurysm repair, and repair or rewiring of sternal wounds.

Table 7.1

Main diagnostic categories (all patients during study period).

Diagnosis	**Number**
Cardiac surgery	1054
Thoracic surgery	97
Vascular surgery	28
Medical	102
Readmissions	17
TOTAL	1298

There were 87 deaths, giving an overall ICU mortality rate of 6.7%.

7.3.1.2 Study patients

There were 1008 patients who met the entry criteria and were entered into this study, out of a total of 1298 ICU admissions during the study period. There were 733 males and 275 females. Data for age, bypass time, ischaemic time, Parsonnet score, APACHE II and III score, and duration of ICU admission, are outlined in Table 7.2.

Table 7.2

Data for age, bypass time, ischaemic time, Parsonnet score, APACHE II score, APACHE III score, and duration of ICU admission.

Variable	Mean	SD	Median	Range
Age (years)	60.29	12.78	63	13-84
Bypass time (minutes)	92.44	36.42	88	0-336
Ischaemic time (minutes)	55.87	25.60	51	0-210
Parsonnet score	7.78	6.69	6	0-33
APACHE II score	11.75	4.17	11	3-33
APACHE III score	42.45	15.23	40	9-132
Duration (hours)	46.94	108.74	20	7-1440

Seventy-two patients had emergency surgery, and 49 patients were failed "fast tracks". There were 45 patients who had chronic ill-health as defined by APACHE II, and 3 as defined by APACHE III. Operations performed with mean APACHE II and APACHE III scores and hospital mortality (%) are outlined in Table 7.3. ICU mortality was 2.68%, while hospital mortality was 3.77%. The overall predicted risk of dying calculated from APACHE II methodology was 5.31%, giving a standardised mortality ratio (actual deaths divided by predicted deaths) of 0.71.

Table 7.3

Operations performed, with mean APACHE II and APACHE III scores and hospital mortality (%).

Operation	Number	AP II	AP III	Mort
CABG	646	11.5	40.8	2.48
CABG + valve surgery	89	14.1	52.6	10.11
CABG + aneurysm/ablation	7	13.0	48.1	14.29
Aortic valve surgery	97	11.2	42.2	2.06
Mitral valve surgery	84	12.1	44.4	2.38
Double valve surgery	23	10.7	42.3	8.70
Congenital heart disease	34	9.15	35.6	8.82
Other valve surgery	6	13.2	48.3	0
Aortic dissection repair	6	19.3	71.8	33.3
Other operation	16	13.1	46.9	6.25
TOTAL	1008			

7.3.2 Selected single variables and mortality

Patients who had emergency surgery had a mortality of 13.89%, while those who had elective surgery had a hospital mortality of 2.99% (relative risk 4.64, 95% confidence interval 2.35-9.18, P=0.0002, Fisher's exact test).

Patients admitted on inotropes had hospital mortality of 16.11%, while those who did not had a mortality of 1.63% (relative risk 9.88, 95% confidence interval 5.23-18.67, P<0.0001, Fisher's exact test).

Patients who had an intra-aortic balloon pump (IABP) had a hospital mortality of 23.26%, while those who did not had a mortality of 2.90% (relative risk 7.74, 95% confidence interval 4.12-14.51, P<0.0001, Fisher's exact test).

Hospital mortality for patients readmitted to ICU (n=26) was 30.8%, and for patients not readmitted was 3.1% (relative risk 10.07, 95% confidence interval 5.12-19.80, P<0.0001, Fisher's exact test).

Failed "fast-track" patients had a hospital mortality of 4.08 %, compared with the other patients' mortality of 3.75% (P=0.71, Fisher's exact test), while patients with a chronic health history as defined by APACHE II had a hospital mortality of 6.67 %, compared with the other patients' mortality of 3.63% (P=0.24, Fisher's exact test).

7.3.3 Scoring systems and mortality

The relationship between Parsonnet and APACHE II and III scores and mortality, together with chi-squared test for trend and P values, is shown in Table 7.4.

Table 7.4

The relationship between Parsonnet, APACHE II, and APACHE III scores and mortality.

Parsonnet score	Number	ICU mortality (%)	Hospital mortality (%)
0 - 4	381	0.5	0.5
5 - 9	254	0	1.2
10 - 14	198	3	3.5
15 - 19	117	10.3	12.8
≥ 20	58	12.1	18.97
TOTAL	1008		
Chi-squared for trend		46.64	64.29
P value		<0.0001	<0.0001

APACHE II score	Number	ICU mortality (%)	Hospital mortality (%)
0 - 5	53	0	0
6 - 10	332	0	0.6
11 - 15	466	0.6	1.5
16 - 20	121	10.7	14.9
21 - 25	28	14.3	14.3
>25	8	87.5	87.5
TOTAL	1008		
Chi-squared for trend		112.9	101.1
P value		<0.0001	<0.0001

(Table 7.4 continued...)

APACHE III score	Number	ICU mortality (%)	Hospital mortality (%)
0 - 20	43	0	0
21 - 40	466	0	0.2
41 - 60	396	1.2	2.8
61 - 80	77	13.0	16.9
81 - 100	23	39.1	43.5
>100	3	100	100
TOTAL	1008		
Chi-squared for trend		135.5	135.1
P value		<0.0001	<0.0001

No patient with an APACHE II score of less than 14 or APACHE III score of less than 47 died in ICU, though two patients with APACHE II scores of 8 and 9 respectively and one with an APACHE III score of 37 died in hospital. In addition there were no survivors with an APACHE II score of more than 26 (7 patients) or APACHE III score greater than 93 (6 patients).

7.3.4 Bypass and ischaemic times and mortality

The relationship between bypass and ischaemic times and mortality, together with chi-squared test for trend and p values, is shown in Table 7.5.

Table 7.5

The relationship between bypass and ischaemic times and mortality.

Bypass (minutes)	Number	ICU mortality (%)	Hospital mortality (%)
0 - 60	155	0.7	1.9
61 - 120	685	1.6	2.3
121 - 180	147	8.8	9.5
181 - 240	14	0	14.3
> 240	7	28.6	42.9
TOTAL	1008		
Chi-squared for trend		26.13	33.14
P value		<0.0001	<0.0001

Ischaemic (minutes)	Number	ICU mortality (%)	Hospital mortality (%)
0 - 30	110	2.7	3.6
31 - 60	545	1.5	2.4
61 - 90	255	3.9	4.7
91 - 120	79	6.3	8.9
121 -150	16	0	0
> 150	3	33.3	66.7
TOTAL	1008		
Chi-squared for trend		6.57	9.89
P value		0.01	0.0017

7.3.5 Simple combinations of variables and mortality

There were no simple combinations of preoperative, intraoperative, and postoperative factors that unfailingly predicted death in any substantial number of patients.

7.3.6 Univariate logistic regression

Results of univariate logistic regression of all variables with hospital outcome are shown in Table 7.6.

Table 7.6

Univariate logistic regression with hospital outcome.

Variable	Odds ratio (95% CI)	P value
Age	1.04 (1.00-1.07)	0.031
Albumin	0.88 (0.81-0.95)	<0.0001
Aortic valve surgery	2.07 (1.02-4.18)	0.043
Acute renal failure	56.05 (20.17-155.72)	<0.0001
Bilirubin	1.03 (1.02-1.04)	<0.0001
Bypass time	1.02 (1.01-1.02)	<0.0001
CABG and valve surgery	4.08 (1.91-8.71)	<0.0001
Creatinine	1.01 (1.10-1.01)	<0.0001
Diabetes	2.55 (1.09-5.97)	0.031

(Table 7.6 continued...)

Emergency surgery	5.23 (2.43-11.26)	<0.0001
Failed "fast track"	1.09 (0.25-4.67)	0.907
Glucose	1.19 (1.08-1.30)	<0.0001
Glasgow Coma Scale	0.505 (0.321-0.794)	<0.0001
Haematocrit	0.94 (0.87-1.03)	0.173
Heart rate	1.01 (1.00-1.02)	0.005
Hypertension	1.13 (0.57-2.25)	0.719
IABP	10.14 (4.55-22.59)	<0.0001
Inotropes	11.59 (5.84-23.00)	<0.0001
Ischaemic time	1.02 (1.01-1.03)	0.001
Potassium	2.00 (1.20-3.33)	0.008
LVEF	2.34 (1.53-3.58)	<0.0001
LV aneurysm	7.64 (1.53-38.09)	0.013
Mean arterial pressure	0.92 (0.90-0.95)	<0.0001
Mitral valve surgery	1.09 (0.42-2.86)	0.854
Morbid obesity	0.74 (0.17-3.13)	0.679
Sodium	1.06 (0.96-1.17)	0.267
Pacing	0.51 (0.07-3.78)	0.508
PAP > 60 mmHg	All survived, therefore not analysed	
Preoperative IABP	8.71 (0.88-85.76)	0.064
PSG > 120 mmHg	2.89 (0.36-23.38)	0.321
Reoperation	1.58 (0.94-2.65)	0.083
Respiratory rate	0.90 (0.85-0.96)	0.001

(Table 7.6 continued...)

Sex (male)	0.81 (0.40-1.62)	0.545
Emergency after cath	10.26 (3.11-33.86)	<0.0001
Temperature	1.02 (0.82-1.27)	0.844
Urea	1.13 (1.08-1.19)	<0.0001
Urine output	1.00 (1.00-1.00)	0.010
White cell count	1.03 (0.96-1.10)	0.419
Worst pH	0.004 (0.00-0.12)	0.002
Acute Physiology Score	1.45 (1.33-1.57)	<0.0001
APACHE II score	1.41 (1.30-1.53)	<0.0001
APACHE III score	1.10 (1.08-1.13)	<0.0001
Parsonnet score	1.16 (1.11-1.21)	<0.0001

7.3.7 Forward stepwise logistic regression

The model for hospital mortality, derived by forward stepwise logistic regression, is shown in Table 7.7. The Hosmer-Lemeshow goodness-of-fit statistic was 1.282 (P=0.733).

Table 7.7

Forward stepwise logistic regression model for hospital mortality

Variable	Coefficient	Odds ratio (95% CI)
Bypass time	-0.01134	0.989 (0.982-0.996)
Inotropes	-1.285	0.277 (0.119-0.645)
Mean arterial pressure	0.0694	1.07 (1.04-1.10)
Urea	-0.09081	0.913 (0.858-0.972)
Glasgow Coma Scale	0.5933	1.81 (1.28-2.55)
Constant	-7.116	

7.3.8 ROC curve analysis

The areas under the ROC curves for hospital mortality and Parsonnet score, APACHE II risk of dying, and the derived model are shown in Table 7.8. The ROC curves themselves are shown in Figures 7.4 -7.6.

Table 7.8

The areas under the ROC curves for hospital mortality.

Test	Area under ROC curve
Parsonnet score	0.8229
APACHE II risk of dying	0.8411
Derived model	0.8708

Figure 7.4

ROC plot of Parsonnet score vs hospital mortality.

```
      .+.....+.....+.....+.....+.....+.....+.....+.....+.....+.....+.
   1.0 +                                                       **//+
       -                                          *  **  *      *//-
       -                                       *            /      -
       -                               *                   /       -
       -                              *                  /         -
       -                                                           -
       -                 *                            /            -
       -                                                           -
P  .80 +               *                               /           +
R      -                                                           -
O      -                                                           -
P      -                                                           -
O      -                                                           -
R      -                                                           -
T      -                                                           -
I      -       *                           /                       -
O  .60 +                                                           +
N      -                                                           -
       -                                                           -
T      -                                                           -
R      -                                                           -
U      -                                                           -
E      -                                                           -
       -                                                           -
P  .40 +                                                           +
O      -                                                           -
S      -                                                           -
I      -                                                           -
T      -                                                           -
I      -                                                           -
V      -                                                           -
E      -                                                           -
S  .20 +                                                           +
       -  *      /                                                 -
       -                                                           -
       -                                                           -
       -                                                           -
       -                                                           -
       -                                                           -
       -                                                           -
   0.0 +/                                                          +
      .+.....+.....+.....+.....+.....+.....+.....+.....+.....+.....+.
            .10         .30         .50         .70         .90
      0.0         .20         .40         .60         .80         1.0

                    PROPORTION FALSE POSITIVES
```

THE AREA UNDER THE POLYGON, FORMED BY CONNECTING THE POINTS (0,0) THROUGH THE ASTERISKS TO (1,1), IS 0.8229

BMDPLR - STEPWISE LOGISTIC REGRESSION Copyright 1977, 1979, 1981, 1982, 1983, 1985, 1987, 1988, 1990, 1993 by BMDP Statistical Software, Inc.

Figure 7.5

ROC plot of APACHE II risk of dying vs hospital mortality.

```
     .+.....+.....+.....+.....+.....+.....+.....+.....+.....+.....+.
  1.0 +                                     *   **  *  *  *   **//+
      -                                     * * *               //-
      -                                 *                       / -
      -                            *                           /  -
      -                  *                                    /   -
      -                                                           -
      -                                                           -
      -                                                           -
P .80 +                                                           +
R     -           *                                 /             -
O     -                                                           -
P     -                                                           -
O     -                                                           -
R     -                                                           -
T     -                                                           -
I     -                                                           -
O .60 +                                                           +
N     -                                                           -
      -                                                           -
T     -                                                           -
R     -                                                           -
U     -                                                           -
E     -                                                           -
      -                                                           -
P .40 +                                                           +
O     -                                                           -
S     -                                                           -
I     -                                                           -
T     -                                                           -
I     -                                                           -
V     -                                                           -
E     -                                                           -
S .20 +                                                           +
      -                                                           -
      -                                                           -
      -                                                           -
      -                                                           -
      -                                                           -
      -                                                           -
      -                                                           -
  0.0 +/                                                          +
     .+.....+.....+.....+.....+.....+.....+.....+.....+.....+.....+.
           .10         .30         .50         .70         .90
     0.0         .20         .40         .60         .80         1.0

                     PROPORTION FALSE POSITIVES

THE AREA UNDER THE POLYGON, FORMED BY CONNECTING THE POINTS (0,0)
THROUGH THE ASTERISKS TO (1,1), IS 0.8411
```

Figure 7.6

ROC plot of the derived model vs hospital mortality.

```
       .+.....+.....+.....+.....+.....+.....+.....+.....+.....+.....+.
    1.0 +                                    * ** ** * ** * *     //+
        -                 *    ** * * * **                        //-
        -                                                           -
        -               *                              /            -
        -                                                           -
        -         *                                /                -
        -                                                           -
        -                                                           -
P   .80 +                                                           +
R       -                                                           -
O       -                                                           -
P       -                                                           -
O       -                                                           -
R       -                                                           -
T       -                                                           -
I       -                                                           -
O   .60 +                                                           +
N       -                                                           -
        -                                                           -
T       -                                                           -
R       -                                                           -
U       -                                                           -
E       -                                                           -
        -                                                           -
P   .40 +                                                           +
O       -                                                           -
S       -                                                           -
I       -                                                           -
T       -                                                           -
I       -                                                           -
V       -                                                           -
E       -                                                           -
S   .20 +                                                           +
        -                                                           -
        -                                                           -
        -                                                           -
        -                                                           -
        -                                                           -
        -                                                           -
        -                                                           -
    0.0 +/                                                          +
       .+.....+.....+.....+.....+.....+.....+.....+.....+.....+.....+.
             .10         .30         .50         .70         .90
       0.0         .20         .40         .60         .80         1.0

                         PROPORTION FALSE POSITIVES
```

THE AREA UNDER THE POLYGON, FORMED BY CONNECTING THE POINTS (0,0) THROUGH THE ASTERISKS TO (1,1), IS 0.8708

7.3.9 Subgroup of CABG patients

A model was developed in a similar manner to predict hospital outcome in the 646 patients undergoing CABG only. It is not shown as it included too large a number of variables (sixteen) for any practical use. The Hosmer-Lemeshow goodness-of-fit statistic was 1.615 (P=0.204)

7.3.10 Complications

Postoperative complications (see above for the definitions of those chosen for evaluation) are shown in Table 7.9.

Table 7.9

Postoperative complications.

Complication	Number of patients
Postoperative myocardial infarct	5
IABP (new)	17
Inotropes (new)	74
Ventilated > 48 hours	81
Reintubation	20
Pneumonia	12
Creatinine x 2 (from preoperative)	17
Dialysis	12
Neurological deficit	15
Cardiac arrest	18
Reoperation	53
Readmission	26

Any complication occurring in less than 2% of the study population was not analysed further. Reoperation and new inotrope administration often occurred during the first 24 hours of admission and the physiology variables may have been at their worst before, during, or afterwards, making analysis difficult. Readmission may be influenced by a number of variables in the ICU and the ward, and therefore was not analysed either. A model was developed from all relevant variables to predict the occurrence of ventilation for more than 48 hours but was not practical as it included 11 variables.

7.3.11 Length of stay in ICU

The mean length of stay of patients with chronic health by APACHE II criteria was 116.3 hours (median 44 hours) while that of patients without chronic health was 43.7 hours (median 20 hours), $P<0.0001$, Mann-Whitney test. The mean length of stay of failed "fast track" patients was 47.8 hours (median 21 hours) while that of the other patients was 31.1 hours (median 16 hours), $P<0.0001$, Mann-Whitney test.

A model was developed from variables available at the time of ICU admission to predict a length of stay of more than 24 hours (Table 7.10). There were 230 patients in this group. The Hosmer-Lemeshow goodness-of-fit statistic was 3.218 ($P=0.920$).

Table 7.10

Model for length of stay more than 24 hours, from variables available at time of ICU admission.

Variable	Coefficient	Odds ratio (95% CI)
Emergency surgery	1.283	3.61 (1.95-3.66)
Failed fast track	1.049	2.85 (1.45-5.61)
Bypass time	-0.01093	1.01 (1.01-1.02)
Inotropes	1.941	6.97 (4.36-11.1)
IABP	3.005	20.2 (2.62-155)
LVEF*	0.3423	1.41 (1.04-1.90)
Constant	-3.402	

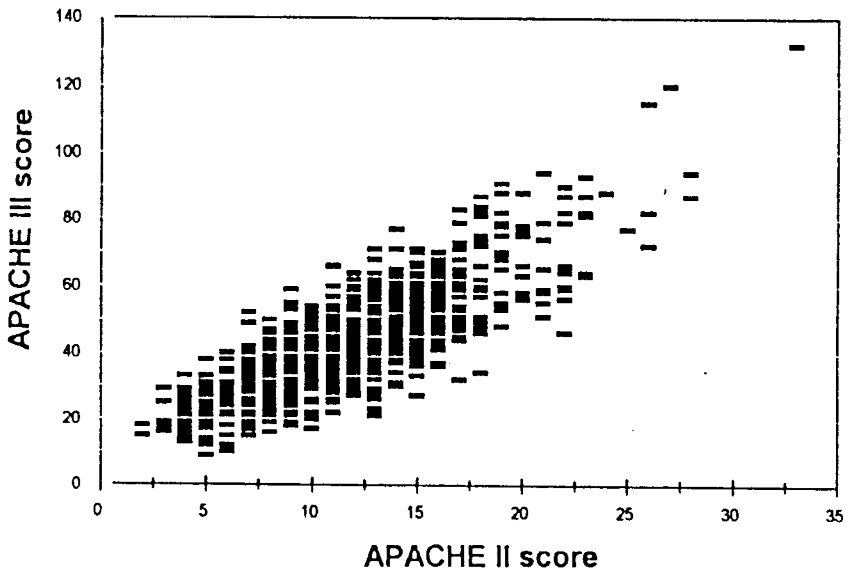
* LVEF scored as 1 if ≥50%, 2 if 30-49%, and 3 if <30%.

7.3.12 APACHE II vs APACHE III scores

APACHE II scoring was compared with APACHE III scoring in all patients entered into the study. The scatterplot is presented in Figure 7.7 ($r^2 = 0.645$).

Figure 7.7

APACHE II vs APACHE III scoring in all patients.



7.3.13 Cardiac surgery patients not admitted to ICU

There were 257 patients who had cardiac surgery procedures but were not admitted to the ICU during the study period ("fast-track" patients). None died and they are not described further.

7.4 Discussion

It is perhaps for good reasons that the major severity of illness scoring systems, in particular the more recent SAPS II (LeGall et al 1993) and MPM II (Lemeshow et al 1993), have excluded all cardiac surgery patients, while APACHE III has excluded those undergoing CABG (Knaus et al 1991a). High scores do not reproducibly transfer into high mortality, and in addition the ability of pharmacological and mechanical support to normalise physiology makes the interpretation of physiology-based scoring systems difficult. We have previously found limitations in the application of APACHE II scoring in cardiac surgery patients (Turner et al 1991) and in this study have attempted to rectify this by collecting preoperative and intraoperative variables, as well as using both APACHE II and III scores and the raw physiological variables that generate them. Thus we are not competing with a number of preoperative risk evaluations, which obviously cannot take into account the intraoperative course of the patient, but rather attempting more accurate prediction by including intraoperative events and the physiological abnormalities of the first 24 hours in ICU. Missing data in this study was minimal. This may be compared with the APACHE II data collection where all 12 physiological variables were available in only 87% of patients (Knaus et al 1985a), and also with the APACHE III data collection where 99% of patients had complete heart rate, respiratory rate, blood pressure, and urine measurements and 85% of patients had sodium, potassium, and haematocrit measurements. However only 65% of patients had arterial blood gas measurements performed (Knaus et al 1991a).

While demographic data were generally similar to our previous study, there was a significant difference in the numbers of patients allocated APACHE II chronic health points (45 out of 1008 patients here vs 145 out of 811 patients). As the patient population is unlikely to have substantially changed, this may show that application of the definitions may be more subjective than previously thought. Hospital mortality was 3.77% in the whole study group and 2.48% in the patients undergoing CABG alone. These numbers are consistent with current mortality rates worldwide (Hannan et al 1990, Geraci et al 1993, Daly et al 1993).

Combination of CABG with valve surgery, aneurysm resection, or ablation of bypass pathway increased mortality appreciably. The standardised mortality ratio of 0.71 looks impressive but it may simply reflect the inaccuracy of current physiology-based severity scoring in cardiac surgery.

The present study was successful in deriving models for mortality and a length of stay of more than 24 hours. Models for complications were less useful. Although the patient numbers are relatively small, these models are powerful by virtue of their derivation, namely forward stepwise logistic regression. This is a very powerful statistical tool which allows variables to be objectively selected and weighted, with each variable in the final model controlled for the other variables (Hosmer & Lemeshow 1989). Of course the choice of which variables to include is a subjective one, but if a large enough number of variables is chosen it is unlikely that any important ones will be omitted. In addition, the variables in the final model may be dependent upon the data set used, as may the size of the coefficients. Thus models so derived should be validated using other data sets; split-halves validation may go some of the way towards this if the numbers are large enough, as in the APACHE III methodology (Knaus et al 1991a). Stepwise forward logistic regression in this study allowed the derivation of a model for hospital mortality. This model included bypass time, need for inotropes, mean arterial pressure, urea, and Glasgow Coma Scale. None of the variables available preoperatively survived the regression process. The model fitted the data well, with an impressive Hosmer-Lemeshow goodness-of-fit statistic and an area under the ROC curve of 0.87. However the numbers of deaths was small, possibly weakening the predictive ability. In addition, the model is complicated to calculate and the probability of death could only be computed after 24 hours.

On a simpler note, single variables which were statistically associated with hospital outcome by 2x2 table analysis included emergency surgery, readmission to ICU, and the need for inotropes or IABP. The presence of chronic health (APACHE II criteria) was not associated with significantly increased mortality, unlike in our previous study, while failed fast-track patients also had no significantly increased mortality. The reasons for this are unclear. However, both these groups had significantly longer ICU stays.

Increasing Parsonnet score, bypass time, ischaemic time, APACHE II score, and APACHE III score were all significantly correlated with mortality. This was shown in another way by univariate logistic regression. The chi-squared for trend values were highest for APACHE III and lowest for ischaemic time. The Parsonnet score overestimated hospital mortality, particularly in the lower ranges, just as it did in a previous UK study (Nashef et al 1992). Low APACHE II and III scores were still associated with hospital mortality in a small number of patients, as were high scores with survival. There was no combination of Parsonnet score, bypass time, or APACHE II and III scores which infallibly predicted death in any but a very small number of patients; in fact simple APACHE II and III scores had more patients beyond a 100% mortality cutoff score.

The APACHE II risk of dying had an area under the ROC curve of 0.84. This is close to the area of 0.85 previously reported for APACHE II (Knaus et al 1991a), but not as good as the 0.90 reported for APACHE III (Knaus et al 1991a) or 0.88 for SAPS II (LeGall et al 1993). These differences may be clinically insignificant.

We were unable to calculate APACHE III risk of dying as the equations are copyright, but as the reasons for admission and treatment locations were similar in all patients, the equations may not have added much to the predictive ability. In any case the relationship between APACHE III and hospital mortality as shown by the chi-squared for trend test was impressive.

In practical terms, however, there are two important reasons for prognostic scoring in cardiac surgery. The first is for both the physician and the patient to get an idea of likely risk i.e. hospital mortality. This prepares the physician for complications and helps stratify the patient into needing ICU admission or not. It also helps the patient and family to weigh up the risks and benefits of the surgery and clarifies their expectations. The second is to be able to compare results for patients using equality of risk groups. The Parsonnet score seems able to perform the first function with reasonable facility, with an area under the ROC curve of 0.82; the overestimation of risk in low-risk groups was small and in this context not necessarily bad.

A further more pragmatic reason for prognostic scoring (and one that is becoming more and more important in the United Kingdom with the ongoing health care reforms), is to provide the potential purchasers of health care with comparative performance figures. These can be used both to show that the standard of care is good and to negotiate prices and contracts for the relevant services.

The relationship between APACHE II and APACHE III scores showed a fairly wide scatter with a relatively low correlation. This is not entirely unexpected, as the scores have markedly different scales for point allocation. There is no "gold standard", so further conclusions would be futile.

The model derived in this study should ideally have been prospectively validated, at least in the same ICU, if not in another cardiac surgery ICU in another country. Unfortunately and regrettably this was not possible for a number of reasons beyond the control of the researcher. It is hoped that the simpler hospital outcome model will be able to be validated in the near future but this will depend upon the availability of resources.

7.5 Conclusions

In conclusion, cardiac surgery remains a very complex area for outcome prediction. Even the selection of preoperative, intraoperative, and postoperative variables does not seem to cover all outcome eventualities although it does improve upon available systems. Severity of illness scoring systems provide good predictive power for groups of patients, but individual patient prediction remains difficult.

CHAPTER 8

Conclusions, future directions, publications, presentations

8.1 Conclusions

There have been important advances in both the growth and development of severity of illness scoring systems in the last decade, and the three studies presented here follow these trends. They also follow the advances in the understanding of the use of these scores and, perhaps more importantly, their limitations.

Severity of illness scoring is certainly an effective means of stratifying critically ill patients, both in general ICUs and in most specialised areas. It is also effective in outcome prediction in patients as a group, though individual outcome prediction is substantially less useful, resulting at times in misconceptions and disappointment. This may change with the development of very large international databases and the mathematical and statistical ability to derive risk prediction equations and thereby improve the ability to predict death or survival with tight confidence intervals.

Cardiac surgery presents particular problems both in the paucity of data regarding the use of general severity of illness scores, and in the predictive difficulties implicit in the discipline. The combination of variables used in the final study proved to be the best mortality predictor evaluated, and may show the way for future predictive instruments.

8.2 Future directions

The models developed in the final study should ideally have been prospectively validated; this was unfortunately impossible due to a variety of reasons. However, depending on the availability of resources, it may be possible to validate the model for prediction of hospital mortality in the same ICU.

It would appear that the combination of preoperative, intraoperative, and postoperative variables, as used in the third study, might be the best option for a scoring system and outcome predictor, balanced as it is between the preoperative risk predictions and the postoperative conventional scoring systems. Future studies might address different combinations of such variables, in larger groups of patients, and with prospective validation of the derived models.

8.3 Publications related to this research

1. Turner JS, Potgieter PD. Scoring severity of illness (Editorial). Southern African Journal of Critical Care 1987;3:25

2. Turner JS, Linton DM, Potgieter PD. Systems for scoring severity of illness in intensive care. South African Medical Journal 1989;76:17-20

3. Turner JS, Potgieter PD, Linton DM. Severity of disease classification system (letter). Critical Care Medicine 1991;19:301-302

4. Turner JS, Mudaliar YM, Chang RWS, Morgan CJ. Acute Physiology and Chronic Health Evaluation (APACHE II) scoring in a cardiothoracic intensive care unit. Critical Care Medicine 1991;19:1266-1269

5. Turner JS, Potgieter PD. Severity scoring in ARDS. In: ARDS (Eds Evans TW, Haslett C), Chapman and Hall, London, (in press)

8.4 Abstracts related to this research

1. Turner JS, Mudaliar YM, Chang RWS, Morgan CJ. APACHE II scoring in a cardiothoracic ICU. American Review of Respiratory Disease 1991;143:A469

2. Turner JS, Morgan CJ. Outcome prediction in cardiac surgery. Clinical Intensive Care 1993;4:314

2. Turner JS, Morgan CJ. A novel approach to outcome prediction in cardiac surgery. Chest (in press)

8.5 Presentations to learned societies related to this research

1. **The application of severity of illness scoring systems in a Respiratory ICU.** Congress of the Critical Care Society of Southern Africa, Port Elizabeth, South Africa, 1987 (Sabax award for best Registrar presentation).

2. **APACHE II scoring in a cardiothoracic ICU.** American Thoracic Society International Conference, Anaheim, USA, 1991.

3. **Outcome prediction in cardiac surgery.** Intensive Care Society (United Kingdom), London, England, 1993.

4. **A prospective study of outcome prediction in cardiac surgery.** Congress of the Critical Care Society of Southern Africa, Johannesburg, South Africa, 1994.

5. **APACHE III scoring in cardiac surgery: comparison with APACHE II.** Congress of the Critical Care Society of Southern Africa, Johannesburg, 1994.

6. **A novel approach to outcome prediction in cardiac surgery.** 60th Annual International Scientific Assembly of the American College of Chest Physicians, New Orleans, USA, 1994.

CHAPTER 9

References

Abbott RR, Setter M, Chan S, Choi K (1991) APACHE II: prediction of outcome of 451 ICU oncology admissions in a community hospital. Annals of Oncology 2:571-574

Acar C, Jebara VA, Portoghese M, et al (1992) Revival of the radial artery for coronary artery bypass grafting. Annals of Thoracic Surgery 54:652-660

Acinapura AJ, Jacobowitz IJ, Krame MD, et al (1992) Internal mammary artery bypass: thirteen years of experience. Influence of angina and survival in 5125 patients. Journal of Cardiovascular Surgery 33:554-559

Adam BA, Kin LC, Wahab ASA (1989) Therapeutic intervention scoring system in medical intensive care. Medical Journal of Malaysia 44:134-139

Akl ES, Ozdogan E, Ohri SK, et al (1992) Early and long term results of re-operation for coronary artery disease. British Heart Journal 68:176-180

American Thoracic Society (1991) Withholding and withdrawing life-sustaining therapy. American Review of Respiratory Disease 144:726-731

Anderson RP (1986) Will the real CASS please stand up? A review and perspective on the Coronary Artery Surgery Study. Journal of Thoracic and Cardiovascular Surgery 91:698-709

Anonymous (1986) TPN and APACHE (Editorial). Lancet i:1478.

Apgar V (1953) A proposal for a new method of evaluation of the newborn infant. Current Researches in Anesthesia and Analgesia 32:260-267

Arregui LM, Moyes DG, Lipman J, Fatti LP (1991) Comparison of disease severity scoring systems in septic shock. Critical Care Medicine 19:1165-1171

Artigas A, Carlet J, Chastang C, et al (1992) Adult respiratory distress syndrome: clinical presentation, prognostic factors and outcome. In: Adult Respiratory Distress Syndrome (eds A. Artigas, F. Lemaire, P.M. Suter, W.M. Zapol), Churchill Livingstone, Edinburgh, 509-525

Baker SP, O'Neill B, Haddon W, Long WB (1974) The injury severity score: a method for describing patients with multiple injuries and evaluating emergency care. Journal of Trauma 14:187-196

Baldwin RT, Slogoff S, Noon GP, et al (1993) A model to predict survival at time of postcardiotomy intraaortic balloon pump insertion. Annals of Thoracic Surgery 55:908-913

Bastos PG, Sun X, Wagner DP, et al (1993) Glasgow Coma Scale score in the evaluation of outcome in the intensive care unit: findings from the Acute Physiology and Chronic Health III study. Critical Care Medicine 21:1459-1465

Baumgartner J, Bula C, Vaney C, et al (1992) A novel score for predicting the mortality of septic shock patients. Critical Care Medicine 20:953-960

Berger MM, Marazzi A, Freeman J, Chiolero R (1992) Evaluation of the consistency of Acute Physiology and Chronic Health Evaluation (APACHE II) scoring in a surgical intensive care unit. Critical Care Medicine 20:1681-1687

Bergner M, Bobbitt RA, Carter WB, et al (1981) The sickness impact profile: development and final revision of a health status measurement. Medical Care 19:787-805

Bickell NA, Pieper KS, Lee KL, et al (1992) Referral patterns for coronary artery disease treatment: gender bias or good clinical judgment? Annals of Internal Medicine 116:791-797

Bion JF, Aitchison TC, Edlin SA, Ledingham IMcA (1988) Sickness scoring and response to treatment as predictors of outcome from critical illness. Intensive Care Medicine 14:167-172

Black PMcL (1978) Brain death. New England Journal of Medicine 299:338-344 and 393-401

Black PMcL (1985) Predicting the outcome from hypoxic-ischemic coma: medical and ethical implications (Editorial). Journal of the American Medical Association 254:1215-1216

Blamey SL, Imrie CW, O'Neill J, et al (1984) Prognostic factors in acute pancreatitis. Gut 24:1340-1346

Bland JM, Altman DG (1986) Statistical methods for assessing agreement between two methods of clinical measurement. Lancet i:307-310

Bohnen JMA, Mustard RA, Oxholm SE, Schouten D (1988) APACHE II score and abdominal sepsis. Archives of Surgery 123:225-229

Bolsin SN, Morgan C, Ray R (1990) A simple, preoperative scoring system for predicting mortality and intensive care unit stay in adult patients undergoing heart surgery with cardiopulmonary bypass. Journal of Cardiothoracic Anesthesia 4 (Supplement 3):139

Boyd CR, Tolson MA, Copes WS (1987) Evaluating trauma care: the TRISS method. Journal of Trauma 27:370-378

Boyd O, Grounds RM (1993) Physiological scoring systems and audit. Lancet 341:1573-1574

Brannen AL, Godfrey LJ, Goetter WE (1989) Prediction of outcome from critical illness. A comparison of clinical judgement with a prediction rule. Archives of Internal Medicine 149:1083-1086

Brawley RK, Merrill W, Gott VL, et al (1980) Unstable angina pectoris. Factors influencing operative risk. Annals of Surgery 191:745-750

Bueno-Cavanillas A, Rodriguez-Contreras R, Lopez-Luque A, et al (1991) Usefulness of severity indices in intensive care medicine as a predictor of nosocomial infection risk. Intensive Care Medicine 17:336-339

Bull JP (1971) Revised analysis of mortality due to burns. Lancet ii:1133-1134

Burgos LG, Ebert TJ, Assiddao C, et al (1989) Increased intraoperative cardiovascular morbidity in diabetics with autonomic neuropathy. Anesthesiology 70:591-597

Castella X, Gilabert J, Torner F, Torres C (1991) Mortality prediction models in intensive care: Acute Physiology and Chronic Health Evaluation II and Mortality Prediction Model compared. Critical Care Medicine 19:191-197

Cerra FB, Negro F, Abrams J (1990) APACHE II score does not predict multiple organ failure or mortality in postoperative surgical patients. Archives of Surgery 125:519-522

Chaitman BR, Rogers WJ, Davis K, et al (1980) Operative risk factors in patients with left main coronary-artery disease. New England Journal of Medicine 303:953-957

Champion HR, Sacco WJ, Carnazzo AJ, et al (1981) Trauma score. Critical Care Medicine 9:672-676

Champion HR, Sacco WJ, Copes WS, et al (1990) A new characterization of injury severity. J Trauma 30:539-546

Chang RWS, Jacobs S, Lee B (1986) Use of APACHE II severity of disease classification to identify intensive-care-unit patients who would not benefit from total parenteral nutrition. Lancet i:1483-1487.

Chang RWS, Jacobs S, Lee B (1988) Predicting outcome among intensive care unit patients using computerised trend analysis of daily APACHE II scores corrected for organ system failure. Intensive Care Medicine 14:558-566

Chang RWS, Lee B, Jacobs S, Lee B (1989) Accuracy of decisions to withdraw therapy in critically ill patients: clinical judgement versus a computer model. Critical Care Medicine 17:1091-1097

Chang RWS, Bihari DJ (1993) Physiological scoring systems and audit (letter). Lancet 342:306

Chen FG, Koh KF, Goh MH (1993) Validation of the APACHE II score in a surgical intensive care unit. Singapore Medical Journal 34:322-324

Chisakuta AM, Alexander JP (1990) Audit in intensive care. The APACHE II classification of severity of disease. Ulster Medical Journal 59:161-167.

Christakis GT, Weisel RD, Fremes SE, et al (1992) Coronary artery bypass grafting in patients with poor ventricular function. Journal of Thoracic and Cardiovascular Surgery 103:1083-1092

Chu DY (1993) Predicting survival in AIDS patients with respiratory failure. Application of the APACHE II scoring system. Critical Care Clinics 9:89-105

Civetta JM, Hudson-Civetta JA, Kirton O, et al (1992) Further appraisal of APACHE II limitations and potential. Surgery, Gynecology and Obstetrics 175:195-203

Civetta JM, Hudson-Civetta JA, Nelson LD (1990) Evaluation of APACHE II for cost containment and quality assurance. Annals of Surgery 212:266-274

Civetta JM (1990) "New and improved" scoring systems (Editorial). Critical Care Medicine 18:1487-1490

Civil ID, Schwab CW (1988) The Abbreviated Injury Scale, 1985 revision: a condensed chart for clinical use. J Trauma 28:87-90

Clement R, Rousou JA, Engelman RM, Breyer RH (1988) Perioperative morbidity in diabetics requiring coronary artery bypass surgery. Annals of Thoracic Surgery 46:321-323

Copeland GP, Jones D, Walters M (1991) POSSUM: a scoring system for surgical audit. British Journal of Surgery 78:356-360

Copes WS, Champion HR, Sacco WJ, et al (1988) The Injury Severity Score revisited. Journal of Trauma 28:69-77

Cullen DJ, Civetta JM, Briggs BA, Ferrara LC (1974) Therapeutic intervention scoring system: a method for quantitative comparison of patient care. Critical Care Medicine 2:57-60

Cullen DJ (1977) Results and cost of intensive care. Anesthesiology 47:203-216

Daly LE, Lonergan M, Graham I (1993) Predicting operative mortality after coronary artery bypass surgery in males. Quarterly Journal of Medicine 86:771-778

Dart R, Patel B, Perez-Alard J, et al (1991) Prognosis of oncology patients receiving intensive care using the APACHE II scoring system. Maryland Medical Journal 40:273-276

Dawes RM, Faust D, Meehl PE (1989) Clinical versus actuarial judgement. Science 243:1668-1674

Day AC, Rankin APN, Judson JA (1993) Grading asthma severity: using the APS component of the APACHE II system. Intensive Care Medicine 19:221-226

Dobkin JE, Cutler RE (1988) Use of APACHE II classification to evaluate outcome of patients receiving hemodialysis in an intensive care unit. Western Journal of Medicine 149:547-550

Dragsted L, Jorgensen J, Jensen NH, et al (1989a) Interhospital comparisons of patient outcome from intensive care: importance of lead-time bias. Critical Care Medicine 17:418-422

Dragsted L, Qvist J (1989b) Outcome from intensive care. I. A 5-year study of 1308 patients: methodology and patient population. European Journal of Anaesthesiology 6:23-37

Dragsted L, Qvist J (1989c) Outcome from intensive care III. A 5-year study of 1308 patients activity level. European Journal of Anaesthesiology 6:385-396

Draper EA, Wagner DP, Russo M, et al (1989) APACHE III study design: analytic plan for evaluation of severity and outcome in intensive care unit patients. Study design-data collection. Critical Care Medicine 17:S186-S193

Dripps RD, Lamont A, Eckenhoff JE (1961) The role of anesthesia in surgical mortality. Journal of the American Medical Association 178:261-266

Ebell MH, Preston PS (1993) The effect of the APACHE II score and selected clinical variables on survival following cardiopulmonary resuscitation. Family Medicine 25:191-196

Edgren E, Hedstrand U, Kelsey S, et al (1994) Assessment of neurological prognosis in comatose survivors of cardiac arrest. Lancet 343:1055-1059

Edmunds LH, Stephenson LW, Edie RN, Ratcliffe MB (1988) Open-heart surgery in octogenarians. New England Journal of Medicine 319:131-136

Edwards FH, Albus RA, Zajtchuk R, et al (1989) A quality assurance model of operative mortality in coronary artery surgery. Annals of Thoracic Surgery 47:646-649

Edwards FH, Bellamy RF, Burge JR, et al (1990) True emergency coronary artery bypass surgery. Annals of Thoracic Surgery 1990;48:603-610

Ephron B, Tibshirani R (1991) Statistical data analysis in the computer age. Science 253:390-395

Escarce JJ, Kelley MA (1990) Admission source to the medical intensive care unit predicts hospital death independent of APACHE II score. Journal of the American Medical Association 264:2389-2394

Fagon JY, Chastre J, Novara A, et al (1993) Characterization of intensive care unit patients using a model based on the presence or absence of organ dysfunctions and/or infection: the ODIN model. Intensive Care Medicine 19:137-144

Faxon DP, Ryan TJ, Davis KB, et al (1982) Prognostic significance of angiographically documented left ventricular aneurysm from the Coronary Artery Surgery Study (CASS). American Journal of Cardiology 50:157-164

Fedullo AJ, Swinburne AJ, Wahl GW, Bixby KRI (1988) APACHE II score and mortality in respiratory failure due to cardiogenic pulmonary oedema. Critical Care Medicine 16:1218-1221

Feller I, Tholen D, Cornell RG (1980) Improvements in burn care, 1965 to 1979. Journal of the American Medical Association 244:2074-2078

Ferraris VA, Propp ME (1992) Outcome in critical care patients: a multivariate study. Critical Care Medicine 20:967-976

Fietsam R, Basset J, Glover J (1991) Complications of coronary artery surgery in diabetic patients. American Surgeon 57:551-557

Firth WJ (1991) Chaos-predicting the unpredictable. British Medical Journal 303:1565-1568

Fischl MA, Pitchenik A, Gardner LB (1981) An index predicting relapse and need for hospitalization in patients with acute bronchial asthma. New England Journal of Medicine 305:783-789

Fisher LD, Kennedy JW, Davis KB, et al (1982) Association of sex, physical size, and operative mortality after coronary artery bypass in the Coronary Artery Surgery Study (CASS). Journal of Thoracic and Cardiovascular Surgery 84:334-341

Fisher LD, Kennedy JW (1983) Operative mortality in coronary bypass grafting (letter). Journal of Thoracic and Cardiovascular Surgery 85:146-147

Foster ED, Fisher LD, Kaiser GC, et al (1984) Comparison of operative mortality and morbidity for initial and repeat coronary artery bypass grafting: the Coronary Artery Surgery Study (CASS) registry experience. Annals of Thoracic Surgery 38:563-570

Gagner M, Franco D, Vons C, et al (1991) Analysis of morbidity and mortality rates in right hepatectomy with the preoperative APACHE II score. Surgery 110:487-492

Geraci JM, Rosen AK, Ash AS, et al (1993) Predicting the occurrence of adverse events after coronary artery bypass surgery. Annals of Internal Medicine 118:18-24

Gersh BJ, Kronmal RA, Frye RL, et al (1983) Coronary arteriography and coronary artery bypass surgery: morbidity and mortality in patients ages 65 years or older. A report from the Coronary Artery Surgery Study. Circulation 67:483-491

Giangiuliani G, Mancini A, Gui D (1989) Validation of a severity of illness score (APACHE II) in a surgical intensive care unit. Intensive Care Medicine 15:519-522

Gilbert J, Schoolfield J (1991) The outcome index and system outcome score: a method of quality assurance through outcome analysis in the special care area. International Journal of Clinical Monitoring and Computing 8:85-94

Glower DD, Christopher TD, Milano CA, et al (1992) Performance status and outcome after coronary artery bypass grafting in persons aged 80 to 93 years. American Journal of Cardiology 70:567-571

Goldman L, Caldera DL, Nussbaum SR, et al (1977) Multifactorial index of cardiac risk in noncardiac surgical procedures. New England Journal of Medicine 297:845-850

Goldstein RL, Campion EW, Thibault GE, et al (1986) Functional outcome following medical intensive care. Critical Care Medicine 14:783-788

Gross PA, DeMauro PJ, Van Antwerpen C, et al (1988) Number of comorbidities as a predictor of nosocomial infection acquisition. Infection Control and Hospital Epidemiology 9:497-500

Gross PA, Stein MR, van Antwerpen C, et al (1991) Comparison of severity of illness indicators in an Intensive Care Unit. Archives of Internal Medicine 151:2201-2205

Grover FL, Johnson RJ, Marshall G, Hammermeister KE (1993) Factors predictive of operative mortality among coronary artery bypass subsets. Annals of Thoracic Surgery 56:1296-1307

Hammermeister KE, Burchfiel C, Johnson R, Grover FL (1990) Identification of patients at greatest risk for developing major complications at cardiac surgery. Circulation 82(Suppl IV):380-389

Hammermeister KE, Kennedy JW (1974) Predictors of surgical mortality in patients undergoing direct myocardial revascularisation. Circulation 49(Suppl 2):112-115

Hanley JA, McNeil BJ (1982) The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 143:29-36.

Hannan EL, Kilburn H, O'Donnell JF, et al (1990) Adult open heart surgery in New York State. An analysis of risk factors and hospital mortality rates. Journal of the American Medical Association 264:2768-2774

Hannan EL, Beranrd HR, Killburn HC, O'Donnell JF (1992) Gender differences in mortality rates for coronary artery bypass surgery. American Heart Journal 123:866-872

Hasibeder W, Haisjackl M, Germann R, et al (1993) Physiological scoring systems and audit (letter). Lancet 342:307

Headley J, Theriault R, Smith TL (1992) Independent validation of APACHE II severity of illness score for predicting mortality in patients with breast cancer admitted to the intensive care unit. Cancer 70:497-503

Hickling KG, Henderson SJ, Jackson R (1990) Low mortality associated with low volume pressure limited ventilation with permissive hypercapnia in severe adult respiratory distress syndrome. Intensive Care Medicine 16:372-377

Higgins TL, Estafanous FG, Loop FD, et al (1992) Stratification of morbidity and mortality outcome by preoperative risk factors in coronary artery bypass patients. Journal of the American Medical Association 267:2344-2348

Holt AW, Bersten AD, Worthley LI, Vedig AE (1993) Physiological scoring systems and audit (letter). Lancet 342:307-308.

Holt AW, Bury LK, Bersten AD, et al (1992) Prospective evaluation of residents and nurses as severity score data collectors. Critical Care Medicine 20:1688-1691

Hopefl AW, Taaffe CL, Herrmann VM (1989) Failure of APACHE II alone as a predictor of mortality in patients receiving total parenteral nutrition. Critical Care Medicine 17:414-417

Horn SD, Horn RA (1986) The Computerized Severity Index. A new tool for case-mix management. Journal of Medical Systems 10:73-78

Horst HM, Mild LJ, Obeid FN, et al (1987) The relationship of scoring systems and mortality in the surgical intensive care unit. American Surgeon 53:456-459

Hosmer DW, Lemeshow S (1989) Applied logistic regression. New York, John Wiley & Sons

Irwin D, Jessop EG (1993) Severity of sickness at admission to hospital in Colchester 1985 and 1990. Public Health 107:171-175

Istre GR, Gustafson TL, Baron RC, et al (1985) A mysterious cluster of deaths and cardiopulmonary arrests in a pediatric intensive care unit. New England Journal of Medicine 313:205-211

Iyer VS, Russell WJ, Leppard P, Craddock D (1993) Mortality and myocardial infarction after coronary artery surgery. A review of 12003 patients. Medical Journal of Australia 159:166-170

Jacobs P, Noseworthy TW (1990) National estimates of intensive care utilization and costs: Canada and the United States. Critical Care Medicine 18:1282-1286

Jacobs S, Arnold A, Clyburn PA, Willis BA (1992) The Riyadh Intensive Care Program applied to a mortality analysis of a teaching hospital intensive care unit. Anaesthesia 47:775-780

Jacobs S, Chang RWS, Lee B, Lee B (1988) Audit of intensive care: a 30 month experience using the APACHE II severity of disease classification system. Intensive Care Medicine 14:567-574

Jacobs S, Chang RWS, Lee B, Lee B (1989) An analysis of the utilisation of an intensive care unit. Intensive Care Medicine 15:511-518

Jennett B, Teasdale G, Braakman R, et al (1979) Prognosis of patients with severe head injury. Neurosurgery 4:283-289

Johnson MH, Gordon PW, Fitzgerald FT (1986) Stratification of prognosis in granulocytopenic patients with hematologic malignancies using the APACHE II severity of illness score. Critical Care Medicine 14:693-697

Jones DR, Copeland GP, de Cossart L (1992) Comparison of POSSUM with APACHE II for prediction of outcome from a surgical high-dependency unit. British Journal of Surgery 79:1293-1296

Jones EL, Craver JM, King SB, et al (1980) Clinical, anatomic and functional descriptors influencing morbidity, survival and adequacy of revascularization following coronary bypass. Annals of Surgery 192:390-402

Jones EL, Weintraub WS, Craver JM, et al (1991) Coronary artery bypass surgery: is the operation different today? Journal of Thoracic and Cardiovascular Surgery 101:108-115

Joshua N, Jena GP, Mletelwa NV (1989) Multidisciplinary intensive care unit at Cecelia Makiwane Hospital, Mdantsane, Ciskei. South African Medical Journal 75:286-289

Junod FL, Harlan BJ, Payne J, et al (1987) Preoperative risk assessment in cardiac surgery: comparison of predicted and observed results. Annals of Thoracic Surgery 43;59-64

Keene AR, Cullen DJ (1983) Therapeutic intervention scoring system: update 1983 Critical Care Medicine 11:1-3

Kelly W, Parkin G, King RWF (1986) Intensive care unit audit. Prince Henry's Hospital 1983-4. Australian Clinical Review 6:12-16

Kennedy JW, Kaiser GC, Fisher LD, et al (1980) Multivariate discriminant analysis of the clinical and angiographic predictors of operative mortality from the Collaborative Study in Coronary Artery Surgery (CASS). Journal of Thoracic and Cardiovascular Surgery 80:876-887

Kennedy JW, Kaiser GC, Fisher LD, et al (1981) Clinical and angiographic predictors of operative mortality from the Collaborative Study in Coronary Artery Surgery (CASS). Circulation 63:793-802

Khan SS, Nessim S, Gray R, et al (1990) Increased mortality of women in coronary artery bypass surgery: evidence for referral bias. Annals of Internal Medicine 112:561-567

King KB, Clark PC, Norsen LH, Hicks GL (1992) Coronary artery bypass graft surgery in older women and men. American Journal of Critical Care 1:28-35

Kirklin JW, Akins CW, Blackstone EH, et al (1991) ACC/AHA guidelines and indications for coronary artery bypass graft surgery. Circulation 83:1125-1173

Knaus WA, Zimmerman JE, Wagner DP, et al (1981) APACHE - Acute physiology and chronic health evaluation: a physiologically based classification system. Critical Care Medicine 9:591-597

Knaus WA, Draper EA, Wagner DP, Zimmerman JE (1985a) APACHE II: A severity of disease classification system. Critical Care Medicine 13:818-829

Knaus WA, Draper EA, Wagner DP, Zimmerman JE (1985b) Prognosis in acute organ-system failure. Annals of Surgery 202:685-693

Knaus WA, Draper EA, Wagner DP, Zimmerman JE (1986) An evaluation of outcome from intensive care in major medical centers. Annals of Internal Medicine 104:410-418

Knaus WA (1989) Prognosis with mechanical ventilation: the influence of disease, severity of disease, age, and chronic health status on survival from an acute illness. American Review of Respiratory Disease 140:S8-13

Knaus WA, Rauss A, Alperovitch A, et al (1990) Do objective estimates of chances for survival influence decisions to withhold or withdraw treatment? Medical Decision Making 10:163-171

Knaus WA, Wagner DP, Draper EA, et al (1991a) The APACHE III prognostic system. Chest 100:1619-1636

Knaus WA, Wagner DP, Lynn J (1991b) Short-term mortality predictions for critically ill hospitalized adults: science and ethics. Science 254:389-394

Knaus WA, Sun X, Nystrom P, Wagner DP (1992) Evaluation of definitions for sepsis. Chest 101:1656-1662

Knaus WA, Harrell FE, Fisher CJ, et al (1993a) The clinical evaluation of new drugs for sepsis. A prospective study based on survival analysis. Journal of the American Medical Association 270:1233-1241

Knaus WA, Wagner DP, Zimmerman JE, Draper EA (1993b) Variations in mortality and length of stay in intensive care units. Annals of Internal Medicine 118:753-761

Koch CG, Estafanous FG (1993) Anesthesia for coronary artery surgery. Current Opinion in Cardiology 8:897-909

Kollef MH, Enzenauer RJ (1992) Predicting outcome from intensive care for patients with rheumatologic diseases. Journal of Rheumatology 19:1260-1262

Kollef MH, Schuster DP (1994) Predicting intensive care unit outcome with scoring systems. Underlying concepts and principles. Critical Care Clinics 10:1-18

Kraus PA, Lipman J, Lee CCJ, et al (1993) Acute lung injury at Baragwanath ICU. An eight-month audit and call for consensus for other organ failure in the Adult Respiratory Distress Syndrome. Chest 103:1832-1836

Kreuzer E, Kaab S, Pilz G, Werdan K (1992) Early prediction of septic complications after cardiac surgery by APACHE II score. European Journal of Cardiothoracic Surgery 6:524-528

Kruse JA, Carlson RW (1993) Severity of illness scoring: East meets West (Editorial). Critical Care Medicine 21:647-648

Kruse JA, Thill-Baharozian MC, Carlson RW (1988) Comparison of clinical assessment with APACHE II for predicting mortality risk in patients admitted to a medical intensive care unit. Journal of The American Medical Association 260:1739-1742

Lanken PN (1994) Critical care at a new crossroads: the intersection of economics and ethics in the intensive care unit. American Journal of Respiratory and Critical Care Medicine 149:3-5

Larvin M, McMahon MJ (1989) APACHE II score for assessment and monitoring of acute pancreatitis. Lancet ii:201-204

Lee KH, Hui KP, Lim TK, Tan WC (1993) Acute physiology and chronic health evaluation (APACHE II) scoring in the Medical Intensive Care Unit, National University Hospital, Singapore. Singapore Medical Journal 34:41-44

Le Gall JR, Loirat P, Alperovitch A (1984) A simplified acute physiology score for ICU patients. Critical Care Medicine 12:975-977

Le Gall JR (1991) Do we need a new severity score? Critical Care Medicine 19:857-858

Le Gall JR, Lemeshow S, Saulnier F (1993) A new Simplified Acute Physiology Score (SAPS II) based on a European/North American multicenter study. Journal of the American Medical Association 270:2957-2963

Lehmkuhl P, Jeck-Thole S, Pichlmayr I (1989) A new scoring system for disease intensity in a surgical intensive care unit. World Journal of Surgery 13:252-258

Lemeshow S, Teres D, Pastides H, et al (1985) A method for predicting survival and mortality of ICU patients using objectively derived weights. Critical Care Medicine 13:519-525

Lemeshow S, Teres D, Avrunin JS, Pastides H (1987) A comparison of methods to predict mortality of intensive care unit patients. Critical Care Medicine 15:715-722

Lemeshow S, Teres D, Avrunin JS, Gage RW (1988) Refining intensive care unit outcome prediction by using changing probabilities of mortality. Critical Care Medicine 16:470-477

Lemeshow S, Teres D, Klar J, et al (1993) Mortality Probability Models (MPM II) based on an international cohort of intensive care unit patients. Journal of the American Medical Association 270:2478-2486

Levy DE, Bates D, Caronna JJ, et al (1981) Prognosis in nontraumatic coma. Annals of Internal Medicine 94:293-301

Levy DE, Caronna JJ, Singer BH, et al (1985) Predicting outcome from hypoxic-ischemic coma. Journal of the American Medical Association 253:1420-1426

Lloyd-Thomas AR, Wright I, Lister TA, Hinds CJ (1988) Prognosis of patients receiving intensive care for lifethreatening medical complications of haematological malignancy. British Medical Journal 296:1025-1029.

Lockrem JD, Lopez E, Gallagher J, et al (1991) Severity of illness: APACHE II analysis of an ICU population. Cleveland Clinic Journal of Medicine 58:477-486

Loop FD, Berrettoni JN, Pichard A, et al (1975) Selection of the candidate for myocardial revascularization. A profile of high risk based on multivariate analysis. Journal of Thoracic and Cardiovascular Surgery 69:40-51

Loop FD, Golding LR, MacMillan JP, et al (1983) Coronary artery surgery in women compared with men: analyses of risks and long-term results. Journal of the American College of Cardiologists 1:383-390

Luce JM, Raffin TA (1988) Withholding and withdrawal of life support in critically ill patients. Chest 94:621-626

Luce JM, Wachter RM (1994) The ethical appropriateness of using prognostic scoring systems in clinical management. Critical Care Clinics 10:229-241

Maher ER, Robinson KN, Scoble JE, et al (1989) Prognosis of critically-ill patients with acute renal failure: APACHE II score and other predictive factors. Quarterly Journal of Medicine 72:857-866

Mancebo J, Artigas A (1987) A clinical study of the adult respiratory distress syndrome. Critical Care Medicine 15:243-246

Manske CL, Wang Y, Rector T, et al (1992) Coronary revascularisation in insulin-dependent diabetic patients with chronic renal failure. Lancet 340:998-1002

McCormick JR, Schick EC, McCabe CH, et al (1985) Determinants of operative mortality and long-term survival in patients with unstable angina. Journal of Thoracic and Cardiovascular Surgery 89:683-688

Marks RJ, Simons RS, Blizzard RA, Brown DRG (1991) Predicting outcome in intensive therapy units-a comparison of APACHE II with subjective assessments. Intensive Care Medicine 17:159-163

Marsh HM, Krishan I, Naessens JM, et al (1990) Assessment of prediction of mortality by using the APACHE II scoring system in intensive-care units. Mayo Clinic Proceedings 65:1549-1557

McAnena OJ, Moore FA, Moore EE, et al (1992) Invalidation of the APACHE II scoring for patients with acute trauma. Journal of Trauma 33:504-507

McCabe WR, Jackson GG (1962) Gram-negative bacteremia: etiology and ecology. Archives of Internal Medicine 110:847-855

McClish DK, Powell SH (1989) How well can physicians estimate mortality in a medical intensive care unit? Medical Decision Making 9:125-132

Meehl PE (1954) Clinical versus statistical prediction. University of Minnesota Press, Minneapolis

Merrill WH, Elkins CC, Stewart JR, et al (1993) Third-time coronary artery bypass grafting: midterm results. Annals of Thoracic Surgery 55:582-585

Merry AF, Ramage MC, Whitlock RM, et al (1992) First-time coronary artery bypass grafting: the anaesthetist as a risk factor. British Journal of Anaesthesia 68:6-12

Meyer AA, Messick WJ, Young P, et al (1992) Prospective comparison of clinical judgement and APACHE II score in predicting the outcome in critically ill surgical patients. Journal of Trauma 32:747-753

Mohan R , Amsel BJ, Walter PJ (1992) Coronary artery bypass grafting in the elderly - a review of studies on patients older than 64, 69, and 74 years. Cardiology 80:215-225

Moreau R, Soupison T, Vauquelin P, et al (1989) Comparison of two simplified severity scores for patients with acute myocardial infarction. Critical Care Medicine 17:409-413

Murray LS, Teasdale GM, Murray GD, et al (1993) Does prediction of outcome alter patient management? Lancet 341:1487-1491

Nashef SAM, Carey F, Silcock MM, et al (1992) Risk stratification for open-heart surgery: trial of the Parsonnet system in a British hospital. British Medical Journal 305:1066-1067

Naunheim KS, Fiore AC, Wadley JJ, et al (1988) The changing profile of the patient undergoing coronary artery bypass surgery. Journal of the American College of Cardiology 11:494-498

Naunheim KS, Barner HB, Fiore AC (1992) Update to "Results of internal thoracic artery grafting over 15 years: single versus double grafts". Annals of Thoracic Surgery 53:716-718

NIH Workshop summary (1986) Withholding and withdrawing mechanical ventilation. American Review of Respiratory Disease 134:1327-1330

Niskanen M, Kari A, Nikki P, et al (1991) Acute physiology and chronic health evaluation (APACHE II) and Glasgow Coma Scores as predictors of outcome from intensive care after cardiac arrest. Critical Care Medicine 19:1465-1473

O'Connor GT, Plume SK, Olmstead EM, et al (1991) A regional prospective study of in-hospital mortality associated with coronary artery bypass grafting. Journal of the American Medical Association 266:803-809

O'Connor GT, Plume SK, Olmstead EM, et al (1992) Multivariate prediction of in-hospital mortality associated with coronary artery bypass graft surgery. Circulation 85:2110-2118

Oh TE, Hutchinson R, Short S, et al (1993) Verification of the Acute Physiology and Chronic Health Evaluation scoring system in a Hong Kong intensive care unit. Critical Care Medicine 1993;21:698-705

Oldham HN, Kong Y, Bartel AG, et al (1972) Risk factors in coronary artery bypass surgery. Archives of Surgery 105:918-923

Osaka S, Barratt-Boyes BG, Brandt PWT, et al (1988) Early and late results of reoperation for coronary artery disease: a 13 year experience. Australian and New Zealand Journal of Surgery 58:537-541

Osborne ML (1992) Physician decisions regarding life support in the intensive care unit. Chest 101:217-224

Osborne ML, Evans TW (1994) Allocation of resources in intensive care: a transatlantic perspective. Lancet 343:778-780

Palazzo M, Patel M (1993) The use and interpretation of scoring systems in ICU. Part 2. British Journal of Intensive Care 7:286-289

Palazzo M, Soni N, Hinds C (1993) Physiological scoring systems and audit (letter). Lancet 342:307

Pallis C (1983) ABC of brain stem death. London, British Medical Journal Press

Parno JR, Teres D, Lemeshow S, et al (1984) Two-year outcome of adult intensive care patients. Medical Care 22:167-176

Parsonnet V, Fisch D, Gielchinsky I, et al (1988) Emergency operation after failed angioplasty. Journal of Thoracic and Cardiovascular Surgery 96:198-203

Parsonnet V, Dean D, Bernstein AD (1989) A method of uniform stratification of risk for evaluating the results of surgery in acquired adult heart disease. Circulation 79(Suppl I):3-12

Pelletier C, Cosette R, Dontigny L, et al (1980) Determinants of mortality following coronary bypass surgery. Canadian Journal of Surgery 23:199-204

Perkins HS, Jonsen AR, Epstein WV (1986) Providers as predictors: Using outcome predictions in intensive care. Critical Care Medicine 14:105-110

Pierpont GL, Kruse M, Ewald S, Weir EK (1985) Practical problems in assessing risk for coronary artery bypass grafting. Journal of Thoracic and Cardiovascular Surgery 89:673-682

Pollack MM, Ruttimann UE, Getson PR (1988) Pediatric risk of mortality (PRISM) score. Critical Care Medicine 16:1110-1116

Poses RM, Bekes C, Copare FJ, Scott WE (1989) The answer to "What are my chances, Doctor?" depends on whom is asked: prognostic disagreement and inaccuracy for critically ill patients. Critical Care Medicine 17:827-833

Poses RM, McClish DM, Bekes C, et al (1991) Ego bias, reverse ego bias, and physicians' prognostic. Critical Care Medicine 19:1533-1539

Prasad US, Walker WS, Sang CT, et al (1991) Influence of obesity on the early and long term results of surgery for coronary artery disease. European Journal of Cardiothoracic Surgery 5:67-72

Rafkin HS, Hoyt JW (1994) Objective data and quality assurance programs. Current and future trends. Critical Care Clinics 10:157-177

Rahimtoola SH, Fessler CL, Grunkemeier GL, Starr A (1993) Survival 15 to 20 years after coronary bypass surgery for angina. Journal of the American College of Cardiology 21:151-157

Ranson JHC, Rifkind KM, Roses DF, et al (1974) Prognostic signs and the role of operative management in acute pancreatitis. Surgery, Gynecology and Obstetrics 139:69-81

Rapoport J, Teres D, Lemeshow S, Harris D (1990a) Timing of intensive care unit admission in relation to ICU outcome. Critical Care Medicine 18:1231-1235

Rapoport J, Teres D, Lemeshow S, et al (1990b) Explaining the variability of cost using a severity-of-illness measure for ICU patients. Medical Care 28:338-348

Relman AS (1988) Assessment and accountability: the third revolution in medical care. New England Journal of Medicine 319:1220-1222

Rhee KJ, Fisher CJ, Willitis NH (1987) The Rapid Acute Physiology Score. American Journal of Emergency Medicine 5:278-282

Rhee KJ, Baxt WG, Mackenzie JR, et al (1990) APACHE II scoring in the injured patient. Critical Care Medicine 18:827-830

Ridley S, Biggam M, Stone P (1993) A cost-benefit analysis of intensive therapy. Anaesthesia 48:14-19

Rocca B, Martin C, Viviand X, et al (1989) Comparison of four severity scores in patients with head trauma. Journal of Trauma 29:299-305

Rockwood K, Noseworthy TW, Gibney RTL, et al (1993) One-year outcome of elderly and young patients admitted to intensive care units. Critical Care Medicine 21:687-691

Roi LD, Flora JD, Davis TM, Wolfe RA (1983) Two new burn severity indices. Journal of Trauma 23:1023-1029

Rosenfeldt FL, Wong J (1993) Current expectations for survival and complications in coronary artery bypass grafting. Current Opinion in Cardiology 8:910-918

Rothwell PM (1993) Physiological scoring systems and audit (letter). Lancet 342:306

Roumen RMH, Schers TJ, de Boer HHM, Goris RJA (1992) Scoring systems for predicting outcome in acute haemorrhagic necrotising pancreatitis. European Journal of Surgery 158:167-171

Rowan KM, Kerr JH, Major E, et al (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. British Medical Journal 307:977-981

Ruark JE, Raffin TA (1988) Initiating and withdrawing life support. Principles and practice in adult medicine. New England Journal of Medicine 318:25-30

Rutledge R, Fakhry SM, Rutherford EJ, et al (1991) 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 1,238 patients. Critical Care Medicine 19:1048-1053

Sage WM, Rosenthal MH, Silverman JF (1986) Is intensive acre worth it?-An assessment of input and outcome for the critically ill. Critical Care Medicine 14:777-782

Salomon NW, Page US, Bigelow JC, et al (1990) Reoperative coronary surgery. Journal of Thoracic and Cardiovascular Surgery 100:250-260

Sarmiento J, Torres A, Guardiola J, et al (1991) Statistical modeling of prognostic indices for evaluation of critically ill patients. Critical Care Medicine 19:867-870

Schafer JH, Maurer A, Jochimsen F, et al (1990) Outcome prediction models on admission in a medical intensive care unit: do they predict individual outcome? Critical Care Medicine 18:1111-1117

Schafer JH, Jochimsen F, Keller F, et al (1991) Outcome prediction of acute renal failure in medical intensive care. Intensive Care Medicine 17:19-24

Schein M, Gecelter G (1989) APACHE II score in massive upper gastrointestinal haemorrhage from peptic ulcer: prognostic value and potential clinical applications. British Journal of Surgery 76:733-736

Selker HP (1993) Systems for comparing actual and predicted mortality rates: characteristics to promote cooperation in improving hospital care (Editorial). Annals of Internal Medicine 118:820-822

Shapiro AR (1977) The evaluation of clinical predictions. A method and initial application. New England Journal of Medicine 296:1509-1514

Shoemaker WC (1989) Methodologic assessment of outcome. Critical Care Medicine 17:S169

Sinclair JR, Magee PT, Gould TH, Collins CH (1991) A comparison of APACHE II and a clinical sickness score. A study of 97 consecutive admissions to a District General Hospital Intensive Care Unit. Anaesthesia 46:442-446

Slatyer MA, James OF, Moore PG, Leeder SR (1986) Costs, severity of illness and outcome in intensive care. Anaesthesia and Intensive Care 14:381-389

Sleigh JW, Labusch D, Loose R, Chinyanga HM (1989) Abdominal sepsis: The identification of a high risk subgroup using the APACHE II system in twenty-seven patients. Central African Journal of Medicine 35:410-413

Sleigh JW, Brook RJ, Miller M (1992) Time-dependent error in the APACHE II scoring system. Anaesthesia and Intensive Care 20:63-65

Smedira NG, Evans BH, Grais LS, et al (1990) Withholding and withdrawal of life support from the critically ill. New England Journal of Medicine 322:309-315

Snyder JV, McGuirk M, Grenvik A, Stickler D (1981) Outcome of intensive care. An application of a predictive model. Critical Care Medicine 9:598-603

Spiegelhalter DJ (1992) Risk stratification for open heart surgery (letter). British Medical Journal 305:1500

Spitzer WO, Dobson AJ, Hall J, et al (1981) Measuring the quality of life of cancer patients: a concise QL-index for use by physicians. Journal of Chronic Diseases 34:585-597

Stern M, Waisbren BA (1978) Comparison of methods of predicting burn mortality. Burns 6:119-123

Suma H, Wanibuci Y, Terada Y et al (1993) The right gastroepiploic artery graft: clinical and angiographic midterm results in 200 patients. Journal of Thoracic and Cardiovascular Surgery 105:615-623

Teasdale G, Jennett B (1974) Assessment of coma and impaired consciousness. A practical scale. Lancet ii:81-84

Teres D, Lemeshow S, Avrunin JS, Pastides H (1987) Validation of the mortality prediction model for ICU patients. Critical Care Medicine 15:208-213

Teres D (1989) Peer review, publication policy, and APACHE. Critical Care Medicine 17:S169-S172

Teres D (1993) Trends from the United States with end of life decisions in the intensive care unit. Intensive Care Medicine 19:316-322

Teskey RJ, Calvin JE, McPhail I (1991) Disease severity in the coronary care unit. Chest 100:1637-1642

The International Neonatal Network (1993) The CRIB (clinical risk index for babies) score: a tool for assessing initial neonatal risk and comparing performance of neonatal intensive care units. Lancet 342:193-198.

Thibault GE, Mulley AG, Barnett GO, et al (1980) Medical intensive care: indications, interventions, and outcomes. New England Journal of Medicine 302:938-942

Tobin JN, Wassertheil-Smoller S, Wexler JP, et al (1987) Sex bias in considering coronary bypass surgery. Annals of Internal Medicine 107:19-25

Tuman KJ, McCarthy RJ, March RJ, et al (1992) Morbidity and duration of ICU stay after cardiac surgery. A model for preoperative risk assessment. Chest 102:36-44

Turner JS, Potgieter PD, Linton DM (1988) Systems for scoring severity of illness in intensive care. South African Medical Journal 76:17-20

Turner JS, Mudaliar YM, Chang RWS, Morgan CJ (1991) Acute Physiology and Chronic Health Evaluation (APACHE II) scoring in a cardiothoracic intensive care unit. Critical Care Medicine 19:1266-1269

Van Lanschot JJB, Feenstra BWA, Vermeij CG, Bruining HA (1988) Outcome prediction in critically ill patients by means of oxygen consumption index and simplified acute physiology score. Intensive Care Medicine 14:44-49

Vassar MJ, Wilkerson CL, Duran PJ, et al (1992) Comparison of APACHE II, TRISS, and a proposed 24-hour ICU point system for prediction of outcome in ICU trauma patients. Journal of Trauma 32:490-499

Verheul HA, Moulijn AC, Hondema S, et al (1991) Late results of 200 repeat coronary artery bypass operations. American Journal of Cardiology 67:24-30

Viviand X, Gouvernet J, Granthil C, Francois G (1991) Simplification of the SAPS by selecting independent variables. Intensive Care Medicine 17:164-168

Wagner DP, Knaus WA, Draper EA (1983) Statistical validation of a severity of illness measure. American Journal of Public Health 73:878-884

Wagner DP, Knaus WA, Draper EA (1987) Identification of low-risk monitor admissions to medical- surgical ICUs. Chest 92:423-428

Wasson JH, Sox HC, Neff RK, Goldman L (1985) Clinical Prediction rules. Applications and methodological standards. New England Journal of Medicine 313:793-799

Waters M, Nightingale P, Edwards JD (1990) A critical study of the APACHE II scoring system using earlier data collection. Archives of Emergency Medicine 7:16-20

Watters DAK, Wilson IH, Sinclair JR, Ngandu N (1989) A clinical sickness score for the critically ill in Central Africa. Intensive Care Medicine 15:467-470

Weinstein ND (1989) Optimistic biases about personal risks. Science 246:1232-1233

Weintraub WS, Craver JM, Cohen CL, et al (1991) Influence of age on results of coronary artery surgery. Circulation 84:(Suppl III)226-235

Wilson C, Heath DI, Imrie CW (1990) Prediction of outcome in acute pancreatitis: a comparative study of APACHE II, clinical assessment and multiple factor scoring systems. British Journal of Surgery 77:1260-1264

Wisner DH (1992) History and current status of scoring systems for critical care. Archives of Surgery 127:352-356

Wong DT, Knaus WA (1991) Predicting outcome in critical care: the current status of the APACHE prognostic scoring system. Canadian Journal of Anaesthesia 38:374-383

Yeung H, Lu M, Martinez EG, Puri VK (1990) Critical Care Scoring System - New concept based on hemodynamic data. Critical Care Medicine 18:1347-1352

Zimmerman JE, Knaus WA, Sharpe SM, et al (1986) The use and implications of do not resuscitate orders in intensive care units. Journal of the American Medical Association 255:351-356

Zimmerman JE, Knaus WA, Judson JA, et al (1988) Patient selection for intensive care: A comparison of New Zealand and United States hospitals. Critical Care Medicine 16:318-326

Zimmerman JE, Shortell SM, Knaus WA, et al (1993a) Value and cost of teaching hospitals: a prospective multicenter, inception cohort study. Critical Care Medicine 21:1432-1442

Zimmerman JE, Shortell SM, Rousseau DM, et al (1993b) Improving intensive care: observations based on organizational case studies in nine intensive care units: a prospective, multicenter study. Critical Care Medicine 21:1443-1451

Zoch G, Schemper M, Kyral E, Meissl G (1992) Comparison of prognostic indices for burns and assessment of their accuracy. Burns 18:109-112

APPENDIX 1

Submission for ethical approval

SUBMISSION TO THE ETHICS COMMITTEE

OF

THE ROYAL BROMPTON NATIONAL HEART AND LUNG HOSPITAL

1. Project Title

OUTCOME PREDICTION IN CARDIOTHORACIC SURGERY.

2. Principle Investigator(s)

John Turner MMed (Cape Town) FCP (SA)

Clifford Morgan FRCA

3. Consultant(s) accepting medical responsibility

Clifford Morgan FRCA

4. Purpose of the study

To develop a prognostic index for cardiothoracic surgery patients which may ultimately be used to help make decisions to withdraw therapy in hopeless cases.

5. Type of study

(new , sequel , modification , local , multicentre)

New.

6. Study design

(i) observational

(ii) statistical and give numbers

(iii) patients

(iv) volunteers

(v) donors, stating whether they are minors (m) or adult (a), in each category

Observational, prospective.

7. Sources

(i) where you will recruit them from

(ii) what are you inclusion criteria

(iii) exclusion criteria (eg age, sex, pregnancy, risk of pregnancy during trial, breast feeding, current disease, previous disease, occupation, ability to comprehend, ability to comply, travel problems, loss of income)

Adult patients

(i) Records of in-patients admitted to the AICU of the Royal Brompton National Heart and Lung Hospital.

(ii) All patients undergoing cardiac surgery (CABG, valve replacements/repair, correction of congenital heart disorder).

(iii) None.

8. Design justification

Introduction

Responsible use of intensive care unit (ICU) facilities and cost containment are becoming more and more important and have even become political issues. We have already shown that although mortality for cardiothoracic surgery patients is relatively low (4.6%), non-survivors stay in the ICU for a long time (mean of 9.8 days vs 2 days for survivors) and use a large amount of resources (1). Current predictors of mortality in cardiothoracic surgery patients are inadequate for making decisions to withdraw therapy in the individual patient (1).

The APACHE II scoring system (2) performed in the first 24 hours of ICU admission cannot reliably estimate prognosis for the individual patient (and is not designed to do so). The Organ-system failure score (3), the APACHE II trend analysis (4) and the mortality prediction model (5) have all been relatively successful in this respect, but at a later stage in the patient's admission. The APACHE III score may be the most successful model yet, but due to the difficulties in predicting mortality in cardiothoracic surgical patients, it excludes them (6, 7), as does the mortality prediction model (5). More and more cardiac surgery is being performed on higher-risk patients, many of whom will do poorly. There is thus a need to develop a better prognostic system for these patients.

Materials and method

This prospective study will be carried out in the Adult ICU at the Royal Brompton National Heart and Lung Hospital, London. The need for informed consent to enter the study will be requested to be waived by the Ethics Committee of the hospital, as this is purely observational data collection. All patients undergoing cardiac surgery will be consecutively entered into the study. Demographic data, medical histories, physiological data (starting at the induction of general anaesthesia), and laboratory investigations will be collected from clinical records, the CareVue 9000 system, and the McDonnell Douglas network and will be entered (or downloaded) into a microcomputer database. The Parsonnet Score (8) will be used to assess preoperative status. Patients will be followed up to hospital discharge to determine survival.

Analysis

Data will be simultaneously analysed in a variety of ways. Firstly tests of association of each of the data variables with mortality will be carried out. Then discriminant function analysis and multiple logistic regression analysis will be used to find the best predictive variables. The systems developed from these analyses will be further refined as the database expands. The area under receiver-operating characteristic curves will be used to determine the accuracy of prediction of each analytical method.

References

1. Turner JS, Mudaliar YM, Chang RWS, Morgan CJ. Acute Physiology and Chronic Health Evaluation (APACHE II) scoring in a cardiothoracic intensive care unit. Crit Care Med 1991;19:1266-1269

2. Knaus WA, Draper EA, Wagner DP, Zimmerman JE. APACHE II - a severity of disease classification system. Crit Care Med 1985;13:818

3. Knaus WA, Draper EA, Wagner DP, et al. Prognosis in acute organ-system failure. Ann Surg 1985;202:685

4. Chang RWS, Jacobs S, Lee B, Pace N. Predicting deaths among intensive care unit patients. Crit Care Med 1988;16:34

5. Lemeshow S, Teres D, Avrunin JS, Gage RW. Refining intensive care unit outcome prediction by using changing probabilities of mortality. Crit Care Med 1988;16:470

6. Draper EA, Wagner DP, Russo M, et al. APACHE III. Study design - data collection. Crit Care Med 1989;17:S186

7. Knaus WA, Wagner DP,Draper EA, et al. The APACHE III prognostic system. Chest 1991;100:1619-1636

8. Parsonnet V, Dean D, Bernstein AD. A method of uniform stratification of risk for evaluating the results of surgery in acquired adult heart disease. Circulation 1989;79:I 3-12

9. Resources

Is any commercial organisation facilitating this study by the provision of:

(i)	staff	(specify)
(ii)	drugs	"
(iii)	equipment	"
(iv)	finance of any kind	"

No to all the above.

10. Pharmacology

(i) the substance

(ii) route of administration

(iii) frequency

(iv) dosage

(v) stage of CSM evaluation

enclose details of any CSM certificates

Not applicable.

12. Dispensing

Not applicable.

13. Risks

What else will be done to your subjects/patients that is extra to normal clinical management? Please comment on the risks and precautions:

(i) venous samples
(ii) arterial samples
(iii) cardiovascular catheters / indwelling lines
(iv) other invasive procedures (eg bronchoscopy)
(v) imaging
(vi) biopsies (sites, method, size, number)
(vii) anaesthetics
(viii) involvement in other concurrent trials
(ix) denial or withholding of conventional treatment
(x) temp/permanent withdrawal of current therapy

None of the above.

14. Indemnities

(i) agreement on indemnity
(ii) agreement on statistical analysis
(iii) agreement on publication

Not applicable.

15. Ethical problems

(i)	your control procedures on normal subjects
(ii)	your test procedures in normal subjects
(iii)	your control procedures in patients
(iv)	your test procedures in patients
(v)	What are the potential benefits to your patients?
(vi)	What are the potential benefits to others?

Not applicable

16. Please enclose a copy of your consent form for healthy volunteers and/or patients

Not applicable.

17. Informing the subjects and patients' general practitioners

Not applicable.

18. Untoward events

Not applicable.

19. Case notes

Not applicable.

20. Date of application

18/01/1993

21. Starting date of study

January 1993

22. Sponsors signature

Not applicable.