Intensive Care Med (1988) 14:558–566

Intensive Care Medicine

Predicting outcome among intensive care unit patients using computerised trend analysis of daily Apache II scores corrected for organ system failure

R. W. S. Chang[1], S. Jacobs[2] and B. Lee[1]

[1] Department of Surgery, and [2] Department of Anaesthesia, Riyadh Armed Forces Hospital, Riyadh, Saudi Arabia

Received: 15 September 1987; accepted: 15 February 1988

Abstract. Daily Apache II scores were determined prospectively on 310 consecutive adult Intensive Care Unit (ICU) patients to reflect the dynamic patho-physiological processes affecting ICU patients. Organ failure scores were derived from the Apache II scores by applying a coefficient which corresponded to the number and duration of organ failures to account for the increased mortality associated with established major organ-system failures. Computerised trend analysis of data from the first 100 patients was used to develop criteria for predicting hospital outcome of the ICU patients. The analysis took into account the absolute value of the daily scores and the rate of change relative to that of the previous day. Allowance was made for changes in scores as a result of surgery or major iatrogenic complications arising after admission to the ICU. The criteria were then tested prospectively on the next 210 consecutive ICU patients. Predictions by Apache II assessments were NOT used to influence clinical decisions during the study period. It was possible to predict with a specificity of 100% 37 out of the 72 deaths in the test group. The predictive power of daily organ failure scores was superior to those obtained from a single Apache II assessment or from daily Apache II scores by a factor of 5.3 and 1.4 respectively. The estimated long-term risk of a false prediction is 1.4% at the 95% confidence level.

Key words: Intensive Care – Prognosis – Predictions

The Apache II system [1] was developed originally for quality assurance purposes and was based on a single Apache II assessment on the day of admission to the Intensive Care Unit (ICU). We reported the first ever use of the Apache II system to predict individual outcome among ICU patients as an aid to clinical decision making [2]. However there are theoretical and practical reasons why the original Apache II system is inadequate for making individual outcome predictions.

Firstly, the patho-physiological processes affecting ICU patients are dynamic and cannot be reflected by a single assessment on the day of admission to the ICU. Secondly, although the Apache II score, with the exception of neurological points, is based on objective data, derivation of the risk of death is based on a subjective choice of a single specific diagnostic category or major organ system as the primary cause of admission to the ICU. The correct choice can sometimes be extremely difficult to make, especially among patients with multi-organ systems failure and high mortality, precisely that group of patients in whom a correct prediction would be useful as an aid to decision-making. An incorrect choice can lead to a wrong estimation of the risk of death and therefore a wrong prediction. Thirdly, the Apache II risk of death is a probability, obtained by applying coefficients to the Apache II score. The coefficients were derived from data obtained from the first day of admission. The Apache II risk of death may therefore not be valid for analysis over time. Fourthly, it is common for ICU patients to develop other major organ system insufficiency or failure during their stay in the ICU with important prognostic implications. A single Apache II assessment and the choice of a single specific diagnostic category or major organ system category does not reflect this aspect of the progress of an ICU patient. Lastly, it would be quite unacceptable to clinicians, patients and relatives to base major clinical decisions on just one assessment. In our original

study, we tried to reflect the dynamic patho-physiological changes in a limited way by using two Apache II assessments. Although we were able to achieve a specificity of 100% in that study, we were unhappy, as too great an element of chance persisted. Out next model was based on trend analysis of daily Apache II scores, but did not take into account the effect of the incidence and duration of major organ system failure on mortality [3]. Despite these limitations, both models had far better predictive power than a single Apache II assessment.

This study reports our use of daily organ failure scores (Apache II scores corrected for major organ system failure) to predict individual hospital outcome among ICU patients.

Table 1. Criteria for defining the presence of acute major organ system failure. Adapted from Knaus [4] and Garden [7]. If the patient fulfilled one or more of the following criteria for each of the organ systems during a 24-h period (regardless of other values), acute organ system failure existed on that day

Cardiovascular failure:
Heart rate $< = 54$/min
Mean arteria blood pressure $< = 49$ mmHg
Occurrence of ventricular tachycardia and/or ventricular fibrillation
Serum pH $< = 7.24$ with $PaCO_2$ of $< = 49$ mmHg
Respiratory failure:
Respiratory rate $< = 5$/min or $> = 49$/min
$PaCO_2 > = 50$ mmHg
$AaDO_2 > = 350$ mmHg
($AaDO_2 = 713\ FIO_2 - PaCO_2 - PaO_2$)
Dependent on ventilator on the 4th day of organ system failure e.g. NOT applicable for the initial 72 h
Renal failure[a]:
Urine output $< = 479$ ml/24 h or $< = 159$ ml/8 h and
(Serum urea $> = 214$ mg/100 ml or
Serum creatinine $> = 3.5$ mg/100 ml)
Haematological failure:
WBC $< = 1000\ mm^3$
Platelets $< = 20000\ mm^3$
Haematocrit $< = 20\%$ (not chronic renal failure)
Neurologic failure:
Glasgow coma score $< = 6$ (in absence of sedation at any one point in day)
Liver failure[b]:
Clinical acute liver failure
AND
$P < 0.66$
where $L_n\ (P/1-P) =$
$10 - (4.3$ Prothrombin ratio)
$-(0.03 \times \text{creat} \times 88.4) - (0.85 \times \text{ENC})$
ENC $= +1$ in presence of encephalopathy
ENC $= -1$ in absence of encephalopathy
Creatinine in mg% (88.4 is the conversion factor to S.I. Units)

[a] Excluding patients on chronic dialysis before hospital admission;
[b] Equation from Garden et al. (1985) Br J Surg 72:91

Patients and methods

The Apache II score is based on deviations from normal of 11 acute physiological variables (rectal temperature, mean blood pressure, heart rate, respiratory rate, oxygenation, arterial pH, serum sodium, potassium, creatinine, haematocrit and white blood count), the Glasgow coma score, the age of the patient and the presence of defined chronic disease of the cardiovascular, respiratory, hepatic, renal and immunological systems.

Correction for organ system failures

We have modified the Apache II score by making a correction for the presence and duration of established major organ systems failure. The presence of major organ system failure is based on criteria developed by Knaus and his colleagues [4] (Table 1). This correction is obtained by the formula:

Organ failure score =

Apache II score $\times$ (1 + organ failure coefficient).

The organ failure coefficients based again on data obtained by Knaus from over 5500 patients [4] are given in Table 2. The rationale for the use of organ failure coefficients is that the mortality of patients is increased by the number and duration of major organ system failures. By applying a correction to the Apache II score, patients with major organ system failures can be differentiated from those without. As the effect of applying an organ failure correction is greatest among patients with high Apache II scores, this improved the identification of those patients who are likely to die (Fig. 1).

Daily Apache II and organ system failure assessment were performed prospectively on 310 consecutive adult ICU patients on the day of admission until their discharge between May 1986 to April 1987. Eighteen

Table 2. Organ failure coefficients used in calculating organ failure score from Apache score. The values were obtained by dividing by 1000, the mortality rates obtained by Knaus [4]

Day of failure	Number of organ system failures		
	1 organ	2 organs	>3 organs
1	0.022	0.052	0.080
2	0.031	0.067	0.095
3	0.034	0.066	0.093
4	0.035	0.062	0.096
5	0.040	0.056	0.100
6	0.042	0.064	0.100
7	0.041	0.068	0.100

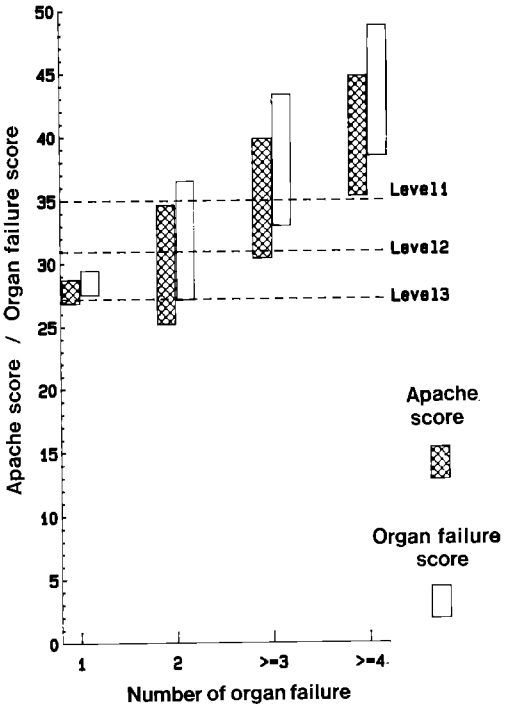

Fig. 1. Apache II score and organ failure score on day 1 of organ failure for non-survivors predicted to die for varying number of organ failures, to illustrate how organ failure score improved the sensitivity of the present predictive model. Each box represents the mean±1 standard deviation. All organ failure score boxes were above the various cut-off levels in comparison with Apache II score boxes

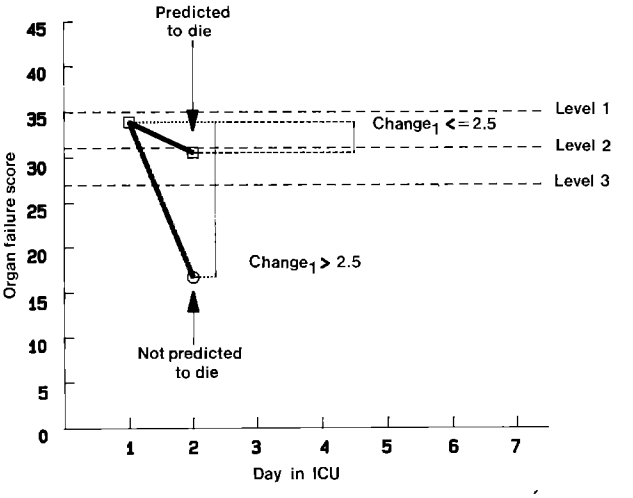

Fig. 2. Rate of change of organ failure scores on day 2 relative to that on day 1 of survivors and non-survivors with day 1 score in the fuzzy zone

hours was the minimum period used in the collection of data for a day 1 assessment. Data from shorter periods were carried over to the following day which was then regarded as day 1. Data on the day of death was not collected. The worst values for acute physiological variables and the best Glasgow coma score were collected for each day to generate Apache II scores and to assess for organ failure. During the period of the study, the results of Apache II assessments were NOT used to influence clinical decisions to withdraw treatment from any of the patients studied. Children under 12 years of age, major burns, post cardiac surgery patients and coronary care patients not mechanically ventilated were not included in the study.

Criteria defining group

The first 100 consecutive adult ICU patients were used to define the prediction criteria to identify patients who would die either in the ICU or during their hospital stay after their discharge from the ICU.

Test group

The following 210 consecutive adult ICU patients were used to test the power of the defined predictive criteria.

Trend analysis of organ failure scores

Trend analysis was used to analyse the curves of daily organ failure scores plotted against ICU days of the 100 patients, survivors and non-survivors, in the criteria defining group to obtain the predictive criteria to identify those patients who would subsequently die. The data was processed in a search for repeatable trends. The analysis took into account both the absolute value of the organ failure score each day and its rate of change relative to the score of the previous day. The trend analysis of the daily organ failure scores produced the following patterns of change in scores and related them to hospital outcome: (a) High scorers: patients admitted with multiple end-stage organ failures and no hope of recovery had very high organ failure scores on day 1 ($score_1$). Our previous experience with over 300 patients and the present 100 patients in the criteria defining group showed that no patient with $score_1$ greater than 35 survived. $Level_1$ was therefore set at 35. If $score_1$ was greater than $level_1$, the patient was predicted to die. (b) Fuzzy banders: There were a few patients with $score_1$ greater than 31 and less than or equal to 35 who survived. The survivors who scored between 31 and $< = 35$ on day 1, responded rapidly to treatment and had lower scores on day 2. Among the non-survivors, if the day 2 score ($score_2$) was less than $score_1$ the difference was never more than 2.5. There was therefore a fuzzy band be-

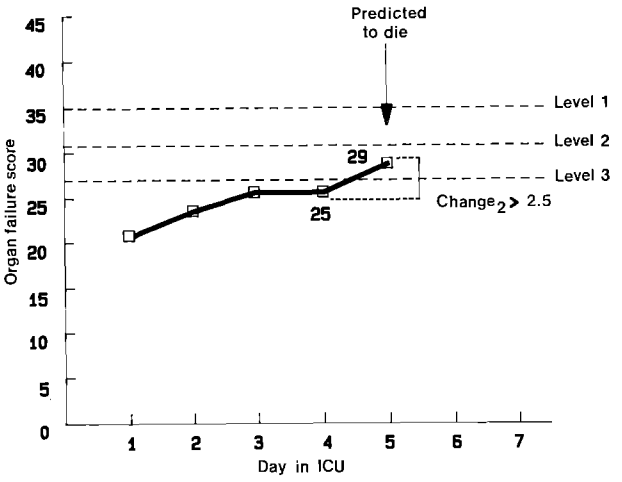

Fig. 3. Typical curve of daily organ failure scores of patients who deteriorated suddenly and died subsequently

tween scores of 31 and 35. $Level_2$ set at 31 defined the lower limit of this fuzzy band. Survivors and non-survivors with $score_1$ within the fuzzy band could be differentiated by the change in score from day 1 to day 2, with $(score_1 - score_2)$ of survivors >2.5 and $(score_1 - score_2)$ of non-survivors $< = 2.5$. This gave us the first criterion $(change_1)$ related to the rate in change in score, which was set at 2.5. Thus if $score_1$ was between $level_1$ and $level_2$, no prediction was made until day 2. If on day 2, $(score_1 - score_2)$ was $< = change_1$, the patient was predicted to die (Fig. 2). Patients with diabetic ketoacidosis who are usually admitted with severe physiological derangements, but respond rapidly to treatment. Their day 1 and day 2 scores are disregarded. (c) Leapers: many patients with reasonable scores on day 1, deteriorated and died. The process of deterioration may be sudden or gradual. The trend in the organ failure score of patients who deteriorated suddenly and died (Leapers) appeared to be related to both the absolute value and the rate of change in the scores. We found that if a patient's score

Table 3. Values of predictive criteria used to predict death among ICU patients using daily organ failure scores, daily Apache scores and a single Apache score on day 1 of admission

Criteria	Daily organ failure scores	Daily Apache scores	Single Apache score
$Level_1$	35	35	35
$Level_2$	31	30	–
$Level_3$	27	27	–
$Change_1$	2.5	3	–
$Change_2$	2.5	2	–

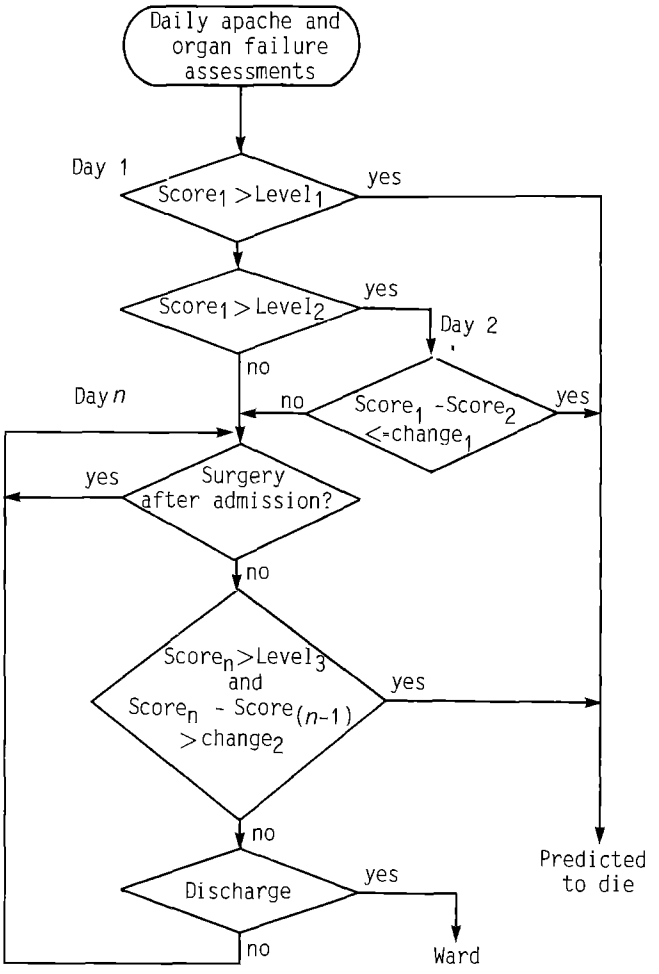

Fig. 4. Algorithm of the computer program to perform trend analysis of the daily organ failure scores in order to predict those patients who would subsequently die

on any day after the day of admission $(score_n)$ increased by more than 2.5, relative to that of the previous day $(score_{n-1})$ AND the absolute value was greater than 27, then the patient died (Fig. 3). Larger changes in score, with an absolute value below 27 were not predictive of death. These observations gave us a value of 27 for $level_3$ and a value of 2.5 for $change_2$ $(score_n - score_{n-1})$. (d) Creepers: patients who deteriorated gradually and died had a slightly different pattern in the organ failure scores. Their scores rose gradually, never increasing by more than 1 or 2 points a day as they passed the 27 mark $(level_3)$ They may still survive if their scores fell. However, if their scores continued to increase gradually beyond the 35 mark $(level_1)$, they subsequently died. (e) The outcome of patients who did not meet the above criteria were unpredictable.

Table 4. Characteristics of the 310 ICU patients studied

	Criteria-defining group	Test group	
Number of patients:	100	210	
Males	67	134	(64%)
Mean age	49	50	
Females	33	76	(36%)
Mean age	51	44	
Major systems affected:			
Respiratory	30	86	(41%)
Cardiovascular	23	54	(26%)
Gastrointestinal	22	24	(11%)
Neurological	19	37	(18%)
Metabolic	3	6	(3%)
Haematological	3	3	(1%)
Renal	0	0	
Positive for chronic health problems:			
Liver	10	16	(8%)
Cardiovascular	8	9	(4%)
Pulmonary	6	8	(4%)
Renal	2	8	(4%)
Immune compromised	6	26	(12%)
Significant categories:			
Non-operative	63	106	(50%)
Post-operative	37	104	(50%)
Emergency	4	36	(17%)
Elective	33	68	(32%)
Operative complications	10	32	(15%)
Ward transfers	34	47	(22%)
Hospital transfers	15	41	(20%)
Emergency room admissions	14	18	(9%)
Outcome analysis:			
Died in ICU	22	52	(25%)
Died in ward	13	20	(9%)
Discharged alive from hospital	65	138	(66%)

Allowance for surgery or major iatrogenic complications after admission to ICU

Surgery or major iatrogenic complications arising after admission to the ICU can result in an increase in score that may not be related to the basic pathological processes affecting the patient. The resulting high score on the day of the insult must be regarded in the same manner as a day 1 score, i.e. no prediction should be made until the following day. The trend analysis algorithm (Fig. 4) also took this eventuality into account. Table 3 gives the values of the criteria used for predicting hospital outcome.

All Apache II assessments were carried out by two of us on special database forms. One of us followed up the outcome of those patients discharged alive from the ICU to determine their hospital outcome. The data were entered into an IBM AT microcomputer by the use of the Riyadh ICU program written in dBASE III Plus (Ashton-Tate) by one of us to generate the Apache II scores, organ failure scores and to carry out trend analysis and other statistical evaluations. In this paper the following formulae were used to evaluate the predictive power of the system:

Specificity (%) = PredAAlive/(PredAAlive+PredDAlive)×100
Sensitivity (%) = PredDDead/(PredDDead+PredADead)×100
Correct (%) = (PredAAlive+PredDDead)/Total patients×100
Predictive value positive (%) = PredDDead/Total predicted to die×100
Predictive value negative (%) = PredAAlive/Total predicted to live×100

where:

PredAAlive = Predicted to live, lived
PredADead = Predicted to live, died
PredDAlive = Predicted to die, lived
PredDDead = Predicted to die, died.

The "rule of three" [5] was used to estimate the long-term reproducibility of the predictions at a level of confidence of 95%.

Results

Table 4 shows the characteristics of the 100 patients in the criteria defining group and the 210 patients in the test group. The mean ages, sex distribution and mortality rates of the two groups were similar. There was a slight difference in the patient mix between the criteria defining and the test group, with a larger pro-

Table 5. Day 1 Apache scores and organ failure scores of the patients studied. There was no significant difference between the scores of the criteria group and the test group

	Criteria group ($n = 100$)				Test group ($n = 210$)			
	Survivors ($n = 65$)	Non-survivors ($n = 35$)	t	p	Survivors ($n = 138$)	Non-survivors ($n = 72$)	t	p
Mean Apache score (SD)	15 (6.51)	23 (8.59)	4.46	<0.001	12 (5.96)	24 (8.96)	10.82	<0.001
Mean organ failure score (SD)	15.7 (6.81)	23.5 (9.38)	4.47	<0.001	12.5 (6.12)	24.5 (9.88)	10.87	<0.001

Table 6. Predicted and actual hospital outcome for 100 patients in criteria defining group

	Predicted to live	Predicted to die	Total
Actual alive	65	0	65
Actual dead	19	16	35
Total	84	16	100

Specificity 100%; Sensitivity 45.7%; Proportion correct 81%; Predictive value +ve 100%; Predictive value −ve 77.4%

portion of patients with respiratory problems and immunocompromised patients in the test group. There was no difference in the mean Apache II score or organ failure score of the criteria defining group and of the test group (Table 5).

Table 6 shows the predicted and actual hospital outcome of the 100 patients in the criteria defining group using the daily organ failure score method for prediction. Sixteen out of 34 deaths were identified. Table 7 compares the predictive powers of daily organ failure scores, daily Apache II scores and a single Apache II score on day 1 to predict hospital outcome among the 210 patients in the test group. Although there were no false predictions of death in all 3 methods, daily organ failure scores, which predicted 37 out of the 72 patients who died, was superior to daily Apache II scores and a single Apache score by a factor of 1.4 and 5.3, respectively. The difference in prediction rate between daily organ failure scores and daily Apache II scores however did not achieve statistical significance. The 7 patients predicted to die by a single Apache score on day 1 (i.e. score greater than 35) were admitted with septic shock (4), after resuscitation from cardiac arrest (2) and severe upper gastrointestinal haemorrhage from oesophageal varices (1). The trend curves of the 37 patients predicted correctly to die were made up of: 10 "High Scorers", 4 in the Fuzzy band, 22 "Leapers", and 1 "Creeper".

Table 7. Predictive power of daily organ failure scores, daily Apache II scores and a single Apache score on day 1 for the 210 patients in the test group

	Predicted to live	Predicted to die	Total
Daily organ failure scores			
Actual alive	138	0	138
Actual dead	35	37	72
Total	173	37	210
Specificity	100%		
Sensitivity	51.4%		
Proportion correct	83.3%		
Predictive value +ve	100%		
Predictive value −ve	79.8%		
Daily Apache scores			
Actual alive	138	0	138
Actual dead	46	26	72
Total	184	26	210
Specificity	100%		
Sensitivity	36.1%		
Proportion correct	78.1%		
Predictive value +ve	100%		
Predictive value −ve	75%		
Single Apache score on day 1			
Actual alive	138	0	138
Actual dead	65	7	72
Total	203	7	210
Specificity	100%		
Sensitivity	9.7%		
Proportion correct	76.4%		
Predictive value +ve	100%		
Predictive value −ve	76.4%		

The effect of organ system failure was examined. Patients were divided into those without organ failure, with 1, 2, and with 3 or more organ failures. The classification process took into account the fact that patients with one organ failure may develop 2 or more organ failures during the course of their stay in the ICU. Thus patients categorized as having 1 organ failure were those who never developed more than one organ system failure, and similarly for 2 and 3 or more organ failures, to avoid double counting. The mortality of patients was increased by the presence of organ system failures. Thirteen out of 211 patients without organ failures died giving a mortality rate of 10.7%. The mortality of 89 patients with 1 organ system failure was 24.7%. The mortality rates of patients with 2 organ failures (50) and 3 or more organ failures (50) were 52% and 92%, respectively. All patients who were predicted to die had one or more organ failures on the day of prediction.

The incidence and types of acute organ failures on the first day that one or more organ failures occurred is shown in Table 8. There was no difference in the incidence of respiratory failure for patients with at least 1 organ failure among the survivors, non-survivors not predicted to die and non-survivors predicted to die. The incidence of acute renal failure and neurological failure increased from around 10% among the survivors to 53% and 38% respectively among the non-survivors predicted to die and 34% and 18% among non-survivors not predicted to die; and as mortality was greater among predicted non-survivors compared to unpredicted non-survivors, they seemed to have a positive predictive value. The frequency of cardiovascular failure among predicted non-survivors was double that of the other two groups. This was also the case with the presence of haematological failure. Although the incidence of acute liver failure among the non-survivors was double that among the sur-

Table 8. The incidence and types of organ failures on day 1 of organ failures for the 310 ICU patients

Organ failures	Incidents	Neurologic	Cardiovascular	Respiratory	Haematologic	Hepatic	Renal
Survivors (n = 203)							
1	107	6	19	64	11	2	5
2	34	2	18	22	10	5	11
> = 3	4	3	3	3	1	1	1
> = 4	0	0	0	0	0	0	0
> = 1	145	11	40	89	22	8	17
(% of incidents)		(8%)	(28%)	(61%)	(15%)	(6%)	(12%)
Non-survivors not predicted to die (n = 54)							
1	58	9	6	27	7	1	8
2	54	12	17	38	13	9	19
> = 3	33	5	25	26	14	8	22
> = 4	1	0	1	1	0	1	01
> = 1	145	26	48	91	34	18	49
(% of incidents)		(18%)	(33%)	(63%)	(23%)	(12%)	(34%)
Non-survivors predicted to die (n = 53)							
1	3	1	0	2	0	0	0
2	21	6	10	13	6	0	7
> = 3	23	11	19	15	8	4	18
> = 4	6	6	6	5	2	1	4
> = 1	47	18	29	30	14	4	25
(% of incidents)		(38%)	(62%)	(64%)	(30%)	(10%)	(53%)

vivors there was little difference in its incidence among predicted non-survivors and unpredicted non-survivors.

The total number of ICU bed days and the corresponding number of hospital bed days after discharge from the ICU, occupied by the 310 patients in the study were 1653 and 6759, respectively. Table 9 shows the day of prediction of death and the hypothetical saving in ICU days and ward days if treatment was withdrawn on the day a patient was predicted to die and the patient were to die within 24 h of termination of therapy. The vast majority of the predictions were made within 2 days with the largest number of saving in ICU days. A total of 249 ICU and 215 ward days would have been saved if the predictions were acted upon. "Leapers" were usually predicted to die within the first week while "Creepers" were predicted to die after 7 days. The ICU and ward days occupied by unpredicted non-survivors were 200 and 370 days respectively.

Table 9. Day of prediction of death and ICU and ward days saved if treatment was withdrawn and patient died within 24 h. The 310 patients spent a total of 1653 days in the ICU and 6762 days on the wards after discharge from the ICU

Day of prediction	Number	ICU days saved	Ward days saved
1st or 2nd day	27	111	81
3rd to 7th day	13	46	2
8th to 14th day	10	83	124
After 14th day	3	9	8

There were no errors in the predictions of death among the 210 consecutive patients in the test group. Applying the "rule of three" it is estimated that the long-term risk of a false prediction is 1.43% at the 95% confidence level.

Discussion

Outcome predictions of ICU patients using trend analysis of daily Organ Failure Scores was superior to all the other predictive models that we have tried so far. Thirty-seven out of the 72 patients who died in the test group were correctly identified. This was despite the fact that there were differences in the patient mix of the criteria defining group and the test group. This is an impressive demonstration of the power and reproducibility of the original Apache II system.

Trend analysis appeared to be an important improvement in methodology as only 10 patients were predicted to die from a high day 1 score, while another 27 were predicted by criteria based on trend analysis. The largest number of predictions were on patients with the "Leapers" curve. This underlines the importance of analysis which takes into account the dynamic changes in the patho-physiological processes during the course of patient's stay in the ICU. The use of a fuzzy band and a simple form of "fuzzy logic" enabled us to overcome one of the common problems associated with the use of digital computers, and added another 4 patients to the total number predicted to die.

Predictions by daily organ failure scores predicted 11 more patients than daily Apache scores. This indicated that the use of a correction factor for the number and duration of organ system failure further improved the sensitivity of the predictions.

We were only able to analyse the relation between organ failure and mortality for the first day of organ failure as the number of patients in our study was small. Our figures of 24.7% for 1, 52% for 2, and 92% for 3 or more organ failures however were remarkably similar to those reported by Knaus (22%, 52% and 80%, respectively) [4]. There was however an important disparity between our results and those reported by Knaus. None of the 310 patients in this series and the more than 300 patients from our previous series survived a day 1 Apache II score of 35. In Knaus series, the mortality of patients scoring above 35 was 84% [6]. It may be that our patient mix was such that there was no patients admitted with conditions compatible with survival with a day 1 Apache II score greater than 35. Another possibility is that we have been extremely cautious with our use of the best Glasgow coma score. We noted early on in our use of the Apache II system, that Glasgow coma scores contributed a major component to the Apache II score and have therefore tended to give the benefit of the doubt to patients who are difficult to assess. It is therefore very likely that while our methodology can improve the predictive power of the Apache II system, the absolute values of the criteria may have to be recalibrated upwards for other units. We are of the opinion that a validation period of at least a year should be undertaken in order to obtain criteria for a specific unit. We are planning a multi-centre study to validate our model.

Knaus [4] did not include criteria for defining acute liver failure. Most of our liver patients suffer from cirrhosis as a consequence of hepatitis B or schistosomiasis infections. Alcoholic cirrhosis is uncommon in Saudi Arabia. They tend to develop acute crisis as a consequence of variceal bleed or infections. In this study, we have tentatively used the Glasgow predictor [7], derived from patients suffering mainly from alcoholic cirrhosis, as a criterion for liver patients, as we have found it to have good predictive value (unpublished data). Its use will require further evaluation.

Analysis of the types and frequency of organ failures and their relation to mortality indicated that respiratory failure as defined in the study, while very common, had little predictive value. This is in contrast to acute renal failure, neurological, cardiovascular and haematological failure. Liver failure, while more common among the non-survivors also did not appear to have any predictive value. It is theoretically possible to further increase the sensitivity of our predictive model if coefficients related to the specific organ failure or combinations of organ failures were available. However, the size of the database required would be enormous.

We did NOT use our predictive model to influence clinical decision to withdraw therapy during this study. However our hypothetical analysis indicated that over 200 ICU and 200 ward days would have been saved had the predictions been acted upon. Saving would be greatest among those predicted to die between the 8th and 14th day, as not only the fixed cost of beds and staff would have been reduced but also the variable cost of therapy, which has been shown to be greatest among patients with prolonged ICU stays [8]. These patients are also those who have the worse mortality rates.

Self-fulfilling predictions and therapeutic nihilism have been raised as dangers of computer based predictive models [9]. If therapy is withdrawn on the strength of computer generated prediction it is certain that the patient will die. But what is the practice at present? Such decisions are based on clinical judgement and experience, which are not only indefinable and unmeasurable but may vary between clinicians. Such decisions are just as self-fulfilling. But are they more accurate and reproducible than predictions based on computerised analysis of objective patho-physiological data? We have estimated that the long-term risk of a false prediction using the present model and based on the results obtained from our small sample size to be 1.4%. Is this estimated error rate acceptable? The answer to this is to compare the error rate of our model with that made by clinicians. Such a study is currently in progress in our unit.

The danger of therapeutic nihilism exists only if doctors were to totally abdicate their clinical responsibility and blindly follow predictions by computers. The other side of the coin is therapeutic abuse; the needless prolongation of the process of dying by unwarranted treatment of ICU patients with hopeless prognosis. Not only the financial consequences, but the unnecessary distress and agony inflicted upon the patient's relatives as a result of therapeutic abuse has also to be considered. It is important to note that on the *one* occasion during this study when the model was used, it was to justify the continuation of maximal efforts in the management of a patient in whom it was judged clinically that his prognosis was hopeless. The patient survived. If the predictive power of our model is validated by a multicentre study, it could be used as an objective criterion to aid decisions to proceed with or continue aggressive therapy or to withdraw treatment. We feel that much of the moral dilemma facing clinicians who have to make these

decisions is based on their acute awareness of their fallibility.

References

1. Knaus WA, Draper EA, Wagner DP, Zimmerman JE (1985) APACHE II: a severity of disease classification system. Crit Care Med 13:818
2. 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
3. Chang RWS, Jacobs S, Lee B, Pace N (1988) Predicting deaths among ICU patients. Crit Care Med 16:34
4. Knaus WA, Draper EA, Wagner DP, Zimmerman JE (1985) Prognosis in acute organ-system failure. Ann Surg 202:685
5. Hanley JA, Lippman-Hand A (1983) If nothing goes wrong. Is everything allright? JAMA 249:1743
6. Knaus WA, Draper EA, Wagner DP, Zimmerman JE (1986) An evaluation of outcome from intensive care in major medical centers. Ann Int Med 104:410
7. Garden OJ, Motyl H, Gilmour WH, Utley RJ, Carter DC (1985) Prediction of outcome following acute variceal haemorrhage. Br J Surg 72:91
8. Madoff RD, Steven MS, Fath JJ, Simmons RL, Cerra FB (1985) Prolonged surgical intensive care. Arch Surg 120:698
9. Anonymous (1986) Editorial TPN and Apache. Lancet I:1478

Dr. R. W. S. Chang
Riyadh Armed Forces Hospital
P.O. Box 7897
Riyadh 11159
Saudi Arabia