Predicting death from initial disease severity in very low birthweight infants: a method for comparing the performance of neonatal units

William Tarnow-Mordi, Simon Ogston, Andrew R Wilkinson, Elisabeth Reid, John Gregory, Mahmoud Saeed, Rosalie Wilkie

Abstract

Objectives—To investigate (*a*) which clinical variables and physiological measures of disease severity best predict death in very low birthweight infants and (*b*) their use in comparing mortality between two neonatal units.

Design—Retrospective study of two cohorts of very low birthweight infants from overlapping time periods who received mechanical ventilation.

Setting—Two neonatal intensive care units (hospitals A and B).

Subjects—262 Very low birthweight infants, 130 in hospital A, 132 in hospital B.

Main outcome measure—Death in hospital.

Results—In hospital A the mean level of oxygenation in the first 12 hours of life, whether measured as inspired oxygen requirement (F_IO_2), arterial/alveolar oxygen (a/AO_2) ratio, or alveolar-arterial oxygen difference ($A\text{-}aDO_2$), was more closely associated with death than any of four "traditional" risk factors: low birth weight, short gestation, the diagnosis of respiratory distress syndrome, and male sex. Mean pH in the first 12 hours was as strongly associated with death as birth weight. Multiple logistic regression models were derived in infants from hospital A using the four traditional risk factors with measures of oxygenation and pH. The validity of each model was then tested in infants from hospital B. The model based on the four traditional risk factors alone predicted death in hospital B with only 31% sensitivity. Adding mean a/AO_2 ratio and mean pH increased its sensitivity to 75%, and when mean a/AO_2 ratio was replaced by mean F_IO_2 its sensitivity increased further to 81%. Based on crude mortality rates alone, the odds of death in hospital A versus hospital B were 0·67 (95% confidence interval 0·37 to 1·23). After correcting for traditional risk factors and mean F_IO_2 and mean pH, however, the odds of death in hospital A increased to 3·27 (1·35 to 7·92; $p<0·01$). This increased risk persisted after adjusting for the time difference between each cohort.

Conclusions—Crude comparisons of hospital mortality can be highly misleading. Reliable assessment of neonatal outcome is impossible without correcting for major risk factors, particularly initial disease severity. International agreement on a minimum core dataset of clinical and physiological information could improve neonatal audit and help to identify effective treatments and policies.

Departments of Child Health and Community Medicine, Ninewells Hospital and Medical School, University of Dundee, Dundee DD1 9SY
William Tarnow-Mordi, MRCP, *senior lecturer in child health*
Simon Ogston, MA, *lecturer in medical statistics*
Elisabeth Reid, SRN, *research sister*
John Gregory, MRCP, *paediatric registrar*
Mahmoud Saeed, MRCP, *paediatric registrar*
Rosalie Wilkie, MRCP, *paediatric senior registrar*

Neonatal Unit, Department of Paediatrics, John Radcliffe Maternity Hospital, Oxford OX3 9DU
Andrew R Wilkinson, FRCP, *consultant paediatrician*

Correspondence to:
Dr Tarnow-Mordi.

Br Med J 1990;300:1611-4

Introduction

The need for reliable audit of neonatal intensive care is widely recognised,[1-5] and several studies have compared the performance of different neonatal intensive care units.[6-9] Any comparison based on crude mortality, however, or on mortality adjusted only for clinical factors like birth weight, gestation, race, or sex, may be open to question because of the bias caused by differences in the initial severity of disease.[9-12]

Much of the variation in disease severity between infants treated in different hospitals stems from differences in the antenatal and postnatal referral of high risk cases. The resulting selection bias can largely be eliminated by comparing big, geographically defined populations.[13,14] This does not, however, fully resolve the problem as there is no reason why disease severity should remain constant between different geographical populations or even within the same population over time.[15] Comparisons of perinatal outcome between hospitals might be made more reliable by first correcting for severity of disease and clinical variables like gestation and birth weight, but this would require an accurate measure of neonatal disease severity.

The most accurate, quantitative measure of disease severity in childhood is the physiologic stability index.[14] This was derived from a multivariate logistic model using data obtained in 822 patients within the first day of admission to a single paediatric intensive care unit in 1980-2. It was tested by comparing observed mortality versus that predicted by the model in 1572 children admitted to eight other intensive care units in 1982-4. Using a probability of >50% mortality as a criterion for predicting death, the model accurately predicted outcome in 67% of hospital deaths and in 99% of survivors. Although crude mortality rates among the nine institutions varied from 3·0% to 17·6%, these differences disappeared after initial disease severity had been corrected for, indicating that the quality of care was the same in each hospital. There was also no difference in corrected mortality in the reference institution between the periods 1980-2 and 1984-5, suggesting that no major advances in paediatric intensive care had occurred.

We wanted to assess whether a predictive model of comparable accuracy might be developed in very low birthweight infants. We therefore studied the prognostic value of several clinical and physiological variables, all derived from routinely collected data. From these we developed and tested various multiple logistic regression models of risk of death, then used them to compare the performance of two neonatal units.

Patients and methods

The infants studied had birth weights <1500 g and all received artificial ventilation via an endotracheal tube at some time in the first 72 hours of life. They were treated in two United Kingdom hospitals from 1 January 1986 to 31 May 1989 (hospital A) and from

1 November 1983 to 14 September 1986 (hospital B) with one type of ventilator (Sechrist IV 100B). Infants who died within the first 12 hours of life or had severe congenital malformations were excluded. Both hospitals used similar criteria to define respiratory distress syndrome,[16] and the concentration of inspired oxygen (F_IO_2) was strictly controlled by monitoring transcutaneous oxygen tension. Desired blood gas tensions were also similar (Pao_2 6·7-10 kPa; Pco_2 4·5-6 kPa). Pao_2 was estimated from umbilical or peripheral arterial blood or transcutaneously. Mean airway pressure was measured at the proximal airway by electronic airway pressure monitors (APM, Sechrist Industries Inc, Anaheim, California, or Insight 2000 EME Ltd, Brighton, England).

Definitions and calculations—The measures of disease severity examined were:

(*a*) arterial oxygen tension (Pao_2);

(*b*) arterial or capillary pH;

(*c*) arterial or capillary carbon dioxide tension (Pco_2);

(*d*) concentration of inspired oxygen, F_IO_2, expressed as 0·21–1·0;

(*e*) alveolar-arterial oxygen difference (A-aDo_2): from $A\text{-}aDo_2 = F_IO_2 \times 713 - Pco_2/R + Pco_2 \times F_IO_2 \times (1-R)/R - Pao_2$ (where R=0·8)[17];

(*f*) arterial/alveolar oxygen ratio (a/Ao_2 ratio): from a/Ao_2 ratio $= Pao_2/(F_IO_2 \times 713 - Pco_2/R + Pco_2 \times F_IO_2 \times (1-R)/R)$[17];

(*g*) mean airway pressure;

(*h*) mean airway pressure–shunt product—that is, mean airway pressure×venous admixture (QS/QT), calculated from Rojas *et al*[18];

(*i*) ventilator efficiency index (VEI), calculated from $VEI = 3800/(f \times (PIP - PEEP) \times Pco_2)$, where f is ventilator rate, PIP is peak inspiratory pressure, and PEEP is positive end expiratory pressure in cm H_2O[19];

(*j*) ventilator index number 1 (VIN1), calculated from VIN1=f×PIP[20]; and

(*k*) ventilator index number 2 (VIN2), calculated from VIN2=mean airway pressure$\times F_IO_2/Pao_2$.[21]

Statistical analysis—The means of continuous variables were compared by unpaired *t* tests and proportions by χ^2 test with Yates's correction. In hospital A the mean value of each measure of initial disease severity was calculated for each infant from blood gas or ventilator data for all blood gas samples in the first 12 hours of life. The worst values of some variables—that is, highest F_IO_2 and A-aDo_2, lowest pH and a/Ao_2 ratio—were also obtained for the first 12 hours of life in each infant. For infants in hospital A the mean and worst values were related to subsequent risk of death by univariate logistic regression[22] and ranked with clinical and demographic variables according to strength of association, denoted by the magnitude of χ^2. We planned in advance to construct multiple regression models[22] for risk of death on the four traditional factors—sex, birth weight, gestation, and respiratory distress syndrome—then to add by stepwise regression all measures of disease severity which contributed independently ($p<0·05$) to each model's explanatory power. When mean values were not available we included a dummy variable indicating a missing value in the regression equation. Each model was tested by comparing the observed versus the calculated risk of death in a group of very low birthweight infants ventilated in hospital B. Infants whose calculated risk of death exceeded 0·5 were predicted to die, the rest to live. The data were then pooled and new regression models derived to compare observed versus predicted mortality between the hospitals after correcting for previously identified risk factors. Finally, a factor corresponding to year of birth was added to correct for any effects due to the time differences between each hospital cohort.

Results

Descriptive variables for infants in each hospital are listed in table I. For every variable except sex the sample of infants in hospital B was at significantly higher risk of death. Table II shows that of 11 measures of disease severity examined in hospital A nine were significantly associated with subsequent risk of death. Of these, high mean A-aDo_2, high mean F_IO_2, and low mean a/Ao_2 ratio were all more strongly associated with death than the four traditional risk factors (low birth weight, short gestation, the presence of respiratory distress syndrome, and male sex). Low mean pH and low birth weight were roughly equivalent in their strength of association with death. For each of the four measures of disease severity in the first 12 hours, A-aDo_2, F_IO_2, a/Ao_2 ratio, and pH, the mean values were consistently more closely associated with death than the worst values (table II).

The performance of models derived from hospital A in predicting death in hospital B is shown in tables III

TABLE I—*Characteristics of very low birthweight infants mechanically ventilated in neonatal intensive care units in two hospitals*

	Hospital A (n=130)	Hospital B (n=132)
Dates of birth	1 Jan 86-31 May 89	1 Nov 83-14 Sep 86
No (%) of boys	68 (52)	74 (56)
No (%) with respiratory distress syndrome	66 (51)	94 (71)***
Mean (SD) birth weight (g)	1143 (220)	1044 (260)***
Mean (SD) gestation (weeks)	29·2 (2·6)	28·3 (2·3)***
Mean (SD) F_IO_2	0·49 (0·21)	0·65 (0·22)***
Mean (SD) a/Ao_2 ratio	0·33 (0·22)	0·23 (0·14)***
Mean (SD) A-aDo_2 (kPa)	32·0 (8·8)	45·5 (21·5)***
Mean (SD) pH	7·36 (0·09)	7·31 (0·08)***
No (%) of hospital deaths	23 (18)	32 (24)

***p<0·001.

TABLE II—*Prognostic factors and death before discharge in hospital A*

Variables	No	χ^2	No	χ^2
	Clinical variables			
Birth weight	130	14·4***		
Gestation	130	7·5**		
Respiratory distress syndrome	130	6·2*		
Sex	130	0		
	Measures of disease severity in first 12 hours			
	Using mean of all values		Using worst single value	
A-aDo_2	114	23·0***	114	13·8***
F_IO_2	120	21·2***	120	17·4***
a/Ao_2 ratio	114	15·6***	114	10·0***
pH	126	13·7***	126	8·3**
Ventilator efficiency index	108	12·5***		
Ventilator index No 1	109	11·5***		
Ventilator index No 2	47	9·1**		
Mean airway pressure	59	7·3**		
Pao_2	114	5·6*		
Pco_2	119	3·0		
Cumulative mean airway pressure-shunt product	88	0·3		

*p<0·05; **p<0·01; ***p<0·001.

TABLE III—*Validation of multiple regression model* derived in hospital A by predicting death in hospital B*

	Predicted		
	Alive	Dead	Total
Observed:			
Lived	74	26	100
Died	6	26	32
Total	80	52	132

Sensitivity 26/32=81%; specificity 74/100=74%; positive predictive value 26/52=50%; negative predictive value 74/80=93%; % correctly classified 100/132=76%.

*Death was predicted using a multiple regression model based on sex, birth weight, gestation, respiratory distress syndrome, mean F_IO_2, and mean pH. For each infant a calculated probability of death >0·5 was interpreted as predicting death, all other probabilities predicted survival.

TABLE IV—*Effect of including various factors on the power of regression models derived from hospital A to predict death in hospital B. Sensitivity and specificity are expressed as percentages (and numbers on which these are based are given in parentheses)*

Factors included in the regression model	Sensitivity (%)	Specificity (%)
Traditional risk factors (sex, birth weight, gestation, respiratory distress syndrome)	31 (10/32)	94 (94/100)
Traditional risk factors + mean a/A_{O_2} ratio	59 (19/32)	87 (87/100)
Traditional risk factors + mean a/A_{O_2} ratio + mean pH	75 (24/32)	76 (76/100)
Traditional risk factors + mean $F_{I}O_2$ + mean pH	81 (26/32)	74 (74/100)

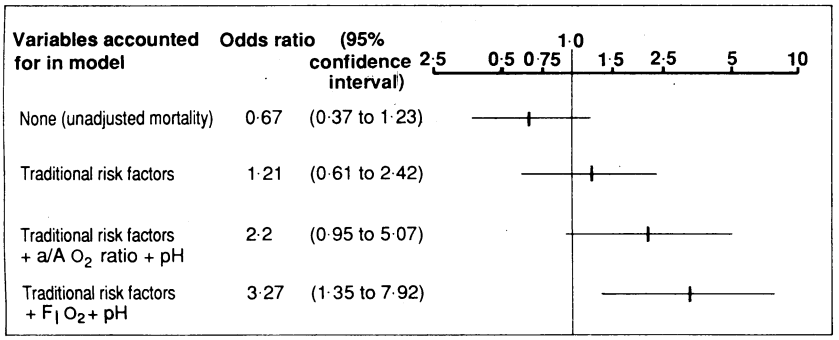
The sensitivity of the models (% successful prediction of outcome among those who died) increased as clinical factors and measures of disease severity were taken into account. The specificity of the models (% successful prediction of outcome among survivors) showed a reciprocal relation to their sensitivity. Had the study included the low risk infants who were not mechanically ventilated, among whom deaths are rare, the specificity of the models would have been increased with little change in sensitivity.

and IV. Of the four traditional risk factors, only birth weight contributed significantly to the models' explanatory power. After adding mean oxygenation and mean pH, no other measures of initial disease severity contributed significantly to any model. The three indices of oxygenation, mean $F_{I}O_2$, mean A-aD_{O_2}, and mean a/A_{O_2} ratio, each had a roughly equivalent effect; when one had been entered in a model, adding another produced little change in explanatory power. The prediction of death in hospital B was consistently slightly more accurate after including the mean rather than the worst values of a/A_{O_2} ratio, $F_{I}O_2$, or pH (the respective sensitivities were: 59% *v* 50%, 81% *v* 56%, 75% *v* 63%). Although mean A-aD_{O_2} was more strongly associated with death than mean $F_{I}O_2$, we examined mean $F_{I}O_2$ in the models because it is simpler to calculate in routine use. The predictive power of the models changed little when cases with missing values for a/A_{O_2} ratio or $F_{I}O_2$ were excluded from the regression equations (in table IV the respective sensitivities before and after exclusion were 59% *v* 50%, 75% *v* 70%, and 81% *v* 80%).

Comparisons of mortality between the hospitals, before and after correcting for major risk factors using appropriate regression models, are shown in the figure. This shows a considerable increase in the odds of death in hospital A after adjusting for clinical variables and initial disease severity, estimated from mean oxygenation and mean pH. The regression equation including year of birth to adjust for the time difference between each hospital cohort is given in the appendix. Year of birth did not contribute significantly to the model, suggesting that the excess risk of death in hospital A was independent of any trend due to time.

Discussion

We have developed a multiple logistic regression model for risk of death in hospital based on four clinical factors—sex, gestation, respiratory distress syndrome, and birth weight (the only significant one)—and two measures of disease severity—mean $F_{I}O_2$ and mean pH. The model was validated in a test population, in which it predicted death with 81% sensitivity (table IV). When the performance of two neonatal units was compared after correcting for important prognostic factors the tendency for the crude mortality figures to favour one hospital was reversed (figure).



Odds for risk of death in hospital A versus hospital B according to variables included in model

This study confirms that crude comparisons of mortality can be highly misleading. This may be true even when important traditional risk factors are taken into account, as in a recent comparison of the outcome of mechanical ventilation in newborn infants in two geographically defined regions.[13,14] Although there was a clear difference in neonatal mortality after adjusting for birth weight, this should not be attributed solely to differences in the policies for mechanical ventilation in each region. The greater mortality in one population might simply have reflected lower severity of disease in the other. An accurate measure of initial disease severity would allow more reliable comparisons of perinatal outcome between populations, whether made by region or hospital or over different periods.

In an earlier analysis of 387 very low birthweight infants admitted to a single intensive care unit Patterson and Halliday developed a multivariate model in 100 infants based on three factors in the first hour of life: gestational age, the 5 minute Apgar score, and the presence or absence of signs of respiratory distress.[23] The model predicted death in the remaining infants (using a probability criterion of >0·5) with 54% sensitivity and 89% specificity. Although the sensitivity of our final model was higher (81%), its specificity was lower (74%). Nevertheless, had we included very low birthweight infants who were not ventilated (among whom mortality is very low) the specificity of our model would have increased, with little change in sensitivity. Both studies confirm that simple predictive models can be developed in very low birthweight infants. A comprehensive model should now be derived and validated in a wider range of neonatal units. One methodological improvement would be to derive the model in a 50% random sample of infants in all units and then test its validity in the other 50%. This would tend to balance unknown factors, such as variations among units in the quality of care, equally between observation and test samples.

Although differences in hospital mortality after correcting for important early prognostic factors are most likely to reflect differences in the subsequent quality of care, they must be interpreted with caution. Early physiological measurements reflect both early treatment and intrinsic disease severity, and the effects of each are inseparable. By increasing the severity of disease bad early treatment is likely to worsen prognosis without affecting the accuracy of a predictive model. On the other hand, when their subsequent mortality is corrected for their greater initial severity of disease neonatal units in which early treatment is bad would benefit in comparison with others. This source of bias could be limited by (*a*) shortening the period in which initial disease severity is measured or (*b*) identifying neonatal units where poor early treatment is suggested by an unexpectedly high mortality after adjusting for clinical risk factors alone. Comparisons between units may also be distorted by deaths unrelated to factors in the model—for example, due to necrotising enterocolitis or epidemic infections—particularly when sample sizes are small. Lastly, reduced mortality may be offset by increased numbers of survivors with severe handicap. Future studies should therefore also measure the quality of survival.

Despite the small sample, our model compares reasonably well with others in parallel disciplines. In older children receiving intensive care the physiologic stability index, based on 34 variables from seven organ systems, predicted death (using a probability criterion of >0·5) with 67% sensitivity and 99% specificity.[12] APACHE II, a classification of disease severity based on 12 variables and validated in 5815 patients admitted to adult intensive care units, similarly predicted death

with a sensitivity of 47% and specificity of 95%.[11] Both these systems used only the worst values for ordered categories of physiological data during the first day of admission. We found that the mean values of variables such as oxygenation or pH in the first 12 hours of life were consistently more closely associated with death than the worst values of the same variables (table II) and that models which included mean rather than worst values uniformly predicted death more accurately. The fact that mean values of ventilator efficiency index, ventilator index number 1, ventilator index number 2, and mean airway pressure were less strongly associated with death on univariate analysis than measures of mean oxygenation and mean pH, and that mean $a/A\bar{o}_2$ ratio was less strongly associated with death than mean F_IO_2, may simply have reflected the greater numbers of missing values (table II). Strict standardisation of the timing of blood gas samples is not possible, so a further refinement might be to calculate the time weighted average for all variables derived from blood gas measurements.[18]

This and other studies[11 12] have emphasised objective clinical variables and quantitative physiological data, thus reducing the problems posed by variations in diagnostic terminology when comparing different populations. To develop an internationally applicable measure of disease severity for very low birthweight infants would require agreement on a minimum core dataset of clinical and physiological information to be collected in routine management. Early measures of oxygenation may play a major part in this. We suggest that $a/A\bar{o}_2$ ratio would be preferable to F_IO_2, because F_IO_2 is at best an indirect measure of oxygenation and not a true physiological variable. Also, while F_IO_2 settings were appropriate in this study because oxygenation was continuously monitored, this may not always be the case in routine practice. The $a/A\bar{o}_2$ ratio is relatively unaffected even when F_IO_2 settings are inappropriate, because it is independent of acute fluctuations in F_IO_2.[24]

Although a reliable predictive model for risk of death could be used to rank hospitals in order of performance, there are other more important applications. Such models could identify potentially effective treatments and policies which merit rigorous assessment by controlled trial or provide more accurate prognoses for individual patients. It is crucial that such information is not misused. Evidence that a particular baby is at increased risk should be used to concentrate attention on that patient not to withdraw it. Evidence that one institution is performing less successfully than others should stimulate action to analyse and improve its performance. We are confident that doctors, parents, health service researchers, and managers will rise to the challenge.

WTM was supported by the National Fund for Research into Crippling Diseases (Action Research) and by travel grants from the Scottish Hospitals Endowment Research Trust and the Lawrence Bequest; ER was supported by the Scottish Chest Heart and Stroke Association and the Scottish Home and Health Department. We particularly thank Douglas Richardson, Iain Chalmers, Malcolm Chiswick, Edmund Hey, Professors Mary Ellen Avery, David Hull, Victor Chernick, and Michael Healy, and the anonymous referees for valuable suggestions.

1 Stewart AL, Reynolds EOR, Lipscomb AP. Outcome for infants of very low birth weight: survey of the world literature. *Lancet* 1981;i:1038-41.
2 Hull D. The viable child. The Croonian lecture 1988. *J R Coll Physicians Lond* 1988;**22**:169-75.
3 Yu VYH, Loke HL, Bajuk B, *et al.* Prognosis for infants born at 23 to 28 weeks' gestation. *Br Med J* 1986;**293**:1200-3.
4 Tarnow-Mordi WO, Wilkinson AR. Mechanical ventilation of the newborn. *Br Med J* 1986;**292**:575-6.
5 Tarnow-Mordi WO. Ventilator care and respiratory audit. In: Harvey D, Cooke RWI, Levitt G, eds. *The baby under 1000 grams.* London: Wright, 1989:65-77.
6 Greenland S, Watson E, Neutra RR. The case-control method in medical care evaluation. *Med Care* 1981;**19**:872-8.
7 Avery ME, Tooley WH, Keller JB, *et al.* Is chronic lung disease preventable? A survey of eight centers. *Pediatrics* 1987;**79**:26-30.
8 Kraybill EN, Bose CL, D'Ercole AJ. Chronic lung disease in infants with very low birth weight. *Am J Dis Child* 1987;**141**:784-8.
9 Horbar JD, McAuliffe TL, Adler SM, *et al.* Variability in 28-day outcomes for very low birth weight infants: an analysis of 11 neonatal intensive care units. *Pediatrics* 1988;**82**:554-9.
10 Horn SD. Validity, reliability and implications of an index of inpatient severity of illness. *Med Care* 1981;**19**:354-62.
11 Knaus WA, Draper EA, Wagner DP, Zimmerman JE. APACHE II: a severity of disease classification system. *Crit Care Med* 1985;**13**:818-29.
12 Pollack MM, Ruttmann UR, Getson PR, *et al.* Accurate prediction of the outcome of pediatric intensive care. A new quantitative method. *N Engl J Med* 1987;**316**:134-9.
13 Maddock CR. A population based evaluation of sustained mechanical ventilation of newborn babies. *Lancet* 1987;ii:1254-8.
14 Maddock CR, Carpenter RG, Gardner A. Mechanical ventilation for the newborn. *Lancet* 1988;i:707.
15 Saigal S, Rosenbaum P, Hattersley B, Milner R. Decreased disability rate among 3 year old survivors weighing 501 to 1000 grams at birth and born to residents of a geographically defined region from 1981 to 1984 compared with 1977 to 1980. *J Pediatr* 1989;**114**:839-46.
16 Tarnow-Mordi WO, Narang A, Wilkinson AR. Lack of association between barotrauma and air leak in hyaline membrane disease. *Arch Dis Child* 1985;**60**:555-9.
17 Horbar JD. A calculator program for determining indices of neonatal respiratory distress syndrome severity. *Am J Perinatol* 1987;**4**:20-3.
18 Rojas J, Green RS, Fannon L, *et al.* A quantitative model for hyaline membrane disease. *Pediatr Res* 1982;**16**:35-9.
19 Kwong MS, Egan EA, Notter RH, Shapiro DL. A double blind clinical trial of a calf lung surfactant extract for the prevention of hyaline membrane disease in extremely premature infants. *Pediatrics* 1985;**76**:585-92.
20 Enhorning G, Shennan A, Possmayer F, *et al.* Prevention of neonatal respiratory distress syndrome by tracheal instillation of surfactant: a randomized clinical trial. *Pediatrics* 1985;**76**:145-53.
21 Hallman M, Merritt TA, Jarvenpaa AL, *et al.* Exogenous human surfactant for treatment of severe respiratory distress syndrome: a randomized prospective clinical trial. *J Pediatr* 1985;**106**:963-9.
22 Baker R, Nelder JA. *The GLIM system, version 3.* Oxford: Numerical Algorithms Group, 1978.
23 Patterson CC, Halliday HL. Prediction of outcome after delivery for the very low birthweight (≤1500 g) infant. *Paediatric and Perinatal Epidemiology* 1988;**2**:221-8.
24 Gilbert R, Keighley JF. The arterial/alveolar oxygen tension ratio. An index of gas exchange applicable to varying inspired oxygen concentrations. *Am Rev Respir Dis* 1974;**109**:142-5.

Appendix

The regression equation derived from pooled data on all infants for risk of death in either hospital was:

$$\begin{aligned}\log(R/1-R) = {} & 54{\cdot}1\ (20{\cdot}4) \times 1 - 0{\cdot}11\ (0{\cdot}11) \times \text{gestation} \\ & - 0{\cdot}36\ (0{\cdot}41) \times \text{sex} \\ & - 0{\cdot}004\ (0{\cdot}001) \times \text{birth weight} \\ & - 0{\cdot}36\ (0{\cdot}47) \times \text{respiratory distress syndrome} \\ & + 0{\cdot}05\ (0{\cdot}01) \times F_IO_2\ (+2{\cdot}71\ (1{\cdot}4) \text{ if } F_IO_2 \text{ missing}) \\ & - 6{\cdot}67\ (2{\cdot}68) \times \text{pH}\ (-47{\cdot}2\ (19{\cdot}5) \text{ if pH missing}) \\ & - 0{\cdot}23\ (0{\cdot}21) \times \text{year of birth} \\ & - 1{\cdot}89\ (0{\cdot}79) \times \text{hospital},\end{aligned}$$

where R is the risk of death in hospital; gestation is in completed weeks; male sex=0, female sex=1; birth weight is in grams; respiratory distress syndrome=0, no respiratory distress syndrome=1; F_IO_2=mean fractional inspired oxygen concentration (21-100) in the first 12 hours of life; pH is mean pH in the first 12 hours of life; year of birth is coded 1983=1, 1984=2, etc; hospital A=0, hospital B=1. The figures outside the brackets represent the estimates, those inside represent the standard errors of the estimates.

(Accepted 2 April 1990)