Predicting outcome in very low birthweight infants using an objective measure of illness severity and cranial ultrasound scanning

P W Fowlie, W O Tarnow-Mordi, C R Gould, D Strang

Abstract

Aim—To investigate the feasibility of developing an objective tool for predicting death and severe disability using routinely available data, including an objective measure of illness severity, in very low birthweight babies.

Method—A cohort study of 297 premature babies surviving the first three days of life was made. Predictive variables considered included birthweight, gestation, 3 day cranial ultrasound appearances and 3 day CRIB (clinical risk index for babies) score. Models were developed using regression techniques and positive predictive values (PPV) and likelihood ratios (LR) were calculated.

Results—On univariate analysis, birthweight, gestation, 3 day CRIB score and 3 day cranial ultrasound appearances were each associated with death. On multivariate analysis, 3 day CRIB score and 3 day cranial ultrasound appearances remained independently associated. A 3 day CRIB score > 4 along with intraventricular haemorrhage (IVH) grade 3 or 4 was associated with a PPV of 64% and an LR of 9.8 (95% confidence limits 3.5, 27.9). Only 3 day CRIB score and 3 day cranial ultrasound appearances were associated with severe disability on univariate analysis. Both remained independently associated on multivariate analysis. A 3 day CRIB score > 4 along with an IVH grade of 3 or 4 was associated with a PPV of 60% and an LR of 24.2 (95% CI 4.4, 133.3).

Conclusion—Incorporating objective measures of illness severity may improve current prediction of death and disability in premature infants.

(Arch Dis Child Fetal Neonatal Ed 1998; 78:F175–F178)

Keywords: severity of illness index; prematurity; CRIB score; positive predictive values; likelihood ratios

Neonatologists frequently estimate the likely prognosis for very premature infants. However, many neonatal deaths occur within the first two or three days of life usually because the baby succumbs despite continuing intensive care or because continuing treatment is electively withdrawn.1 As a result, these babies do not tend to present the same prognostic dilemmas as those who survive the immediate newborn period. For those who do survive the first three days, two questions are crucial: what are the chances that this baby will still die; and if this baby survives, what are the chances of severe disability? These questions usually arise when discussing prognosis with parents2 or when considering withdrawal of intensive treatment.3

Currently, clinicians amalgamate prognostic data from a variety of sources when predicting the outcome for any given baby.4–6 Simple objective risk factors such as birthweight and gestational age at birth are often used6 as well as objective evidence of brain damage obtained from the appearances on cranial ultrasound scans.7 Over the past decade several models and scoring systems have also been described.8–10 However, recent evidence suggests that birthweight may not be an accurate guide to prognosis beyond the first few days of life5 and most of the currently available scoring systems are too complex for use in routine clinical practice. Illness severity can also be an independent predictor of mortality2 and although neonatologists routinely use their clinical experience to assess subjectively just how sick any given baby might be, little work has been done to examine how objective measures of illness severity such as SNAP (Score for Neonatal Acute Physiology) or CRIB8,9 might be used in this context. We therefore investigated the possibility of developing a simple predictive model incorporating a measure of illness severity that might provide clinicians with an objective estimate of (a) the risk of death before discharge; or (b) the risk of severe disability at 18 months in premature babies surviving the first three days of neonatal intensive care. We deliberately explored only the use of data that are routinely available and simple: birthweight; gestational age at birth; CRIB score on day 3 of life, a simple valid and reliable measure of clinical risk and illness severity8,9; and appearances on cranial ultrasound scans on or around day 3 of life.

Method

This was a retrospective cohort study in infants weighing ≤ 1500 g at birth or born at < 31 weeks of gestation between 1988 and 1990 in six Scottish hospitals. Infants born by caesarean section were included but not those born outside the target hospitals and subsequently transferred into one of them. Three of the hospitals were regarded as tertiary referral centres. Data were abstracted from the medical records of each infant: birthweight; gestation; presence of non-lethal congenital malformation; the lowest appropriate concentration of inspired oxygen between 48 and 72 hours; the highest
appropriate concentration of inspired oxygen between 48 and 72 hours; and the worst base deficit between 48 and 72 hours. A CRIB score relating to day 3 of life, CRIB72, was then calculated according to the original scoring system. Where they were available, data reporting the cranial ultrasound appearances on or around day 3 of life were also recorded: a variety of ultrasound machines were used across the study centres although all images were reported according to the Papile classification. Data on disability were collected for each surviving infant at a corrected age of 18 months by mailing a questionnaire to the child’s health visitor. For the purposes of this study, we have defined a child with severe disability as one who cannot sit unsupported by 18 months and/or is blind, and/or is deaf, and/or is more than 12 months behind in any other field of development. Specific aspects of fine motor control, communication, or respiratory disease were not considered because of difficulties in case definition. Disability was therefore a dichotomous outcome: a child was either severely disabled at 18 months according to our definition or was not.

Predictive variables considered were birthweight, gestation, appearances on cranial ultrasound scans and CRIB score on day 3 of life. Birthweight, gestation, and CRIB score were treated either as continuous variables. Appearances on cranial ultrasound scans were treated either as ordinal data (grades 0–4) or, for simplicity, converted into a dichotomous outcome (grade 0–2 vs grades 3 or 4).

In those infants surviving until the end of day 3, predictive models for (i) subsequent death before hospital discharge and (ii) severe disability at 18 months in survivors were fitted to the data using regression techniques. Model performance was assessed by creating receiver operating characteristic (ROC) curves and calculating the area below the curve, A Z, as well as by the Hosmer-Lemeshow test of goodness of fit. Models were compared by comparing the areas under the respective ROC curves.

For use in clinical practice, models were transformed into $8 \times 2$ tables and positive predictive values and likelihood ratios were calculated along with their 95% confidence intervals. The positive predictive value here represents the number of infants who have a positive “test” and actually experience the outcome of interest (true positives) divided by the total number of infants with a positive “test.” The likelihood ratio is defined as the ratio of true positive “tests” to false positive “tests.”

### Results

Four hundred and twenty eight infants were born in the six units over the study period. Forty six infants died in the first 72 hours of life, leaving 382 babies to contribute data to the analyses. Eighty five of these infants had either missing CRIB data or missing ultrasound data, leaving 297 infants contributing to the models predicting death. Of 254 infants surviving to age 2 with CRIB scores and ultrasound data, 14 were not assessed at 18 months, leaving 240 infants for analysis of models predicting severe disability. The characteristics of babies with and without missing data are shown in table 1. Forty three babies died after 3 days of life: median age at death day 13, range day 4 to day 308. The common pathologies leading to death included respiratory distress syndrome (n=22), intraventricular haemorrhage (n=15), and pulmonary interstitial emphysema (n=5). Of the 14 severely disabled children included in the analysis, five were blind, four were deaf, seven could not sit unsupported at 18 months and nine had clinically significant delay in another area of development. Most cranial ultrasound scans were performed on day 3 of life: median age at death day 13, range day 2 of life, median day 3 of life (quartiles day 2 of life, day 5 of life). Cases were excluded because the notes could not be traced (n=12), the physiological data required to calculate the CRIB score were missing (n=21), no cranial ultrasound scan had been performed, or follow up data were not available (n=22).

### Table 1  Characteristics of study infants (those infants excluded because of missing data are compared with those with complete data)

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Complete data</th>
<th>Missing data</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male:female</td>
<td>1:1</td>
<td>1:1.1</td>
<td>85</td>
</tr>
<tr>
<td>Median gestation (range)</td>
<td>29 (24–37)</td>
<td>297 31 (24–38)</td>
<td>85 &lt;0.001†</td>
</tr>
<tr>
<td>Median birth weight (range)</td>
<td>1170 (560–1500)</td>
<td>297 1340 (770–1500)</td>
<td>85 &lt;0.001†</td>
</tr>
<tr>
<td>Median 5 minute Apgar score (range)</td>
<td>9 (0–10)</td>
<td>297 9 (3–10)</td>
<td>85 &lt;0.001†</td>
</tr>
<tr>
<td>Median day 3 CRIB score, CRIB72 (range)</td>
<td>2 (0–17)</td>
<td>297 1 (1–5)</td>
<td>64 &lt;0.001‡</td>
</tr>
<tr>
<td>Number of infants with grade III or IV IVH on day 3</td>
<td>27 (9%)</td>
<td>297 1 (7%)</td>
<td>14 0.54**</td>
</tr>
<tr>
<td>Died</td>
<td>43 (14%)</td>
<td>297 6 (7%)</td>
<td>85 0.10**</td>
</tr>
<tr>
<td>Severe disability at 2 years in survivors‡</td>
<td>14 (5%)</td>
<td>283 1 (1%)</td>
<td>77 0.21**</td>
</tr>
</tbody>
</table>

* The ultrasound appearances have been treated as dichotomous (IVH grade 0, 1 or 2 vs grade 3 or 4) in the model reported for simplicity. The relation holds when the variable is treated as ordinal (IVH grade 0–4).

### Table 2  Accuracy and validity of models predicting death before discharge

<table>
<thead>
<tr>
<th>Predictive model</th>
<th>Area under the ROC curve $A_Z$ (SE)</th>
<th>Hosmer-Lemeshow $y^2$ ($p$ value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log odds death = -2.0044 + (0.0034 x birth weight)</td>
<td>0.70 (0.05)</td>
<td>0.32 (0.50)</td>
</tr>
<tr>
<td>Log odds death = -10.7097 + (0.4423 x gestation)</td>
<td>0.74 (0.08)</td>
<td>0.27 (0.51)</td>
</tr>
<tr>
<td>Log odds death 4.0334 - (1.6584 x USS*) - (0.4004 x CRIB72)</td>
<td>0.89 (0.05)</td>
<td>5.57 (0.85)</td>
</tr>
</tbody>
</table>

* The ultrasound appearances have been treated as dichotomous (IVH grade 0, 1 or 2 vs grade 3 or 4) in the model reported for simplicity. The relation holds when the variable is treated as ordinal (IVH grade 0–4).
Predicting outcome in very low birthweight infants

Table 3  Positive predictive values (PPV) and likelihood ratios (LR) associated with model predicting death before discharge

<table>
<thead>
<tr>
<th>“Test”</th>
<th>Died</th>
<th>Survived</th>
<th>PPV</th>
<th>LR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRIB72 &gt; 4 plus IVH grade 3 or 4</td>
<td>9</td>
<td>5</td>
<td>64%</td>
<td>9.8 (3.5-27.9)</td>
</tr>
<tr>
<td>CRIB72 = 2-4 plus IVH grade 3 or 4</td>
<td>3</td>
<td>3</td>
<td>50%</td>
<td>5.5 (1.1-26.2)</td>
</tr>
<tr>
<td>CRIB72 = 1 plus IVH grade 3 or 4</td>
<td>1</td>
<td>6</td>
<td>14%</td>
<td>0.91 (0.11-7.4)</td>
</tr>
<tr>
<td>CRIB72 = 0 plus IVH grade 3 or 4</td>
<td>0</td>
<td>0</td>
<td>0%</td>
<td>0.0 (0.0-1.0)</td>
</tr>
<tr>
<td>CRIB72 &gt; 4 plus IVH grade 0, 1 or 2</td>
<td>24</td>
<td>44</td>
<td>35%</td>
<td>3.0 (2.0-4.3)</td>
</tr>
<tr>
<td>CRIB72 = 2-4 plus IVH grade 0, 1 or 2</td>
<td>5</td>
<td>74</td>
<td>6%</td>
<td>0.37 (0.16-0.86)</td>
</tr>
<tr>
<td>CRIB72 = 1 plus IVH grade 0, 1 or 2</td>
<td>2</td>
<td>74</td>
<td>3%</td>
<td>0.15 (0.04-0.58)</td>
</tr>
<tr>
<td>CRIB72 = 0 plus IVH grade 0, 1 or 2</td>
<td>0</td>
<td>34</td>
<td>0%</td>
<td>0.08 (0.01-1.3)</td>
</tr>
</tbody>
</table>

Table 4  Positive predictive values (PPV) and likelihood ratios (LR) associated with model predicting severe disability at 2 years

<table>
<thead>
<tr>
<th>“Test”</th>
<th>Severe disability</th>
<th>No severe disability</th>
<th>PPV</th>
<th>LR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRIB72 &gt; 4 plus IVH grade 3 or 4</td>
<td>3</td>
<td>2</td>
<td>60%</td>
<td>24.2 (4.4-133.3)</td>
</tr>
<tr>
<td>CRIB72 = 2-4 plus IVH grade 3 or 4</td>
<td>0</td>
<td>3</td>
<td>0%</td>
<td>2.2 (0.12-41.2)</td>
</tr>
<tr>
<td>CRIB72 = 1 plus IVH grade 3 or 4</td>
<td>0</td>
<td>6</td>
<td>0%</td>
<td>1.2 (0.07-20.3)</td>
</tr>
<tr>
<td>CRIB72 = 0 plus IVH grade 3 or 4</td>
<td>0</td>
<td>0</td>
<td>0%</td>
<td>0.0 (0.0-1.0)</td>
</tr>
<tr>
<td>CRIB72 &gt; 4 plus IVH grade 0, 1 or 2</td>
<td>5</td>
<td>39</td>
<td>11%</td>
<td>2.1 (1.0-4.4)</td>
</tr>
<tr>
<td>CRIB72 = 2-4 plus IVH grade 0, 1 or 2</td>
<td>3</td>
<td>71</td>
<td>4%</td>
<td>0.70 (0.25-1.90)</td>
</tr>
<tr>
<td>CRIB72 = 1 plus IVH grade 0, 1 or 2</td>
<td>1</td>
<td>73</td>
<td>1%</td>
<td>0.20 (0.03-1.50)</td>
</tr>
<tr>
<td>CRIB72 = 0 plus IVH grade 0, 1 or 2</td>
<td>2</td>
<td>32</td>
<td>6%</td>
<td>1.00 (0.27-3.8)</td>
</tr>
</tbody>
</table>

While birthweight (p = 0.9345) and gestation (p = 0.1778) were no longer significantly associated with the risk of death. The univariate models using birthweight and gestational age at birth and the multivariate model using CRIB72 and ultrasound appearances are described in table 2. All three models achieved satisfactory goodness of fit using the Hosmer-Lemeshow χ² statistic. In addition, the area under the curve associated with the model combining CRIB72 and ultrasound appearance was significantly greater than the area associated with both the model based on birth weight alone (p = 0.001) and the model based on gestation alone (p = 0.0436).

Table 3 shows the positive predictive values and likelihood ratios associated with the model based on CRIB72 and ultrasound appearances.

MODEL PREDICTING SEVERE DISABILITY AT 18 MONTHS CORRECTED AGE

On univariate analysis, CRIB72 (p = 0.0051) and ultrasound appearance (p = 0.0193) were both significantly associated with the risk of being severely disabled at 18 months. Neither birthweight (p = 0.0755) nor gestation (p = 0.0719) were associated with this outcome. On multiple regression analysis, both CRIB72 (p = 0.0293) and ultrasound appearance (p = 0.0405) remained independently associated with the risk of severe disability: Log odds severe disability = 3.5871 –(0.1882×CRIB72)–(1.5241×USSS); Az(SE) = 0.675411 (0.133221); Hosmer-Lemeshow χ² = 5.089515; p = 0.89. The corresponding predictive values and likelihood ratios for this model are shown in table 4.

Discussion

In this cohort 48% of the babies who died before discharge from the neonatal unit did so within the first 72 hours of life. This is a smaller proportion than was found in the study by Meadow et al probably because our population was heavier and more mature, but still justifies our decision to concentrate on developing a model(s) predicting the outcome for very low birthweight or preterm babies only if they have survived the first three days of life. We also concentrated on developing models that might help provide answers to highly relevant clinical questions: (1) what are the chances that this baby will survive to be discharged home; and (2) if this baby does survive to go home, what are the chances that s/he will be severely disabled? This approach to predicting outcome has not to our knowledge been reported elsewhere, although others have tried to predict outcome before birth, at or shortly after birth, and at discharge home.

We believe we have gone some way towards demonstrating the feasibility of developing simple clinically useful models. We specifically tried to identify those infants we considered to have severe disability as it is a concept that clinicians, parents, and lay people readily understand, and it seems to be clinically and practically most relevant when considering prognosis. Cranial ultrasonography is a valid and reliable tool for assessing brain damage and is an accepted routine investigation in neonatal practice; and the CRIB score is a simple, robust instrument for measuring clinical risk and illness severity which is calculated easily using only six variables, all of which are usually routinely available. However, certain methodological issues prevent the immediate application of our models. Although the infants with missing data were similar in their characteristics and outcomes, we cannot be sure what effect their exclusion had on the overall results. Neonatal practice has also changed dramatically over the past 10 years with the introduction of antenatal steroids and surfactant treatment, and although CRIB remains a valid score (G Parry, personal communication), it is not clear what effect this might have on the models presented here. It is also recognised that treatment may sometimes be electively withdrawn and that this practice can be influenced by cranial ultrasound appearances. Without being able to adjust for this practice, our results might, therefore, overestimate the predictive capacity of the models presented. A prospective study that takes these issues into account and validates the models described is therefore needed.

As there is good evidence that neonatal survival has improved, the positive (and negative) predictive values from this study will have changed and there would be advantages to considering the likelihood ratio rather than the predictive value when trying to predict outcome. Assuming the validity of the models is maintained and the current death rate and prevalence of disability for any given population are known, the risk of either outcome can be extrapolated using the likelihood ratio along with a simple nomogram for Bayes’s theorem.
respectively) to alter significantly the post test probability.

Our results suggest that assessment of illness severity permits a more accurate estimate of prognosis than use of birthweight, gestation, or cranial ultrasound scans alone. However, no single piece of evidence can provide a definitive estimate of prognosis for individual babies. Any score is based on population characteristics that may not pertain to the individual baby, relates to outcome data that are by definition already “out of date,” and, when derived from relatively small cohorts, will give estimates which lie within wide confidence limits. Clinicians’ clinical judgement is a powerful tool in making clinical decisions, particularly about withdrawal of active treatment, and we would encourage clinicians to continue to use an holistic approach.

14 Patterson CC, Halliday HL. Predictions of outcome shortly after delivery for the very low birth weight (less than or equal to 1500g) infant. *Pediatr Perinat Epidemiol* 1988;2:221-8.
Predicting outcome in very low birthweight infants using an objective measure of illness severity and cranial ultrasound scanning

P W Fowlie, W O Tarnow-Mordi, C R Gould and D Strang

Arch Dis Child Fetal Neonatal Ed 1998 78: F175-F178
doi: 10.1136/fn.78.3.F175

Updated information and services can be found at:
http://fn.bmj.com/content/78/3/F175

These include:

References
This article cites 36 articles, 9 of which you can access for free at:
http://fn.bmj.com/content/78/3/F175#BIBL

Email alerting service
Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Topic Collections
Articles on similar topics can be found in the following collections

- Clinical diagnostic tests (720)
- Radiology (692)
- Radiology (diagnostics) (637)
- Child health (1515)
- Infant health (857)
- Neonatal health (928)
- Epidemiologic studies (929)

Notes

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/