Editorial comment

Predictions in the intensive care unit – in search of an “unholy grail”!

I. A. Murdoch¹ and D. J. Bihari²

¹ Paediatric Intensive Care
² Adult Intensive Care, Guy’s Hospital, St. Thomas’ Street, London SE1 9RT, UK

Received July 3, 1993; received in revised form July 20, 1993; accepted August 4, 1993

Key words: Intensive care – Scoring system – Patient assessment

Various scoring systems have been developed for the classification of adult and paediatric critically ill patients admitted to the intensive care unit (ICU) [1, 2]. These systems have been introduced for a number of reasons. Firstly, it is important to have an objective measure of severity of illness so that the casemix of an individual ICU can be described accurately. This is essential for the purposes of medical audit and becomes vital in the comparison of cohorts of patients entered into controlled clinical trials. Secondly, in the process of "resource management" that is so fashionable in the present financial climate it is necessary to measure both patient dependency and levels of therapeutic intervention in relation to severity of illness so as to obtain some understanding of the cost-effectiveness of intensive care in general, and more specifically to allow some comparison of individual ICUs. Finally, and perhaps the "holy" or "unholy" grail for those who care to speculate about life and death in the ICU, is the prediction of outcome and particularly death in individual cases [3, 4]. This last objective is somewhat emotive and has resulted in much debate within and outside the medical community [5, 6].

Whilst scoring systems for the measurement of the severity of a specific condition, e.g. head injury [7] or acute pancreatitis [8, 9], have been available for some time, the first attempts at developing some general system for the classification of patients in the ICU were those of Cullen et al. [10] in 1974. The clinical classification system (CCS) is a qualitative assessment of severity of illness made on admission, whereas the therapeutic intervention scoring system (TISS) is a quantitative assessment of the amount of therapy administered to each patient [11]. These systems were initially developed for the assessment of adult critically ill patients, but evaluation of both CCS and TISS has been performed successfully in paediatric ICUs in North America [12] and Europe [13]. CCS is inherently unsatisfactory since it relies upon a subjective assessment of the patient’s condition. In contrast, an updated version of TISS is extremely useful for the measurement of therapeutic intervention which is related both to outcome (Table 1) and cost. Indeed, in our adult practice we have found that it is possible to recharge the costs of intensive care within an internal market, in our hospital, according to the daily TISS score of individual patients [14]. Nevertheless, the intensity of treatment of patients in the ICU also reflects local custom and practice and there is an undesirable tendency for individual ICUs to adapt the TISS to their own needs. This makes cross-unit comparison – one of the goals of the system – difficult, if not impossible!

These difficulties together with the qualitative nature of the CCS argue for the development of scoring systems which are based upon well defined abnormalities in physiological function. In general there have been two approaches to the problem; firstly, some investigators have carefully defined "organ failure" or "organ dysfunction" and related mortality to the number and duration of organ systems in failure [15–18]. Secondly, systems have been built around deviations from normal of physiological and biochemical variables in relation to diagnosis, age and chronic health status. Whilst the former "multiple organ failure" scores are easy to use, they lose sensitivity and specificity by ignoring diagnosis [19]. Of the latter, the Simplified Acute Physiological Score (SAPS) [20] and the Acute Physiologic and Chronic Health Evaluation score (APACHE II) [21], both designed to measure severity of illness in adults, are now widely used. The APACHE II system, the acute physiology score being based upon abnormalities in 12 simple variables occurring within the first 24 h of admission to the ICU, evolved from the APACHE I score [22]. The latter (APACHE I) was found to be too complex (34 variables) for use in the general clinical setting. The APACHE II system has now been validated.
adapted the acute physiology score from APACHE II to use in children. In this system, extra points for severity of illness were allocated to children with a raised creatinine and an abnormal urine output. Given these limitations, it is not surprising that PRISM is unhelpful in the assessment of children with acute renal failure who will not benefit from dialysis. For indeed, it has to be emphasised that the vital prediction to make is not survival but death! A prediction of death with a 100% specificity and no false positive diagnosis would allow the early withholding or withdrawal of life support.

The tools for the measurement of severity of illness in the ICU are at the present time undoubtedly crude! They are useful for the assessment of cohorts of patients and they are continually being refined. There is now considerable interest in “dynamic” scoring – the assessment of day to day changes in acute physiology and therapeutic intervention, e.g. the APACHE III system, the Riyadh Intensive Care Program – but the ”holy grail” of individual outcome prediction continues to allude us.

References


---

Table 1. The outcome, intensive care unit (ICU) mortality, based upon the mean daily therapeutic intervention scoring system (TISS) point score incurred by individual patients seen in our combined paediatric general and cardiothoracic ICU over the last 2 months

<table>
<thead>
<tr>
<th>TISS points</th>
<th>0–20</th>
<th>21–30</th>
<th>31–40</th>
<th>&gt;40</th>
</tr>
</thead>
<tbody>
<tr>
<td>All patients</td>
<td>0/30</td>
<td>0/22</td>
<td>4/34</td>
<td>6/13</td>
</tr>
<tr>
<td>Cardiac surgery patients</td>
<td>0/5</td>
<td>0/12</td>
<td>0/17</td>
<td>1/7</td>
</tr>
<tr>
<td>General ICU patients</td>
<td>0/25</td>
<td>0/10</td>
<td>4/17</td>
<td>5/6</td>
</tr>
</tbody>
</table>

---

world wide in critically ill patients more than 16 years of age and has become the "gold standard" in severity scoring of adult patients. The APACHE III system has become available [23] and appears to be even more predictive of outcome in individual cases than its older brother, especially when it is used dynamically from day to day. The problem in adult practice has become not "how to do it" but "which system to use", made more difficult by the various commercial interests that surround the marketing of software packages!

Similarly, a physiological scoring system, the Physiologic Stability Index (PSI) was developed for infants and children by Pollack et al. [24] and validated by Yeh et al. [25]. The PSI examined seven organ systems with 34 variables combined with 75 ranges. As was the case with APACHE I, it soon became obvious that the system included too many variables, many not being measured routinely in paediatric ICUs, since they were not thought clinically indicated. Therefore the same group developed the Paediatric Risk of Mortality score (PRISM) from their original PSI [26]. This score involves 14 commonly measured variables with 23 ranges and covers five organ systems – but ignores the kidney! In reducing the number of variables, the authors claim that neither the specificity nor the sensitivity to predict mortality is lost. Others have gone further and suggest that only 5 variables are required to maintain the specificity to predict mortality [27].

It goes without saying that all scoring systems have problems and none are 100% specific in the prediction of outcome of individual patients, especially those static scores obtained within the first 24 h of ICU admission. Nevertheless, the British Paediatric Intensive Care Society in association with the Intensive Care Society of Great Britain have recommended the use of the PRISM score as one part of the assessment of critically ill children. It is our opinion, however, that it will never be possible to use PRISM or a similar "static" scoring system on its own to determine the appropriateness or otherwise of life support, such as dialysis. Fargason and Langman, in this issue, point out the limitations of PRISM in assessing children with acute renal failure. It is of interest to note that the PRISM score only records the serum potassium concentrations as an indirect estimation of renal function. The simplification of the original PSI has lead to the loss of the blood urea, serum creatinine and hourly urine output as contributing factors to this measure of severity of illness. This point has been previously emphasised by Zobel et al. [28] who