End points for Phase II trials in intensive care: recommendations from the Australian and New Zealand Clinical Trials Group consensus panel meeting

Paul Young, Carol Hodgson, Joel Dulhunty, Manoj Saxena, Michael Bailey, Rinaldo Bellomo, Andrew Davies, Simon Finfer, Peter Kruger, Jeffrey Lipman, John Myburgh, Sandra Peake, Ian Seppelt, Stephen Streat, Rhiannon Tate, Steven Webb and the ANZICS Clinical Trials Group

ABSTRACT

Background: There is uncertainty about which end points should be used for Phase II trials in critically ill patients.

Objective: To systematically evaluate potential end points for Phase II trials in critically ill patients.

Design and setting: A report outlining a process of literature review and recommendations from a consensus meeting conducted on behalf of the Australian and New Zealand Intensive Care Society Clinical Trials Group (ANZICS CTG) in October 2011.

Results and conclusions: The consensus panel concluded that there are no adequately validated end points for Phase II trials in critically ill patients. However, the following were identified as potential Phase II end points: hospital-free days to Day 90, ICU-free days to Day 28, ventilator-free days to Day 28, cardiovascular support-free days to Day 28, and renal replacement therapy-free days to Day 28. We recommend that these end points be evaluated further.

Crit Care Resusc 2012; 14: 211–215

Mortality rates for patients managed in the intensive care unit are decreasing. Consequently, the sample sizes required to detect clinically meaningful reductions in mortality are increasing. Few Phase III trials have had sufficient power to detect small absolute reductions[1-4] in mortality. Many ICU trials seek to detect unrealistically large reductions in mortality, thereby reducing the required sample size but increasing the risk of discarding effective treatments and reporting non-significant results.[5] As a consequence, potentially useful interventions are either never tested in an appropriately powered Phase III trial[5-7] or require repeated studies before an adequately powered trial is conducted. One reason to use mortality as a primary end point in small trials is the lack of proven surrogate outcomes that can be used as primary end points for Phase II trials.[8-10]

Despite this, conducting a rigorously designed Phase II trial is an important step in the progression to an appropriately designed Phase III trial.[11] Ideally, primary end points for a Phase II trial should be proven surrogates for differences in mortality or other long-term clinically meaningful outcomes and should be easily measurable in an ICU setting. In addition, the primary end point in a Phase II trial should require a smaller sample size than a Phase III end point such as mortality.[10]

As the only existing guidelines that specify appropriate primary end points for Phase II trials in critically ill patients are for studies involving patients with adult respiratory distress syndrome (ARDS),[8] the Australian and New Zealand Intensive Care Society Clinical Trials Group (ANZICS CTG) established a working group to systematically evaluate potential end points in Phase II trials in critically ill patients and to provide recommendations to researchers undertaking such trials.

Methods

The recommendations were developed in three phases. In the first phase, the working group (PY, CH, JD, MS) conducted a detailed literature review. The abstracts of all randomised controlled trials (RCTs) in critically ill adults published in the three highest impact medical journals (*JAMA*, *The Lancet*, and the *New England Journal of Medicine*) between 1 January 2005 and 4 July 2011 were reviewed to identify all of the reported end points. Next, a comprehensive review was undertaken of all abstracts included in Medline or Embase on 4 July 2011 from RCTs involving critically ill adults, limited to studies published after 1 January 2000. Keywords used to identify relevant trials were "intensive care" or "critical care" or "critical illness". The search was limited to "English language" and "randomised controlled trial" and "all adult (19 plus years)". Full articles were reviewed if it was unclear from the abstract which end points were used or whether a trial was, in fact, an RCT.

From this search, a list of 50 potential end points for Phase II trials in the following six populations were identified: general ICU patients, patients with acute lung injury, patients with acute kidney injury, patients with sepsis, patients who had just undergone cardiac surgery, and patients receiving nutritional interventions.

The second phase involved short-listing the 50 potential end points and exploring the short list in more detail. To create the short list, 10 Australian and New Zealand critical care trialists, including a biostatistician, were asked to identify what they regarded as the 10 most relevant end points for ICU trials in general and for each of the six subgroups. A second survey of 10 critical care trialists (eight from the first survey and two others) was undertaken before the consensus meeting to document their opinions of the potential end points' reliability, validity, importance to clinicians, and relationship with mortality or other long-term outcomes. This survey and the subsequent consensus meeting discussions were framed around the COSMIN (Consensus-based standards for the selection of health measurement instruments) guidelines.[12] Practicality was considered in terms of inter- and intra-rater reliability and logistical issues around data collection. Face validity, structural validity, content validity, criterion validity and cross-cultural validity were considered as described in the COSMIN guidelines.[12]

For the short-listed end points, the working group examined articles identified by the earlier search strategy and extracted data that could be used to aid sample size calculations (eg, group means and standard deviations) and data pertaining to statistical methods used. Information relating to end point definition, practicality, reliability, validity and trial sample size was then summarised.

The third phase involved a consensus panel meeting to discuss the strengths and weaknesses of the short-listed and alternative candidate end points and to produce recommendations for Phase II trial end points. In preparation for this meeting, the working group circulated the panel with a summary of information from the literature review and survey. The meeting was attended by nine delegates and four working group members.

Results

The literature review process

The materials provided to the consensus panel are included in a supplementary appendix available online at <http://www.cicm.org.au/journal.php>. They include the framework for assessment of end points, the summary of critical care RCTs published in *JAMA*, *The Lancet* and the *New England Journal of Medicine*, and all identified RCTs that reported on the short-listed end points.

The survey

The five most commonly identified end points placed on the short list were (i) hospital length of stay; (ii) "ICU-free days"; (iii) "hospital-free days"; (iv) functional end points, including independent functional status at hospital discharge and 6-minute walk distance at hospital discharge; and (v) end points that measured duration of organ support, including cardiovascular support and renal replacement therapy.

The ANZICS CTG Phase II end points consensus panel meeting

On reviewing the information provided, panellists agreed that *there are no adequately validated primary end points for Phase II critical care trials*. Seven emerged during further discussions.

The first theme was that many critical care trials have used mortality as the primary end point *while postulating unrealistic differences in outcome*. The panel discussed the importance of mortality as an end point in Phase II trials. While recognising that "effective" therapies may be associated by chance with increased mortality in small Phase II trials, it was agreed that for potential candidate interventions to be considered suitable by funding agencies for study in a Phase III trial, they would usually show a signal *towards* reduced mortality in a Phase II study, along with possible benefit in other end points.

The second theme was the issue of how to deal with patients who die. This issue was exemplified through the discussion of hospital length of stay as a potential primary end point. It was recognised that interventions that increase mortality may reduce hospital length of stay, thereby creating the false impression that an intervention is having a beneficial effect. It was argued that, as interventions will generally only cause small changes in mortality, this issue will only be a minor consideration. In certain circumstances, competing risks regression analysis may be used to model the subhazard function of a failure event of primary interest in the presence of a competing failure event.[13] However, it was agreed that when using end points like hospital length of stay, it is preferable to report outcomes separately for both survivors and non-survivors. Other options for dealing with this problem were discussed, including (i) arbitrarily assigning patients who die a length of stay that is longer than the maximum length of stay observed in survivors, or (ii) using "free-days" (eg, ventilator-free days and ICU-free days) and assigning all patients who die zero free-days. The panel thought it logical to use the same approach for all end points of this type. As there is some statistical modelling supporting the use of ventilator-free days in acute lung injury,[14] and as this end point is increasingly used in a number of settings in high-profile articles,[15-23] it was decided that the preferred approach was to use free-days. Free-days end points may offer power advantages over mortality end points,[14] because they bring additional information from the ranking of outcomes among survivors: they are composite end points that combine the duration of support for survivors with mortality. Smaller sample sizes are therefore required for interventions that both reduce mortality and shorten the duration of support for survivors. It is reasonable to expect, although untested, that interventions that reduce mortality will do this.

The third theme was the importance of interventions that reduce overall cost. It was agreed that health research

funders are, to some degree, aligned to health service funders and that trial interventions that may be both effective and cost-saving are likely to be more attractive to fund in Phase III trials than interventions that may be effective but not cost-saving. It was agreed that using a Phase II end point that gives some insight into whether an intervention might reduce cost is important. If observed mortality differences are small, ICU-free days, ventilation-free days and hospital-free days may provide surrogate measures of health care resource utilisation and some costs.

The fourth theme was the possibility of bias due to dependence of the favoured end points on the operator. It was recognised that issues including exit block (delayed ICU discharge), access block (delayed ICU admission) and individual operational variations between institutions and countries could affect length of stay and may confound interpretation of measured variables. Although options such as using careful definitions (through protocols and guidelines) of "readiness for ICU discharge" or "readiness for extubation" were discussed, the associated data collection was considered too labour-intensive for these estimates to be practical. Of the end points discussed, it was thought that (acute) hospital length of stay or (acute) hospital-free days were least likely to be influenced by ICU staff and therefore the most free from potential bias. Of course, these outcomes may also be more greatly affected by confounders unrelated to the tested intervention.

The fifth theme was the importance of good functional survival. While it was agreed that functional outcomes were important, a number of issues precluded the group from recommending their general use as primary end points in Phase II trials. The use of survival plus functional independence as an outcome usually does not add statistical power compared with mortality, because it is often a dichotomous outcome. It may allow the postulation of greater effect sizes that would reduce sample size requirements, but a larger effect and a reduced sample size will not occur if there are significant numbers of "new" survivors as a result of an intervention and those survivors are not independent at the time of assessment. This is a particular issue when time frames for follow-up are short, as is likely to be the case in a Phase II trial. Generally, functional end points have only been used in ICU trials of physical therapies.

A sixth theme was the use of biomarkers as primary end points in Phase II trials. The consensus panel did not favour this, and noted that biomarkers have been very difficult to validate (as they may not be in the causal pathway of a disease) and have been shown to be unreliable in many studies.[24-26] They might be important for Phase IIa trials — for example, to show biological activity — but the panel agreed they should not be used alone to determine whether to progress to a Phase III trial. Phase IIb trials in critically ill patients should primarily look at clinical effects.

Finally, composite end points were discussed. Generally, these were felt to be problematic because of the difficulty in assigning different weightings to the different end point components. These difficulties were felt to make the clinical implications of these end points difficult to establish.

Recommendations of the ANZICS CTG consensus panel

Recommendation #1

Primary end points for Phase II trials should be surrogates for one or more of the following: (i) 90-day mortality, (ii) medium or long-term health-related quality of life, and (iii) cost of treatment.

Recommendation #2

The following are recommended as potential and testable primary end points for Phase II trials:

- Hospital-free days to Day 90
- ICU-free days to Day 28
- Ventilator-free days to Day 28
- Cardiovascular support-free days to Day 28
- Renal replacement therapy-free days to Day 28

Recommendation #3

The choice of a primary end point in a particular trial will depend on the population being studied and, in particular populations (eg, patients with brain injury), alternative end points may be more appropriate than those recommended above.

Recommendation #4

The main purpose of specifying a single primary end point for a Phase II trial is to justify sample size. However, Phase II trials should also report on a number of secondary end points relevant to a Phase III trial, including 90-day mortality. As sample size calculations are based solely on the primary outcome, secondary outcomes in Phase II trials may be underpowered. However, in order to fully inform readers, it will often be beneficial to provide power calculations for secondary outcomes.

Recommendation #5

When designing Phase II trials, investigators should consider that the ultimate decision as to whether to progress to a Phase III trial depends on the results of a range of end points (not just the primary end point) measured in a Phase II trial.

Recommendation #6

If a "free-day" method (a composite of survival and absence of the receipt of that particular resource) is used as the primary end point, duration of support or length of stay

should be reported for all patients and separately for survivors and non-survivors.

Recommendation #7

For trials investigating interventions involving mobility, functional end points may be used for Phase II studies, but data are currently insufficient to support their use as primary end points for trials of other interventions.

Recommendation #8

Because functional end points should often be assessed in Phase III trials, use of these end points in Phase II trials is recommended as a preliminary data collection method.

Recommendation #9

Findings from Phase II trials can be misleading and should not be primarily used to change clinical practice. In particular, "free-day" end points are problematic when used to infer clinically important differences, because they encompass both duration of support for survivors and mortality, which are not of equal importance to patients. However, in circumstances where mortality rates are similar, ICU-free days, ventilation-free days and hospital-free days may provide surrogate measures of some costs.

Discussion

A Phase II trial is an important step in the evolution of an appropriately designed Phase III trial. However, there are currently no end points that are well validated or that adequately predict subsequent differences in mortality or long-term clinically meaningful outcomes.

Although the process of validating Phase II end points has been hindered in part by the relative paucity of Phase III trials that report a treatment effect, four landmark studies in critically ill patients (CRASH,[3] CRASH-2,[4] ARDSNet low tidal volume study[27] and NICE-SUGAR[2]) have reported differences between the control and intervention groups, representing an opportunity to explore our recommendations using their data. Interrogation of these study databases may provide valuable information regarding the relationship between our recommended Phase II end points and Phase III outcomes. Furthermore, future Phase III trials showing a difference between the control and intervention groups could potentially help define the relationship between Phase II and Phase III end points, ideally by incorporating a pre-planned substudy in their design. While statistical modelling exists to support the use of ventilator-free days as an end point,[14] similar data are required for more generalisable endpoints, such as ICU-free days and hospital-free days.

Conclusions

Following a process that combined a structured literature review, a survey of the opinions of critical care trialists and a consensus panel meeting, we have made recommendations regarding end points for use in Phase II trials. The consensus panel recommends hospital-free days to Day 90, ICU-free days to Day 28, ventilator-free days to Day 28, cardiovascular support-free days to Day 28, and renal replacement therapy-free days to Day 28 as end points to be considered and tested for Phase II trials.

Acknowledgements

Working group: Paul Young (Chair), Carol Hodgson, Joel Dulhunty and Manoj Saxena, with assistance and input from Andrew Davies and Steven Webb.

Survey 1: Michael Bailey, Rinaldo Bellomo, Robert Boots, Andrew Davies, Simon Finfer, Peter Kruger, Jeffrey Lipman, John Myburgh, Sandra Peake and Steven Webb.

Survey 2: Michael Bailey, Rinaldo Bellomo, Andrew Davies, Simon Finfer, Peter Kruger, Jeffrey Lipman, Alistair Nichol, Sandra Peake, Ian Seppelt and Steven Webb.

Consensus panel: Paul Young (Chair), Michael Bailey, Andrew Davies, Joel Dulhunty, Simon Finfer, Carol Hodgson, Peter Kruger, Jeffrey Lipman, Sandra Peake, Manoj Saxena, Ian Seppelt, Stephen Streat and Rhiannon Tate (administrative support).

Competing interests

None declared.

Author details

Paul Young, Intensive Care Specialist,[1] and Honorary Senior Research Fellow[2]
Carol Hodgson, Senior Research Fellow,[3] and Senior Physiotherapist[4]
Joel Dulhunty, Research Fellow[5,6]
Manoj Saxena, Senior Research Fellow,[7] and Intensive Care Specialist[8]
Michael Bailey, Statistician[3]
Rinaldo Bellomo, Co-Director,[3] and Intensive Care Specialist and Director of Research[9]
Andrew Davies, Intensive Care Specialist[4]
Simon Finfer, Professorial Fellow,[7] Professor of Medicine,[10] and Specialist[11]
Peter Kruger, Associate Professor,[6] and Deputy Director[12]
Jeffrey Lipman, Director,[5] and Executive Director[6]
John Myburgh, Director,[7] and Intensive Care Specialist[8]
Sandra Peake, Intensive Care Specialist[13]
Ian Seppelt, Honorary Fellow,[7] and Intensive Care Specialist[14]
Stephen Streat, Intensive Care Specialist[15]
Rhiannon Tate, Executive Officer[16]
Steven Webb, Intensive Care Specialist[17]
and the ANZICS Clinical Trials Group

1 Intensive Care Unit, Wellington Regional Hospital, Capital and Coast District Health Board, Wellington, New Zealand.

2 Medical Research Institute of New Zealand, Wellington, New Zealand.
3 Australian and New Zealand Intensive Care Research Centre, School of Public Health and Preventive Medicine, Monash University, Melbourne, VIC, Australia.
4 Intensive Care Unit, Alfred Hospital, Melbourne, VIC, Australia.
5 Department of Intensive Care Medicine, Royal Brisbane and Women's Hospital, Brisbane, QLD, Australia.
6 Burns, Trauma and Critical Care Research Centre, University of Queensland, Brisbane, QLD, Australia.
7 Division of Critical Care and Trauma, George Institute for Global Health, Sydney, NSW, Australia.
8 St George Clinical School, University of New South Wales, Sydney, NSW, Australia.
9 Intensive Care Unit, Austin Hospital, Melbourne, VIC, Australia.
10 Sydney Medical School, University of Sydney, Sydney, NSW, Australia.
11 Intensive Care Unit, Royal North Shore Hospital, Sydney, NSW, Australia.
12 Department of Intensive Care, Princess Alexandra Hospital, Brisbane, QLD, Australia.
13 Department of Intensive Care, Queen Elizabeth Hospital, Adelaide, SA, Australia.
14 Intensive Care Unit, Nepean Hospital, Sydney, NSW, Australia.
15 Department of Critical Care Medicine, Auckland City Hospital, Auckland, New Zealand.
16 Australian and New Zealand Intensive Care Society Clinical Trials Group, Melbourne, VIC, Australia.
17 Intensive Care Unit, Royal Perth Hospital, Perth, WA, Australia.

References

1 Finfer S, Bellomo R, Boyce N, et al. A comparison of albumin and saline for fluid resuscitation in the intensive care unit. *N Engl J Med* 2004; 350: 2247-56.
2 Finfer S, Chittock DR, Su SY, et al. Intensive versus conventional glucose control in critically ill patients. *N Engl J Med* 2009; 360: 1283-97.
3 Edwards P, Arango M, Balica L, et al. Final results of MRC CRASH, a randomised placebo-controlled trial of intravenous corticosteroid in adults with head injury-outcomes at 6 months. *Lancet* 2005; 365: 1957-9.
4 Shakur H, Roberts I, Bautista R, et al. Effects of tranexamic acid on death, vascular occlusive events, and blood transfusion in trauma patients with significant haemorrhage (CRASH-2): a randomised, placebo-controlled trial. *Lancet* 2010; 376: 23-32.
5 Aberegg SK, Richards DR, O'Brien JM. Delta inflation: a bias in the design of randomized controlled trials in critical care medicine. *Crit Care* 2010; 14: R77.
6 Ospina-Tascon GA, Buchele GL, Vincent JL. Multicenter, randomized, controlled trials evaluating mortality in intensive care: doomed to fail? *Crit Care Med* 2008; 36: 1311-22.
7 Rubenfeld GD, Abraham E. When is a negative phase II trial truly negative? *Am J Respir Crit Care Med* 2008; 178: 554-5.
8 Spragg RG, Bernard GR, Checkley W, et al. Beyond mortality: future clinical research in acute lung injury. *Am J Respir Crit Care Med* 2010; 181: 1121-7.
9 Reade MC, Angus DC. The clinical research enterprise in critical care: what's right, what's wrong, and what's ahead? *Crit Care Med* 2009; 37 (1 Suppl): S1-9.
10 Sevransky JE, Checkley W, Martin GS. Critical care trial design and interpretation: a primer. Critical care medicine 2010; 38: 1882-9.
11 McAuley DF, O'Kane C, Griffiths MJ. A stepwise approach to justify phase III randomized clinical trials and enhance the likelihood of a positive result. *Crit Care Med* 2010; 38 (10 Suppl): S523-7.
12 Mokkink LB, Terwee CB, Knol DL, et al. The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: a clarification of its content. *BMC Med Res Methodol* 2010; 10: 22.
13 Fine J, Gray JR. A proportional hazards model for the subdistribution of a competing risk. *J Am Stat Assoc* 1999; 94: 496-509.
14 Schoenfeld DA, Bernard GR. Statistical evaluation of ventilator-free days as an efficacy measure in clinical trials of treatments for acute respiratory distress syndrome. *Crit Care Med* 2002; 30: 1772-7.
15 Strom T, Martinussen T, Toft P. A protocol of no sedation for critically ill patients receiving mechanical ventilation: a randomised trial. *Lancet* 2010; 375: 475-80.
16 Ferrer M, Sellares J, Valencia M, et al. Non-invasive ventilation after extubation in hypercapnic patients with chronic respiratory disorders: randomised controlled trial. *Lancet* 2009; 374: 1082-8.
17 Schweickert WD, Pohlman MC, Pohlman AS, et al. Early physical and occupational therapy in mechanically ventilated, critically ill patients: a randomised controlled trial. *Lancet* 2009; 373: 1874-82.
18 Pandharipande PP, Pun BT, Herr DL, et al. Effect of sedation with dexmedetomidine vs lorazepam on acute brain dysfunction in mechanically ventilated patients: the MENDS randomized controlled trial. *JAMA* 2007; 298: 2644-53.
19 Mercat A, Richard JC, Vielle B, et al. Positive end-expiratory pressure setting in adults with acute lung injury and acute respiratory distress syndrome: a randomized controlled trial. *JAMA* 2008; 299: 646-55.
20 Papazian L, Forel JM, Gacouin A, et al. Neuromuscular blockers in early acute respiratory distress syndrome. *N Engl J Med* 2010; 363: 1107-16.
21 Wiedemann HP, Wheeler AP, Bernard GR, et al. Comparison of two fluid-management strategies in acute lung injury. *N Engl J Med* 2006; 354: 2564-75.
22 Wheeler AP, Bernard GR, Thompson BT, et al. Pulmonary-artery versus central venous catheter to guide treatment of acute lung injury. *N Engl J Med* 2006; 354: 2213-24.
23 Roquilly A, Mahe PJ, Seguin P, et al. Hydrocortisone therapy for patients with multiple trauma: the randomized controlled HYPOLYTE study. *JAMA* 2011; 305: 1201-9.
24 Barter PJ, Caulfield M, Eriksson M, et al. Effects of torcetrapib in patients at high risk for coronary events. *N Engl J Med* 2007; 357: 2109-22.
25 Calfee CS, Ware LB, Eisner MD, et al. Plasma receptor for advanced glycation end products and clinical outcomes in acute lung injury. *Thorax* 2008; 63: 1083-9.
26 Ware LB, Matthay MA, Parsons PE, et al. Pathogenetic and prognostic significance of altered coagulation and fibrinolysis in acute lung injury/acute respiratory distress syndrome. *Crit Care Med* 2007; 35: 1821-8.
27 Ventilation with lower tidal volumes as compared with traditional tidal volumes for acute lung injury and the acute respiratory distress syndrome. Acute Respiratory Distress Syndrome Network. *N Engl J Med* 2000; 342: 1301-8. ❑