Therapeutic hypothermia after cardiac arrest — once again

John L Moran and Patricia J Solomon

Two recent articles in this Journal on therapeutic hypothermia after cardiac arrest,[1,2] along with a systematic review[3] and an editorial,[4] have again raised the question of the appropriate place of this new treatment modality. Let us not mistake the import of this debate: it is proposed that "hypothermia is now standard care for some types of cardiac arrest".[4] This position is based largely on three trials — the Hypothermia after Cardiac Arrest (HACA) trial,[5] and trials by Bernard et al[6] and Hachimi-Idrissi et al[7] — and a meta-analysis of the three by Holzer et al.[3]

Appropriate cautions were argued by Skowronski[1] and ourselves in a previous review.[8] What then are the outstanding issues?

Selection bias

As pointed out by both Skowronski[1] and the advisory statement from the Advanced Life Support Task Force,[9] 92% of patients screened in the HACA study[5] were excluded from the trial. What were the consequences? The most important appears to be the exact type of patients enrolled, which is revealed by the mortality of the control group: in this case, for comparability, the hospital mortality, which was 50%. This was considerably less than the 68% found in the study of Bernard et al,[6] and the 83% found in the comparator Melbourne 2002 study for the subgroup with initial ventricular fibrillation rhythm, quoted by Skowronski.[1] The advisory statement from the Advanced Life Support Task Force, reviewing eligibility in the HACA study, noted that exclusions "had persistent hypotension . . . and causes of coma other than cardiac arrest".[9] However, a concurrent "population-level" study (OPALS) reported a mortality of 87% from the ventricular fibrillation subgroup ($n = 1819$).[10] The implications of these marked disparities in control-group survival have not been appropriately addressed by the proponents of hypothermia.

Efficacy

The recent meta-analysis of individual patient data by Holzer et al reported a risk ratio for hospital discharge with minimal or no neurological damage of 1.68 (95% CI, 1.29–2.07) from "three rather small trials".[3] Flather et al have suggested, regarding the size of meta-analyses, that "the results of small meta-analyses should be regarded with caution, even if the *P* value shows extreme statistical significance. Larger meta-analyses (ie, those with several hundred events) are likely to be more reliable and may be clinically useful".[11] Not having access to individual patient data, we repeated the meta-analysis using the standard grouped data approach (random effects estimator):[12] the treatment effect was comparable at a relative risk (RR) of 1.52 (95% CI, 1.19–1.95; $P = 0.001$). Of note was the weight of 84% given to the HACA study in the meta-analysis.[3] Only the HACA study followed up patients to 6 months, where the RR for being alive with favourable neurological recovery in the hypothermia group was 1.44 (95% CI, 1.11–1.76).[3] After controlling for baseline variables (age, sex and time from collapse to return of spontaneous circulation), this RR was reduced to 1.37 (95% CI, 1.02–1.72). Holzer et al commented on the low power of interaction tests;[12] approximately four times more patients are needed for satisfactory power in interaction assessments.[13,14] Thus, the question of modifying influences on treatment effects, a focal point of the Holzer et al meta-analysis ("we could assess whether hypothermia interacts with clinical variables"[3]), still remains uncertain.

Trial conduct

Issues remain unresolved regarding the trial conduct in the HACA[5] and Bernard et al[6] reports, and have been discussed previously.[8] In particular, the treatment effect estimates appeared subject to potential bias:

- In the Bernard et al study,[6] which used "pseudorandomisation",[9] an apparently unblinded interim analysis was conducted at enrolment of 62 patients, but no information was provided regarding the further conduct of the trial (for instance, were investigators blinded to the outcome of the final 15 patients?), and no adjustment was made in the final treatment estimate for this "interim" analysis.[15]
- The report of the HACA trial provided no details of prospective sample size, minimum clinically important difference, nor the mechanisms for stopping (other than that the trial was stopped because of low enrolment and lack of funding).[5] The most important (unanswered) question was: were the reasons for stopping informative of, or dependent on, the (favourable) outcome? A simple test of this would be the answer to the question: if the treatment effect at trial stopping had been $P = 0.09$ (instead of $P = 0.009$), would the HACA trialists have sought to continue the trial? If so, then the reasons for stopping would be informative, and the estimates biased.

Parenthetically, the Bernard et al trial[6] does not pass this test. From the perspective of the clinician, the question may be reasonably asked: if the two trials under question were drug trials, on the information presented, would the drug have been licensed by regulatory authorities?

Appropriate end-points

The primary outcome in the Bernard et al and HACA trials was favourable neurological outcome, at hospital discharge and 6 months post-discharge, respectively, but the effect on hospital mortality may be assessed for the three trials reported in the meta-analysis of Holzer et al.[3] Again, pooled data were used, and, perhaps not surprisingly, the RR for hospital mortality was favourable for hypothermia at 0.79 (95% CI, 0.66–0.94). As previously pointed out,[8] hospital deaths were fully characterised only in the Bernard et al trial, where no difference in the proportion of patients dying of cardiac versus cerebral causes was evident (P = 0.72). No information was provided regarding these events in the HACA trial,[5] and insufficient detail was provided by Hachimi-Idrissi et al.[7]

This raises the question of the appropriate analysis of favourable or unfavourable neurological events when the proportion of patients who die differs between the treatment and control groups. We detailed these matters previously,[8] and also suggested, in a comment on a recent review by Polderman of therapeutic hypothermia in the intensive care unit,[16] that there may be bias in the assessment of the primary outcome estimates (of neurological function) in the Bernard et al and HACA studies because of the problem of competing risks.[17] That is, a dead patient cannot qualify for assessment of neurological outcome. This position was vigorously contested by Polderman and Ware, who maintained that "both studies used well defined outcome measures".[18] Polderman and Ware further suggested that our position — which they characterised as "what is the probability of good neurological outcome in cooled patients compared with controls *given that the patient is not dead* [original emphasis]" — could not be addressed statistically, "because the number of surviving patients with severe disability was low in both studies".[18] The retrospective review of Mullner et al was also cited to provide evidence that "unfavourable neurological recovery is a very strong predictor of death".[19]

It is instructive to review the report by Mullner et al of a retrospective study which looked at the effect of prior comorbidity on outcome in 411 survivors of cardiac arrest.[19] Sixty percent of survivors had an adverse outcome, where the latter was defined as "mortality and unfavourable neurological recovery at 6-month follow-up". That is, it was a *composite* outcome.[20] Eighty-eight percent of patients with an initially unfavourable neurological recovery died within the 6-month follow-up, but no analysis of these deaths was provided. In Mullner et al's multivariate model of clinical conditions, only New York Heart Association functional class and age were associated with adverse outcome. If the cause of deaths in hospital survivors with an initially unfavourable neurological recovery was "cardiac" — which was likely in this series, given that 14%–45% of the cohort had at least one prior comorbidity (diabetes, congestive heart failure, hypertension and/or coronary heart disease) — then this is a classic scenario of competing risks, which, as we pointed out, may be appropriately analysed in a uni- or multivariate context.[17] If it is considered that the events of interest were not independent (a defensible position), then a multivariate survival approach may be used.[21]

A large and controversial body of literature attends the question of multiple and/or composite end-points,[22] as we discussed previously in this Journal.[23] Composite outcomes (the combination of multiple end-points) are frequently used in trials, exemplified in the above discussion. Freemantle et al surveyed composite primary outcomes that incorporated all-cause mortality in 167 trials published in major journals between 1997 and 2001.[24] The enumerated advantages were increased statistical precision and trial efficiency. The disadvantages were decreased trial efficiency if not all components are substantially affected by therapy, lack of benefit relating to the most important constituent end-point, and clinician-driven end-points being more amenable to change. They recommended that components of composite outcomes "always be defined as secondary outcomes and [be] reported alongside the results of the primary analysis". This appears consistent with our original recommendation that: "Future studies should report separate analyses for both deaths and scaled cerebral outcomes; the latter should be analysed as ordinal data",[8] and with a current review of analytic methods for composite end-points.[25]

In an editorial accompanying Freemantle et al's report, Lauer and Topol suggested that: "A composite end-point that includes death as well as non-fatal events is subject to biases relating to competing risks",[26] which claim also resonates with our critique cited above.[17] Lauer and Topol further conceptualised a scenario whereby "A treatment that leads to an increased risk of death may therefore appear to reduce the risk of non-fatal events". Such a scenario (or rather, its inverse) is *not* necessarily being argued; in the context of the three trials being considered,[5-7] the *expectation* is that individual end-points (hospital mortality, neurological performance) would be consistent in terms of treatment effects, albeit that there may be inequalities in estimates, as noted above. Our primary concern is the unbiased estimation of the overall primary outcome and, not unreasonably, individual end-points; this may only be obtained by appropriate analysis.

Where do we stand?

The clinician is faced with "three rather small" heterogeneous trials reporting atypical patient samples:[5-7] two trials were non-conclusive on primary analysis[6,7] (see Holzer et al, Table 2[3]); and the third, although demonstrating a favourable treatment effect, had an unsettling dearth of reported trial detail.[5] Faced with such uncertainty, clinicians may properly suspend judgement and say that they "just do not know". From the analytical perspective, we maintain our original position that "Class I evidence is not obtained for this intervention".[17]

Author details

John L Moran, Senior Consultant[1]
Patricia J Solomon, Associate Professor[2]

1 Department of Intensive Care Medicine, The Queen Elizabeth Hospital, Adelaide, SA.

2 School of Mathematical Sciences, The University of Adelaide, Adelaide, SA.

Correspondence: john.moran@nwahs.sa.gov.au

References

1 Skowronski GA. Therapeutic hypothermia after cardiac arrest — not so fast. *Crit Care Resusc* 2005; 7: 322-4.

2 Bernard SA. Hypothermia improves outcome from cardiac arrest. *Crit Care Resusc* 2005; 7: 325-7.

3 Holzer M, Bernard SA, Hachimi-Idrissi S, et al. Hypothermia for neuroprotection after cardiac arrest: systematic review and individual patient data meta-analysis. *Crit Care Med* 2005; 33: 414-8.

4 Bernard SA. Therapeutic hypothermia after cardiac arrest. *Med J Aust* 2004; 181: 468-9.

5 The Hypothermia After Cardiac Arrest Study Group. Mild therapeutic hypothermia to improve the neurologic outcome after cardiac arrest. *N Engl J Med* 2002; 346: 549-56.

6 Bernard SA, Gray TW, Buist MD, et al. Treatment of comatose survivors of out-of-hospital cardiac arrest with induced hypothermia. *N Engl J Med* 2002; 346: 557-63.

7 Hachimi-Idrissi S, Corne L, Ebinger G, et al. Mild hypothermia induced by a helmet device: a clinical feasibility study. *Resuscitation* 2001; 51: 275-81.

8 Moran JL, Peake SL, Solomon PJ. Hypothermia as therapy in cerebral injury. *Crit Care Resusc* 2002; 4: 93-102.

9 Nolan JP, Morley PT, Wanden Hoek TL, Hickey RW, ALS Task Force. Therapeutic hypothermia after cardiac arrest. An advisory statement by the Advanced Life Support Task Force of the International Liaison Committee on Resuscitation. *Resuscitation* 2003; 57: 231-5.

10 Stiell IG, Wells GA, Field B, et al. Advanced cardiac life support in out-of-hospital cardiac arrest. *N Engl J Med* 2004; 351: 647-56.

11 Flather MD, Farkouh ME, Pogue JM, Yusuf S. Strengths and limitations of meta-analysis: larger studies may be more reliable. *Control Clin Trials* 1997; 18: 568-79.

12 Bradburn MJ, Deeks J, Altman DG. metan-sbe24 an alternative meta-analysis command. *Stata Technical Bulletin Reprints* 1998; 8: 100.

13 Greenland S. Tests for interaction in epidemiologic studies: a review and a study of power. *Stat Med* 1983; 2: 243-51.

14 Greenland S. Basic problems in interaction assessment. *Environ Health Perspect* 1993; 101 Suppl 4: 59-66.

15 Moran JL, Peake SL, Solomon P. Reporting of clinical trials using group sequential methods. *Crit Care Resusc* 2001; 3: 146-7.

16 Polderman KH. Application of therapeutic hypothermia in the ICU: opportunities and pitfalls of a promising treatment modality. Part 1: Indications and evidence. *Intensive Care Med* 2004; 30: 556-75.

17 Moran JL, Solomon PJ. Application of therapeutic hypothermia in the intensive care unit [author reply]. *Intensive Care Med* 2004; 30: 2288.

18 Polderman KH, Ware JH. Therapeutic hypothermia following cardiac arrest: class I evidence. *Intensive Care Med* 2004; 30: 2287.

19 Mullner M, Sterz F, Behringer W, et al. The influence of chronic prearrest health conditions on mortality and functional neurological recovery in cardiac arrest survivors. *Am J Med* 1998; 104: 369-73.

20 Moye LA. Multiple analyses in clinical trials: fundamentals for investigators. New York: Springer-Verlag, 2003.

21 Therneau TM, Grambsch PM. Chapter 8: Multiple events per subject. Modeling survival data: extending the Cox model. New York: Springer-Verlag, 2000: 169-229.

22 Sankoh AJ, D'Agostino RB, Huque MF. Efficacy endpoint selection and multiplicity adjustment methods in clinical trials with inherent multiple endpoint issues. *Stat Med* 2003; 22: 3133-50.

23 Moran JL, Solomon PJ. Selective digestive decontamination: once again. *Crit Care Resusc* 2003; 5: 241-6.

24 Freemantle N, Calvert M, Wood J, et al. Composite outcomes in randomized trials: greater precision but with greater uncertainty? *JAMA* 2003; 289: 2554-9.

25 Glynn RJ, Rosner B. Methods to evaluate risks for composite end points and their individual components. *J Clin Epidemiol* 2004; 57: 113-22.

26 Lauer MS, Topol EJ. Clinical trials — multiple treatments, multiple end points, and multiple lessons. *JAMA* 2003; 289: 2575-7.

27 The National Heart Lung and Blood Institute Acute Respiratory Distress Syndrome Clinical Trials Network. Efficacy and safety of corticosteroids for persistent acute respiratory distress syndrome. *N Engl J Med* 2006; 354: 1671-84.

28 Gray RJ. A class of K-sample tests for comparing the cumulative incidence of a competing risk. *Ann Stat* 1988; 16: 1141-54.

29 Alberti C, Metivier F, Landais P, et al. Improving estimates of event incidence over time in populations exposed to other events: application to three large databases. *J Clin Epidemiol* 2003; 56: 536-45.

30 Klein JP, Bajorunaite R. Chapter 16. Inference for competing risks. In: Balakrishnana N, Rao CR. Advances in survival analysis: Handbook of statistics Vol 23. Amsterdam: Elsevier, 2004: 291-312.

31 Le CT. Chapter 15. Statistical methods for the comaprison of crossing survival curves. In: Balakrishnana N, Rao CR. Advances in survival analysis: Handbook of statistics Vol 23. Amsterdam: Elsevier, 2004: 277-90.

32 Dignam J, Bryant J, Sam Wieand H. Chapter 17. Analysis of cause-specific events in competing risks survival data. In: Balakrishnana N, Rao CR. Advances in survival analysis: Handbook of statistics Vol 23. Amsterdam: Elsevier, 2004: 313-30.

33 Fine JP, Gray RJ. A proportional hazards model for the subdistribution of a competing risk. *J Am Stat Assoc* 1999; 94: 496-509.

Addendum

After submitting this essay, we reviewed the recent publication by the Acute Respiratory Distress Syndrome (ARDS) Clinical Trials Network on the efficacy of corticosteroids in persistent ARDS.[27] We were struck by the coincidence of the analytic strategy used in that article and the issues discussed above

with respect to competing risks. In particular, the authors used Gray's test[28] to compare the cumulative incidence rates (at 180 days of follow-up) of death, discharge home, resumption of assisted ventilation, and freedom from the need for assisted ventilation among patients whose assisted-breathing history differed (see page 1674 of the trial report, "Figure 1 box"). These time-to-event end-points were thus conceptualised as competing risks and analysed using the cumulative incidence function,[29] as opposed to, say, cause-specific hazards using the log-rank test[30] (assuming that the curves did not cross[31]). The time-to-event end-points in this ARDS trial are comparable with those utilised in the two therapeutic hypothermia trials considered above:[5,6] death and the cerebral performance categories. The cumulative incidence function is a function of all the (*K*) crude hazard rates, and cumulative incidence depends on all (*K*) of the competing crude incidence functions. Thus, differences (as revealed by testing: for example, log-rank versus Gray's test) in crude hazard rates for a particular risk do not translate directly into differences between cumulative incidence curves; such differences also extend to regression analyses and the effect of covariates.[30,32,33] Whereas cause-specific hazards may help to elucidate failure mechanisms, the (testing of the) equality of cause-*K*-specific hazard rates across groups cannot be interpreted as a test of the equality of survival functions (nor of cumulative incidence functions).[32] Cause-specific hazards do not have a direct interpretation in terms of survival probabilities (for a particular failure type), and, in terms of the assessment of the probability of different end-points ("treatment utility") in a balanced trial, the cumulative incidence function would appear to be apposite.[33]

We thus reiterate that:

- the various time-to-event end-points in the trials of therapeutic hypothermia after cardiac arrest are subject to competing risks;
- analysis must reflect this circumstance; and
- the cumulative incidence function has merit in the assessment of treatment utility. ❑