RESEARCH ETHICS

Controversial choice of a control intervention in a trial of ventilator therapy in ARDS: standard of care arguments in a randomised controlled trial

H Mann

When evaluating an innovative intervention in a randomised controlled trial (RCT), choosing an appropriate control intervention is necessary for a clinically meaningful result. An RCT reported in 2000 addressed the relative merits of two tidal volume ventilatory strategies, 6 ml/kg (innovative) and 12 ml/kg (control), in patients with acute respiratory distress syndrome. Critics claim that the 12 ml/kg volume did not represent the clinical practice standard at that time, and that lower tidal volumes had been used in some patients prior to randomisation. The trialists responded that current practice involved the use of a broad range of tidal volumes, including 12 ml/kg. Appropriate control interventions for RCTs can be ensured by: a systematic review of the relevant literature; a formal survey of expert clinicians; and publication of the proposed research protocol to solicit critical appraisal. A global survey of experts during the RCT’s design stage would have been of probative value in determining the appropriate control intervention. Hypothetical, but plausible, results of such a survey are presented and examined to demonstrate the value of this method.

A randomised controlled trial (RCT) is performed to resolve uncertainty about the relative merits of medical interventions. When an innovative intervention is studied, an appropriate choice of a control intervention is necessary to ensure a clinically meaningful result. An RCT to evaluate the relative merits of different tidal volumes used to ventilate patients with acute respiratory distress syndrome (ARDS) was conducted by the ARDS Network (ARDSNet) and published in 2000. The Network was established in 1994 as a Contract Program associated with the National Institutes of Health, and is a consortium of 19 clinical centres. A steering committee comprised of a principal investigator from each centre reviews and approves proposed trials in patients with ARDS. With respect to the ventilator

The benefits, risks, burdens and effectiveness of a new method should be tested against those of the best current prophylactic, diagnostic, and therapeutic methods. (World Medical Association, Declaration of Helsinki1).

A chronology of the salient events in this controversy is provided in table 1. In this essay I will evaluate the contending standard of care arguments in the ARDSNet trial and propose that they reflect substantially disparate interpretations of information available at the time the trial was designed. I will review specific conceptualisations of standard care in medical practice and how these relate to the

Correspondence to: H Mann, Department of Radiology, Program Associate, Division of Medical Ethics, 1A/1 University Hospital, 50 North Medical Drive, Salt Lake City, UT 84132, USA; Howard.Mann@hsc.utah.edu

Received 26 September 2004
In revised form 14 November 2004
Accepted for publication 17 November 2004


www.jmedethics.com
choice of a control intervention for an RCT in general, and for the ARDSNet trial in particular. I will clarify specific methods whereby an appropriate choice of control intervention for such a trial may be pursued: a systematic review of the relevant literature and surveying expert clinicians in the field of interest.

UNCERTAINTY AND TRIAL DESIGN IN THE ARDSNet TRIAL

Critically ill patients with ARDS require the application of multiple medical interventions. ARDS experts believed that one component of assisted ventilation—the degree of alveolar distension—may significantly affect clinical outcome. The ARDSNet trialists proposed to resolve this uncertainty by isolating the effects of tidal volume in an internally valid efficacy RCT, the results of which would be convincing to the medical community. Sackett\(^\text{15}\) has explained how the confidence in the conclusion of an RCT is the ratio of the magnitude of the signal to the magnitude of the noise times the square root of the sample size. Because signal derives from differences between the effects of the interventions being compared (control event rates), the trialists chose 12 ml/kg as the control volume to ensure a “strong” signal, after considering the range of tidal volumes used in clinical practice. In general, the greater the anticipated differences between the specified outcome measures, the smaller the required sample size, all else equal. To minimise noise, non-uniform application of non-experimental cointerventions (typically multiple in critical care trials) was controlled by the use of a predefined computerised management protocol, the application of which had been previously described.\(^{16-18}\)

The trial’s results demonstrated the relative superiority of the low volume strategy: the primary outcome of mortality in this group was 31% compared with 39.8% in the control arm.

CRITICISMS OF THE TRIAL

The critics contended that the trial results demonstrated the superiority of 6 ml/kg over 12 ml/kg, but not the relative merits of the lower tidal volume compared with “routine” care, which allegedly involved the use of lower tidal volumes.\(^{1}\) This reference to routine care followed a publication in 2000 which revealed that, prior to randomisation, patients in the participating centres were ventilated with a mean tidal volume of 10.3 (SD 2) ml/kg, and that the mortality in a group of eligible non-participating patients was lower than in those randomised to the 12 ml/kg arm—31% versus 39.8%.\(^{19}\) The critics also claimed that the mortality in this non-participating group was equivalent to the mortality reported in ARDS patients in 1996,\(^{20}\) the year the ARDSNet trial began. Finally, it was asserted that “Overall, this study tended that each one of these criticisms was invalid, and I discuss their responses in the following sections.

CONTENSIONS ABOUT STANDARD OF CARE IN MEDICAL PRACTICE

The ARDSNet trialists contended that the trial was not designed to compare the outcomes of patients who received lower tidal volumes with those who received routine care, because routine care encompassed a broad range of tidal volumes and inspiratory pressures. They contended that, “at the time the trial was designed, nobody knew if the traditional or lower tidal volume approach was superior”.\(^{8}\) The unimodal distribution of tidal volumes used by clinicians at each of the participating centres\(^7\) (fig 1), with a peak centred at 10.5 ml/kg, was cited by the critics as evidence of a practice standard. The trialists responded that the distribution may be interpreted as an expression of uncertainty on the part of a sufficient number of clinicians concerning the relative merits of lower and higher volume strategies: in anticipation of the results of the ARDSNet trial, many clinicians at the participating centres might have chosen to use a tidal volume intermediate between those employed in the trial.

In published correspondence following the publication of the critics’ article, some expert clinicians defended the choice of the 12 ml/kg tidal volume: Carmichael, coauthor of a report of a 1992 survey of critical care specialists, disputed the
existence of a “best” or “current” practice standard by referring to a “wide disparity in the selection of tidal volumes …” and the fact that “more than half of respondents reported using a tidal volume larger than 10 ml/kg”.

Petty, an acknowledged ARDS expert and author of a state of the art review of ARDS in 1990, also disputed the contention that intermediate tidal volumes represented a standard of care by stating “I don’t believe this is true, based on my own observations while visiting many intensive care units around the country and in Europe over many years, including the decade of the nineties when the ARDS trial was done. Most units employed high tidal volumes according to our original recommendations”. It is notable that these defenders’ units employed high tidal volumes according to our original decade of the nineties when the ARDS trial was done. Most the country and in Europe over many years, including the stating “I don’t believe this is true, based on my own
aknowledged ARDS expert and author of a state of the art www.jmedethics.com

equipoise. The RCT is a formal means of resolving this interventions. This state has been described as one of clinical
ing the relative merits (benefits and harms) of comparator
uncertainty in the community of expert physicians concern-

The impetus for an RCT originates in a state of epistemic
RCTs

ACKNOWLEDGMENT OF UNCERTAINTY AND

SECURING PARTICIPANTS’ MEDICAL INTERESTS IN
RCTs
The impetus for an RCT originates in a state of epistemic
uncertainty in the community of expert physicians concerning
the relative merits (benefits and harms) of comparator
interventions. This state has been described as one of clinical

The RCT is a formal means of resolving this uncertainty and should explicitly be designed to disturb equipoise. This state of uncertainty is also of interest to research ethics committees, which must approve a proposed trial before it may be offered to physicians and patients. Once it is available, participating physicians have to decide whether to offer it to particular patients. Finally, each patient (or a proxy) has to decide whether to accept an offer of enrolment. When a patient, informed about what is known about the benefits and harms of the comparator interventions, is also “maximally uncertain” about their relative merits—when the patient is in a state of equipoise—randomisation represents a reasonable treatment choice under conditions of uncertainty. Thus, the goals of the trialist (the acquisition of scientific knowledge) and the patient-participant (the best treatment) may concurrently be accomplished.

RCTs may involve investigational or commonly used interventions. When an innovative intervention is being evaluated, the appropriate choice of a control intervention is critical. An inappropriate choice may nullify the clinical value of a trial that otherwise meets the requirement of internal validity. For example, in an RCT in which the control intervention is a drug, an inappropriate low dose or method of administration of the control agent will violate the equipoise principle. Comparator bias describes a situation in which a new treatment is compared with an existing treatment that is known to be inferior or when the treatment is compared against nothing when effective treatments already exist. What methods should be used to avoid comparator bias? In a previous communication, Djulbegovic and I elaborated on this issue, and the methods that may be used include:

- systematic review of the relevant literature
- formal survey of relevant clinical experts
- publication of the trial’s protocol to solicit critical appraisal of its rationale and design.

A SYSTEMATIC REVIEW OF THE RELEVANT LITERATURE
The choice of a control intervention should be justified by such a review. In drug RCTs the proper dose and method of administration of the control intervention are relevant, and these may be determined from a systematic literature review. A review is also pertinent when a placebo-only control arm is proposed: research synthesis should establish that there is no known effective therapy for the medical condition involved. The ARDSNet trial differed in an important respect: the intervention being evaluated was a continuous variable—tidal volume. How would such a review have informed the design of the ARDSNet trial? A systematic review of the ARDS related literature would have yielded the aforementioned survey of expert clinicians, conducted in 1992, concerning the management of patients with ARDS. The survey questionnaire was sent to 3264 members of the American Thoracic Society Assembly of Critical Care Medicine. The majority (65%) of the 1023 respondents were board certified internists and certified in pulmonary and/or critical care medicine, practising in the USA and 24 countries abroad. Equal numbers practised in university teaching (43%) and community hospitals (42%). The survey revealed a division of opinion in the expert clinical community concerning the choice of tidal volumes: 45% and 48% of respondents thought tidal volumes of 5–9 ml/kg and 10–13 ml/kg, respectively, most appropriate in patients with ARDS. Although not cited in the trial’s protocol, it was apparently used by the ARDSNet trialists when determining a control tidal volume.

THE PROBATIVE VALUE OF A FORMAL SURVEY OF EXPERT CLINICIANS
Formal surveys of expert clinicians have been used to confirm the existence of uncertainty necessary to propose an RCT, and to document physicians’ prior beliefs about an intervention’s relative efficacy when bayesian methods were used to conduct interim analysis in an ongoing RCT. In a cardiology RCT in which the efficacy of a platelet inhibitor (eptifibatide) was compared with a placebo, a survey of a group of interventional cardiologists to evaluate the existence of a putative standard of care was apparently determinative in persuading the Food and Drug Administration to lift the “clinical hold” the agency initially put on the proposed trial, because of its concern about the appropriateness of the placebo controlled design.

Allegations about breaches of the standard of care are a central feature of medical malpractice actions. In the courtroom, attorneys elicit testimony from expert witnesses for the plaintiffs and defendants according to formal rules of evidence intended to establish facts salient to the action. Evidence is considered “probative”, and hence admissible, if it logically tends to prove the proposition for which it offered. A survey of relevant clinicians has been proposed as a method of securing evidence for or against a putative standard of care in such situations.

Analogously, ascertaining the informed beliefs and practices of expert clinicians may enable an appropriate choice of control intervention for an RCT, particularly when, as in the
ARDSNet trial, the notion of a “standard of care” may arise. Indeed, a requirement for an antecedent survey of the current practice of clinicians at participating centres was suggested by the trial’s critics, and the OHRP. This would certainly help to discern the current practice patterns of physicians at participating centres, but it is not clear that this is necessarily an appropriate strategy for determining a putative standard of care against which to compare the low volume strategy. As previously described, the salience of a survey limited to practices in the participating centres only, is disputable. Is there another way to conceptualise standard of care arguments that may be pertinent to the ARDSNet trial?

The notion of a standard of care in clinical research has been particularly controversial in the design of placebo controlled trials conducted in developing countries by investigators and sponsors from developed nations. London has explored the controversy by distinguishing between two reference points—local versus global; and two conceptual interpretations of standard of care arguments—de facto and de jure. In the de facto interpretation, the standard of practice is set by the actual medical practices of the relevant reference community. In the de jure interpretation, the standard is determined by the judgement of medical experts in the reference community as to which interventions have been proved effective. The notion of proof relates to a body of scientific evidence, particularly from controlled clinical trials, supporting an intervention’s efficacy in a population of interest. The epistemic character of the de jure standard is particularly relevant to the notions of uncertainty and clinical equipoise, and it is also consistent with the greater evidential weight generally accorded to the results of controlled trials compared with uncontrolled observational data.

Prior to the conduct of the ARDSNet trial, a body of evidence supportive of a certain tidal volume, or a certain narrow range of tidal volumes, as a de jure standard of care did not exist. Because the tidal volumes employed by clinicians were known to be diminishing over time, and the last survey had been conducted about four years prior to the initiation of the ARDS trial, the trialists may have repeated the survey of a representative number of clinicians in the global community of ARDS experts. The results may have been of substantial probative value from the de jure perspective: in addition to enumerating the tidal volumes they actually employed in practice, the respondents should have been asked to clarify the reasons for their choices. A survey should attempt to discern whether the choice potentially represents preference neutrality or an explicit choice grounded in an assessment of published evidence pertaining to the clinical issue at hand. This distinction has been made by Ashcroft, as follows: “Equipoise is not simply preference neutrality. The physician may not prefer treatment A to treatment B, or vice versa. But what is important is whether this is his or her treatment preference, but whether he or she has reason to prefer one treatment to the other. Preferences for treatment should be rationally corrigible.”

Hypothetical results of a survey of applied tidal volumes are presented in fig 2. I will now present reasonable interpretations of these results. The results of survey A do not reveal any tendency to a favoured tidal volume in the range evaluated. The proposed choice of 12 ml/kg may be affirmed as acceptable. Survey B reveals a notable and tight concentration of applied tidal volumes around 10 ml/kg: 84% of respondents used tidal volumes between 9 ml/kg and 11 ml/kg. This result would compel the trialists to reconsider the choice of control volume. Use of 10 ml/kg as a control volume would still permit the conduct of a trial with sufficient signal magnitude to address the uncertainty of interest and, if the results convincingly demonstrate the superiority of the lower tidal volume, an evidential claim for

Figure 2 Results of hypothetical surveys of applied tidal volumes.

the latter as a de jure standard of care is established. The community of experts would assess the results in the light of other evidence—including other controlled trials—and such research synthesis and accompanying debate would properly address the merits of the claim. Survey C closely mirrors the range of tidal volumes actually used in the ARDSNet trial participants before randomisation and does not affirm 10 ml/kg as a standard of care against which the 6 ml/kg strategy should be compared: 54% of respondents used tidal volumes between 9 ml/kg and 11 ml/kg, and higher volumes were used by 23% of respondents.

SHOULD A THIRD ARM HAVE BEEN USED IN THE ARDSNet TRIAL?

Whether a third arm using 10 ml/kg, reflective of “routine care”, should have been included in the trial is another issue that has been raised. This proposal should seriously have been considered if, based on other information, it was plausible that an intermediate tidal volume might have been superior to both volumes being compared. Miller and Silverman have discussed this possibility, and the pros and
consequences of a third arm, in the ARDSNet trial. Inclusion of a “routine care” control group in the trial would have necessitated enrolling additional participants, increasing the cost and complexity of the trial, and complicating its analysis. Because the likelihood that a “routine care” group would fare better than either the 6 ml/kg or the 12 ml/kg group could not have been predicted with any certainty antecedent to the initiation of the trial, assertions about this are necessarily influenced by hindsight bias.

However, in an additional informative communication concerning control group selection in critical care RCTs in general, Silverman and Miller consider further the merits of RCTs that, like the ARDSNet trial, compare two contrasting strategies. They argue that “as the contrasting strategies become less representative of the broad range of standard care practices, the less assurance is provided that both study groups will have mortality rates that do not exceed standard practices. Accordingly, the greater becomes the need to include a representative standard of care control group to permit a trial to detect whether research subjects are being exposed to excessive harms”. In judging the extent to which the contrasting strategies are representative of standard practices, they state that “specification of any numerical percentage would be arbitrary”. The results of a survey, as mentioned above, would substantially inform a judgement about the need for inclusion of a “representative standard of care control group”.

Determining the implications of a survey’s results will always entail judgement, which may engender disagreement. Nevertheless, I believe its value outweighs such concerns. The results of the survey should be incorporated in the research protocol submitted for review by research ethics committees, whose actions in this matter have been criticised. A relevant summary of de facto practices may also be presented to potential trial participants in the information sheet for the trial.

OBJECTIONS TO THE CONDUCT OF SURVEYS
I anticipate that concerns about the expense and practicability of conducting such surveys will arise. However, unlike the requirement for a systematic literature review, they will not be necessary for most RCTs. In others, a survey limited to a smaller number of experts may suffice. In any event, the adverse consequences of an inappropriate choice of control intervention justify this method.

CONCLUDING COMMENTS
Conducting informative controlled trials in a critical care environment is complex and challenging. Because mortality is relatively high and potential participants commonly have impaired decisional capacity, appropriate acknowledgment of uncertainty is critical prior to randomising patients among interventions in an RCT. Guidance for conducting trials is available, but this does not typically address uncertainty and standard of care issues in sufficient detail. Systematic methods are available that may be used to inform the choice of a control intervention, and the rationale for the choice should be explicated in the published report of the trial. Miller and Rosenstein have addressed the issue of explicitly reporting pertinent ethical issues to promote public accountability for clinical research. Editors and reviewers should require this of author-investigators. If a survey is used to inform the design of the trial and the choice of interventions, it should also be published.

ACKNOWLEDGEMENTS
J Mann and A London provided helpful critiques of draft versions of this essay.

AUTHOR’S STATEMENT AND CONTRIBUTIONS
HM conceived this essay, made necessary revisions, and approved the final version submitted.

REFERENCES
1 World Medical Association. Declaration of Helsinki, 2000
4 Carone MA. Human research subject protections under multiple project assurances (MPA) M-1331, M-1363, and M-1138 and the OHRP-approved assurances for all ARDS Network institutions. Letter (7 October 2002) to Ronald S Newbower, Massachusetts General Hospital, Lee E Limbird, Vanderbilt University, and Robert Kay, the Cleveland Clinic Foundation. Rockville: Office for Human Research Protections.
14 Carone MA, Borradori L. Human Research Subject Protections Under Federalwide Assurance (FWA) 3126, Multiple Project Assurances (MPA) M-1331, M-1363, and M-1388 and the OHRP-approved assurances for all ARDS Network Institutions. Letter (3 July 2003) to Ronald S Newbower, Massachusetts General Hospital, Lee E Limbird, Vanderbilt University; Robert Kay, the Cleveland Clinic Foundation.
15 Sackett DL. Why randomized controlled trials fail but needn’t 2. Failure to employ physiological statistics, or the only formula a clinician-trialist is ever likely to need (or understand)!. Can Med Assoc J 2001;165:1266–7.
47 Oba Y. Where are the institutional review boards? Am J Respir Crit Care Med 2002;166:1317.
Controversial choice of a control intervention in a trial of ventilator therapy in ARDS: standard of care arguments in a randomised controlled trial

H Mann

*J Med Ethics* 2005 31: 548-553
doi: 10.1136/jme.2004.010736

Updated information and services can be found at:
http://jme.bmj.com/content/31/9/548

These include:

**References**
This article cites 35 articles, 2 of which you can access for free at:
http://jme.bmj.com/content/31/9/548#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 trials (epidemiology) (88)

**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/