Usual Care as the Control Group in Clinical Trials of Nonpharmacologic Interventions

B. Taylor Thompson¹ and David Schoenfeld²

¹Pulmonary and Critical Care Unit and ²Biostatistics Center, Department of Medicine, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts

We discuss the pros and cons of including usual care as a control arm in clinical trials of nonpharmacologic interventions. Usual care is a term used to describe the full spectrum of patient care practices in which clinicians have the opportunity (which is not necessarily seized) to individualize care. The decision to use usual care as the control arm should be based on the nature of the research question and the uniformity of usual-care practices. The use of a usual-care arm in a two-arm trial should be considered for trials of investigational drugs or devices, for trials that propose to test interventions that lie well outside usual-care practices, or for trials where the research question *per se* is to compare a strategy against usual care. Examples of the latter include pragmatic effectiveness trials of clinical pathways or protocolized-care versus usual-care practices. Randomized intervention trials can be safely conducted and monitored using two treatments that lie within the range of usual-care practices if both approaches are considered prudent and good care for the target population.

Keywords: clinical trials; human experimentation; human subjects protection; control group selection

Recent controversy has centered on control group selection for intensive care unit (ICU) trials of nonpharmacologic interventional strategies, specifically Acute Respiratory Distress Syndrome Network (ARDSnet) studies of low tidal volume (1). Criticism has focused on the absence of a control arm that reflected “best current care” to detect potential harm to research subjects with rapidly fatal diseases (2). Critics argue that such arms, which allow for individualized treatment decisions unrestricted by study protocols or rules, require “no assumptions to determine whether or not an experimental therapy is resulting in harm during a trial” because the experimental arm is being compared with *de facto* usual care (3).

We will examine the design considerations for using usual care as a control group in interventional trials and discuss three-arm designs that include two interventional strategies with usual care as a control. We will conclude by emphasizing trial designs and review processes that will help minimize risk and provide for individualized care within the research context.

TERMINOLOGY

The Declaration of Helsinki states that the “benefits, risks, burdens, and effectiveness of a new method should be tested against those of the best current prophylactic, diagnostic, and therapeutic methods” (4). However, in much of medicine, and in ICU medicine in particular, “best current therapy” has not yet been

identified and many “standard” practices have not been validated in clinical trials. The type and volume of intravenous fluids for resuscitation, the duration of antibiotics for many infections, or the use of newer modes of mechanical ventilation are but a few of many examples of nonstandardized usual-care practices. Without good pathophysiologic or clinical evidence for superiority of one practice over another, usual care may encompass a wide variety of practice styles that are difficult to explain. Thus, the terms “best current” therapy or “standard of care” are problematic as they imply a uniform or proven practice standard. We prefer the descriptive term “usual care” to describe *de facto* clinical care without any value judgment. Usual care may be standardized around high-level evidence and thus represent best current therapy (e.g., aspirin therapy for acute coronary syndromes) but may also be highly variable and inclusive of both prudent and undesirable practices.

INCLUSION OF A USUAL-CARE GROUP FOR DETERMINING SAFETY AND EFFECTIVENESS

Proponents of usual-care groups in clinical trials argue that the incremental risk of participating in clinical trials, particularly in trials of rapidly fatal diseases, can only be measured by timely comparisons to the outcomes of patients randomized to usual care even if usual care is variable and not based on a consensus (2, 3). Advocates argue that physicians make individualized treatment decisions based on personal experience, results of preclinical studies or observational trials, and expert opinion so that even in the absence of randomized controlled trials, such care should be the safety standard (2). Furthermore, studies that reveal a superior interventional strategy among only two that are tested (the explanatory or efficacy question) cannot claim that the superior treatment should be the new standard because this strategy has not been shown to be superior to usual care (the pragmatic or effectiveness question). The limitation of efficacy trials without usual-care arms has been termed a “fundamental flaw” and has led to a reexamination of the traditional randomized intervention trials in the ICU (4, 5). Ethical analyses of these issues have been recently published (4, 6).

RANDOMIZED ICU INTERVENTION TRIALS WITHOUT USUAL-CARE CONTROL ARMS

Inclusion of usual care as the control or comparator arm in ICU intervention trials would represent a departure from the traditional approach. A convenience sample of trials without usual-care arms is shown in Table 1. In general, these trials randomize subjects to two or three competing interventional strategies but do not include random assignments to usual care where a physician would make treatment decisions independent of the trial’s goals. For example, in a trial of supranormal oxygen delivery in patients with severe sepsis, a third arm that allowed for individualized hemodynamic therapy was not included (7). In trials of weaning from mechanical ventilation and lung-protective ventilation, no comparisons to unrestricted approaches to

(Received in original form June 22, 2007; accepted in final form August 10, 2007)

Supported by National Institutes of Health, NHLBI contract NO1-HR 46064.

Correspondence and requests for reprints should be addressed to B. Taylor Thompson, M.D., Massachusetts General Hospital, 55 Fruit Street, Boston, MA 02114. E-mail tthompson1@partners.org

Proc Am Thorac Soc Vol 4, pp 577–582, 2007

DOI: 10.1513/pats.200706-072JK

Internet address: www.atsjournals.org

TABLE 1. EXAMPLES OF INTENSIVE CARE UNIT INTERVENTION TRIALS WITHOUT A USUAL-CARE ARM*

Randomized Groups	
Brochard (1994) (9) —Weaning strategies	Three weaning strategies
Esteban (1995) (10) —Weaning strategies	Four weaning strategies
Gattinoni (1995) (7) —Goal-directed therapy in sepsis	CI 4.5 versus $Sv_{O_2} > 70$ versus CI normal
Hebert (1999) (15) —Transfusion threshold in ICU	7–9 versus 10–12 g/dl
ARDSnet (2000) (8) —Lung-protective ventilation	6 versus 12 ml/kg PBW tidal volume
Van den Berghe (2001) (14) —Glycemic control in CSICU	Glucose 80–110 versus < 210
Gattinoni (2001) (13) —Lung-protective ventilation	Supine versus prone ventilation
Chastre (2003) (12) —Antibiotics for VAP	Antibiotics for 8 versus 15 days

Definition of abbreviations: CI = cardiac index; ICU = intensive care unit; PBW = predicted body weight; Sv_{O_2} = systemic venous oxygen saturation; VAP = ventilator-associated pneumonia.

* Includes pharmacologic and nonpharmacologic interventions available in usual-care practice outside the research context. An unrestricted usual-care arm refers a study arm where clinical care is delivered by clinicians without study goals or protocolized treatments.

weaning and ventilator management were done (8–10). None of these trials used usual care as the control arm.

For “controls,” some of the trials in Table 1 used treatment protocols or goals that approximated the mean or median usual-care practice before the study (e.g., the mean tidal volume selected by clinicians in prior trials) or the practice favored by clinicians based on surveys (2 wk of antibiotics for ventilator-associated pneumonia) (11, 12). Others selected control groups that were widely regarded as the standard approach outside the research context, such as supine ventilation as a control for prone positioning or control of blood glucose below 210 mg/dl as a control for normalization of glucose (13, 14). Many of these trials answered important questions without usual care as a control group. Would they have been more informative (or safer) with usual care as a control? To address these questions, we will explore the limitations of usual care as a treatment comparison and explore why investigators chose not to use it.

SELECTION OF TWO INTERVENTIONS WITHIN USUAL CARE: A CASE STUDY

ICU clinicians must balance risks of transfusion versus the risk of anemia, including inadequate delivery of oxygen to tissues. A survey of Canadian ICU clinicians revealed variability in practice. A subsequent randomized trial by Hebert and colleagues of two transfusion strategies showed no change in the primary endpoint (28-d mortality), an important finding that has led to a reduction in unnecessary blood transfusions (15). How were these two transfusion strategies determined?

In a prestudy survey, Canadian clinicians were asked to describe the level of hemoglobin at which they would transfuse a unit of red blood cells in four clinical scenarios (16). For each scenario, wide variation in practice was noted (Figure 1). In one scenario, some clinicians waited until the hemoglobin fell to 6.5 g/dl or less to transfuse, whereas others transfused at the previously recommended threshold of 10 g/dl. The distribution of transfusion thresholds was normally distributed for this scenario around a mean of approximately 8 g/dl. Thus, clinicians caring for the identical patient made very different treatment decisions. This is an example of clinically unexplained practice variation and it gives a sense of the magnitude of this variation. Some of this variation was explained by nonclinical factors, such as the province in which the clinician practiced and the practice setting (community or academic).

Not all variation in usual care is unexplained. When clinicians were presented with another transfusion scenario—a critically ill patient who was actively bleeding—clinicians said they would

transfuse at higher hemoglobin thresholds (anticipatory transfusion) (16). This shift to a different practice is evidence of customization, although there was still wide variation in practice for the bleeding scenario (Figure 1). Thus, the range of decisions for transfusion in the ICU reflects both clinically unexplained variation (or practice styles) and customization.

Hebert and colleagues chose a two-arm trial design (A vs. B in Figure 1) (15). “A” was a restrictive transfusion strategy (hemoglobin, 7 g/dl) and “B” was a traditionally recommended transfusion threshold of 10 g/dl. Both had a strong pathophysiologic rationale for benefit and both were represented within usual-care practices.

Because the prestudy survey showed that very few clinicians would transfuse at 7 g/dl in a bleeding patient, such patients were excluded from the trial. However, the prestudy survey also revealed that age and severity of illness also influenced transfusion decisions. How can we control for this form of customization? One solution would be to model physician behavior around the known variables used in clinical decision making (protocolized usual care). For transfusion, clinicians said they weighed these additional factors in transfusion decisions, but unexplained variation remained and the approaches were far from prescriptive (16). Thus, protocolizing this approach was not possible. This is a general problem with *de facto* usual care because it is often difficult to describe, it is inconsistent, and it may not be fully explained by clinical factors.

Another solution to control for customization is to randomize to *de facto* usual-care practices (unprotocolized usual care). However, if usual care is inferior or superior to a more fixed intervention, it may be not be possible to ever know why. Thus, clinicians will not be able to interpret the study in a way that informs their future practice. As noted by Claude Bernard, “The experimenter who does not know what he is looking for will not understand what he finds.” For these and other reasons, such trials may not be compelling and thus unlikely to change practice (17).

Hebert and colleagues chose a two-arm trial after extensive pretrial preparation. This trial provided valuable information that has changed practice, reduced unnecessary transfusions, and helped stimulate research on the potential toxicities of red cell transfusion. It did not answer the customization (effectiveness) question, however, and the trial design remains controversial in some quarters (18). We ask the reader to consider the potential impact of the various options. For example, would a study showing no difference between a 7-g/dl transfusion threshold and usual care (with customization within a wide range of practices [6–11 g/dl; see Figure 1] and no clear explanation for how clinicians

Figure 1. Hemoglobin thresholds for packed red blood cell (PRBC) transfusions for 254 intensive care unit clinicians varied from less than 6.5 g/dl to over 10.5 g/dl. C_1 indicates the entire range of hemoglobin transfusion thresholds (*de facto* usual care). C_2 indicates a narrower range of practices that might be considered prudent care. Different treatment decisions for the identical patient scenario is an example of clinically unexplained practice variation. The province and practice setting (academic or community) explained some of the variation in practice (practice styles). Clinician practices were different for an actively bleeding patient scenario (*hatched bars*) in comparison to three nonbleeding scenarios (*open bars*), indicating evidence of customization, although practices still varied widely. Severity of illness and age also influenced practice (data not shown), which is further evidence of customization. The entire range of practices represents *de facto* usual care. Within these practices lies what a consensus of the investigators likely determined to be prudent care and they proposed to randomize two strategies within these bounds (A vs. B) (15). This proposal was subjected to peer review (Canadian Clinical Trials Group and the Medical Research Council of Canada), and the participating institutional review boards that approved the study (15). Both approaches were judged not to be prudent care for bleeding patients, so such patients were excluded. Data from Reference 16.

were making these decisions) have provided more or less compelling evidence that a restrictive approach is safe?

SECULAR TRENDS IN USUAL CARE AND THE HAWTHORNE EFFECT

Usual care in the absence of high-level clinical evidence is relatively nondirective and may be susceptible to the influence of the research environment (the so-called Hawthorne effect). Such an effect may have been seen for tidal volume selection during the ARDSnet's lower tidal volume trial (19). Prestudy surveys of clinician preferences and examination of actual clinician practices before the study were used to select the 12-ml/kg predicted body weight traditional arm (11, 20). The distribution of clinician-selected tidal volumes in one of these studies, a 750-patient trial of exogenous surfactant therapy, is shown in Figure 2, as well as clinician-selected tidal volumes during the ARDSnet trial (11, 19). Note that the lower tidal volumes (in the 5–7.5-ml/kg range) were more likely to be used during the trial and the shift of median tidal volume toward the left. It is unclear if movement toward lower tidal volumes represented a secular trend in usual-care practices in response to expert opinion (21) or the influence of the study itself (Hawthorne effect).

The migration of usual care over time has important implications for trial design. Had the ARMA trial used *de facto* usual care as the control for the lower tidal volume intervention, the degree of separation of tidal volumes (the study variable) between randomized groups would have been unpredictable and likely smaller than anticipated based on prestudy data. This, in turn, would have reduced power and may have led to a type II error. Evidence in support of this is provided by an analysis of eligible nonparticipants of ARMA who received usual care. These patients had outcomes that were not significantly different from either the traditional or lower tidal volume groups in the centers in which eligible nonparticipant outcomes were available (22). Thus, a trial of low tidal volume versus usual care may have concluded that tidal volume reduction was not important.

The ephemeral nature of usual care puts clinical trialists in a quandary. If the goal of a control group is to emulate usual care, protocolizing usual care based on prestudy information is no guarantee that this group will reflect usual care *during the conduct of the trial* as usual care may change. Randomizing to

Figure 2. Clinician-selected tidal volumes before the initiation of study procedures are shown for the Exosurf trial (Aerosolized Surfactant for ARDS, conducted in nine countries, 63 hospitals, with 726 patients enrolled between 1992 and 1993) (11) and the Acute Respiratory Distress Syndrome Network (ARDSnet) lower tidal volume trial (conducted in the United States with 17 hospitals, and 861 patients enrolled between 1996 and 1999) (8). Arrow "A" represents the lower tidal volume group, selected after the completion of a phase II trial and the publication of another positive trial, and arrow "B" the tidal volume used in the traditional approach (32, 33). The Exosurf trial results were considered during the design of the ARDSnet trial. *De facto* usual care would be the entire range of clinician-selected tidal volumes. Note the shift in usual-care practices toward the use of lower tidal volumes during the trial, which may have represented adoption of lower tidal volumes in response to preclinical studies and expert opinion or the influence of the trial *per se* (Hawthorne effect). *Hatched bars*, Exosurf; *solid bars*, ARMA.

unrestricted usual care runs the risk that usual care may merge with the intervention arm during the trial, narrowing differences between groups, and resulting in loss of power to detect a meaningful difference.

OTHER LIMITATIONS OF *DE FACTO* USUAL CARE AS THE CONTROL

Even in areas where high-level evidence is available to guide clinicians, substantial variability and inconsistency are present in usual care (17, 23). Inconsistencies are also apparent in an individual's decision making when large amounts of data must be processed simultaneously to make decisions. For example, our short-term memory can simultaneously retain only five to nine data constructs, and when we attempt to make decisions with more data, inconsistent decision making results (24). It is very likely that ICU clinicians processing hundreds of data elements while caring for a dozen or more patients confront these limitations and that inconsistent decision making and unnecessary variation in practice follow (25). Such inconsistencies have prompted calls for protocolization and computerized decision support to improve both usual care and clinical research practices (26, 27). Inconsistent practice makes usual care difficult to understand and describe, thus limiting its value as a comparator arm in clinical trials.

THREE-ARM TRIALS INCLUDING USUAL CARE: A WIN-WIN?

Silverman and Miller suggested that a three-arm trial, comparing both tidal volume strategies with a representative usual-care group, has the potential to offer the most clinical value by providing rigorous evidence to guide what should be considered the standard of care (2). For example, comparison of usual care with either of the tidal volume groups in the ARDSnet trial may have answered the efficacy and effectiveness questions in a single trial and would have the added benefit of protecting research participants from unanticipated harm should either of the two experimental arms demonstrate inferiority to usual care. However, this design has some problems, one of which is the counterintuitive finding that such designs may be less safe under many assumptions.

If safety is defined as the additional deaths in the inferior treatment arm at the time it can be determined that this arm is inferior (i.e., when the study stops), three-arm trials with a usual-care control arm usually result in more "additional deaths" (they are less safe) compared with two-arm trials (28). This is due to the increase in sample size. Sample size is increased because of the addition of the third arm but also because power is reduced from multiple comparisons during sequential interim monitoring, thus driving up the sample size even further to maintain power (29). Safety is only improved if usual care is clearly superior to both of the experimental arms (28). Thus, a three-arm design with a usual-care arm might be a reasonable approach if the new therapies under study are not well represented in usual care and/or have the potential for incremental risk over usual-care practices.

A three-arm design with usual care would speed knowledge development if one of the experimental approaches was superior to the other (the efficacy question), thus making the effectiveness question relevant. Because nearly all large ICU trials over the years have been negative, requiring all ICU interventions to have a third usual-care arm would result in larger, more costly trials with little or no gain. It should also be noted that nearly all of the effectiveness trials with usual care show protocolized interventions to be equivalent or superior to usual care (27, 30,

31). These observations call into question the potential benefit of adding *de facto* usual care to efficacy trials.

Adequate separation of groups is another counterintuitive safety consideration. For linear "treatment-response" relationships, trials with a larger separation are actually safer than trials with a smaller separation because the sample size for a clinical trial is inversely proportional to the square of the separation, whereas harm increases linearly with the separation. For example, imagine we are studying the impact of the duration of antibiotics. If a trial with wide separation (7 vs. 14 d) designed to detect a 10% difference (500/arm) actually ended with a 10% mortality difference (e.g., 50 vs. 40%), then this trial would have resulted in approximately 50 additional deaths in the inferior arm before this important question could be answered. A trial with less separation (10 vs. 14 d) designed to detect a smaller mortality benefit (50 vs. 45%; 1,250/arm) could be expected at study completion to result in 63 deaths, assuming a linear response of duration of antibiotics to outcome. Greater separation of antibiotic duration may exceed what clinicians consider to be competent care and may challenge the presumed linear relationship of duration to outcome. The art of clinical trial design is to balance the desire to have adequate separation to detect a clinically important difference with the need to assure that both treatments lie soundly within the boundaries of prudent care.

WHEN IS UNRESTRICTED USUAL CARE AN APPROPRIATE CONTROL GROUP?

The design should fit the purpose. Usual care could be considered for blinded trials of experimental drugs and devices or for treatment strategies that are not part of usual-care practices. Pragmatic effectiveness trials where the research question *per se* is to determine if a treatment approach is superior to usual care should obviously have *de facto* usual care as the control group. Such trials often follow earlier explanatory or efficacy trials that identify the best practice to be tested against usual care. For example, the two explanatory trials of weaning from mechanical ventilation (Table 1) identified the spontaneous breathing trial as the superior approach. Subsequent effectiveness trials then compared weaning protocols to usual care. This logical sequence of knowledge development guards against adopting a strategy found to be superior in a two-arm trial without a usual-care arm. Accordingly, many of the interventional trials outlined in Table 1 marked the beginning, not the end, of continued research to improve clinical care.

ELEMENTS OF ICU INTERVENTION TRIALS IMPORTANT TO MINIMIZE RISK

How can investigators design clinical trials to advance knowledge, minimize risk, and address the concerns of those who champion individualized care? We propose that many of the design elements and good clinical practices currently in place accomplish many of these important goals. These design considerations are outlined in Table 2.

THE INDIVIDUAL PATIENT AND CLINICAL JUDGMENT IN CLINICAL TRIALS

An essential condition for all these studies is that there exists uncertainty, or equipoise, about which of the interventions is superior for the identified patient population. The inclusion and exclusion criteria should be crafted to identify such a population but cannot cover every clinical scenario where uncertainty is preserved. Accordingly, physician consent to allow his or her eligible patient to be approached for research is a standard practice and serves as an additional safeguard. This requires the

TABLE 2. MINIMIZING RISK IN INTENSIVE CARE UNIT INTERVENTION TRIALS

1. Both interventional strategies in a two-arm trial should have a well-grounded pathophysiologic rationale to benefit all eligible patients. Determination of nonlinear treatment responses is best done with relatively fixed treatments along the range of treatments rather than fixed treatments bracketing variable, unexplained usual-care practices.
2. Experts in the field should develop the experimental treatment strategies, much as consensus recommendations for practice are developed when high-level clinical evidence is lacking. Both treatments *must be considered to be good or competent care*.
3. Experimental interventions should be guided by explicit guidelines or protocols so that the results of the clinical trials can be clearly communicated. This allows for use of superior arms to serve as control for future trials and for clinicians to fully understand the intervention when they consider changing their practices.
4. Intervention protocols should, when possible, titrate or customize the intervention to meet individual patient needs. Many dynamic protocols do this: one example is the titration of tidal volume to plateau airway pressures (or dyspnea) in the ARDSnet tidal volume protocols (8). Such protocols meet changing patient needs, thus providing individualized care within the given treatment objective.
5. Treatment protocols and guidelines should be designed to minimize risks in comparison to anticipated benefits. Treatment protocols should also consider incorporating elements that minimize risk *in comparison to usual care*. One example is the addition of protocolized weaning in trials of lung protection.
6. If there has been limited clinical experience with one of the clinical approaches under consideration, then phase II assessments of the approach should be done before proceeding with phase III.
7. Sequential stopping rules should be carefully designed to stop the trial as soon as its goals are reached, while still controlling the error rate and maintaining power. Any continuation of a trial beyond the point where the results can be ascertained with certainty exposes the patients to unnecessary risk. It should be realized, however, that trials that stop early for futility will provide weaker evidence for equivalence in comparison to a trial that continues to its full accrual.
8. Independent and qualified protocol review committees and/or DSMBs should undertake review of the research protocols.
9. DSMBs should review as close to real-time data as possible in ICU studies of rapidly fatal diseases. When possible, outcomes in otherwise comparable historical controls should be provided to DSMBs as an additional safety measure.

Definition of abbreviations: DSMB = data safety monitoring board; ICU = intensive care unit.

attending physician to be familiar with the research protocols to judge the risks and benefits of random assignment to either treatment. The physician must also make another decision: to weigh the risks and benefits of either treatment arm in comparison to his or her own usual-care practices. Uncertainty should exist for all three conditions before an attending physician allows a patient or his or her surrogate to be approached for participation. Attending physician oversight of research-directed care after randomization is also important to assure that individual patient needs are met throughout the study. To the degree that clinicians already do this, the care delivered in the research environment may not be as disparate, nor the research risks as great, as critics fear.

CONCLUSIONS

Randomized intervention trials of competing treatment strategies have made substantial contributions to current ICU practice. This approach provides the greatest chance for discovering which of our ICU practices are superior, without benefit, or even harmful. In our view, clinical trials can be safely conducted and monitored using two treatments that lie within variable usual-care practices if both approaches are considered reasonable and prudent care. This determination should be made in an open, peer-reviewed process by experts in the field and supported with data on the nature of usual-care practices. Separation of the two

treatment arms is important both for detecting a treatment effect and also for minimizing risk. Explicit trial methodology, including the use of guidelines or protocols, is important for interpreting, generalizing, and reproducing (pooling) trial results. Such protocols should be responsive to patient needs and, whenever possible, contain elements of care to reduce risk in comparison to usual-care practices.

The addition of a usual-care arm to a two-arm trial will, under most circumstances, result in additional harm to research subjects, substantially increase the cost and complexity of clinical trials, and delay the acquisition of knowledge. The use of a usual-care arm in a two-arm trial should be considered for trials of investigational drugs or devices, if an intervention lies well outside usual-care practices, or if the research question *per se* is to test a strategy against usual care, such as in effectiveness trials of clinical pathways or protocolized care.

Conflict of Interest Statement: Neither author has a financial relationship with a commercial entity that has an interest in the subject of this manuscript.

References

1. Steinbrook R. How best to ventilate? Trial design and patient safety in studies of the acute respiratory distress syndrome. *N Engl J Med* 2003;348:1393–1401.
2. Silverman HJ, Miller F. Control group selection in critical care randomized controlled trials evaluating interventional strategies: an ethical assessment. *Crit Care Med* 2004;32:852–857.
3. Eichacker PQ, Gerstenberger EP, Banks SM, Cui X, Natanson C. Meta-analysis of acute lung injury and acute respiratory distress syndrome trials testing low tidal volumes. *Am J Respir Crit Care Med* 2002;166:1510–1514.
4. World Medical Association. Declaration of Helsinki: ethical principles for medical research involving human subjects. Paragraph 29 [Internet] [last accessed Aug 9, 2007]. Available from: www.wma.net/e/policy/b3.htm
5. Eichacker PQ, Gerstenberger EP, Banks SM, Cui X, Natanson C. Meta-analysis of acute lung injury and acute respiratory distress syndrome trials testing low tidal volumes. *Am J Respir Crit Care Med* 2002;166:1510–1514.
6. Wise R. Ethical issues confronted in pulmonary clinical trials. *Proc Am Thorac Soc* 2007;4:200–205.
7. Gattinoni L, Brazzi L, Pelosi P, Latini R, Tognoni G, Pesenti A, Fumagalli R. A trial of goal-oriented hemodynamic therapy in critically ill patients. SvO₂ Collaborative Group. *N Engl J Med* 1995;333:1025–1032.
8. The Acute Respiratory Distress Syndrome Network. Ventilation with lower tidal volumes as compared with traditional tidal volumes for acute lung injury and acute respiratory distress syndrome. *N Engl J Med* 2000;342:1301–1308.
9. Brochard L, Benito S, Conti G, Mancebo J, Rekić N, Gasparetto A, Lemaire F. Comparison of three methods of gradual withdrawal from ventilatory support during weaning from mechanical ventilation. *Am J Respir Crit Care Med* 1994;150:896–903.
10. Esteban A, Tobin MJ. A comparison of four methods of weaning patients from mechanical ventilation. *N Engl J Med* 1995;332:345–350.
11. Anzueto ABR, Guntapali K, Weg J, Weidemann H, Raventos AA, Lemaire F, Long W, Zaccardelli DS, Pattishall EN. Aerosolized surfactant in adults with sepsis-induced acute respiratory distress syndrome. *N Engl J Med* 1996;334:1417–1422.
12. Chastre JWM, Fagon JY, Chevret S. Comparison of two durations of antibiotic therapy to treat ventilator-associated pneumonia (VAP) [abstract]. *Am J Respir Crit Care Med* 2003;167:A21.
13. Gattinoni L, Tognoni G, Pesenti A, Taccone P, Mascheroni D, Labarta V, Malacrida R, Di Giulio P, Fumagalli R, Pelosi P, *et al*. Effect of prone positioning on the survival of patients with acute respiratory failure. *N Engl J Med* 2001;345:568–573.
14. van den Berghe G, Wouters P, Weekers F, Verwaest C, Bruyninckx F, Schetz M, Vlasselaers D, Ferdinande P, Lauwers P, Bouillon R. Intensive insulin therapy in the critically ill patients. *N Engl J Med* 2001;345:1359–1367.
15. Hebert P, Wells G, Blajchman MA, Marshall J, Martin C, Pagliarello G, Tweeddale M, Schweitzer I, Yetisir EA. Multicenter, randomized, controlled clinical trial of transfusion requirements in critical care. *N Engl J Med* 1999;340:409–417.

16. Hebert P, Wells G, Martin C, Tweeddale M, Marshall J, Blajchman M, Pagliarello G, Schweitzer I, Calder L. A Canadian survey of transfusion practices in critically ill patients. *Crit Care Med* 1998;26:482-487.
17. Berwick D. Disseminating innovations in health care. *JAMA* 2003;289:1969-1975.
18. Deans KJ, Minneci PC, Suffredini AF, Danner RL, Hoffman WD, Ciu X, Klein HG, Schechter AN, Banks SM, Eichacker PQ, *et al*. Randomization in clinical trials of titrated therapies: unintended consequences of using fixed treatment protocols. *Crit Care Med* 2007;35:1509-1516.
19. Brower R, Matthay M, Morris A, Schoenfeld D, Thompson BT, Wheeler A. Ventilation with lower tidal volumes as compared with traditional tidal volumes for acute lung injury and the acute respiratory distress syndrome. *N Engl J Med* 2000;342:1301-1308.
20. Carmichael LC, Dorinski P, Higgins SB, Bernard GR, Dupond WD, Swindell B, Wheeler A. Diagnosis and therapy of acute respiratory distress syndrome in adults: an international survey. *J Crit Care* 1996;11:9-18.
21. Dreyfuss D, Saumon G. Evidence-based medicine or fuzzy logic: what is best for ARDS management? *Intensive Care Med* 2002;28:230-234.
22. Krishnan JA, Hayden D, Schoenfeld D, Bernard GR, Brower R, for the NHLBI ARDS Network Investigators. Outcome of participants vs. eligible nonparticipants in a clinical trial of critically ill patients [abstract]. *Am J Respir Crit Care Med* 2000;161:A210.
23. Committee on Quality of Health Care in America; Institute of Medicine. To err is human: building a safer health system. Washington, DC: National Academy of Sciences; 2000.
24. Miller GA. The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychol Rev* 1956;63:81-97.
25. Morris AH. Developing and implementing computerized protocols for standardization of clinical decisions. *Ann Intern Med* 2000;132:373-383.
26. Morris AH, East TD, Wallace CJ, *et al*. Standardization of clinical decision making for the conduct of credible clinical research in complicated medical environments. *Proc AMIA Annu Fall Symp* 1996;418-422.
27. Morris AH. Developing and implementing computerized protocols for standardization of clinical decisions. *Ann Intern Med* 2000;132:373-383.
28. Thompson BT, Hayden D, Schoenfeld D. Safety is not improved with the inclusion of a usual care arm in ICU trials [abstract]. *Am J Respir Crit Care Med* 2004;169:A256.
29. Follman DA, Proschan M, Geller, NL. Monitoring pairwise comparisons in multi-armed clinical trials. *Biometrics* 1979;50:325-336.
30. Morris AH. Decision support and safety of clinical environments. *Qual Saf Health Care* 2002;11:69-75.
31. Morris AH. Perceptions, confessions: a systems response is needed. *Am J Respir Crit Care Med* 2003;167:1302-1303.
32. Brower RG, Shanholtz CB, Fessler HE, Shade DM, White P Jr, Wiener CM, Teeter JG, Dodd-o JM, Almog Y, Piantadosi S. Prospective, randomized, controlled clinical trial comparing traditional versus reduced tidal volume ventilation in acute respiratory distress syndrome patients. *Crit Care Med* 1999;27:1492-1498.
33. Amata MB, Barbas CS, Medeiros DM, Magaldi RB, Schettino GP, Lorenzi-Filho G, Kairalla RA, Deheinzelin D, Munoz C, Oliveira R, *et al*. Effect of a protective-ventilation strategy on mortality in the acute respiratory distress syndrome. *N Engl J Med*. 1998;338:347-354.