Comment

Richard Royall

This study is deeply disturbing to me. I believe that no babies should have been randomized to "conventional medical therapy." And given that some were, I believe the decision to use the "randomized consent" procedure, whereby their parents were not informed of that fact, was a grave error. It is impossible to present, in this discussion, all of the ethical and statistical arguments that, in my opinion, fully justify these judgments; I can only offer some observations that I hope will convince the readers of *Statistical Science* to give them serious consideration.

Jim Ware is careful to defend the decision to stop randomizing after 19 patients. His calculations show that "the posterior probability that ECMO is inferior to conventional therapy, given experience in 19 randomized patients, is only .01 when one chooses a prior that ignores recent experience with conventional therapy and falls to .0004 when one bases the prior on that experience." He also finds that "the posterior probability that ECMO has a lower mortality rate is at least .90 over a wide range of priors," and concludes that "Given this analysis it is difficult to defend further randomization ethically."

Recall that this is not the first randomized clinical trial of ECMO versus CMT. The Michigan group (Bartlett et al., 1985) reported on a trial using a play-the-winner rule that randomized 10 patients, 9 to ECMO and 1 to CMT. All 9 ECMO patients survived and the CMT patient died. Ware and Epstein (1985) published comments on that trial, finding that "... the results are not completely convincing. Why not? Primarily because only one patient received the standard therapy." They concluded that "Further randomized clinical trials using concurrent controls and addressing the ethical aspects of consent, randomization, and optimal care will be difficult but remain necessary."

Thus before the Harvard study began, there were 10 cases available from a carefully conducted randomized controlled trial. The main problem was that there were not enough babies in the control (CMT) group. But were there? Using the same prior probability distributions that Ware does, we find that, before the Harvard study was begun, the posterior probability that ECMO is inferior to conventional therapy, given experience in the 10 randomized patients at Michigan, was only .01 when using the prior that ignores Harvard's recent (before their own trial) experience with conventional therapy and falls to .00003 when one uses the prior based on that experience. The corresponding probabilities that ECMO is superior were .90 and .9996 respectively. In Ware's own words "Given this analysis it is difficult to defend further randomization ethically," but now "further randomization" refers to the *first baby* entered in the Harvard trial. The argument that the Michigan results were not convincing because of insufficient concurrent randomized controls is refuted by the same reasoning and calculations that Jim Ware uses to defend the Harvard group's decision to stop randomizing babies to their own control group.

It might be argued that conditions were different at Harvard from those at Michigan, with differences in patient populations as well as in how the two therapies were defined and applied. Thus, although the probability that conventional treatment is better at Michigan is only .01, one cannot be sure that the same is true at Harvard. Note that this argument would apply *regardless of the size of the control group* in the Michigan study. It would mean that no matter how strong the evidence that ECMO sharply improves survival at Michigan, the Harvard trial would still have been necessary. But then the Harvard evidence would not be applicable at Johns Hopkins where yet *another* randomized trial would be necessary ... and so on at hospitals throughout the land. Moreover, because populations, personnel and procedures change over time, one cannot be sure that last year's results at Harvard are relevant to tomorrow's therapeutic decisions there. Now caution in extrapolating findings from one time or place to another is an important aspect of scientific rigor. Thus the present argument can lead to the conclusion that historical controls, or even nonrandomized concurrent controls, are of little scientific value, that good science demands that we avoid extrapolating and settle for nothing less than concurrent randomized controls, that good science demanded the Harvard study.

But if the argument leads to that conclusion, it also leads to the conclusion that the randomized trial itself, being conducted at a unique time and place, has little value. If we are unwilling to extrapolate, to make the assumptions about similarity, stability and consistency required for the judgment that clinicians at other times and places have something to learn from our

Richard Royall is Professor, Department of Biostatistics, The Johns Hopkins University, 614 N. Wolfe St., Baltimore, Maryland 21205.

experience, then randomized clinical trial results are only of temporary, local interest. Thus if we accept the above argument that the only scientifically valid studies are those with concurrent randomized controls, then we are led to the false conclusion that the clinical trial itself is of little value. I think the opinion leading to the Harvard study, the opinion that after the Michigan study "Further randomized clinical trials using concurrent controls . . . remain necessary," was wrong.

I believe there were indeed grave ethical problems involved in randomizing the first baby to "conventional medical therapy" in the Harvard study. Surely there were grave problems in randomizing the last. These problems might have been largely avoided if truly informed consent had been obtained from all of the patients. But it was not. Zelen's randomized consent procedure produced a situation in which parents of a critically ill infant, for whom conventional therapy held little hope of survival,[1] were not even informed that a highly promising[2] alternative therapy was available for their baby, but that by chance conventional therapy had been selected instead. Would they have chosen to let their baby remain in the study and receive conventional therapy, if they had been given the information and the choice? We cannot know. Should they have been given the information and the choice? A negative answer might be justified on the grounds that the randomized trial should not have been done at all, so that no parent would be put in that position. But given that the study was to be done, the decision to withhold the information seems very hard to justify. Ware argues that "The randomized-consent design is attractive . . . because a standard approach to informed consent would require that parents of infants near death be approached to give informed consent for an invasive surgical procedure that would then, in some instances, not be administered." This may be true, but it does not address the key question of why CMT parents were not informed, *after* randomization, about the study, about the therapeutic alternatives, about their physician's considered judgment (uncertain, as all judgments are) concerning what would be best for their baby, about the fact that CMT had been randomly chosen, and about their right not to participate in the study (and not to accept the randomly chosen therapy).

It might be objected that if CMT parents were so informed, then many might have declined to participate, making it impossible to obtain a control sample that was adequate in terms of both size and freedom from selection bias. Angell (1984) discussed a similar quandary in a clinical trial comparing therapies for breast cancer: "What can be done when nonrandomized designs are considered inadequate but randomization would be difficult because of patients' preferences for one treatment or the other? Not all problems have solutions. It simply may not be ethically possible to conduct a valid randomized clinical trial under these circumstances."

Were there no adequate nonrandomized designs for comparing ECMO with CMT? I am sure the Harvard group considered their options carefully and chose to proceed only after a thorough evaluation convinced them of the inadequacy of nonrandomized alternative designs in which every baby received what, in the considered and informed judgment of his physician and parents, was the best therapy available for that baby. But I wonder what perceived flaws precluded a trial using both historical controls and concurrent cases at other institutions where ECMO was not available, along with concurrent cases at Harvard who, by simple bad luck, arrived for therapy at a time when ECMO was unavailable there because of limited facilities and personnel. Five of these latter cases appeared during the Harvard study. Ware reports their outcomes, but they were not included in the study proper because, since ECMO was not available, they could not be randomized.

Of course, even with the best possible protocol to ensure high quality data, carefully developed and applied criteria for inclusion or exclusion of cases, as well as for diagnosis and evaluation, etc., a nonrandomized study will always entail extrapolations. It will entail comparing babies treated with ECMO to babies treated with conventional therapy at a different time or at a different place or under different conditions of ECMO availability. It can never be entirely free from the possibility of selection bias. It can nevertheless be of great scientific value. Randomization is desirable, but as Fried (1974, page 160) so aptly put it ". . . we should not proceed on the fallacious assumption that where there is no randomization there is no truth."

[1] Besides reporting 85% mortality in the 35 pretrial patients at Harvard, Ware chooses a prior distribution for one analysis that ". . . is consistent with reports from other centers of 80% mortality in conventionally treated patients." He describes the infants as "near death," and describes the 60% CMT survival rate as "unexpectedly high."

[2] Ware states that "Several centers have reported excellent results in the use of ECMO . . . . Some investigators reported survival rates in excess of 80% in populations of infants believed to have survival rates of 20% or less in the absence of ECMO treatment." He also reports that at Harvard "the investigator's prior gave greater weight to the superiority of ECMO than of CMT."

ADDITIONAL REFERENCES

ANGELL, M. (1984). Patients' preferences in randomized clinical trials. *New England J. Med.* **310** 1385–1387.

FRIED, C. (1974). *Medical Experimentation: Personal Integrity and Social Policy.* Elsevier, New York.