

Statistics and ethics: Models for strengthening protection of human subjects in clinical research

Christine K. Cassel¹

American Board of Internal Medicine, 510 Walnut Street, Suite 1700, Philadelphia, PA 19106

In this issue of PNAS, Press (1) proposes a method that could substantially advance ethical approaches to clinical trials and comparative effectiveness research. Such advances are needed to harvest the potential benefit of electronic medical information, allowing for greater numbers and far greater specificity of comparisons among diagnostic or treatment modalities in demographic, phenotypic, and genetically distinct comparative populations.

Press (1) deals with the optimization of patient selection for clinical trials that have large or infinite patient horizons. The objective is to assign patients to one of two treatment groups (“arms”) with total cost of treatment being kept to a minimum. “Cost” is not just simple dollar cost, but instead, a combination of measures of loss associated with assignment to the inferior treatment. To make the method relevant to comparative effectiveness research, dollar cost is tallied only when the inferior treatment led to a failed patient outcome. The method is the “two-armed bandit problem” of decision analysis, using a Bayesian approach to estimating success rates (2, 3).

The Equipoise Ideal

Press (1) is absolutely correct in his citation of Royall (4) that there is an inherent ethical problem in placing patients in clinical trials where one arm of the treatment is assumed to be inferior to the other. This reality has plagued the world of bioethics scholarship for decades and generated substantial debate, analysis, and policy. A seminal document, *The Belmont Report* (5), outlined that one reason for informed consent was that patients needed to understand fully when engaging in research that the physician had their best interest in mind and did not know which of the randomized arms was better (6). This agnostic state has been termed “equipoise” and in much of the early ethics literature was a requirement for any comparative clinical trial to be considered ethical. But both psychologists and ethicists have pointed out that scientific progress does not really work that way; that the only reason for doing a clinical trial is because one thinks that one arm of the study will be better, and one is doing the trial to find out whether indeed that

is the case. Any clinical trial is going to be based on early pilot data, perhaps not as thoroughly randomized, suggesting a therapeutic or diagnostic advantage to the “new” arm of the study. Thus, while the ideal of ethics posits equipoise, the reality is a struggle in the minds of both the investigators and the patients themselves. The investigators are told to be careful not to posit the “new” treatment as “better,” and yet the patients coming into the trial often agree to the trial precisely because they

Press raises the vitally important issue of cost of treatment as a variable.

hope to find a better treatment for their condition. The equipoise ideal was further complicated by potential conflict of interest when the researcher was motivated toward academic or financial success and patient benefit.

The policy prescription for this problem was the requirement for informed consent, with clear guidelines about what kinds and levels of information were to be provided and what attributes of setting and patient condition constituted adequate freedom from undue influence for valid consent. Federal regulations (7) established policy for federally funded research, and a substantial and costly infrastructure of Institutional Review Boards (IRBs) was established, both academic and free-standing. The scope of this ethics enterprise grew as the scope of the clinical trials industry grew, to the point where there are ~4,000 IRBs registered with the federal government (G. Drew, Office for Human Research Protections, personal communication), and >26,000 clinical trials were published in 2008 alone (8), including increasing numbers conducted internationally.

The Equipoise Hazard

One could view this larger enterprise as a response to the uncertainty of true equipoise in the minds of researchers and patients. At the same time, the

double-blind randomized clinical trial has been seen as the “gold standard” of methodologically valid clinical research. Enrolling thousands of patients in blinded trials with fixed endpoints created the need for data safety monitoring committees that would monitor cumulative results unseen by the investigators, to stop the trial early if statistically significant results showed clear benefit or harm from one of the arms. Even this form of monitoring means that some number of patients (subjects) will be subjected to inferior treatment before the “stopping rule” is invoked. The statistical solution proposed by Press (1) would allow more realistic approaches to assumptions going into the trial and expose fewer subjects to random assignment as superiority of one arm becomes clear.

One of the most intensely publicly debated examples of this disconnect about equipoise occurred in the 1980s in the early days of the AIDS epidemic where patients suffering from AIDS were eager to get the new treatments and did not want to wait for the completion of a clinical trial, knowing that they might not survive that long (9). Activism actually led to more flexibility from the Food and Drug Administration (FDA) in using innovative approaches and ethical flexibility for people who could be offered the new treatment for “compassionate” reasons, meaning you could try the new drug on them because their condition was so far advanced and no other treatment was available (10, 11). Very little of this debate examined the kind of importantly creative and sophisticated statistical perspectives that Press (1) presents. His solution provides a direction in which decisions to enroll each next individual subject will be made in light of analysis of all of the subjects treated to date. This approach to compassionate statistics might have been just what those early HIV patients needed.

Cost Effectiveness and Personalized Medicine

Press (1) raises the vitally important issue of cost of treatment as a variable

Author contributions: C.K.C. wrote the paper.

The author declares no conflict of interest.

See companion article on page 22387.

¹E-mail: ccassel@abim.org.

and states “there is little dispute that comparative effectiveness research should seek to find interventions that are cheaper and more effective than alternatives.” He goes on to point out that the controversy arises around treatments that are “slightly more effective” but “significantly more expensive.” Even this distinction, between slight and significant, however, has been lost in the kind of fear that is generated by opponents of comparative effectiveness research worried that more expensive treatments may not “get a fair trial” (12) or even more likely that comparative effectiveness will never be able to truly determine whether this particular treatment will work on this particular individual given their unique genetic and biological attributes (13). The real promise for this work would be to make publicly defensible and understandable improvements in addressing both of these issues (14).

As clinical trials become increasingly Bayesian it will rarely be the case that effectiveness is totally unknown. So this heuristic could be expanded to leverage partial results from clinical trials to make it more useful in the real world. In doing this, any optimization based on both effectiveness and cost will have to explicitly trade off one for the other to some extent. Because the recursive analysis of exact methods becomes increasingly computationally intense as the number of patients selected for the study increases, heuristic methods are a reasonable alternative.

While cost effectiveness will inevitably lead to controversial public reactions and require an astute policy voice, the more information is available, including genetic information, the more this adaptive Bayesian approach is likely to get a better answer that is more specific for more specific subgroups of the population. If that is the case, then this kind of knowledge and the clinical recommendations that it leads to will significantly enhance public trust in evidence-based medicine.

To keep the algorithms easy to understand, Press (1) assumes that the success or failure of each assignment becomes known before the assignment of the next patient. In the first phase of his article, he bases selections on an optimal (exact) method, which is more restrictive in the kinds of research to which these methods are able to be applied. In the second phase, he bases selection on an innovative, near-optimal (heuristic) method for comparison. The methods produce similar results, yet there is a distinct advantage to using the non-recursive heuristic method in very large studies or in comparative effectiveness research with multiple variables and endpoints.

The Promise of Comparative Effectiveness

The promise of “personalized medicine” is to use a combination of genetic and clinical information about larger populations of people with the power of computational biology and data-based mod-

eling (15) to both streamline the clinical research process and allow the results to be more directly applicable to individuals. With the growing use of personal health records and electronic medical records, it is within reach for real-time monitoring of patients treated with new modalities to generate results more quickly and with subset analysis that can “personalize” the data to apply to a patient with other complex problems or who might not have been included in a conventional trial for other reasons. The same aggregated data approach also promises much more effective and timely approaches to monitoring drugs and devices for adverse events after FDA approval (16, 17).

All in all, Press’s proposal (1) is a remarkable contribution to a positive direction for the field of comparative effectiveness research and the heuristic should be expanded with the help of policy experts, clinical trials experts, and statistical modelers and evaluated further. Heuristic methods do have a reputation for being robust enough to handle large “horizons” and a number of constraints that gives great promise at a moment in biomedical history when personalized medicine is being touted to the public as right around the corner, and large-scale aggregated databases may result from advances in health information technology. Yet the research methodologies to make good on those promises are still extremely limited by traditional statistical methods.

1. Press WH (2009) Bandit solutions provide unified ethical models for randomized clinical trials and comparative effectiveness research. *Proc Natl Acad Sci USA* 106:22387–22392.
2. Girshick MA (1954) An elementary survey of statistical decision theory. *Rev Edu Res* 24:448–466.
3. Berry DA (2004) Bayesian statistics and the efficiency and ethics of clinical trials. *Stat Sci* 19:175–187.
4. Royall RM (1991) Ethics and statistics in randomized clinical trials. *Stat Sci* 6:52–88.
5. U.S. Department of Health, Education, and Welfare, The National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research (1979) *The Belmont Report*. Available at <http://ohsr.od.nih.gov/guidelines/belmont.html>. Accessed Nov. 19, 2009.
6. Faden RR, Beauchamp TL, King NM (1986) *A History and Theory of Informed Consent* (Oxford Univ Press, New York).
7. U.S. Department of Health and Human Services (2005) Protection of Human Subjects, Title 45, Public Welfare, Code of Federal Regulations, Part 46. Available at www.hhs.gov/ohrp/humansubjects/guidance/45cfr46.htm. Accessed Nov. 19, 2009.
8. Heneghan C (2009) How many randomized trials are published each year? Available at <http://trusttheevidence.net/carl-heneghan/how-many-randomized-trials-are-published-each-year>. Accessed Nov. 19, 2009.
9. Shilts R (1987) *And the Band Played On* (St. Martin’s Press, New York).
10. Gould SJ (1989) The AIDS consumer movement and the FDA: A potential paradigm shift in health care policy. *J Public Policy Marketing* 8:40–52.
11. Food and Drug Administration (1997) Modernization Act of 1997. Available at www.fda.gov/RegulatoryInformation/Legislation/FederalFoodDrugandCosmeticAct/FDCA/SignificantAmendmentsToTheFDCA/FDAMA/FullTextofFDAMALaw/default.htm. Accessed Dec. 1, 2009.
12. Feldstein M (August 18, 2009) Obamacare is all about rationing. *Wall Street Journal*. Available at <http://online.wsj.com/article/SB10001424052970204683204574358233780260914.html#>. Accessed Nov. 19, 2009.
13. Partnership to Improve Patient Care (2009) Individualized patients, personalized care. Available at www.improvepatientcare.org/issue/individualized-patients-personalized-care. Accessed Nov. 19, 2009.
14. Garber AM, Tunis SR (2009) Does comparative-effectiveness research threaten personalized medicine? *N Engl J Med* 360:1925–1927.
15. Archimedes (2009) What is the Archimedes model? Available at <http://archimedesmodel.com/archimedes-model.html>. Accessed Nov. 19, 2009.
16. PricewaterhouseCoopers (2009) Transforming healthcare through secondary use of health data. Available at www.pwc.com/us/en/healthcare/publications/secondary-health-data.jhtml. Accessed Dec. 1, 2009.
17. Institute of Medicine, Board on Health Sciences Policy (2007) Challenges for the FDA: The Future of Drug Safety, Workshop Summary. Available at www.iom.edu/en/Reports/2007/Challenges-for-the-FDA-The-Future-of-Drug-Safety-Workshop-Summary.aspx. Accessed Nov. 19, 2009.