American Journal of Epidemiology

Vol. 161, No. 2
Printed in U.S.A.
DOI: 10.1093/aje/kwi016

Bacchetti et al. Respond to "Ethics and Sample Size—Another View"

Peter Bacchetti[1], Leslie E. Wolf[2], Mark R. Segal[1], and Charles E. McCulloch[1]

[1] Division of Biostatistics, Department of Epidemiology and Biostatistics, School of Medicine, University of California, San Francisco, CA.
[2] Program in Medical Ethics, Department of Medicine, School of Medicine, University of California, San Francisco, CA.

Received for publication September 22, 2004; accepted for publication October 14, 2004.

We thank Dr. Prentice (1) for taking the time to respond to our article (2). We explain here why we do not believe that he has provided a meaningful challenge to our argument. We see possible objections related to unappealing implications, use of power to measure value, implications for series of trials, how value per participant is calculated, and participants' altruistic satisfaction.

To emphasize the implications of our argument, we deliberately chose in our article (2) an example for figure 1 with very high participant burden, noting that the maximum ethical sample size fell below 80 percent power. We realize that some may find this "quite unappealing and counterintuitive" (1, p. 111), but this is not in itself a counterargument. We reemphasize that our argument does not imply that large studies are never ethical or that small studies are better, only that a small study is ethically acceptable whenever a larger one is.

We focused on the assumption that projected value is proportional to power because this is implicit in the argument for ethical condemnation of studies with low power. Other measures of projected value may be more sensible, and we also explored alternatives based on confidence intervals, reaching the same conclusions.

We did not address the more complicated situation of a series of trials, but we do not believe that our argument imposes a cap at the outset on the combined sample size of trials contributing to a meta-analysis, as Prentice suggests. Value assessments for later studies will depend on the results of earlier studies. If formal coherence across multiple studies is really desired, then fully Bayesian approaches might be needed (3). We have examined some expected utility calculations proposed for such methods (3, 4), and they exhibit decreasing value per participant as sample size increases. Thus, our argument continues to hold if those measures are used.

We see little support for what Prentice claims "one might expect" (1, p. 111) about the shape of the value per participant curve. Vague, qualitative reasoning about this is unreliable, just as it was for the existing argument we sought to refute. For example, how are we to know that a "modest benefit" divided by a small sample size really does produce a "low" value per participant? He seems to doubt that value per participant can really be calculated as the study value divided by the number of participants, but it is a simple mathematical fact that comparing study value with total participant burden is equivalent to comparing value per participant (as we defined it) with burden per participant. In addition, we believe that participants' altruistic satisfaction cannot be included as part of the study's value, as Prentice suggests. There must already be a net benefit to justify the participants' altruism. Potential participants certainly weigh altruistic motives when deciding whether to volunteer, but this is distinct from the assessment of ethical acceptability at the planning and approval stages. Indeed, participants are entitled to assume that experts have already judged that the potential scientific or clinical benefits outweigh the burdens they are about to shoulder. If volunteers' own satisfaction must be added to tip the balance, then the study is not ethical.

Scientific inquiry strives for the ideal of following wherever logic and evidence lead, and statistical guidelines should not be exempt from challenge, not even when they are "traditional" and "longstanding." We urge readers to consider our argument on its own merits, without concern for what is conventional, what they have been told by authorities, or other such unscientific considerations. We are confident that those who approach the issue with an open mind will see value in our argument.

REFERENCES

1. Prentice RL. Invited commentary: ethics and sample size—another view. Am J Epidemiol 2005;161:111–12
2. Bacchetti P, Wolf LE, Segal MR, et al. Ethics and sample size. Am J Epidemiol 2005;161:105–10.
3. Lindley DV. The choice of sample size. The Statistician 1997; 46:129–38.
4. Bernardo JM. Statistical inference as a decision problem: the choice of sample size. The Statistician 1997;46:151–3.

Correspondence to Dr. Peter Bacchetti, Box 0560, University of California, San Francisco, CA 94143-0560 (e-mail: peter@biostat.ucsf.edu).