

## Essays and Commentaries

---

# Prisoners' Dilemma: The Importance of Negative Results

Janice C. Probst, PhD

In a desert prison, an older prisoner befriends a new arrival. The young prisoner talks constantly about escape, spinning plan after plan. After a few months, he makes a break. He's gone a week; then the guards drag him back. He's half dead, crazy with hunger and thirst. He wails how awful it was to the old prisoner: endless stretches of sand, no oasis, failure at every turn. The old prisoner listens for a while, then says, "Yep. I know. I tried those escape plans myself, 20 years ago." The young prisoner says, "You did? Why didn't you tell me?" The old prisoner shrugs: "So who publishes negative results?"<sup>1</sup>

The current issue of *Family Medicine* includes a brief report of an educational innovation that *didn't* work.<sup>2</sup> It is useful to review the reasons why research with negative results, such as that reported by Chumley and colleagues,<sup>2</sup> deserves publication and serves as a model for others. First, the research itself explores an important topic and was appropriately done. More broadly, findings from good research are the subject matter of science, regardless of whether the results

support particular hypotheses. Finally, unpublished research wastes resources and leads to poor policy decisions.

### Good and Bad Educational Research

There is a difference between negative findings that result from poor research design and negative findings from good research. In poor research, failure to find effects can stem from an insufficient number of observations, leading to lack of adequate statistical power.<sup>3</sup> Another common problem leading to negative results in educational research is inadequate experimental design, including lack of randomization and lack of control groups. Frequently, learners self-select into an educational intervention or experimental program and thus cannot validly be compared to nonparticipants. In other cases, all learners receive the intervention, and the researchers rely on before and after comparisons to measure effectiveness.

Chumley and associates took advantage of a scheduling situation that allowed creation of case and comparison groups of students who had not self-selected into the intervention, and they also have a sample size that provided adequate power to detect differences. Thus, the research that assessed the intervention was adequately designed to determine whether the intervention would work. Most importantly, the investigators were willing to share their results, even though these re-

sults suggest that either the theory or the implementation they used was "wrong."

### Publication Bias

Publication bias describes the disproportionate representation in the medical literature of papers with positive (ie, statistically significant) findings over those with no significant findings.<sup>4</sup> This phenomenon is so well documented that specific statistical techniques have been developed to compensate for the relative underrepresentation of negative studies when conducting meta-analyses.<sup>5</sup>

Failure to publish negative findings has multiple causes, including bias on the part of reviewers and editors, sponsor influence, and investigator self-censorship. One suspects that the latter—"No one wants to read this, it didn't work"—is a powerful contributor to silence. Regardless of cause, failure to report validly designed, appropriately powered study findings is inappropriate from the standpoint of ethics, common sense, and public policy.

### Ethics

If research involves human subjects, even when those subjects are our trainees, ethical behavior demands that the research outcome be communicated. Individuals who have consented to take part in a study deserve to have their contribution used. One might argue that this is particularly the case when the intervention did not work. A

See related article on pages 696-9.

Fam Med 2006;38(10):742-3.

From the Department of Health Services Policy and Management, University of South Carolina.

known “good” intervention, duplicated, spreads benefit. A seemingly sound but ineffective intervention, however, if never contradicted through publication of negative findings, might be replicated over and over, wasting time for generations of students.

### *Scientific Integrity*

One key scientific reason for sharing “no difference” research were illustrated in the parable at the beginning of this paper. Family medicine educators are constantly working to improve the quality and effectiveness of teaching. The details of an apparently sound intervention that fails to deliver may inspire other educators to revisit both the theory and the implementation of the “failed” concept, leading to educational improvement. Science is a community of shared knowledge. To participate in that community while sharing information selectively devalues the knowledge in circulation.

### *Policy*

Finally, failure to publish educational research can affect public policy. Title VII funds have been made available since 1978 to promote the development of primary care clinicians.<sup>6</sup> An assessment of the effects of this funding during the 1978–1993 period, based on data from 2000, concluded that schools receiving the approximately \$290 million allocated for family medicine had produced more family physicians than other schools and that these physicians were more likely to practice in underserved areas. Nonetheless, differences between high and low performance schools were slight, and as a result, Title VII primary care medical and dental education funds were eliminated from the

FY2006 federal budget and from the FY2007 budget request submitted by the Health Resources and Services Administration (HRSA).<sup>7</sup> Bluntly, HRSA states that these programs were evaluated and found “ineffective.”

Assessing the effects of a national policy is difficult when there is no control group. However, the money disbursed under Title VII supported a significant amount of educational experimentation across more than 25 years. Research objectives among current grantees in residency education, for example, address improvements in teaching for clinical preventive services, chronic disease care, rural health, quality improvement, and patient safety.<sup>8</sup> All of these topics focus on national health priorities and have the potential for improving physician education. In theory, Title VII training grants, along with other educational research and internally supported innovations, could have been generating multiple published studies, potentially changing the way medicine is taught and practiced. In addition to noting small differences in discipline choice at programs that did and did not receive funds, policy analysts might have been able to point to changes in practice and improvements in patient outcomes over time as evidence of Title VII effectiveness. This optimistic scenario, however, would have required educators who rigorously assessed their innovations and shared results even when interventions did not work. Unfortunately, this scenario did not happen.

### **Conclusions**

It is difficult to move forward with submission of a manuscript when a theoretically terrific intervention fails to yield statisti-

cally significant results. I confess to being a member of “great idea” teams that did not publish when our interventions did not work, and I concede that mustering the energy to share “We goofed” is a lot harder than shouting “We won.”

Nonetheless, justifying our position in an increasingly competitive world will require family medicine educators to upgrade our commitment to science. Sharing all of our knowledge, and hoping that others learn from and improve on it, is key to a 21st century discipline.

*Correspondence:* Address correspondence to Dr Probst, University of South Carolina, Arnold School of Public Health, Department of Health Services Policy and Management, South Carolina Rural Health Research Center, Columbia, SC 29208. 803-777-7426. jprobst@gwm.sc.edu.

### REFERENCES

1. Crichton M. A case of need. New York: Signet, Penguin Books, 1968:217-8. [Edited by the author (Probst) to conserve space.]
2. Chumley HS, Dobbie A, Pollock M, Delzell JE Jr. Teaching medical students to prioritize preventive services. *Fam Med* 2006; 38(10):696-9.
3. Chung KC, Kalliainen LK, Spilson SV, Walters MR, Kim HM. The prevalence of negative studies with inadequate statistical power: an analysis of the plastic surgery literature. *Plast Reconstr Surg* 2002;109(1): 1-6.
4. Dickersin K. How important is publication bias? A synthesis of available data. *AIDS Educ Prev* 1997;9(1 suppl):15-21.
5. Pham B, Platt R, McAuley L, Klassen TP, Moher D. Is there a “best” way to detect and minimize publication bias? An empirical evaluation. *Eval Health Prof* 2001;24(2):109-25.
6. Fryer GE Jr, Meyers DS, Krol DM, et al. The association of Title VII funding to departments of family medicine with choice of physician specialty and practice location. *Fam Med* 2002;34(6):436-40.
7. Health Resources and Services Administration. Fiscal year 2007 justification of estimates for Appropriations Committees. [www.hrsa.gov/about/budgetjustification07/TraininginPrimaryCareMedicineDentistry.htm](http://www.hrsa.gov/about/budgetjustification07/TraininginPrimaryCareMedicineDentistry.htm). Accessed September 3, 2006.
8. Health Resources and Services Administration. Residency training in primary care: FY 2005 grant summaries. <http://bhpr.hrsa.gov/medicine-dentistry/05summaries/residency.htm>. Accessed September 3, 2006.