

Scientific Bias Arising from Sampling, Selective Recruitment, and Attrition: The Case for Improved Reporting

Wendy Demark-Wahnefried¹, Deborah J. Bowen², Jennifer M. Jabson², and Electra D. Paskett³

Abstract

In human research, the ability to generalize study findings is incumbent not only on an accurate understanding of the study protocol and measures but also on a clear understanding of the study population. Differential recruitment and attrition has the potential of introducing bias and threatening the generalizability of study findings; yet, relatively few scientific publications report data on sampling, subject exclusion, and dropout. A 4-month census sampling (September–December 2009) of research articles and short communications in this journal ($n = 116$) was no exception. Among articles in which such data were appropriate to report, only 44% documented response rates, 53% described subjects who were excluded, and 10% performed analyses on enrollee versus nonenrollee differences; moreover, of the 17 longitudinal or intervention studies evaluated, only 3 of 17 reported dropout rates, and of those, only 2 of 3 reported reasons for dropout or an analysis that compared the characteristics of dropouts with those of completers. Given *Cancer Epidemiology, Biomarkers and Prevention's* mission to enhance the dissemination of unbiased scientific findings, we propose that guidelines regarding sample description, as defined by CONSORT, STROBE, or STREGA, be adopted by our journal for both observational and interventional studies that accurately describe the study population from the point of contact. *Cancer Epidemiol Biomarkers Prev*; 20(3); 415–8. ©2011 AACR.

Almost a half a century ago, Campbell and Stanley proposed the following description of external validity, "External validity asks the question of generalizability: to what populations, settings, treatment variables, and measurement variables can this effect be generalized?" (1). This definition has been widely accepted, and careful consideration of sampling technique is considered one of the benchmarks of solid research design. Various types of sampling and the pros and cons of each are featured in Table 1. As we conceive of our research protocols, the methods by which we recruit our sample can comprise one or more of these strategies and are based largely on the research question at hand. However, recruitment of human subjects into research studies is complicated and often more complex than initially conceived. In addition, the Health Insurance Portability and Accounting Act (HIPAA) regulations and the downstream effects on Institution Review Boards have made it increasingly more difficult, time-consuming, and costly to recruit

research participants and representative samples into observational studies and clinical trials (2).

Poor study accrual is one of the most common causes of failure for clinical trials, and even for the trials that are completed, subject accrual is often slower than anticipated and the number of subjects who refuse study participation is substantial (3, 4). Selective accrual has the potential of introducing bias into the study results, and can influence findings of not only behavioral and clinical research, but also epidemiologic investigations (5, 6). Likewise, for longitudinal studies (intervention or observation) attrition also can jeopardize external validity, as well as threaten other scientific constructs, such as statistical power (7).

During a session on recruitment and retention convened at the 2009 Annual American Association of Cancer Research Frontiers in Prevention Meeting, it was noted that while careful attention to these topics was usually apparent during the time of study conception, at the time of publication, such information often is lacking. To confirm this observation, an inventory of all research articles and short communications published in *Cancer Epidemiology, Biomarkers and Prevention* (CEBP) from September through December, 2009 (volume 18, numbers 9–12) were reviewed. We reviewed each of the 116 articles for author's inclusion of the following 6 criteria: (i) overall response rate; (ii) number of individuals excluded and explanation for individual exclusion; (iii) analysis on enrollee and nonenrollee demographic (or other) differences; (iv) report of participant dropout or attrition;

Authors' Affiliations: ¹Department of Nutrition Sciences, University of Alabama at Birmingham, Birmingham, Alabama; ²Department of Community Health Sciences, Boston University, Boston, Massachusetts; and ³Department of Internal Medicine, The Ohio State University, Columbus, Ohio

Corresponding Author: Wendy Demark-Wahnefried, UAB Comprehensive Cancer Center, 1675 University Blvd. # 346, Birmingham, AL 35294-3360. Phone: 205-975-4022; Fax: 205-934-7049. E-mail: demark@uab.edu

doi: 10.1158/1055-9965.EPI-10-1169

©2011 American Association for Cancer Research.

Table 1. Recruitment strategies: pros and cons

Recruitment method	Representative of a larger population?	Sampled from definable pool	Application of standard methods for calculation of sample size	Cost
Population based	Yes	Yes	Yes	Expensive
Purposive recruitment	Targets subjects with specific characteristics	No	No	Inexpensive
Convenience sampling	No	No	No	Inexpensive

(v) explanation for participant dropout; and (vi) analysis on differences between dropouts and completers on demographic characteristics or other parameters. Results of this inventory are provided in Table 2. Although some of these studies were secondary analyses and therefore were excluded or not considered applicable, of those that were deemed evaluable, 44% reported response rates, 53% listed the proportion excluded and reasons for exclusion, and only 10% reported analyses on enrollee versus nonenrollee differences; for the 17 longitudinal or intervention studies evaluated, only 18% reported dropout rates (3 of 17), and of those 3, only 2 reported reasons for dropout or an analysis that compared the characteristics of dropouts with those of completers. Thus, our perceptions were confirmed; indeed, the reports appearing within our flagship journal often lack the information needed to generalize findings and to potentially identify sources of bias. This is unfortunate, because this information is crucial for appropriate interpretation, and to weigh the contribution of the current findings against the backdrop of previous research to identify gaps that warrant further investigation.

Roughly 2 decades ago, similar concerns were expressed specifically as they pertained to published reports on clinical trials. (8–11). These discussions gave rise to the Consolidated Standards of Reporting Trials (CONSORT) which now provides guidelines, as well as a 22-item checklist and a flow diagram for use in reporting

the results of clinical trials in a systematic fashion (12). While the CONSORT guidelines focus exclusively on clinical trials and include criteria that are far-reaching, (e.g., criteria for randomization, blinding, study conduct, power calculations, etc.) the framework they provide in areas of sampling and attrition is helpful [e.g., the study flow diagram, and criteria items 13 (participant selection and retention) and 21 (generalizability); ref. 12–14].

A systematic review by Plint and colleagues (2006) explored whether the CONSORT guidelines were associated with improvements in the quality of reports of randomized controlled trials (RCT) and compared report quality in journals before and after the publication of CONSORT guidelines, as well as report quality in journals that adopted CONSORT guidelines versus those that were nonadopters (15). Overall, these guidelines have been associated with an improvement in report quality, though the comparison of CONSORT adopting versus nonadopting journals show greater improvements in criteria in areas such as allocation concealment, than in areas associated with study flow, i.e., sampling, eligibility screening, and retention. That being said, among the subset of journals that have adopted CONSORT guidelines, there is a significantly higher relative risk = 8.06 (95% CI, 4.10–15.83) of attaining quality reporting on study flow-related components when comparing post- to preadoption articles. The authors concluded that indeed

Table 2. Results of sampling and attrition inventory of CEBP research and brief report articles, September 2009–December 2009 ($n = 116$)

Criteria	Status, n (%)		
	Present	Absent	Not Applicable
Response rate	37 (32)	47 (40)	32 (28)
Exclusion criteria	55 (47)	49 (42)	12 (11)
Enrollee vs. nonenrollee differences	10 (9)	94 (80)	12 (11)
Attrition	3 (3)	14 (12)	99 (85)
Description of reasons for dropout	2 (2)	15 (13)	99 (85)
Dropout vs. completer differences	2 (2)	15 (13)	99 (85)

improvements have been realized with the implementation of the CONSORT checklist, and speculate that further improvements are likely as more journals adopt them and editorial efforts are brought to bear.

A review by Mills and colleagues (16) of reports published from 2002 to 2003 in 5 leading medical journals suggests that adherence to study flow criteria is 86% (95% CI, 82–90%) in the post-CONSORT era. Systematic reviews of current scientific reports in the various focus areas, such as urologic surgery, however, still show that relatively few reports meet the criteria established for quality reporting (17). Interestingly, a recent study that assessed trial quality from main outcomes papers, found that commercially-sponsored RCTs appeared to be of superior quality as compared to trials which were sponsored by the government (18). The reporting of attrition, or the lack thereof, was one of the major criticisms of the reports emanating from government-sponsored research. These results provide sad commentary for the bulk of studies which are likely reported in CEBP, that is, behavioral and chemopreventive RCTs that are more likely to be supported by federal funding than by commercial entities.

Although the CONSORT guidelines were the first to address the quality of research reporting and were designed specifically for clinical trials, the charge to improve reporting has been taken up by other research communities, as well. For example, in 2004, a checklist of items was drafted to improve the reporting of observational studies, i.e., Strengthening the Reporting of Observational Studies in Epidemiology (STROBE; ref. 19) and a few years later, guidelines were issued to address the reporting of genetic association studies, i.e., Strengthening the Reporting of Genetic Association studies (STREGA; ref. 20). Like CONSORT, the checklists of STROBE and STREGA address issues beyond recruitment and retention, and generalizability; however, these are common elements across all three. Thus, for the purposes of CEBP, which attracts a diverse readership of epidemiologists and behavioral scientists, the adoption of transparent guidelines to describe sampling and retention may represent "common ground" and a good place to start, especially because the means to evaluate adherence is far more objective than determining potential measurement error, adequate control for confounding, the

appropriateness of design, overinterpretation of the data, etc. Although some have criticized the imposition of guidelines as a measure which will stifle creativity and promote "cookbook science," we view it strictly as a means of reporting which will improve transparency and ultimately lead to superior research (21).

Given CEBP's mission to enhance the dissemination of unbiased scientific findings—a commitment which has spurred the creation of novel platforms, such as "Null Results in Brief" (22, 23), we propose that the guidelines espoused for CONSORT, STROBE, or STREGA (depending on the nature of the study) be adopted by our journal, to transparently convey not only a description of the population from which the sample is drawn, but chronicle numbers, as well as basic characteristics (race, age, cancer type, and stage, etc.) from the point of study contact. Although HIPAA is an acknowledged and formidable barrier to human research, and could be conceived as a barrier to such collection (2), the maintenance of databases with de-identified data on basic demographic and clinical variables should still be possible. Such information could be excerpted from tumor registries or during the point of initial contact if purposive or convenience sampling is performed and stored apart from any identifying information. Such data would help in determining whether the sample that is eventually screened and recruited differs from the pool of origin. Likewise, careful and accurate monitoring of dropout and basic analyses of differences between dropouts and completers also is important for longitudinal investigations. Such action could indeed improve the quality of reporting, allow for better transparency in assessing sources of bias, encourage the publication of studies with more representative samples, and ultimately improve our understanding of the generalizability of findings. The adoption of more rigorous criteria for publication will ultimately allow us to move science forward faster and more clearly.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Received November 4, 2010; revised December 21, 2010; accepted December 22, 2010; published online March 7, 2011.

References

- Campbell DT, Stanley JC. *Experimental and quasi-experimental designs for research*. Skokie, IL: Rand McNally; 1966.
- Ness RB; Joint Policy Committee, Societies of Epidemiology. Influence of the HIPAA Privacy Rule on health research. *JAMA* 2007; 298:2164–70.
- Burchard EG, Ziv E, Coyle N, Gomez SL, Tang H, Karter AJ, et al. The importance of race and ethnic background in biomedical research and clinical practice. *N Engl J Med* 2003;348:1170–1175.
- Fayter D, McDaid C, Eastwood A. A systematic review highlights threats to validity in studies of barriers to cancer trial participation. *J Clin Epidemiol* 2007;60:990–1001.
- Joseph G, Dohan D. Diversity of participants in clinical trials in an academic medical center: the role of the 'Good Study Patient?'. *Cancer* 2009;115:608
- Ruffin MT, Baron J. Recruiting subjects in cancer prevention and control studies. *J Cell Biochem* 2000;34S:80–83.
- Fewtrell MS, Kennedy K, Singhal A, Martin RM, Ness A, Hadders-Algra M, et al. How much loss to follow-up is acceptable in long-term randomised trials and prospective studies? *Arch Dis Child* 2008;93:458–61.
- Pocock SJ, Hughes MD, Lee RJ. Statistical problems in the reporting of clinical trials. A survey of three medical journals. *N Engl J Med* 1987;317:426–32.

9. Altman DG, Doré CJ. Randomisation and baseline comparisons in clinical trials. *Lancet* 1990;335:149–53.
10. The Standards of Reporting Trials Group. A proposal for structured reporting of randomized controlled trials. The Standards of Reporting Trials Group. *JAMA* 1994;272:1926–31.
11. Working Group on Recommendations for Reporting of Clinical Trials in the Biomedical Literature. Call for comments on a proposal to improve reporting of clinical trials in the biomedical literature. Working Group on Recommendations for Reporting of Clinical Trials in the Biomedical Literature. *Ann Intern Med* 1994;121:894–5.
12. CONSORT group. [accessed October 29, 2010]. Available from: <http://www.consort-statement.org>.
13. Begg C, Cho M, Eastwood S, Horton R, Moher D, Olkin I, et al. Improving the quality of reporting of randomized controlled trials. The CONSORT statement. *JAMA* 1996;276:637–9.
14. Moher D, Schultz KF, Altman DCONSORT Group. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. *JAMA* 2001;285:1987–91.
15. Plint AC, Moher D, Morrison A, Schulz K, Altman DG, Hill C, et al. Does the CONSORT checklist improve the quality of reports of randomised controlled trials? A systematic review. *Med J Aust* 2006;185:263–7.
16. Mills EJ, Wu P, Gagnier J, Devereaux PJ. The quality of randomized trial reporting in leading medical journals since the revised CONSORT statement. *Contemp Clin Trials* 2005;26:480–7.
17. Agha R, Cooper D, Muir G. The reporting quality of randomised controlled trials in surgery: a systematic review. *Int J Surg* 2007;5:413–22.
18. Jones R, Younie S, Macallister A, Thornton J. A comparison of the scientific quality of publicly and privately funded randomized controlled drug trials. *J Eval Clin Pract* 2010;16:1322–5.
19. Von Elm E, Altman DG, Egger M, Pocock SJ, Götzsche PC, Vandembroucke JP. STROBE Initiative. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Epidemiology* 2007;18(6):800–4.
20. Little J, Higgins JPT, Ioannidis JPA, Moher D, Gagnon F, von Elm E, et al. Strengthening the Reporting of Genetic Association Studies (STREGA) – an extension of the STROBE statement. *Genet Epidemiol* 2009;33, 581–98.
21. Vandembroucke JP. The making of STROBE. *Epidemiol* 2007;18:797–9.
22. Shields PG. Publication bias is a scientific problem with adverse ethical outcomes: the case for a section for null results. *Cancer Epidemiol Biomarkers Prev* 2000;9:771–2.
23. Shields PG, Sellers TA, Rebbeck TR. Null results in brief: meeting a need in changing times. *Cancer Epidemiol Biomarkers Prev* 2009; 18:2347.