Published in final edited form as:

Psychol Rep. 2007 February ; 100(1): 191–194.

Bioethical Issues in Biostatistical Consulting: Development of a Survey

MIN QI WANG,

University of Maryland

RALPH V. KATZ,

New York University

DONNA HOWARD,

University of Maryland

B. MICHELLE HARRIS, and

University of Maryland

FANG YAN

University of Maryland

Summary

To develop a survey, the Bioethical Issues in Biostatistical Consulting Questionnaire, for investigating bioethical issues in analysis, a comprehensive literature review was conducted to specify areas of bioethics. Through a focus group study and the evaluation by 10 biostatisticians, the questionnaire was constructed. Validation must involve a panel of experts. Now, test-retest reliability and factor analysis should be conducted on a group of eligible subjects.

Although there are no data directly to show the prevalence of unethical behaviors which arise in biostatistical consultations by researchers, scientists' self-reported observations of perceived misconduct suggest that these problematic practices occur with a greater frequency than might be expected by the general public (Swazey, Anderson, & Louis, 1993). Marco and Larkin (2000) identified 16 ethical issues of reporting data, including "failing to report all pertinent data," "failing to report negative results," "reporting differences when statistical significance is not reached," and "inflation of research results for the media," to name just a few.

The purpose of this study was two-fold, (1) to conduct a comprehensive literature review on bioethical issues that arise in relation to data analysis and (2) to develop a survey, the Bioethical Issues in Biostatistical Consulting Questionnaire, with which to assess bioethical issues that arise during biostatistical consultations between researchers and their biostatisticians.

Method

This study involved four phases: (1) conduct of a comprehensive literature review to specify areas of breaches of bioethics related to data analysis and biostatistical consulting as identified in the literature. This automated search, using the keywords or combination of keywords "ethics, bioethics, statistics and ethics, and scientific misconduct," identified over 300 articles primarily covered in the journals presented in Table 1 and a review of articles published

Address correspondence to Min Qi Wang, Ph.D., Department of Public and Community Health, University of Maryland, College Park, MD 20742 or e-mail (mqw@umd.edu)

This study was supported by Grant No. U54 DE14257 from NIDCR/NIH Grant to the NYU Oral Cancer Research on Adolescent and Adult Health Promotion Center.

between 1995 and 2002 which addressed, in some manner, key areas of bioethics related to data analysis; (2) the conduct of a focus group with five university-based biostatisticians to clarify the range of bioethical issues and appropriate language and wording to describe the issues; (3) the development of the Bioethical Issues in Biostatistical Consulting Questionnaire, a preliminary survey; and (4) the conduct of a pilot test of the questionnaire via structured feedback from 10 university-based biostatisticians. These 10 biostatisticians were initially in a poll of 20 biostatisticians who were identified through the research network of the first author. Ten were eliminated for an eligibility reason (less than five years employed as a biostatistician) and availability (on sabbatical or on vacation).

Results

The tentative questionnaire was assessed both qualitatively and quantitatively. The qualitative assessment included the review and the comments from two biostatisticians, two epidemiologists, and one bioethics expert. The evaluation involved their comments about wording, obvious, content validity, and violation of ethics in each item. Only the items which received unanimous agreement that the item was relevant were included in the questionnaire. Quantitative assessment involved responses from the ten selected biostatisticians as to the extent of violation of bioethical principles. In the evaluation respondents were asked to rate the severity of bioethical violation for each item using response categories with anchors of 1 (most severe) and 6 (least severe). Given the small number of bioethical evaluations, no statistical test was applied. Instead, the responses were tabulated on each item and only those items rated as most severe (including the moderately severe) were included in the questionnaire (see Appendix, p. 194).

Discussion

Using these five steps, the Bioethical Issues in Biostatistical Consulting Questionnaire was developed to survey biostatisticians about ethical issues which arose in their biostatistical consulting experience. The questionnaire has 17 items in Section I concerning how often they had directly and personally experienced that specific ethical issue, i.e., first-hand experience, and 17 items in Section II concerning how often they had indirectly heard of that specific ethical issue, i.e., second-hand experience. Although the purpose of this pilot study was to develop a questionnaire for research, it is now important to validate this questionnaire as well as to assess the various reliabilities. One validation should involve a panel of experts. Test-retest reliability and factor analysis should be conducted with a group of eligible subjects.

APPENDIX: Items in Bioethics Issues in Biostatistical Consulting Questionnaire

Section I : During the past 5 years, how many times have you, personally, directly been asked by other researchers to (response options are: 0 1-2 3-4 5-9 10+)

1. Stress only the significant findings while underreporting nonsignificant findings?
2. Remove or alter some data records (observations) to support better an hypothesis?
3. Not report the presence of key missing data which could bias results?
4. Not report the model statistics (including effect size in analysis of variance or R^2 in linear regression) because it appeared too small to indicate meaningful findings?

5. Conduct too many *post hoc* tests but purposefully not adjust alpha levels to make results look more impressive than they are?
6. Modify a measurement scale to achieve some desired results rather than adhering to original scale as validated?
7. Remove categories of a variable to report more favorable results?
8. Interpret the statistical findings based on expectations rather than actual results?
9. Change data to achieve the desired outcome (such as prevalence rate of cancer or other disease)?
10. Falsify the statistical significance (such as the *p* value) to support a desired result?
11. Report results before data have been cleaned and validated?
12. Ignore violations of assumptions since results may change from positive to negative?
13. Not discuss duration of follow-up since it was not consistent?
14. Not fully describe the treatment under study since protocol was not followed exactly?
15. Not mention interim analyses to avoid the problem of “too much testing”?
16. Report power based on a *post hoc* calculation but make it appear as an *a priori* statement?
17. Omit a plot since this did not show as strong an effect as you hoped for?

Section II : During the past 5 years, how many times have you observed—or heard about—other biostatisticians being asked, in their consultations, by other researchers to

1. Not notify the editor when they realized that their published manuscript contained misleading statistical results?
2. Report only results of data analysis performed on subsets of the data?
3. Modify the stated hypothesis when the original null hypothesis of the study is not supported by the data?
4. Report findings significant only by virtue of high statistical power but not practically important?
5. Suppress current findings which contradict previous research?
6. Split the manuscript into smaller parts for the purpose of having more publications which could result in incomplete findings being published?
7. Overstate the statistical findings well beyond what is supported by the data so readers are misled?
8. Alter raw data to support the research hypothesis?
9. Mislead readers by interjecting bias when interpreting a statistical finding?
10. Analyze data fabricated or partially fabricated and intended for publication?
11. Analyze data collected without human subjects review (IRB approval)?
12. Analyze or report analysis that led to breach of confidentiality either incidental or intentional?

13. Propose a study with a flawed design including insufficient power?
14. Break the blinding procedures of the study and not report that?
15. Overstate the sampling design, such as a random sample when the sample was not randomly taken?
16. Change statement of randomization in a study where an investigator does not like the way the randomization has worked out?
17. Set these values aside if the outcome turns on a few outlying values?

References

- Marco CA, Larkin GL. Research ethics: ethical issues of data reporting and the quest for authenticity. *Academic Emergency Medicine* 2000;7:691–694. [PubMed: 10905651]
- Swazey JP, Anderson MS, Louis KS. Ethical problems in academic research. *American Scientists* 1993;81:542–553.

TABLE 1

List of Journals Reviewed

Cancer Epidemiology, Biomarkers & Prevention	Science and Ethics in Medical Research
American Journal of Epidemiology	Science and Engineering Ethics
Epidemiologic Reviews	Accountability in Research
American Journal of Preventive Medicine	Science, Technology and Human Values
Quality of Life Research	Journal of Clinical Epidemiology
Infection Control and Hospital Epidemiology	Academic Medicine
Journal of Clinical Epidemiology	Public Health Reports
Journal of Medical Screening	Journal of Medical Ethics
International Journal of Epidemiology	Human Research Report
Environmental Research	Science
Annals of Epidemiology	Ophthalmologica
Journal of Epidemiology and Community Health	Professional Ethics
The American Statistician	Journal of Nervous and Mental Disorders
British Medical Journal	Studies in Health Technology and Informatics
Journal of International Medical Research	Ethics Report
Controlled Clinical Trials	Journal of Scientific Exploration
Journal of Dental Research	The American Journal of Pediatric Hematology/Oncology
Biometrics	Quintessence International
Statistics in Medicine	Child's Nervous System
British Journal of Clinical Pharmacology	Statistical Methods in Medical Research
	BMC Medical Research Methodology