BMJ Open

Identifying bioethical issues in biostatistical consulting: findings from a US national pilot survey of biostatisticians

Min Qi Wang,[1] Alice F Yan,[2] Ralph V Katz[3]

To cite: Wang MQ, Yan AF, Katz RV. Identifying bioethical issues in biostatistical consulting: findings from a US national pilot survey of biostatisticians. *BMJ Open* 2017;**7**:e018491. doi:10.1136/bmjopen-2017-018491

► Prepublication history for this paper is available online. To view these files please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2017-018491).

Received 2 July 2017
Revised 5 September 2017
Accepted 18 September 2017

[1]Department of Behavioral and Community Health, University of Maryland School of Public Health, College Park, Maryland, USA
[2]Department of Behavioral and Community Health, University of Wisconsin-Milwaukee School of Public Health, Milwaukee, Wisconsin, USA
[3]Department of Epidemiology & Health Promotion, New York University College of Dentistry, New York, New York, USA

Correspondence to
Dr Ralph V Katz;
ralph.katz@nyu.edu

ABSTRACT

Objectives The overall purposes of this first US national pilot study were to (1) test the feasibility of online administration of the Bioethical Issues in Biostatistical Consulting (BIBC) Questionnaire to a random sample of American Statistical Association (ASA) members; (2) determine the prevalence and relative severity of a broad array of bioethical violations requests that are presented to biostatisticians by investigators seeking biostatistical consultations; and (3) establish the sample size needed for a full-size phase II study.

Design A descriptive survey as approved and endorsed by the ASA.

Participants Administered to a randomly drawn sample of 112 professional biostatisticians who were ASA members.

Primary and secondary outcome measures The 18 bioethical violations were first ranked by perceived severity scores, then categorised into three perceived severity subcategories in order to identify seven 'top tier concern violations' and seven 'second tier concern violations'.

Results Methodologically, this phase I pilot study demonstrated that the BIBC Questionnaire, as administered online to a random sample of ASA members, served to identify bioethical violations that occurred during biostatistical consultations, and provided data needed to establish the sample size needed for a full-scale phase II study. The No. 1 top tier concern was 'remove or alter some data records in order to better support the research hypothesis'. The No. 2 top tier concern was 'interpret the statistical findings based on expectation, not based on actual results'. In total, 14 of the 18 BIBC Questionnaire items, as judged by a combination of 'severity of violation' and 'frequency of occurrence over past 5 years', were rated by biostatisticians as 'top tier' or 'second tier' bioethical concerns.

Conclusion This pilot study gives clear evidence that researchers make requests of their biostatistical consultants that are not only rated as severe violations, but further that these requests occur quite frequently.

Strengths and limitations of this study

- First study to quantify bioethics violations in US biostatistical consulting.
- Verified that the Bioethical Issues in Biostatistical Consulting Questionnaire detected differences in frequency and severity of bioethical violations.
- Established sample size needed for full-sized study.
- Established feasibility of recruitment and data collection methods.
- Small sample size of pilot study.
- Limited capability to conduct analysis of co-factors.

INTRODUCTION

This pilot study is the first US national survey to quantitatively identify a wide array of bioethical violations that arise between scientific investigators and their biostatistical consultants, a collaborative research consultation that underpins virtually all scientific studies. This descriptive survey quantifies, for the first time, the frequency of requests for 'inappropriate data manipulation or practices' by investigators via consultations with biostatisticians on a national level. While this phenomenon has been known to exist, the extent to which it exists has simply not been adequately studied, and this lack of research on bioethical research violations has been lamented by several authors.[1–13]

While six previous studies that attempted to quantify aspects of bioethical violations in research have suggested that violations levels were 'of concern', each study has major limitations that preclude the drawing of firm and clear conclusions.[1 2 6 9 11 12] One early study in 1993 only reported on the rate of exposure of doctoral students to perceived misconduct,[1] while another study of that era that evaluated 23 possible ethical research violations reported that 10% of the membership of three surveyed professional research societies had observed data falsification or fabrication.[2] Two later studies targeted research coordinators and asked only a very limited number of questions and achieved low response rates, one 31%, one 37%.[11 12] A fifth study, a survey seeking the opinion of scientific meeting programme chairs from

their annual international research meeting, which focused only on scientific abstracts submitted to their annual meeting, assessed 26 problematic research practices, achieved a response rate of 78% and reported that 30% had observed falsification of data and 54% had observed plagiarism one or more times.[6] The sixth study, which sought to assess scientific fraud experienced by an international group of biostatisticians, reported that 51% were aware of at least one fraudulent study but only achieved a response rate of 37%.[9]

The overall purposes of this pilot study, conducted in collaboration with the American Statistical Association (ASA), were threefold: (1) to administratively pilot test the research methods proposed for use in a full-scale study using the newly developed Bioethical Issues in Biostatistical Consulting (BIBC) Questionnaire as administered to a random sample of US biostatisticians; (2) to establish, for the first time, the prevalence and relative severity of a broad array of bioethical violations requests that are presented to biostatisticians by investigators seeking biostatistical consultations; and (3) to gain estimates of the prevalence and relative severity of those bioethical violations to permit the planning and conducting of a full-scale, phase II study.

METHODS

This phase I pilot national survey used a validated, pretested 18-item BIBC Questionnaire as previously developed within an NIH/NIDCR Oral Health Disparities Center (U54 DE14257) in collaboration with the National Center for Bioethics for Research and Health Care at Tuskegee University.[13] In this phase I pilot study, the 18-item BIBC Questionnaire was administered to a randomly drawn sample of 112 professional biostatisticians who were members of the ASA, as drawn from their national membership list.

Each questionnaire item represents a different bioethical violation event. Specifically, the 18 items ask what bioethical violations the respondent has personally and directly been asked to do during their bioethical consultations over the past 5 years. Respondents were asked to make two assessments for each of the 18 items: (1) the total number of times they had been asked to do that specific bioethical violation over the past 5 years (using a five-point ordinal scale: 0, 1, 2–4, 5–9 and 10+) and (2) their own professional opinion on the 'bioethical violation severity' of that specific bioethical violation (using a five-point ordinal scale ranging from least to most severe: 0–5).

Of the approximately 18 000 total ASA members, approximately 5000 members who are categorised as 'working statisticians' (frequently performing data management and data analysis, consulting to other researchers in data analysis and statistics) comprised the available sample pool. They met the following eligibility criteria: (1) self-identified on their ASA annual registration forms as specialising in biomedical research consulting activities and (2) have at least 2 years of experience as biostatisticians. Our goal for this pilot study was draw a sample of 112 and to achieve a high response rate (>70%) via the use of an endorsement by the ASA and the use three specific incentives to participate: (1) a $99 Amazon gift certificate for completing the estimated 30 min BIBC survey; (2) a web tool online data collection system that avoided the use of any personal identifier for the respondent; and finally for this novel line of inquiry in reporting of violations, (3) the use of the concept of 'requests made to biostatisticians' as its dependent variable (as opposed to the alternative high-risk dependent variable of 'actually committed violations') to ensure higher participation rates, as well as greater participant candour, in this first exploratory study. Data analysis for this initial pilot study consisted of descriptive analysis of the demographic variables, as well as for both the 'perceived severity' rankings and 'frequency' rankings of the 18 listed possible bioethical violations.

RESULTS

First, from an initial working list of 800 emails as provided by the ASA, a random selection process was used progressing in subsets of n=50 to obtain an n=112 while avoiding an overenrolment which would exceed budgetary limits for incentives for enrolled subjects. The final response rate for randomly drawn ASA members was 67%. The demographic data on the respondents revealed that respondents self-reported working as biostatisticians between 2 and 55 years (median number of years=13), and 86.4% were employed full time, 7.3% were self-employed, 2.7% part-time employed with 3.6% not currently working. Of those currently working, 41.8% worked at a university (73.3% at a first tier research university and 11.1% at a second tier research university) while 58.2% were employed at non-university jobs.

Table 1, on its left side, shows the 18 bioethical violations items from the BIBC Questionnaire in ranked order by percentage of respondents rating the item as a '5' (most severe) in 'perceived severity', and then subcategorised into severity group I (the top three most egregious violations), severity group II (the next eight most egregious violations) and severity group III (the seven least egregious violations). The bolded 'q#'s'—within severity groups I and II—are marked by an asterick (eg, **q#2***) and indicate 'top tier violations' (ie, have a 'perceived severity' score of 4–5 for at least 65% of the respondents AND a 'no. of times asked in last 5 years' of 1–10+ times for>20% of the respondents). There were seven identified 'top tier concern violations'.

The unbolded 'q#'s' marked with a cross sign (†) (eg, q#7†)—all these are within severity group III—are labelled as 'second tier concern violations' (ie, have a 'perceived severity' score of 4–5 for at least 33%–64% of the respondents AND a 'no. of times asked in last 5 years' of 1–10+ times for >20% of the respondents). There were also seven identified 'second tier concern violations'.

Table 1 Ranking of bioethical violations by 'perceived severity' and 'number of times directly asked to do it over the past 5 years': BIBC phase I (n=112) findings for q#1–18 which asked biostatisticians 'to estimate the number of times—during the past 5 years—that you, personally, have been DIRECTLY asked to do this'

	Perceived severity score		No. of of times asked		
	'Most severe' or 'high end'		Over past 5 years		
Severity group I: top three bioethical violations as ranked on 'perceived severity'	a '5'	'a 4 or 5'	Never	1–9	10+
q#10. Falsify the statistical significance to support a desired result	91%	92%	96%	3%	1%
q#9. Change data in order to achieve the desired outcome	85%	90%	96%	4%	–
q#2.* Remove or alter some data records in order to better support the research hypothesis	70%	87%	64%	35%	1%
Severity group II: next eight ranked bioethical violations on 'perceived severity'					
q#8.* Interpret the statistical findings based on expectation, not based on the actual results	44%	71%	69%	30%	1%
q#3.* Not report the presence of key missing data that could bias the results	35%	77%	73%	25%	2%
q14. Did not fully describe the treatment under study since protocol was not exactly followed	33%	65%	83%	17%	–
q12.* Ignored violations of assumptions since results may change from positive to negative	33%	69%	68%	29%	3%
q15. Not to mention interim analyses to avoid the problem of 'too much testing'	30%	64%	84%	15%	1%
q16.* Report power based on a post hoc calculation but make it appear as a priori statement	30%	65%	73%	25%	2%
q18.* Request not to properly adjust for multiple testing when 'a priori, originally planned secondary outcomes' get shifted to a 'a posteriori primary outcome status'	29%	66%	72%	27%	1%
q6.* Modify a measurement scale in order to achieve some desired results rather than adhering to the original scale as validate	25%	65%	73%	26%	1%
Severity group III: Lowest seven bioethical violations as ranked on 'perceived severity'					
q7.† Remove categories of a variable in order to report more favourable results	20%	60%	60%	40%	–
q11.† Reporting results before data has been cleaned and validated	18%	49%	40%	51%	9%
q5.† Conduct too many post hoc tests but purposefully fail to adjust alpha levels in order to make results look more impressive than they really are	17%	61%	39%	48%	13%
q13.† Did not discuss duration of follow-up since it was not consistent	16%	39%	74%	74%	–
q1.† Stress only the significant findings	14%	45%	35%	55%	10%
q4.† Not report the model statistics (including effect size in ANOVA or R^2† in linear regression) because it appeared too small to indicate any meaningful changes	12%	39%	66%	32%	2%
q17.† Fail to show plot since it did not show as strong as effect as you would have hoped for	8%	33%	51%	45%	4%

*First top tier concern violations, that is, perceived severity score of 4–5 for at least 65% of sample + 'no. of times asked in last 5 years' of 1–10+ times for at least 20% of sample.
†Second tier concern violations, that is, perceived severity score of 4–5 for 33%–64% of sample + 'no. of times asked in last 5 years' of 1–10+ times for at least 20% of sample.
ANOVA, analysis of variance.

Based on these pilot study findings that the observed effect size of most of the variables in relation to the demographic factors were moderate (ie, in the range of 0.3–04), our follow-up phase II study will seek a sample of 400 ASA members which will have a statistical power above 80% while being able to detect a minimum of 10% difference of the dependent variable between demographic and environmental variables.

DISCUSSION AND CONCLUSIONS

Thus, 14 of the 18 BIBC Questionnaire items, as judged by a combination of 'severity of violation' and 'frequency of occurrence over past 5 years', were rated by biostatisticians as 'top tier' or 'second tier' bioethical concerns, that is, minimally having the characteristics of a 'perceived severity' score in the high range (ie, a score of 4 or 5) for at least 33% of the respondents AND having been 'been asked during a biostatistical consultation' over the past 5 years for at least 20% of the respondents. Inevitably, if unfortunately, the limited sample size of this pilot study prevents detailed subanalyses of the findings by demographic and work environmental factors. Finally, given that these findings are from a pilot study designed to answer methodologic issues, any detailed comparisons of our bioethical violations findings with prior studies would be inappropriate; those comparisons must await the findings from our funded—and now underway—full-sized, phase II study.

Nevertheless, there are clear public health implications from the findings of this phase I pilot study. First, the pilot US national survey quantitatively identified a wide array of bioethical violations that arise between scientific investigators and their biostatistical consultants, giving clear evidence that researchers make requests of their biostatistical consultants that are not only rated as severe violations, and that these requests occur quite frequently. Second, these phase I pilot findings provide strong evidence in support of future studies that will (1) provide replication of these findings in a large sample of subjects and (2) allow a more refined analysis of the findings by demographic variables.

Following our successful completion of this phase I pilot study, our research team submitted a phase II grant that was funded by the Office of Research Integrity at the US Department of Health and Human Services to conduct a follow-up phase II full-sized study which is currently underway, again in collaboration with the ASA. The findings from that phase II full-sized study will serve to more definitively describe both the frequency and severity of bioethical violations requested during biostatistical consultations, as well as guide the development of future educational bioethical training modules targeted at university-based clinical research training programme and their directors as well as to encourage and develop means for research universities and companies to improve their institutional environmental efforts regarding job and publication pressures to reduce the frequency of these bioethical violation requests.

Contributors MQW led the research team in the writing of phase I grant; designed the online data collection system; conducted the data analysis; reviewed and approved the final draft of this manuscript. AFY contributed to the writing of the phase I grant; administered and managed the online data collection system; contributed to the data analysis; reviewed and approved the final draft of this manuscript. RVK initially conceived of the project and research design method; contributed to the writing of this phase I grant; wrote the first draft of this manuscript.

Funding This work was supported by a grant from the Office of Research Integrity (ORI) of the U.S. Department of Health and Human Services (DHHS). Grant No. 1 ORIIR150017-01-00.

Competing interests None declared.

Ethics approval This pilot study was approved by the IRB at the University of Maryland School of Public Health and by the IRB at New York University as an Expedited Review category involving minimal risk for the subjects.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement Technical appendix, statistical code and dataset are available from MQW (Department of Behavioral and Community Health, University of Maryland School of Public Health) via his email: mqw@umd.edu.

REFERENCES

1. Swazey JP, Anderson MS, Louis KS. Ethical problems in academic research. *AmSci* 1993;81:542–53.
2. Greenberg M, Greenberg L. Ethical challenges to risk scientists: an exploratory analysis of 271 survey data. *Science, technology and human values*. , 1994:19, 223–41.
3. Marco CA, Larkin GL. Research ethics: ethical issues of data reporting and the quest for authenticity. *Acad Emerg Med* 2000;7:691–4.
4. *Responsible Science: Ensuring the Integrity of the Research Process. Vol. 1: 17, National Academy of Science*, 1992.
5. Eastwood S, Derish P, Leash E, *et al*. Ethical issues in biomedical research: perceptions and practices of postdoctoral research fellows responding to a survey. *Sci Eng Ethics* 1996;2:89–114.
6. Bebeau MJ, Davis EL. Survey of ethical issues in dental research. *J Dent Res* 1996;75:845–55.
7. Frankel MS. Ethics in research: current issues for dental researchers and their professional society. *J Dent Res* 1994;73:1759–65.
8. Ellenberg SS. Fraud is bad, studying fraud is hard. *Control Clin Trials* 2000;21:498–500.
9. Ranstam J, Buyse M, George SL, *et al*. Fraud in medical research: an international survey of biostatisticians. *Control Clin Trials* 2000;21:415–27.
10. Broome ME, Pryor E, Habermann B, *et al*. The scientific misconduct questionnaire-revised (SMQ-R): validation and psychometric testing. *Account Res* 2005;12:263–80.
11. Pryor ER, Habermann B, Broome ME. Scientific misconduct from the perspective of research coordinators: a national survey. *J Med Ethics* 2007;33:365–9.
12. Habermann B, Broome M, Pryor ER, *et al*. Research coordinators' experiences with scientific misconduct and research integrity. *Nurs Res* 2010;59:51–7.
13. Wang MQ, Katz RV, Howard D, *et al*. Bioethical issues in biostatistical consulting: development of a survey. *Psychol Rep* 2007;100:191–4.