PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (**http://bmjopen.bmj.com/site/about/resources/checklist.pdf**) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Identifying bioethical issues in biostatistical consulting: findings from a U.S. national pilot survey of biostatisticians
AUTHORS	Wang, Min; Yan, Alice; Katz, Ralph

VERSION 1 – REVIEW

REVIEWER	Rueben C. Warren, Professor of Bioethics and Director of the National Center for Bioerthics in Research and Health Care Tuskegee University United States
REVIEW RETURNED	28-Jul-2017

GENERAL COMMENTS	Good work. You clearly indicate that this is a pilot study, so your findings will not be generalized, inappropriately. Your questionnaire is an unique attempt to quantify what many scientists perceived to be a major ethical challenge with limited ability to quantify their perceptions.

REVIEWER	Allison Nugent National Institute of Mental Health National Institutes of Health Bethesda, MD USA
REVIEW RETURNED	09-Aug-2017

GENERAL COMMENTS	In this report, the authors describe a study to assess how often biostatisticians are requested to perform actions that they view as unethical. While this is an extremely worthwhile research question, and the actual data collection was performed appropriately, there are issues with the manuscript as presented, detailed below. 1) Abstract. Under "primary and secondary outcome measures" the authors describe methods - ranking of 18 violations, grouping into different tiers, etc. What is the actual outcome measure? 2) Introduction. The exploration of prior research and motivations for the study is inadequate. The authors mention two prior studies, but not the results of those studies. A sentence that states that these violations are "known to exist" but "not adequately studied" has seven references, the content of these references should be explored. Is there a reference for the development of the BIBC Questionnaire? How was it devised? What does it measure? 3) Did the survey respondents sign informed consent, or did the IRB judge the study exempt from the consent requirement? 4) There is absolutely nothing in the methods section which describes the data analysis.

	5) Results - why were retired statisticians included? If they hadn't been working for some of the prior five years, that would potentially skew the number of requests received. 6) Were any other demographic variables collected? 7) How were the cutpoints of "score of at least 4 or 5 for 65% of respondents" etc. derived? How were "severity groups" defined? 8) Many interesting statistics are left out. How many of the respondents reported any of the 18 violations? This is impossible to tell due to overlap of the violation requests. Were there more requests at university vs. non-university statisticians? Was any data collected regarding field of specialization? 9) The authors state that part of the purpose of this study was to determine sample size for a larger study. Was this done? 10) The discussion should put the current results in the context of the extant literature. 11) Limitations should be discussed in the Conclusions.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
Reviewer Name: Rueben C. Warren, Professor of Bioethics and Director of the National Center for Bioerthics in Research and Health Care
Institution and Country: Tuskegee University, United States
Please state any competing interests: None declared

Please leave your comments for the authors below

Good work. You clearly indicate that this is a pilot study, so your findings will not be generalized, inappropriately. Your questionnaire is an unique attempt to quantify what many scientists perceived to be a major ethical challenge with limited ability to quantify their perceptions.
NO RESPONSE needed, as no requests/suggestions made.

Reviewer: 2
Reviewer Name: Allison Nugent
Institution and Country: National Institute of Mental Health, National Institutes of Health, Bethesda, MD USA
Please state any competing interests: None

Please leave your comments for the authors below

In this report, the authors describe a study to assess how often biostatisticians are requested to perform actions that they view as unethical. While this is an extremely worthwhile research question, and the actual data collection was performed appropriately, there are issues with the manuscript as presented, detailed below.

Comment 1) Abstract. Under "primary and secondary outcome measures" the authors describe methods - ranking of 18 violations, grouping into different tiers, etc. What is the actual outcome measure?

Response: The ‘actual outcome measure’ was, in point of fact, the ‘ranking’ of these 18 bioethical violations by the respondents, as stated….nothing more. Therefore, no revision made to the manuscript.

Comment 2) Introduction. The exploration of prior research and motivations for the study is inadequate.
The authors mention two prior studies, but not the results of those studies.

Response: As per the request of Reviewer #2, the literature review summary in the Introduction section has been expanded with an entire paragraph added providing more details given in this revised submitted manuscript (lines #: 78-94)

Comment 2a) A sentence that states that these violations are "known to exist" but "not adequately studied" has seven references, the content of these references should be explored.
--- see response to #2 above, as detailed content of the References are now provided in the newly added paragraph (lines #: 80-94)

Comment 2b) Is there a reference for the development of the BIBC Questionnaire? How was it devised? What does it measure?

Response: Yes, the reference to the development of the BIBC Questionnaire is Reference #13, and was cited in the first paragraph of the Methods section of the originally submitted manuscript. (line #: 109….see Ref #)

Comment 3) Did the survey respondents sign informed consent, or did the IRB judge the study exempt from the consent requirement?

Response: This study was approved by the IRB as an 'Expedited Review category #7' study involving 'Minimal Risk'. The IRB Consent Forms were part of the 'online data collection system', as approved by the IRB, and invited subjects read the Informed Consent Form and then checked a box indicating that they "agree to participate' (or 'not'), …and if the former, they then proceeded to the survey questionnaire itself. These words were added to Methods section of the Revised Manuscript "…as an Expedited Review category involving minimal risk for the subjects"
(lines #: 139 & 140).

Comment 4) There is absolutely nothing in the methods section which describes the data analysis.

Response: A sentence was added in the Methods section to address this concern:
"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." (lines #: 135-137)

Comment 5) Results - why were retired statisticians included? If they hadn't been working for some of the prior five years, that would potentially skew the number of requests received.
In fact, no 'retired biostatisticians' were included in the sample drawn for Pilot Study….so none of the

Response: 112 respondents were 'retired'. That 3.6% should have been labeled as 'not currently working/employed'. So, this 3.6% were relabeled as 'not currently working' (line #": 150).

Comment 6) Were any other demographic variables collected?

Response: No, no other demographic variables were collected.

Comment 7) How were the cutpoints of "score of at least 4 or 5 for 65% of respondents" etc. derived? How were "severity groups" defined?

Response: In this Pilot Study—with its small sample size—the descriptive data patterns were reviewed by the research team and cut-off points were created by concensus opinion that best—and most meaningfully—summarized the findings…both for 'frequency' and for 'severity'.

Comment 8) Many interesting statistics are left out. How many of the respondents reported any of the 18 violations? This is impossible to tell due to overlap of the violation requests. Were there more requests at university vs. non-university statisticians? Was any data collected regarding field of specialization?

Response: As this was a Pilot Study, our report is focused on answering our primary points of interest that were 'being piloted', namely: 1) Did the BIBC Questionnaire work to discern bioethical violations (in frequency and in severity) in surveyed biostatisticians?; 2) Did the BIBC reveal broad patterns among the respondents?; 3) Did our online data collection work and work efficiently for the respondents?: and, 4) what sample size would be needed in the Phase II full-sized study? Those are our targeted questions from this true 'pilot study'.

The detailed analyses inquired about by the reviewer, while 'of interest', will have to wait to be answered in our full-sized Phase II study (now underway) which will be sized to meaningfully explore the data more deeply by demographic and work-environment sub-groups. To address this for the readership, the following statement: "Inevitably, if unfortunately, the limited sample size of this pilot study prevents detailed sub-analyses of the findings by demographic and work-environmental factors." (lines #: 180-182)

Comment 9) The authors state that part of the purpose of this study was to determine sample size for a larger study. Was this done?

Response: Yes. Our Phase II grant application for the funding agency (ORI, DHHS) used the early analyses of our Phase I Pilot Study to estimate a sample size for our Phase II full-sized study which led to our getting approved to conduct the Phase II study, now funded and underway.

This detail has now been added to Results section in our revised manuscript:
"Based upon these pilot study findings that the observed effect size of most of the variables in relation to the demographic factors were moderate (i.e., 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 able to detect a minimum of 10% difference of the dependent variable between demographic and environmental variables." (lines 168-172)

Comment 10) The discussion should put the current results in the context of the extant literature. Given that this was only a pilot study designed to answer a few key methodological issues, the

Response: following sentence was added to remind the readership of this inherent limitation of pilot studies: "Finally, given that these findings are from a pilot study designed to answer methodologic issues, any detailed comparison of our bioethical violations findings with prior studies would be inappropriate; those comparisons must await the findings from our funded—and now undeway—full sized, Phase II study.
(lines #: 182-186).

11) Limitations should be discussed in the Conclusions.
This is now fully addressed (see inserted new sentences in #8 and #10 above).