Author’s response to reviews

Title: Data Access Committees

Authors:
Phaikyeong Cheah (phaikyeong@tropmedres.ac)
Jan Piasecki (jan.piasecki@uj.edu.pl)

Version: 1 Date: 17 Nov 2019

Author’s response to reviews:

17 Nov 2019

Dear Editor

Response to reviewers’ comments
Title: Data Access Committees
Authors: Phaik Yeong Cheah and Jan Piasecki

We would like to thank the reviewers for their helpful comments. Here are our responses as well as a list of changes in comparison with the previous version of the manuscript. Pages refer to the pages in the track changed document.

Response to Dr Angela Ballantyne (Reviewer 1):

Comment 1: “I think it would help to be really explicit at the beginning that the scope of the paper is 'research data'; you are not talking about sharing health data collected in a clinical context or health administrative data, insurance data etc.”
Response. We concur with the recommendation.
We added a sentence (Introduction, page 5, last second paragraph): “The sharing data collected in a clinical context or health administrative data, insurance data is beyond the scope of this paper.”

Comment 2: “In the "Introduction" you describe positions in the existing literature that advocate for a public health ethics approach to data sharing. From the rest of the paper, it seems you endorse these arguments but this is never explicitly stated. So it would be easier for the reader if you could clarify your position here.”
Response: Indeed we adopt this position, however, especially at the beginning of the paper, it is not stated expressis verbis. We added a short passage to introduction that clarifies our position (Introduction, page 4 and 5, paragraph 2): “This is because public health ethics focuses on public benefit, proportionality, accountability, equity and trust while research ethics tends to focus on consent and individual interests. (14) There are significant similarities between public health activities, epidemiological research and data sharing in terms of goals, benefits and risks associated with these activities. Therefore in this paper we defend a position that data sharing should be guided by the principles of public health ethics than by the research ethics. We provide reasons for this position in the later part of the paper.”
Comment 3: Can you please also clarify whether you are assuming that any research data potentially available for sharing already has broad consent for future use from the data subjects at the point of data collection; or whether the DAC can also grant consent waivers for the re-use of research data where subjects did not give consent for future use? If both types of data sharing are to be considered, can you specify whether the same standards should apply (social value, minimal harm, institutional priorities etc) to both?

Response: We added a passage that clarifies this issue (page 9, paragraph 2):

“Protection of data subjects also entails protecting their rights. DACs should make sure that the shared data do not contain any personally identifiable information, and that data will be used within the scope of broad consent provided by subjects. In the case of old datasets where broad consent for sharing was not obtained, DACs should adhere to the criteria set out by CIOMS 2016: the secondary use offer important otherwise unobtainable information, has social value, and poses minimal risks to the subjects, and that it would be impracticable or prohibitively expensive to contact subjects for their consent (7).”

Comment 4: “I think there is more debate about the 'proper' role of RECs than the paper suggests. See for example ‘The job of ethics committees’ by Andrew Moore, https://jme.bmj.com/content/44/7/481.short; and the responses by Schaefer and Holm and others. Moore argues that the proper role of ethics committees is to assess whether proposed research complies with national policy. Furthermore, I agree with you that a primary role of RECs is to protect research subjects, but personally I would argue they should also play a role in promoting and facilitating socially valuable research. I suggest that you could be more explicit about the degree of debate concerning the appropriate role of RECs; and that some of what you suggest DACs should do differently to RECs, is consistent with what other writers think that RECs should actually be doing in the first place.”

Response: Indeed there is a broad discussion on the role of RECs and we agree that RECs can also play a role in promoting and facilitating socially valuable research. However, we are not focusing on how the REC should work, but merely stating they have been actually working, that is why we stressed protective and adversarial side of the institution. At the reviewer’s suggestion, we have edited the passage and added the suggested references:

See page 13, paragraph 2

“However, there is no a universal standard of how exactly the RECs should function. Some suggest that RECs should check, if research protocols are consistent with certain ethical or legal codes rather than with abstract ethical principles; (38) when others argue that RECs should “perform ethically informed code consistency review”.(39)”

Comment 5: In "Functions" you describe the primary question that DACs should consider. I think it would help to clarify here that you see DACs as only assessing consequences (harms and benefits), and not assessing rights. This will set you up for the conversation on privacy later in the paper, where you state that you are not interested in privacy breaches per say, but whether any harm arises from such rights violations

Response: Thank you for that comment that helps us to clarify our stance. In fact to some extent DACs are also in a position to protect participants’ rights. We added a short passage, see the response to the comment 3 above. This happens when DACs check, if sharing is within the scope of broad consent. The consent form sometimes specifies exactly what is the scope of data sharing, and DACs should provide assurance that the participants’ will is respected.
Comment 6: “In "protecting data subjects" - here you talk about the risks to communities and data subjects, I think it is worth acknowledging that part of the epistemic challenge with data sharing is that we don't necessarily know what these risks will be if we don't get consult with communities. In the end, it often won't be practical to require consultation, but it is worth being explicit about the risk of underestimating or misidentifying risks to data subjects if we don't consult. On this same point, in "Composition" it may be worth noting that DACs can always consult on an ad-hoc basis with people familiar with the community or data subjects where necessary; they don't need to be standing members of the DAC.”

Response: We have added this sentence under the “Protecting data subjects, their communities, data producers, their institutions and the scientific enterprise” (page 9, paragraph 1)

“We recognize that that DACs do not necessarily know what these group harms will be. There will be risks of underestimating or misidentifying potential harms to data subjects. These risks can be minimized by careful engagement with research communities during primary research,(28, 29) e.g. consultation with community leaders and community advisory boards (30, 31).”

In addition, we have added a sentence in the “Composition” section (page 11, paragraph 3, last sentence).

“DACs can also consult on an ad-hoc basis with people familiar with the community or data subjects where necessary.”

Comment 7: “Also in "protecting data subjects" - you make the point that the magnitude of risks to participants in research is different to the magnitude of risks to data subjects in data sharing. I agree, but I think it is not simply the magnitude but also the nature of the risks in research v data sharing that differ.”

Response: We concur and have now made our point clearer. See “Promotion of data sharing in the interest of science, data producers, data subjects and their communities” (page 7, paragraph 1)

“The nature and magnitude of risks arising from secondary data use is different from the nature and magnitude of risks of original studies. This fact is reflected in many existing international and national regulations.”

Comment 8: “In the abstract you mention that DACs should assess the social value of the research and ensure research is low risk. In the body of the paper you argue that benefit should be proportional to potential harm. Does this mean there is no upper limit on the potential harm a study could involve? Or are you assuming that all data sharing is low risk? You also argue that DAC should promote institutional interests; what would happen when there is a conflict between the interests of the institution and value of the research - e.g. a really value data sharing arrangement is proposed, but the institution fails to recognise this value, or the value doesn't correspond to institutional priorities - what takes precedent here?”

Response: Thank you for pointing this out. We do not assume that all data sharing is low risk. We have now clarified our position.

See under “Promotion of data sharing in the interest of science, data producers, data subjects and their communities” (page 6, paragraph 2)
“However, unlike primary research, data access should be granted as long as the data reuse fulfils the criterion of having even a minimal social value (7), and minimal risk to data subjects and other stakeholders.”

We indeed argue that DAC should encourage secondary uses that promote the interests of the institution. This is to motivate primary researchers to share their data and promote data sharing overall. It does not mean that data should not be shared sharing if it conflicts with institutional priorities.

See page 7, paragraph 2, last sentence.

“However, it does not mean that data should not be shared if the objective of the secondary use conflicts with institutional priorities.”

Comment 9: "Goal of review"…"the main goal of an ethics committee review…" here you state the researchers may overlook aspects of a study because of COIs, but more often this happens because of research inexperience or lack of familiarity with the specific law or research ethics policy for a certain type of research.

Response: We agree. We have revised the sentence to:

“The goals of an ethics review are to discern those aspects of a study that were overlooked or misjudged by a researcher whose perspective might be skewed by conflicting interests, and to ensure that the research complies with specific laws and research ethics guidelines.”

See page 13, paragraph 1

Comment 10: Can you please clarify whether you think DACs should assess the methodology (scientific integrity, data analysis plan) of the secondary analysis, or is this outside the scope of DAC review?

Response: We think that this is outside the scope of a DAC.

See paragraph 3 under “Goals of Review”.(this also addresses a comment by Reviewer 2). See page 14-15.

“RECs review new research studies and any secondary data research that requires ethical approval which can differ from jurisdiction to jurisdiction. DACs review data access requests for secondary uses. These uses may be for secondary data research but could also be for teaching purposes, to confirm the findings of the original analyses or other purposes. DACs roles should not include full review of the secondary research such as methodology of the secondary research and the statistical approaches. That is the job of RECs.”

Response to Dr Cason Schmit (Reviewer 2):

Comment 1: page 5 (page numbers relate to PDF package document, where the main document begins at page 5) Line 7: the authors' focus on de-identified data is curious. Most laws only restrict identifiable data, so in most contexts de-identified data faces far fewer barriers to disclosure. It seems to me that the proposed framework would also help address some of the real and perceived barriers for releasing data that is still "identifiable" (e.g., contains at least some direct or indirect identifiers). I would like to see some justification/rationale for this limitation somewhere.

Response: This is true that in some jurisdictions research projects with de-identified data are not considered research involving human beings (e.g. US). However data are protected before being de-identified and according to influential Guidelines (e.g. CIOMS 2016 as well as European GDPR) one cannot de-identify to avoid the restrictions of consent. The use of data has to be within the scope of previous (broad) consent. We clarify this issue and we added a passage.

See page 9, paragraph 2.
“Protection of data subjects also entails protecting their rights. DACs should make sure that the shared data do not contain any personally identifiable information, and that data will be used within the scope of broad consent provided by subjects. In the case of old datasets where broad consent for sharing had not been obtained, DACs should adhere to the criteria set out by CIOMS 2016: the secondary use offer important otherwise unobtainable information, has social value, and poses minimal risks to the subjects, and that it would be impractical or prohibitively expensive to contact subjects for their consent (7).”

Comment 2: Page 5 Line 40-43: "new research methodologies tend to pose fewer risks to participants and at the same time promise benefits at a larger scale," this statement is crazy broad, and consequently, I am highly skeptical of its accuracy. Moreover, the supporting example is so vague that it provides little support for the statement relating to benefits, and does not support the statement relating to risks. The manuscript should either remove this statement or provide stronger evidence for why new research methodologies pose fewer risks and larger scale benefits. A more thorough description of the methodologies would be helpful.

Response: Thank you for pointing this out. We have removed this sentence.

Comment 3: Page 5 line 50-60: the manuscript makes an argument for why public health ethics is more appropriate than research ethics. At some point, the manuscript should tell the reader why public health ethics is meaningfully different from the status quo bioethics/research ethics. The manuscript does highlight the common good v. individual protection argument, but this focuses on only one principle from each ethical framework. As the manuscript indicates, public health ethics is comprised of several different ethical principles. Without discussion, some readers might be mistaken that the only meaningful difference between public health ethics and research ethics is the focus on common good.

Response: We have added a sentence in Introduction (paragraph on pages 4 and 5) “This is because public health ethics focuses on public benefit, proportionality, accountability, equity and trust while research ethics tends to focus on consent and individual interests.(14) There are significant similarities between public health activities, epidemiological research and data sharing in terms of goals, benefits and risks associated with these activities. Therefore in this paper we defend a position that data sharing should be guided by the principles of public health ethics than by the research ethics. We provide reasons for this position in the later part of the paper.”

Comment 4: Page 8 line 7: The "goals" of institutions vary wildly. Could this include profit, political objectives, etc? As a reader, I would be interested to know if these types of "goals" would fit in this framework. However, I would respect a decision to not provide specificity on this point.

Response: Thank you for this comment. We have clarified in the Introduction that we are primarily talking about public funded research.

Page 5, paragraph 3
“While many previous discussions have centred around genomic data, our paper discusses DACs which operate as custodians of all types of health data generated from public funded health research.”

Comment 5: Page 17 line 10-13: The scope limitation is a bit late in the document. I am curious why this limitation at all. Surely there are socially beneficial uses of education data, or government services data. After all, there are many non-clinical datasets that contain information on various social determinants of health. Are there anticipated problems or pitfalls that the authors are trying to avoid?
Response: Thank you. We have removed this limitation.

Comment 6: General. The DAC is proposed to have a complementary role to RECs. One could argue that this only adds a bureaucratic layer to data use. Moreover, since 1) both DACs and RECs involve ethical reviews and 2) the reviews use different ethical frameworks that both analyze different aspects of risk, then DACs could create an additional barrier to data disclosure and use. I would like to see the manuscript address this argument.

Response: Thank you for this comment. We have clarified the roles of RECs and DAC. See page 14-15, under “Goals of Review”.

“RECs review new research studies and any secondary data research that requires ethical approval which can differ from jurisdiction to jurisdiction. DACs review data access requests for secondary uses. These uses may be for secondary data research but could also be for teaching purposes, to confirm the findings of the original analyses or other purposes. DACs roles should not include full review of the secondary research such as methodology of the secondary research and the statistical approaches. That is the job of RECs.”

Minor Issues:

MI 1: page 5 line 7: The manuscript uses various terms to refer to data that is not identifiable (e.g., de-identified, anonymized), in some cases these are terms of art that might relate to legal definitions. It is not clear to me if these terms are intended to be interchangeable or if some are referring to a term of art. I would encourage consistent terminology, and if relevant, some clarification if they are talking about specific legal definitions, statistical determinations, or if the term is just used generally.

Response: We have now clarified what we mean by anonymised data and added a reference. See first sentence in “Protecting data subjects….“ (page 8)

Anonymised data - “personally identifiable information (PII) is irreversibly altered in such a way that a PII principal can no longer be identified directly or indirectly”

MI 2.: Page 7 lines 16-19. Clearly, minimal social value is intended to be a low bar. Does economic benefit count as social value? This question comes up in several places in this document.

Response: We refer to the CIOMS 2016 guidelines for definition of social value ie.the prospect of generating the knowledge and the means necessary to protect and promote people’s health. We have now referenced CIOMS 2016.

MI 3: Page 7 line 33: Moral equivalence is strange in this context, particularly since the comparison involves broad categories of research. I don't know how research using data of child victims of incest fits within this moral equivalence framework. Suggest "Although data research carries risks, it typically involves less risk than enrolling subjects in new clinical or observational research studies."

Response: We have removed the word “moral”

MI 4: Page 7 Line 33-48: Also relevant GDPR Art. 5(1)(b).

Response: We have added this article.

IM 5: Page 7 lines 50-57: Replace HIPPA with HIPAA. Please separate the legal citations. Currently, it appears as if the HIPAA provision is part of the Common Rule, which is not accurate. Note that BOTH HIPAA and the Common Rule permit the disclosure and use of IDENTIFIABLE data for research. In
In addition, we have made minor changes throughout the manuscript to improve the clarity of our arguments. We have also included some additional references.

We hope that our paper is now suitable for publication.

Many thanks.

Phaik Yeong Cheah
Corresponding Author