



RESEARCH NOTE

REVISED A review of data sharing statements in observational studies published in the BMJ: A cross-sectional study [version 2; referees: 2 approved]

Laura McDonald¹, Anna Schultze², Alex Simpson², Sophie Graham², Radek Wasiak², Sreeram V. Ramagopalan ¹

¹Centre for Observational Research and Data Sciences, Bristol-Myers Squibb, Uxbridge, UB8 1DH, UK

²Real-World Evidence, Evidera, London, W6 8DL, UK

v2 First published: 19 Sep 2017, 6:1708 (doi: [10.12688/f1000research.12673.1](https://doi.org/10.12688/f1000research.12673.1))
 Latest published: 06 Nov 2017, 6:1708 (doi: [10.12688/f1000research.12673.2](https://doi.org/10.12688/f1000research.12673.2))

Abstract

In order to understand the current state of data sharing in observational research studies, we reviewed data sharing statements of observational studies published in a general medical journal, the British Medical Journal. We found that the majority (63%) of observational studies published between 2015 and 2017 included a statement that implied that data used in the study could not be shared. If the findings of our exploratory study are confirmed, room for improvement in the sharing of real-world or observational research data exists.

Open Peer Review

Referee Status:

	Invited Referees	
	1	2
REVISED version 2 published 06 Nov 2017		 report
version 1 published 19 Sep 2017	 report	 report

- 1 **Laurie Tomlinson** , London School of Hygiene & Tropical Medicine, UK
- 2 **Robin E Champieux** , Oregon Health & Science University, USA

Discuss this article

Comments (0)

Corresponding author: Sreeram V. Ramagopalan (sreeram@ramagopalan.net)

Author roles: **McDonald L:** Formal Analysis, Investigation, Methodology, Project Administration, Writing – Original Draft Preparation; **Schultze A:** Writing – Review & Editing; **Simpson A:** Writing – Review & Editing; **Graham S:** Writing – Review & Editing; **Wasiak R:** Supervision, Writing – Review & Editing; **Ramagopalan SV:** Conceptualization, Supervision, Writing – Review & Editing

Competing interests: LM and SR are employees of Bristol-Myers Squibb Company. AS, AS, SG and RW are employees of Evidera Inc.

How to cite this article: McDonald L, Schultze A, Simpson A *et al.* **A review of data sharing statements in observational studies published in the BMJ: A cross-sectional study [version 2; referees: 2 approved]** *F1000Research* 2017, 6:1708 (doi: [10.12688/f1000research.12673.2](https://doi.org/10.12688/f1000research.12673.2))

Copyright: © 2017 McDonald L *et al.* This is an open access article distributed under the terms of the [Creative Commons Attribution Licence](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. Data associated with the article are available under the terms of the [Creative Commons Zero "No rights reserved" data waiver](#) (CC0 1.0 Public domain dedication).

Grant information: The author(s) declared that no grants were involved in supporting this work.

First published: 19 Sep 2017, 6:1708 (doi: [10.12688/f1000research.12673.1](https://doi.org/10.12688/f1000research.12673.1))

REVISED Amendments from Version 1

In response to reviewer's comments we have expanded the Methods to include a precision calculation, edited the Results to more accurately reflect what was done and expanded the Conclusion to talk about future work that is needed to assess the impact of journal data sharing requirements.

See referee reports

Introduction

Over the recent years, a number of articles and movements have called for the sharing of clinical trial data¹⁻³. However the access to and sharing of real-world/observational data receives little and arguably insufficient attention. In this study we sought to assess the current state of data sharing in published observational studies in a general medical journal, namely the British Medical Journal (BMJ). The BMJ was chosen as the journal does not enforce a commitment of data sharing for observational studies⁴, but all research articles are required to contain a data sharing statement⁵.

Methods

All observational research articles published in the BMJ between 1st January 2015 and 31st August 2017 were investigated. These dates were chosen as it provides over 100 articles for analysis and recent data post-dating some of the articles regarding clinical trial data sharing². Assuming a proportion of articles not being shared of between 20–40%, 100 articles would enable a precision of 7.8 to 9.6%. Observational research articles were defined as cohort, case-control and cross-sectional studies, as well as case series. Meta-analyses, systematic reviews, randomised controlled trials and genetic/Mendelian randomisation studies were excluded. The data sharing statements of these studies were reviewed. If the statement written was “no additional data available”⁵ or that data was not publicly available, the study data was classed as not shared. Statements alluding to data being available from the corresponding author, or a referral to access policies of the data source the study data was classed as shared. Where statements alluded to code or technical appendix being available, but no reference to data specifically being available, the study data was classed as not shared.

Results

Two hundred and thirty seven observational studies were included. A review of the data sharing statements of these studies revealed that 149 (63%) studies did not share data.

Dataset 1. Raw data showing the studies identified from the BMJ and whether the data sharing statements indicated data was not/were available

<http://dx.doi.org/10.5256/f1000research.12673.d177871>

Conclusions

In our review of the data sharing statements of observational studies published in the BMJ, we identified that the majority of studies did not share data. Whilst there are likely many reasons for this, including patient confidentiality concerns and the access possibilities of the data source, the lack of data sharing in observational research is a potential cause for concern. The key limitation of our study was the scope of the data. We only reviewed the data sharing statements of one medical journal and therefore the generalisability of our results is unclear. The data sharing policy of the BMJ is relatively general (sharing is encouraged, but specific requirements are not communicated). Follow on work should assess how the robustness of a journal's data sharing policy could influence the rate of data sharing. Our analysis should be seen as exploratory rather than definitive. Further studies are needed, in greater depth, to confirm or refute our findings. If consistent findings are seen, the lack of sharing of observational research data is an area that warrants further attention.

Data availability

Dataset 1: Raw data showing the studies identified from the BMJ and whether the data sharing statements indicated data was not/were available. doi, [10.5256/f1000research.12673.d177871](https://doi.org/10.5256/f1000research.12673.d177871)⁶

Competing interests

LM and SR are employees of Bristol-Myers Squibb Company. AS, AS, SG and RW are employees of Evidera Inc.

Grant information

The author(s) declared that no grants were involved in supporting this work.

References

- Goldacre B, Lane S, Mahtani KR, *et al.*: **Pharmaceutical companies' policies on access to trial data, results, and methods: audit study.** *BMJ.* 2017; **358**: j3334.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Krumholz HM, Peterson ED: **Open Access to Clinical Trials Data.** *JAMA.* 2014; **312**(10): 1002–1003.
[PubMed Abstract](#) | [Publisher Full Text](#)
- AllTrials: **All Trials Registered. All Results Reported.** *AllTrials.*
[Reference Source](#)
- Research | The BMJ.** (Accessed: 2nd September 2017).
[Reference Source](#)
- Groves T: **Sharing the raw data from medical research [corrected].** *BMJ.* 2009; **338**: b1252.
[PubMed Abstract](#) | [Publisher Full Text](#)
- McDonald L, Schultze A, Simpson A, *et al.*: **Dataset 1 in: A review of data sharing statements in observational studies published in the BMJ: A cross-sectional study.** *F1000Research.* 2017.
[Data Source](#)

Open Peer Review

Current Referee Status:  

Version 2

Referee Report 07 November 2017

doi:[10.5256/f1000research.14218.r27649](https://doi.org/10.5256/f1000research.14218.r27649)



Robin E Champieux 

OHSU Library, Ontology Development Group, Oregon Health & Science University, Portland, OR, USA

The authors sufficiently addressed my original concerns. Overall, this study offers a valuable jumping off point for further investigations in this area.

Competing Interests: No competing interests were disclosed.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Version 1

Referee Report 02 November 2017

doi:[10.5256/f1000research.13724.r26439](https://doi.org/10.5256/f1000research.13724.r26439)



Robin E Champieux 

OHSU Library, Ontology Development Group, Oregon Health & Science University, Portland, OR, USA

This exploratory analysis contributes to the literature on biomedical data sharing practices, and demonstrates the importance of the further understanding of data sharing policies, practices, and barriers for domain and type specific data sets, such as those associated with observational studies.

While it does not impact the soundness of this short report, I would recommend that the authors expand on their reasoning for choosing BMJ for their analysis. Specifically, while BMJ requires data sharing statements, its data sharing policy is relatively general (sharing is encouraged, but specific requirements are not communicated). It would be helpful for the authors to discuss (potentially as a follow-up analyses), how the robustness of a journal's data sharing policy could influence the rate of data sharing.

Regarding the sampling methodology, the authors write, "All observational research articles published in the BMJ between 1st January 2015 and 31st August 2017 were investigated. These dates were chosen as it provides a reasonable sample size for analysis and recent data post-dating some of the articles regarding clinical trial data sharing." In my view, this statement is insufficient. The authors should provide more detailed and quantitative information about the strength and the nature of the sample.

Finally, the authors write "A review of the data sharing statements of these studies revealed that 149 (63%) studies had a statement implying that the data underlying the study could not be share". However the methodology does not indicate if the data "could" be shared, only if it was shared. I would urge the authors to revise this statement be be more precise.

Is the work clearly and accurately presented and does it cite the current literature?

Yes

Is the study design appropriate and is the work technically sound?

Yes

Are sufficient details of methods and analysis provided to allow replication by others?

Yes

If applicable, is the statistical analysis and its interpretation appropriate?

Partly

Are all the source data underlying the results available to ensure full reproducibility?

Yes

Are the conclusions drawn adequately supported by the results?

Yes

Competing Interests: No competing interests were disclosed.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 02 Nov 2017

Sreeram Ramagopalan,

We thank the reviewer for their comments. We focussed on the BMJ because of the journal being a key backer of the all trials initiative but we agree with the reviewers point and have updated the discussion.

We have updated the methods to include a precision calculation.

We corrected the results sentence.

Competing Interests: LM and SR are employees of Bristol-Myers Squibb. AS, AS, SG and RW are employees of Evidera Inc.

Referee Report 26 September 2017

doi:10.5256/f1000research.13724.r26085



Laurie Tomlinson 

Faculty of Epidemiology and Population Health, London School of Hygiene & Tropical Medicine, London, UK

This is an interesting study that draws attention to an important area - open access to the data underpinning research papers allowing assessment of reproducibility. It is of course limited by its focus on one journal as data-sharing policies vary between journals. In addition, requested statements by journals often become formulaic as the tendency is to copy what other papers have said previously. Some journals have tried new approaches in an effort to encourage data-sharing such as the NEJM 'Sprint Challenge'.

However, the major limitation is a lack of further data on why data was not shared. Some providers such as the UK primary care data repository CPRD do not allow data-sharing, even if the authors of articles are keen proponents of open data-access. More in depth work will be required to understand the true barriers and way ahead for data-sharing.

Is the work clearly and accurately presented and does it cite the current literature?

Yes

Is the study design appropriate and is the work technically sound?

Yes

Are sufficient details of methods and analysis provided to allow replication by others?

Yes

If applicable, is the statistical analysis and its interpretation appropriate?

Not applicable

Are all the source data underlying the results available to ensure full reproducibility?

Yes

Are the conclusions drawn adequately supported by the results?

Yes

Competing Interests: No competing interests were disclosed.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more

- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com

F1000Research