

ORIGINAL ARTICLE

Authors of trials from high-ranking anesthesiology journals were not willing to share raw data

Mirko Gabelica^a, Jakica Cavar^b, Livia Puljak^{c,*}

^aDepartment of Otorhinolaryngology, University Hospital Split, Spinciceva 1, 21000 Split, Croatia

^bDepartment of Neuroscience, University of Lethbridge, EP1249 Exploration Place 4401 University Drive, Lethbridge, AB, Canada

^cCenter for Evidence-Based Medicine and Health Care, Catholic University of Croatia, Ilica 242, 10000 Zagreb, Croatia

Accepted 29 January 2019; Published online 6 February 2019

Abstract

Objectives: To analyze data sharing practices among authors of randomized controlled trials (RCTs) published in seven high-ranking anesthesiology journals from 2014 to 2016.

Study Design and Setting: We analyzed data sharing statements in 619 included RCTs and contacted their corresponding authors, asking them to share de-identified raw data from trial.

Results: Of the 86 (14%) authors who responded to our query for data sharing, only 24 (4%) provided the requested data. Only one of those 24 had a data sharing statement in the published manuscript. Only 24 (4%) of manuscripts contained statements suggesting a willingness to share trial data; only one of those authors actually shared data. There was no difference in proportion of data sharing between studies with commercial and nonprofit funding. Among the 62 authors who refused to provide data, reasons were seldom provided. When reasons were provided, common themes included issues regarding data ownership and participant privacy. Only one of the seven analyzed journals encouraged authors toward data sharing.

Conclusion: Willingness to share data among anesthesiology RCTs is very low. To achieve widespread availability of de-identified trial data, journals should request their publication, as opposed to only encouraging authors to do so. © 2019 Elsevier Inc. All rights reserved.

Keywords: Randomized controlled trial; Raw data; Data sharing; Publication; Anesthesiology; Authors

1. Introduction

The terms “open data” and “data sharing” in biomedical research refer to the availability of raw data sets that are fully open. This enables transparency in research, reproducibility, and conduct of new studies using already generated data to expand on previous discoveries [1].

In 2006, Vickers observed that data sharing can be seen as a trivial logistical problem because of technological advancements, namely, the Internet [2]. Some journals have been early adopters of data sharing policies, requesting authors to either make their raw data sets publicly available in a public data sharing repository or to make them available on request [3,4].

Despite these encouraging developments, it has been repeatedly indicated that study authors are not willing to share

their data [5–7]. In 2018, Polanin reported that in 121 meta-analyses of individual patient data, authors managed to obtain 61% of eligible data sets [8]. Therefore, there is still need to promote transparency of research and data, as well as to conduct further studies that will help achieve that goal [9–22].

The aims of our study were to analyze data sharing policies of high-ranking journals in the field of anesthesiology from the years 2014 to 2016; quantify data sharing trends in randomized controlled trials (RCTs) of interventions published in those journals; test the willingness of RCT authors to share their raw data sets; analyze whether source of funding was associated with willingness to share data; and finally, to explore whether publicly available raw data sets are available in a way that would enable reanalysis.

2. Methods

2.1. Study design

We conducted a retrospective cohort study of published trials and journals’ instructions for authors, combined with survey of trial authors.

Funding: No extramural funding.

Conflict of interest: None.

* Corresponding author: Tel.: +385-0-1-370-66-00; fax: +385-0-1-370-66-01.

E-mail addresses: livia.puljak@gmail.com; livia.puljak@unicath.hr (L. Puljak).

What is new?

Key findings

- Among 619 randomized controlled trials published in seven high-impact anesthesiology journals, only 24 (4%) had data sharing statements in the manuscript.
- When asked to share de-identified raw data from their trial, authors of only 24 (4%) manuscript shared data. Among 24 trials with data sharing statements in the manuscript, only one author actually shared raw data.

What this adds to what was known?

- It has been suggested that data sharing plans should be provided in manuscripts, as a part of open data sharing movement. Our study indicates that providing a data sharing plan, or data sharing statements is not a guarantee that trial authors will actually share their data on request.

What is the implication and what should change now?

- The authors should be asked to deposit their raw data in publicly accessible repositories as a condition for publication.

2.2. Ethics

The Ethics Committee of the University of Split School of Medicine approved the study protocol. Scanned approval of the Ethics Committee was available to the study participants on request.

2.3. Inclusion criteria

We included RCTs of interventions published from January 1, 2014, to December 31, 2016, in seven journals from the Journal Citation Reports category Anesthesiology that belong to Q1, that is, the top 25% of the journal impact factor distribution. Based on the 2015 Journal Citation Reports impact factor and in alphabetical order, the journals were *Anaesthesia*, *Anesthesia and Analgesia*, *Anesthesiology*, *Pain*, *British Journal of Anaesthesia*, *European Journal of Anaesthesiology*, and *Regional Anesthesia and Pain Medicine*.

2.4. Search

We searched MEDLINE (via PubMed) via advanced search by using journal names, a filter for RCTs and a filter for publication dates 2014–2016. We exported titles and abstracts into a reference management software. Two authors independently screened titles and abstracts, and if

necessary, full texts to verify that those studies were indeed RCTs. We resolved any discrepancies in opinion via discussion. Full texts from all included RCTs were downloaded for further analysis.

2.5. Data extraction

We analyzed full texts of the included RCTs and extracted data into a Microsoft Excel worksheet (Microsoft Inc., Redmond, WA, USA). Worksheet was piloted with five studies to make sure that it is suitable to extract target data. We extracted the following data: name of the first author, year of publication, email of the corresponding author, source of funding (commercial/industry or nonprofit), presence of data sharing information (statement that raw data set is available in a certain repository, or available on request, or no such statement), and names of repositories where raw data sets were made available.

2.6. Author survey

All RCT authors that did not indicate that data were available in a publicly available repository (i.e., those that indicated that data were available on request, and those that did not have any statements regarding open data sharing) were contacted via email. First author (M.G.) sent personalized emails to each potential participant from personal email account. A de-identified copy of the email sent to the RCT authors is available in [Appendix 1](#). If the authors did not respond after initial email, they received only one reminder. Likewise, if the authors respond positively with willingness to share raw data sets but did not provide data within 2 weeks, they received an additional email reminder. Corresponding emails were obtained directly from the manuscripts of the examined studies. All emails and initial reminders were sent between January 26 and February 27 of 2018. Communication with several authors who requested additional information continued throughout March and April of 2018.

If corresponding authors suggested we should contact another team member to obtain data and provided their email addresses, we contacted those persons. If the authors indicated that additional regulatory or approval procedures were required for obtaining raw data sets, we did not engage in that processes. This was due to our previous experience of such requests taking years to receive responses, and even then, there is potential for data to be refused without explanation [23]. If message sent to corresponding authors was returned undelivered, we did not attempt to find their alternative email address. If the corresponding authors did not respond, we did not attempt to contact other co-authors.

2.7. Characteristics of raw data sets that would enable reanalysis

After accessing raw data sets, we checked whether data were available in a way that enabled reanalysis, that is,

published in a file that enabled data use, and whether relevant metadata were included.

2.8. Data sharing policies in journals

On June 17, 2018, we analyzed open–data sharing policies for authors of the included journals and noted whether such policy was in place, and if so, how it was formulated. Data sharing policies were compared against data sharing practices in analyzed RCTs.

2.9. Statistical analyses

We used a convenience sample of the most recently published RCTs within 3 years, considering that this would be a large enough sample to notice the current state of open data sharing. We presented descriptive data as frequencies and percentages. Differences in proportions were analyzed using a chi-square test. Analyses were conducted using MedCalc statistical software, v 15.2.1 (MedCalc Software bvba, Ostend, Belgium). Statistical significance was set at $P < 0.05$.

3. Results

We included 619 RCTs, published in the seven analyzed journals between January 1, 2014, and December 31, 2016. The highest number of RCTs was published in the *Anesthesia and Analgesia* ($N = 112$), followed by the *British Journal of Anaesthesia* ($N = 103$), *Pain* ($N = 97$), *Anesthesiology* ($N = 90$), *Anaesthesia* ($N = 86$), *European Journal of Anaesthesiology* ($N = 66$), and *Regional Anesthesia and Pain Medicine* ($N = 65$).

Most of the RCTs had nonprofit funding ($N = 439$; 71%); the remaining used commercial funding ($N = 80$; 13%) or declared no funding ($N = 51$; 8%). In 49 (8%) trials, statements regarding funding were not reported.

3.1. Data sharing in manuscripts describing RCTs

Among the 619 analyzed RCTs, not a single one provided raw data in the manuscript or provided a link to the repository where raw data are available. There were 24 studies with data sharing statements. Of those 24 studies, 20 were published by 20 different research groups, whereas four studies were published by two different research groups (each of the two groups published two of those studies). It is worth noting that one of those research groups with more papers had published a total of six studies from our cohort, but they provided data sharing statement only in two of those six studies, and on request, they did not provide data at all.

3.2. Requesting data from trial authors

We contacted the 619 corresponding authors of our cohort of RCTs. There were 31 emails that bounced back

as undelivered. For the remaining 588 manuscripts, we received responses from 86 (14%); 502 (82%) authors did not answer our query. From the 86 responses, further raw data were only obtained from 24 (3.9%), whereas 62 declined to share raw data.

Seventeen (2.8%) corresponding authors responded that they do not have raw data for sharing and gave us email address of another person to whom we should address our inquiry. Six of those 17 other individuals responded. Twelve (2.0%) automatic messages indicated that the recipients were temporarily away. These authors were contacted again on April 21, 2018; three responded.

Among 62 authors who refused to provide data, not all provided a reason. Table 1 summarizes responses of those authors. When reason was provided, the most common two responses were that they do not own data and participant privacy concerns. All those who responded that data were not theirs to share indicated that this was stipulated by the rules of their country, institution, or study sponsor. Three authors responded that their research was not an RCT, although it was described as such in manuscripts.

Twelve corresponding authors who mentioned privacy concerns as reasons for not sharing raw data had the following explanations: local research ethics approval prevents or might prevent disclosure of data in any way ($N = 3$); country-specific rules prevent sharing of raw data—this was the case for Norway ($N = 1$) and Denmark ($N = 3$); participants were not asked for consent to share their anonymized data with other researchers ($N = 2$);

Table 1. Responses received by authors of randomized controlled trials who did not share raw data on request

Responses	N (%)
Request for raw data refused immediately without reason provided	18 (3)
Corresponding author requested more data about our study, but after receiving it, they declined to provide raw data	13 (2.1)
Data are not mine, and therefore, I cannot share raw data	12 (1.9)
Participant privacy concern	12 (1.9)
Our research was not a randomized controlled trial	3 (0.5)
Trials are still ongoing, and the data are still in use, so therefore they cannot share the data	2 (0.35)
An author wanted co-authorship in exchange for raw data	1 (0.16)
Corresponding author wrote back that another person will provide raw data and gave us an email address; we contacted this other person but never received response	1 (0.16)
Total	62 (10)

university regulations prevent raw data sharing as these data are not public ($N = 2$); and the raw data are not anonymized ($N = 1$).

3.3. Response from authors who had data sharing statement

Of the 24 corresponding authors with manuscripts that provided statement regarding raw data availability, only one provided raw data after we requested it. Eighteen of those corresponding authors did not reply, three replied they cannot share their data, one stated they will get back to us in 4 months (but did not), and one sent affirmative email that they are sending data, but without data attached, and did not reply on request to kindly provide the attachment.

3.4. Raw data that were shared

In total, we received 24 raw data sets from 24 manuscripts, provided by 19 authors; however, two data sets were sent in pdf format; 22 (3.6%) usable data sets from a total of 619 manuscripts were retrieved. These data sets were sent in the SPSS and Microsoft Excel format. Median response time was 6 days.

Among the 24 manuscripts whose authors provided raw data sets, 4 (17%) had commercial funding, whereas the remaining 20 (83%) had nonprofit funding. There was no difference in proportion of data sharing between manuscripts with commercial funding and those with nonprofit funding ($\chi^2 = 0.030$, $df = 1$, $P = 0.816$).

3.5. Data sharing policies in analyzed journals

Analysis of instructions for authors of the seven analyzed journals showed that only one of the analyzed journals, *Anaesthesia*, had any mention of raw data, that is, data sharing policies. This journal indicated in their instructions the following (quote): “*Anaesthesia encourages authors to share the data and other artefacts supporting the results in the paper by archiving it in an appropriate public repository. Authors should include a data accessibility statement, including a link to the repository they have used, in order that this statement can be published alongside their paper*” [24]. Among the 24 manuscripts that had a data sharing statement and whose authors provided raw data on request, not a single one was published in the journal *Anaesthesia*.

4. Discussion

After attempting to obtain anonymized raw data sets from 619 RCTs that were published in anesthesiology journals from 2014 to 2016, we managed to receive data from only 3.6% of the trials, in which data would allow for reanalysis. The majority of the contacted corresponding authors

ignored our query for sharing raw data, and the most common reasons for refusing to share raw data were issues related to data ownership and participant privacy concerns.

A study that analyzed trials published in BMJ and PLoS One, journals that mandate data sharing, concluded that authors' data sharing behavior was “*not optimal*,” despite the journals' “*strong policy for data sharing*” [25]. This research groups also tried to reanalyze data obtained from those trials and concluded that they managed to reanalyze and replicate the original results for the majority of analyzed studies. They concluded that data sharing should be more widespread, as well as streamlined, so that other independent author teams could reanalyze and reuse the data collected in clinical trials [25].

Although International Committee of Medical Journal Editors (ICMJE) in 2016 suggested that responsible data sharing is an ethical obligation of trialists because “*participants have put themselves to risk*” [26], they now require that “*as of 1 July 2018, manuscripts submitted to ICMJE journals that report the results of clinical trials must contain a data sharing statement as described below*” [27]; however, our study shows that having a data sharing statement is not indicative in actually providing data. Data were received from one of 24 manuscripts that incorporated a data sharing statement in their manuscript, indicating that authors' behavior is not in line with their statements. Judging by the instructions for authors we analyzed, if we assume that these instructions did not change, the authors of those 24 manuscripts were not obliged by journal's instructions to provide data sharing statement.

The studies we analyzed were published in seven high-impact journals in the field of anesthesiology, and of those seven journals, only one journal, *Anaesthesia*, had raw data sharing policy, which “*encourages*” authors to share data and indicates that authors “*should*” share their data; however, among manuscripts that had data sharing statement in the manuscript or manuscripts for which authors provided data, not a single one was published in the *Anaesthesia* journal.

Rathi et al. surveyed 317 corresponding authors of RCTs published in 2010–2011 in six high-impact general medicine journals; 74% of authors indicated that sharing of de-identified data via repositories should be required, whereas 72% indicated that investigators should be required to share their de-identified data after receiving individual request [28]. Among 47% of authors who indicated that they had received requests for sharing their trial data, 77% indicated that they had indeed granted the data to those who requested them [28].

Analysis of data sharing predictors, using data from the same survey, showed no significant differences between trialists who did or did not support data sharing [29]. In 2018, Polanin and Terzian reported results of an RCT that was conducted among authors of studies included in recent meta-analyses via web-based survey [30]. They found that participants who were randomly assigned to receive a data

sharing agreement were more willing to share individual patient data of their primary study [30]; however, in all those studies, the trialists were asked a hypothetical question. The real-world data indicate that authors' words and their behavior do not necessarily match. A high number of authors who indicated that they provided their data on their request may simply be socially acceptable answers.

In the first study of Rathi et al., a survey asked authors about to comment on their concerns regarding data sharing. Major concerns were related to appropriate use of shared data. These authors were less concerned with the interests of investigators and funders, whereas protection of research participants was among their least important concerns [28]. The two most common reasons for refusing data sharing in our study were issues of data ownership, as corresponding authors indicated that they do not own the data and issues of concerns for participants' privacy. Although we highlighted to the contacted authors that we were asking for de-identified data, some authors still cited privacy concerns as reason for refusal of data.

Because it is assumed that recruitment of participants into clinical trials is an altruistic act that will contribute to advancement of medicine and medical knowledge, it is easy to see why many authors argue that failure to publish trials and lack of data sharing are considered to be a violation of trust of trial participants [31]. Spence et al. recently analyzed individual consent forms (ICFs) to see whether trial participants were informed about the investigators' plans related to contributing to medical knowledge, publishing trial results, and sharing de-identified trial data. Their study showed that ICFs seldom provide trialists' intentions regarding sharing of de-identified data or trial publication, and 91% of the ICFs did not indicate information regarding ownership of the trial data [31]. This finding is important in light of arguments suggesting that by refusing to share data, the trialists are protecting privacy of participants.

In our study, we did not have any experience with authors requesting us to cover expenses related to data sharing, as described by Naduet et al. When they asked corresponding authors of selected trials to share data, one research team, that authored two target manuscripts, requested a sum of £607 (equivalent to \$857 or €694) as a condition for sharing data, but the study authors refused to cover these expenses, as the other teams shared their data without any charge [25]; however, we did have one corresponding author in our cohort who mandated co-authorship in exchange for data.

Data sharing transparency could be achieved by involving relevant stakeholders that can influence behavior of authors, such as editors, organizations such as ICMJE, and research funders. Our study indicates that simply requiring a data sharing plan is not sufficient. Instituting mandatory data repositories and requesting higher accountability from corresponding authors are also potential practical solutions [21,22].

Limitations of our study include limited sample of journals, limited time frame analyzed, and potential nonresponder bias. It is possible that sending a data sharing agreement along with our email could have yielded additional responses [30]. We aimed to contact only corresponding authors and other individuals that were specifically suggested by corresponding authors; we did not attempt to contact all study authors. Furthermore, in our invitation to the corresponding authors, we emphasized that we are studying open data sharing in RCTs from the field of anesthesiology and that our team is interested in re-examining RCT raw data sets. It is possible that this type of request can be considered too general and blinded, and our results may not generalize to more targeted requests.

In conclusion, authors should be required to disclose their de-identified trial data publicly. Journal encouragement for data sharing is not enough to elicit willingness to share when approached. Whether the authors should be required to make their trial data available at the time of manuscript submission, or manuscript publication or sometime after publication, is something that can be debated. Left to their own devices, authors would likely rather refuse to share their data, unless they are required to.

CRedit authorship contribution statement

Mirko Gabelica: Data curation, Formal analysis, Investigation, Writing - original draft, Writing - review & editing.
Jakica Cavar: Data curation, Formal analysis, Investigation, Writing - original draft, Writing - review & editing.
Livia Puljak: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Supervision, Writing - original draft, Writing - review & editing.

Acknowledgments

The authors truly appreciate the effort of all the authors who took time to respond to our queries and particularly those who shared their raw data.

Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jclinepi.2019.01.012>.

References

- [1] Watson M. When will 'open science' become simply 'science'? *Genome Biol* 2015;16:101.
- [2] Vickers AJ. Whose data set is it anyway? Sharing raw data from randomized trials. *Trials* 2006;7:15.
- [3] Groves T. BMJ policy on data sharing. *BMJ* 2010;340:c564.
- [4] Godlee F, Groves T. The new BMJ policy on sharing data from drug and device trials. *BMJ* 2012;345:e7888.

- [5] Reidpath DD, Allotey PA. Data sharing in medical research: an empirical investigation. *Bioethics* 2001;15(2):125–34.
- [6] Savage CJ, Vickers AJ. Empirical study of data sharing by authors publishing in PLoS journals. *PLoS One* 2009;4(9):e7078.
- [7] Rowhani-Farid A, Barnett AG. Has open data arrived at the British Medical Journal (BMJ)? An observational study. *BMJ Open* 2016;6(10):e011784.
- [8] Polanin JR. Efforts to retrieve individual participant data sets for use in a meta-analysis result in moderate data sharing but many data sets remain missing. *J Clin Epidemiol* 2018;98:157–9.
- [9] Knottnerus JA, Tugwell P. Promoting transparency of research and data needs much more attention. *J Clin Epidemiol* 2016;70:1–3.
- [10] Hopkins C, Sydes M, Murray G, Woolfall K, Clarke M, Williamson P, et al. UK publicly funded clinical trials units supported a controlled access approach to share individual participant data but highlighted concerns. *J Clin Epidemiol* 2016;70:17–25.
- [11] West R. Data and statistical commands should be routinely disclosed in order to promote greater transparency and accountability in clinical and behavioral research. *J Clin Epidemiol* 2016;70:254–5.
- [12] Bouter LM. Open data are not enough to realize full transparency. *J Clin Epidemiol* 2016;70:256–7.
- [13] Ioannidis JP. Anticipating consequences of sharing raw data and code and of awarding badges for sharing. *J Clin Epidemiol* 2016;70:258–60.
- [14] Koers H. How do we make it easy and rewarding for researchers to share their data? A publisher's perspective. *J Clin Epidemiol* 2016;70:261–3.
- [15] Attwood AS, Munafo MR. Navigating an open road. *J Clin Epidemiol* 2016;70:264–6.
- [16] Goldacre B, Brown T. Fixing flaws in science must be professionalized. *J Clin Epidemiol* 2016;70:267–9.
- [17] Knottnerus JA. Research data as a global public good. *J Clin Epidemiol* 2016;70:270–1.
- [18] Shamseer L, Roberts J. Disclosure of data and statistical commands should accompany completely reported studies. *J Clin Epidemiol* 2016;70:272–4.
- [19] Kotz D, Tugwell P, Knottnerus JA. How to promote transparency and accountability in clinical and behavioral research. *J Clin Epidemiol* 2016;70:275.
- [20] West R. The end of scientific articles as we know them? *J Clin Epidemiol* 2016;70:276.
- [21] Manca A, Cugusi L, Dvir Z, Deriu F. Non-corresponding authors in the era of meta-analyses. *J Clin Epidemiol* 2018;98:159–61.
- [22] Teunis T, Nota SP, Schwab JH. Do corresponding authors take responsibility for their work? A covert survey. *Clin Orthop Relat Res* 2015;473(2):729–35.
- [23] Puljak L, Marin A, Vrdoljak D, Markotic F, Utrobicic A, Tugwell P. Celecoxib for osteoarthritis. *Cochrane Database Syst Rev* 2017;5:CD009865.
- [24] Lawson SN, Harper AA, Harper EI, Garson JA, Anderton BH. A monoclonal antibody against neurofilament protein specifically labels a subpopulation of rat sensory neurones. *J Comp Neurol* 1984;228(2):263–72.
- [25] Naudet F, Sakarovich C, Janiaud P, Cristea I, Fanelli D, Moher D, et al. Data sharing and reanalysis of randomized controlled trials in leading biomedical journals with a full data sharing policy: survey of studies published in the BMJ and PLOS Medicine. *BMJ* 2018;360:k400.
- [26] Taichman DB, Backus J, Baethge C, Bauchner H, de Leeuw PW, Drazen JM, et al. Sharing clinical trial data—A proposal from the international committee of medical journal editors. *N Engl J Med* 2016;374:384–6.
- [27] Taichman DB, Sahni P, Pinborg A, Peiperl L, Laine C, James A, et al. Data sharing statements for clinical trials. *BMJ* 2017;357:j2372.
- [28] Rathi V, Dzara K, Gross CP, Hrynaszkiewicz I, Joffe S, Krumholz HM, et al. Sharing of clinical trial data among trialists: a cross sectional survey. *BMJ* 2012;345:e7570.
- [29] Rathi VK, Strait KM, Gross CP, Hrynaszkiewicz I, Joffe S, Krumholz HM, et al. Predictors of clinical trial data sharing: exploratory analysis of a cross-sectional survey. *Trials* 2014;15:384.
- [30] Polanin JR, Terzian M. A data-sharing agreement helps to increase researchers' willingness to share primary data: results from a randomized controlled trial. *J Clin Epidemiol* 2019;106:60–9.
- [31] Spence O, Onwuchekwa Uba R, Shin S, Doshi P. Patient consent to publication and data sharing in industry and NIH-funded clinical trials. *Trials* 2018;19(1):269.