



## LETTERS

## DATA SHARING IN MEDICAL RESEARCH

## Why Cochrane should prioritise sharing data

Farhad Shokraneh *information specialist*<sup>1</sup>, Clive E Adams *professor*<sup>2</sup> *chair of mental health services research*<sup>2</sup>, Mike Clarke *professor*<sup>2</sup>, Laura Amato *psychiatrist*<sup>3</sup>, Hilda Bastian *former coordinator*<sup>4</sup>, Elaine Beller *associate professor*<sup>5</sup>, Jon Brassey *director*<sup>6</sup>, Rachelle Buchbinder *professor*<sup>7</sup>, Marina Davoli *director*<sup>3</sup>, Chris Del Mar *professor*<sup>5</sup>, Paul Glasziou *professor and director*<sup>5</sup>, Christian Gluud *head of department*<sup>8</sup>, Carl Heneghan *professor of evidence-based medicine and director*<sup>9</sup>, Tammy Hoffmann *professor of clinical epidemiology*<sup>5</sup>, John PA Ioannidis *professor of medicine and metrics co-director*<sup>10</sup>, Mahesh Jayaram *senior lecturer*<sup>11</sup>, Joey Kwong *research assistant professor*<sup>12</sup>, David Moher *director*<sup>13</sup>, Erika Ota *professor*<sup>14</sup>, Rebecca Syed Sheriff *honorary lecturer*<sup>15</sup>, Luke Vale *professor*<sup>16</sup>, Ben Goldacre *senior clinical research fellow*<sup>9</sup>

<sup>1</sup>Institute of Mental Health, University of Nottingham, Nottingham, UK; <sup>2</sup>Centre for Public Health, Queen's University Belfast, Northern Ireland; <sup>3</sup>Department of Epidemiology, Lazio Region-ASL Rome, Italy; <sup>4</sup>Cochrane Consumer Network; <sup>5</sup>Centre for Evidence Based Practice, Bond University, Gold Coast, Australia; <sup>6</sup>Trip Database; <sup>7</sup>Monash University and Cabrini Institute, Australia; <sup>8</sup>The Copenhagen Trial Unit, Centre for Clinical Intervention Research, Rigshospitalet; Copenhagen University Hospital, Copenhagen, Denmark; <sup>9</sup>Centre for Evidence Based Medicine, University of Oxford, UK; <sup>10</sup>Stanford University, USA; <sup>11</sup>Department of Psychiatry, University of Melbourne, Australia; <sup>12</sup>JC School of Public Health and Primary Care, Faculty of Medicine, The Chinese University of Hong Kong, Hong Kong; <sup>13</sup>Centre for Journalism, Ottawa Hospital Research Institute, Canada; <sup>14</sup>Global Health Nursing, Graduate School of Nursing Science, St. Luke's International University, Tokyo, Japan; <sup>15</sup>Child and Adolescent Psychiatric Unit, Department of Psychiatry, Faculty of Medicine, Chulalongkorn University, Bangkok, Thailand; <sup>16</sup>Cochrane Incontinence, Newcastle University, UK

Packer<sup>1</sup> says that the one who submits a research for public good should be ready to receive a request for data sharing for examination and re-analysis and that tax payers assume that a national agency is checking such data and analysis. Here we discuss Cochrane's practice on data sharing.

Open science, as endorsed by the G7,<sup>2</sup> includes sharing data, computer code, and materials. It is essential for reproducibility, collaboration, and innovation. We support the work of Cochrane, but are concerned that Cochrane is not sharing all its reviews' data. These data should be fully accessible for reuse by third parties.

Cochrane, a non-profit private company<sup>3</sup> and registered charity, produces and maintains systematic reviews in health and social care. Its work is undertaken by a global network of thousands of people,<sup>4</sup> and its support largely comes from public funding.<sup>5</sup> Most people producing Cochrane reviews are volunteers not specifically funded for this work,<sup>6,7</sup> and Cochrane encourages "crowdsourcing" of work.<sup>8-10</sup>

Cochrane editorial bases help volunteers obtain study reports and manually extract the wealth of data needed to generate

systematic reviews.<sup>11-13</sup> Cochrane teams use RevMan software<sup>14</sup> to produce files in standard format (XML), storing information on the studies, their methods, and results for publication in the Cochrane Library.

Benefits of sharing extracted data from trials and systematic reviews are well known, as are the costs of not sharing.<sup>13-17</sup> Sharing maximises transparency, reliability of data extraction, and syntheses. It improves access to data—saving time and money—and opens new avenues of inquiry.<sup>18</sup> Sharing is associated with increased citations,<sup>19</sup> more publications,<sup>20</sup> and reuse for new purposes.<sup>16</sup>

Structured data from Cochrane should be fully accessible for download, reuse, and review (box 1). Currently, they are not. Although Cochrane supports transparency initiatives such as AllTrials<sup>21</sup> and is explicit about this in its policy,<sup>22</sup> it has no similar clear principles on opening full access to the data in Cochrane reviews. Cochrane does provide access to results data from reviews but, crucially, these cannot be readily reused, and the available information is an incomplete set of the data generating these reviews, comes in a technically problematic

format, and can only be viewed by those with access to the full content of the Cochrane Library.<sup>23-25</sup>

#### Box 1: Structured data and associated metadata

##### Reference data

All data in the Cochrane Central Register of Controlled Trials (CENTRAL) excluding copyrighted abstracts (so creating OPEN CENTRAL)

All data in the Cochrane Register of Studies (CRS) excluding copyrighted abstracts (so creating OPEN CRS)

Links to "parent" study

Links to "parent" reviews

##### Study data

Links to "child" references

Links to "parent" reviews

##### Characteristics of studies

Methods, participants, interventions, outcomes

Qualitative data on risk of bias

Quantitative data on outcomes

Qualitative and quantitative derived data

Meta-analysis results, grading of quality of outcomes

Small amounts of Cochrane data have been released with bespoke arrangements for specific individuals. This sharing is welcome, but organisational culture, policy, and process regarding data release are lacking; there is no appeals process. For example, OpenTrials aggregates all accessible documents on all trials in an open database and makes it free for public reuse.<sup>26,27</sup> Thus far, OpenTrials has been unable to persuade Cochrane to share data for reuse. The Trip Database<sup>28</sup> is a searchable library of evidence that asked if it could re-present structured data from Cochrane and also encountered barriers to access.<sup>29</sup> Open sharing could foster collaborative ecosystems of digital innovation going beyond academic publications, with outputs that might include live, interactive presentations of summaries and results of trials produced by teams around the world, interactive decision support tools, and many more.

Cochrane's non-release of data is unlikely to reflect the preferences of funders, publishers, the thousands of Cochrane volunteers, participants in trials, or patients. When asked, 83% of the members of the Cochrane Individual Participant Data Meta-analysis Methods Group supported sharing systematic review data through a central repository (recognising that these data might require some form of moderated access).<sup>30</sup> Many funders now require that data arising from their grants are shared.<sup>31-34</sup> Cochrane volunteer authors give tacit consent for use of their work in reviews but may not be aware of the restrictions placed on access to the data they worked so hard to prepare.<sup>25</sup> This is morally and ethically questionable, potentially eroding public trust.<sup>16,35</sup>

This issue of open science is now pressing, after recent moves by Cochrane to create more information and become a hub for systematic review data. This has the potential to improve evidence and patient care, but although the Cochrane Linked Data Project aims to share reusable data in some form,<sup>36,37</sup> there is not yet any information on how or when this will happen.<sup>38,39</sup> Furthermore, Cochrane is working towards "living" systematic reviews, with updates from data in real time.<sup>40</sup> This is important work, but progress is slow. Opening up this work with shared data resources and in collaboration with the open source software community—where all can contribute—would accelerate progress and best reflect the culture of collaboration in science.

Open data offers a transformative, collaborative future for the systematic review community. Cochrane has enabled a vast workforce to painstakingly extract information for great benefit. It could act as a hub, harmonising data collected across groups and sharing these widely, reflecting the collective funding and volunteer workforce that produces them. This could include converting the morass of free text trial reports into machine readable curated data, in archived, citable, accessible, interoperable and reusable formats, as set out in the FAIR principles.<sup>41,42</sup> Cochrane could show leadership in supporting innovation and open science for clinical trials with full credit to all data extractors before<sup>43</sup> and after review publication<sup>44</sup> and, in this way, harness the greatest broadest impact. This reflects the exciting current move towards better use of data to produce digital tools of direct value to clinicians, rather than academic publications alone.

We have raised these issues with Cochrane and understand that the organisation is considering whether to start reviewing its approach to sharing data (D Tovey, personal communication, 2017). We hope that our setting out the benefits of open data is a helpful contribution to open that discussion.

We appreciate that Cochrane must focus on making itself sustainable and that open data sharing may be commercially sensitive.<sup>45</sup> But making Cochrane a champion for openness, transparency, and sharing can only be beneficial for the organisation's reputation—and finances. We encourage Cochrane leadership to create a policy that allows open data sharing and to make explicit any concerns they have on open data sharing so that these can be resolved.

#### Key messages

Cochrane could lead and set standards for open data sharing from systematic reviews

Availability of data from Cochrane reviews would give opportunities for collaboration, innovation, scientific replication, novel research, and clinical decision making

It would also reduce the considerable waste of the current duplication of effort in systematic reviewing

We thank the coordinating editors at Cochrane groups who supported internal discussion within Cochrane on sharing data. We are grateful for the time they spent studying and commenting on earlier versions of this manuscript and replying to our communications. We thank David Tovey, editor in chief of the Cochrane Library, for his thoughtful and helpful response to our written communication. We are grateful to the Cochrane coordinating editors: Gianni Virgili, Carlos Grillo Ardila, Juan-Pablo Casas, Jos Verbeek, Richard Wormald, and editor of Cochrane Methodology Review Group, Karen Robinson, for supporting the idea of Cochrane sharing the data.

Competing interests. FS is the information specialist of Cochrane Schizophrenia and has voluntarily extracted data for 12 Cochrane groups. CEA promotes Cochrane extensively to the public and policy makers; trains hundreds of reviewers per year, is coordinating editor of Cochrane Schizophrenia, is principal investigator on randomised trials testing the effects of disseminating Cochrane reviews in different forms and on the National Institute for Health Research infrastructure grant for Cochrane Schizophrenia. MC promotes Cochrane to the public, practitioners, and policy makers; provides training in the conduct of randomised trials and systematic reviews, is coordinating editor of the Cochrane Methodology Review Group, and seeks funding and conducts research into the methods using in systematic reviews and other evaluations of health and social care. BG has promoted Cochrane extensively to the public and policy makers; is principal investigator on OpenTrials.net, which has had a data sharing request rejected by Cochrane; has received research funding from the Laura and John Arnold Foundation, the Wellcome Trust, the NHS NIHR School of Primary Care, the Health Foundation, NHS England, the NIHR Oxford Biomedical Research Centre, and the World Health Organization; receives personal income from speaking and writing for lay audiences

on the misuse of science; and has a longstanding commitment to open science. LA promotes Cochrane to the public and policy makers; is coordinating editor of Cochrane Drugs and Alcohol Group; has received grant funding from WHO, the European Monitoring Centre for Drugs and Drug Addiction (EMCDDA), the Italian National Institute of Health, and AIFA (Italian Medicines Agency). HB has received access to Cochrane data for projects and services. JB is director and shareholder in the Trip Database, a limited company, and is actively involved in evidence synthesis. Trip has the potential to benefit from better access to the data Cochrane currently restricts. RB promotes Cochrane extensively to the public, clinicians, and policy makers; trains several reviewers a year, is joint coordinating editor of Cochrane Musculoskeletal, is principal investigator on grants developing two living Cochrane reviews and on National Health and Medical Research Council (NHMRC) editorial base funding for Cochrane Musculoskeletal, and has received research funding from the NHMRC, Cabrini Foundation, Medical Research Council, and Patient Centered Outcomes Research Institute. She is funded by an NHMRC Senior Principal Research Fellowship. CDM has received consultancy fees/honorariums from National Prescribing Service MedicineWise, the Royal Australian College of General Practitioners "red book" preventive guidelines committee; Therapeutic Guidelines (eTG); Remote Primary Health Care Manuals Editorial Committee for expert advice; editorial work (deputy editor of the *Medical Journal of Australia*; American College of Physicians' journal club; *The BMJ*); Consultation work for Bupa (UK) on shared decision making: Australian Medicine Handbook; royalties for three books (Wiley and BMJ Books) on evidence based medicine and clinical thinking; grants from NHMRC (Australia) two centres for research excellence; NIHR (UK); Human Tissue Authority (UK); from a private donor (for the Cochrane Collaboration Acute Respiratory Infections Group); Australian Commission on Safety and Quality in Health Care. MD is coordinating editor of Cochrane Drugs and Alcohol Group; has received grant funding from WHO, EMCDDA, the Italian National Institute of Health and AIFA, and disseminates Cochrane review results to the public and policy makers. PG is a member of editorial group of the Cochrane Acute Respiratory Infections group. CH has received grant funding from WHO, NIHR and the NIHR School of Primary Care. MJ is an editor at Cochrane Schizophrenia Group. DM is on Cochrane Oversight Committee. RSS is joint coordinating editor of the Cochrane Schizophrenia group. LV holds an NIHR systematic reviews grant for the Cochrane Incontinence. He holds grants from: EU2020, Wellcome, Economic and Social Research Council, MRC, Health Foundation, NIHR for research using systematic review methods. EB, CG, TH, JPAI, JK, and EO have declared no conflict of interests.

- 1 Packer M. Data sharing in medical research. *BMJ* 2018;360:k510. 10.1136/bmj.k510 29444885
- 2 Annex 4: G7 Expert Group on Open Science: executive summary. Italy, 2017 <http://www.g8.utoronto.ca/science/2017-G7-Annex4-Open-Science.pdf>.
- 3 Companies House. The Cochrane Collaboration Companies House. 2018 <http://bit.ly/2GmVlko>.
- 4 Cochrane Collaboration. About us. 2018. <http://bit.ly/2sFZRZn>.
- 5 Cochrane Collaboration. Our funders. 2018. <http://bit.ly/2lnh6TL>.
- 6 Tharyan P. Evidence of the people, by the people, and for the people. *Cochrane Database Syst Rev* 2010;2011:ED000013.21833936
- 7 Wilson M. Help us improve the health of people everywhere. Cochrane Collaboration. 2018. <http://join.cochrane.org/>.
- 8 Cochrane Collaboration. Cochrane Crowd. 2018. <http://crowd.cochrane.org/index.html>.
- 9 Cochrane Collaboration. Cochrane TaskExchange. 2018. <http://taskexchange.cochrane.org/>.
- 10 Wallace BC, Noel-Storr A, Marshall IJ, Cohen AM, Smalheiser NR, Thomas J. Identifying reports of randomized controlled trials (RCTs) via a hybrid machine learning and crowdsourcing approach. *J Am Med Assoc* 2017;24:1165-8. 10.1093/jamia/ocx053 28541493

- 11 Shokraneh F, Adams CE. Study-based registers of randomized controlled trials: Starting a systematic review with data extraction or meta-analysis. *Bioimpacts* 2017;7:209-17. 10.15171/bi.2017.25 29435428
- 12 Shokraneh F, Adams CE. Increasing value and reducing waste in data extraction for systematic reviews: tracking data in data extraction forms. *Syst Rev* 2017;6:153. 10.1186/s13643-017-0546-z 28778216
- 13 Wolfenden L, Grimshaw J, Williams CM, Yoong SL. Time to consider sharing data extracted from trials included in systematic reviews. *Syst Rev* 2016;5:185. 10.1186/s13643-016-0361-y 27809924
- 14 Manager R. <https://community.cochrane.org/help/tools-and-software/revman-5>.
- 15 Mayo-Wilson E, Li T, Fusco N, Dickersin KMUDS investigators. Practical guidance for using multiple data sources in systematic reviews and meta-analyses (with examples from the MUDS study). *Res Synth Methods* 2018;9:2-12. 10.1002/jrsm.1277 29057573
- 16 Nordic Trial Alliance Working Group on Transparency and Registration. *Transparency and registration in clinical research in the Nordic countries*. Nordic Trial Alliance, 2015.
- 17 Uhlir PF, Schröder P. Open Data for global science. *Data Sci J* 2007;6:OD36-53.
- 18 Agency for Healthcare Research and Quality. About the Systematic Review Data Repository. 2015. <https://srd.ahrq.gov/about>
- 19 Angraal S, Ross JS, Dhruva SS, Desai NR, Welsh JW, Krumholz HM. Merits of Data Sharing: The Digitalis Investigation Group Trial. *J Am Coll Cardiol* 2017;70:1825-7. 10.1016/j.jacc.2017.07.786 28958337
- 20 Piwowar HA, Day RS, Fridsma DB. Sharing detailed research data is associated with increased citation rate. *PLoS One* 2007;2:e308. 10.1371/journal.pone.0000308 17375194
- 21 Brown T. It's time for all trials registered and reported. *Cochrane Database Syst Rev* 2013;(4):ED000057.23728702
- 22 Cochrane Collaboration. Access to data from AllTrials. An official Cochrane policy: Cochrane Supports Access to Data from All Trials. 2018. <http://bit.ly/2FzS9qN>.
- 23 Cochrane Collaboration. Open Access. 2018. <http://bit.ly/2FBjUPC>.
- 24 Cochrane Library. Access options for Cochrane Library. 2018. <http://bit.ly/1WskJvg>.
- 25 Soares-Weiser K. Cochrane and conflict of interest: Addendum—28 June 2017. 2017. <http://bit.ly/2pbNC2J>.
- 26 Goldacre B, Gray J. OpenTrials: towards a collaborative open database of all available information on all clinical trials. *Trials* 2016;17:164. 10.1186/s13063-016-1290-8 27056367
- 27 Goldacre B, Turner EOpenTrials team. You can now search FDA approval documents easily at [fda.opentrials.net](http://fda.opentrials.net). *BMJ* 2017;356:j677. 10.1136/bmj.j677 28179258
- 28 Database T. Trip Medical Database 2018. <https://www.tripdatabase.com/>.
- 29 Brassey J. Cochrane's EMBASE screening project—did you participate? 2016. <https://www.jiscmail.ac.uk/cgi-bin/webadmin?A2=evidence-based-health:998c7b0.1605>
- 30 Tudur Smith C, Dwan K, Altman DG, Clarke M, Riley R, Williamson PR. Sharing individual participant data from clinical trials: an opinion survey regarding the establishment of a central repository. *PLoS One* 2014;9:e97886. 10.1371/journal.pone.0097886 24874700
- 31 Bill and Melinda Gates Foundation. Open access policy. 2018. <http://gates.ly/2Ab4hHw>.
- 32 Medical Research Council. Data sharing. 2018. <http://bit.ly/2x1Seq>.
- 33 National Health and Medical Research Council. NHMRC statement on data sharing. 2017. <http://bit.ly/2FLjMfL>.
- 34 National Institute for Health Research. Data sharing. 2018. <http://bit.ly/2FP973G>.
- 35 Coyne JC. Why I am formally requesting the data set from a Cochrane review 2017 <http://bit.ly/2FD5v1a>.
- 36 Li T, Vedula SS, Hadar N, Parkin C, Lau J, Dickersin K. Innovations in data collection, management, and archiving for systematic reviews. *Ann Intern Med* 2015;162:287-94. 10.7326/M14-1603 25686168
- 37 Slaughter L, Bernsten CF, Brandt L, et al. Enabling living systematic reviews and clinical guidelines through semantic technologies. *D-Lib* 2015. 10.1045/january2015-slaughter 10.1045/january2015-slaughter .
- 38 Cochrane Collaboration. Cochrane strategy to 2020. 2017. <https://community.cochrane.org/organizational-info/resources/strategy-2020>
- 39 Cochrane Collaboration. Trusted evidence. Informed decisions. Better health 2018. <http://www.cochrane.org/>.
- 40 Elliott JH, Turner T, Clavisi O, et al. Living systematic reviews: an emerging opportunity to narrow the evidence-practice gap. *PLoS Med* 2014;11:e1001603. 10.1371/journal.pmed.1001603 24558353
- 41 Data Citation Synthesis Group. Joint Declaration of Data Citation Principles. 2014. <https://www.force11.org/datacitationprinciples>.
- 42 Wilkinson MD, Dumontier M, Aalbersberg IJ, et al. The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* 2016;3:160018. 10.1038/sdata.2016.18 26978244
- 43 Chalmers I, Glasziou P. Should there be greater use of preprint servers for publishing reports of biomedical science? *F1000Res* 2016;5:272. 10.12688/f1000research.8229.1 26998238
- 44 Bierer BE, Crossas M, Pierce HH. Data authorship as an incentive to data sharing. *N Engl J Med* 2017;376:1684-7. 10.1056/NEJMs1616595 28402238
- 45 Cochrane Collaboration. Senior management team. Cochrane Strategy to 2020—in 2017: Definition of success by 2020, an assessment of progress, and a framework for work remaining. 2017. <https://community.cochrane.org/news/strategy-2020-2017-targets>

Published by the BMJ Publishing Group Limited. For permission to use (where not already granted under a licence) please go to <http://group.bmj.com/group/rights-licensing/permissions>