DATA SHARING IN MEDICAL RESEARCH

Why Cochrane should prioritise sharing data

Farhad Shokraneh information specialist1, Clive E Adams professor2 chair of mental health services research7, Mike Clarke professor8, Laura Amato psychiatrist9, Hilda Bastian former coordinator6, Elaine Beller associate professor5, Jon Brasseys director4, Rachelle Buchbinder professor7, Marina Davoli director3, Chris Del Mar professor5, Paul Glasziou professor and director2, Christian Gluud head of department8, Carl Heneghan professor of evidence-based medicine and director10, Tammy Hoffmann professor of clinical epidemiology4, John PA Ioannidis professor of medicine and metrics co-director7, Mahesh Jayaram senior lecturer11, Joey Kwong research assistant professor12, David Moher director13, Erika Ota professor14, Rebecca Syed Sheriff honorary lecturer15, Luke Vale professor16, Ben Goldacre senior clinical research fellow9

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

Packer1 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 G7,7 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 company1 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, and its support largely comes from public funding. Most people producing Cochrane reviews are volunteers not specifically funded for this work,6,7 Cochrane encourages “crowdsourcing” of work.6,7

Cochrane editorial bases help volunteers obtain study reports and manually extract the wealth of data needed to generate systematic reviews.11-13 Cochrane teams use RevMan software14 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.11-13,17 Sharing maximises transparency, reliability of data extraction, and syntheses. It improves access to data—saving time and money—and opens new avenues of inquiry.18 Sharing is associated with increased citations,19 more publications,20 and reuse for new purposes.16

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 AllTrials21 and is explicit about this in its policy,22 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
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. Cochrane could show leadership in supporting innovation and open science for clinical trials with full credit to all data extractors before and after review publication 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. 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.

**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. Thus far, OpenTrials has been unable to persuade Cochrane to share data for reuse. The Trip Database is a searchable library of evidence that asked if it could re-present structured data from Cochrane and also encountered barriers to access. 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). Many funders now require that data arising from their grants are shared. 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. This is morally and ethically questionable, potentially eroding public trust.

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, there is not yet any information on how or when this will happen. Furthermore, Cochrane is working towards “living” systematic reviews, with updates from data in real time. 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.
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’ preventative guidelines committee; Therapeutic Guidelines (eTG); Remote Primary Health Care Manuals for the Royal Australian College of General Practitioners; and the National Clinical Practice Guidelines Program for the Australian College of General Practitioners. She is also a registered participant in the Cochrane Incontinence Grant Program. PG is a member of editorial group of the Cochrane Acute Respiratory Infections Group. He holds grants from: NIHR School of Primary Care Research; WHO; the Cochrane Collaboration; MHRA; AstraZeneca; Bupa (UK) on shared decision making: Australian Medicine Management, and archiving for systematic reviews. CH 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, NHMRC and the NHIR School of Primary Care. MJ is an editor of Cochrane Schizophrenia Group. DM is on Cochrane Oversight Committee. RSS is joint coordinating editor of the Cochrane Schizophrenia group. LV holds an NHIR systematic reviews grant for the Cochrane Incontinence. He holds grants from: EU2020, Wellcome, Economic and Social Research Council, MRC, Health Foundation, NHMRC for research using systematic review methods. EB, CG, TH, JPAl, 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 2018;360:k503. 10.1136/bmj.k503 24944885
10.1093/jamia/ocx053 28541493