Why data sharing from clinical trials really matters

PhUSE European Computational Science Symposium
19 June, London

Dr Trish Groves
tgroves@bmj.com twitter @trished
Director of academic outreach BMJ, Editor in chief BMJ Open,
Honorary deputy editor The BMJ

Competing interests

I'm editor in chief of BMJ Open and Director of academic outreach at BMJ Publishing Group, owned by the British Medical Association (BMA)

Part of the revenue for BMJ comes from drug & device manufacturers through advertising, reprint sales, & sponsorship. The BMJ (British Medical Journal) and BMJ Open are open access journals that charge article publishing fees for research. I'm editorial lead for the BMJ Research to Publication eLearning programme (by subscription).

My annual bonus scheme is based partly on the overall financial performance of both BMJ and BMJ Research to Publication.

What I'll cover

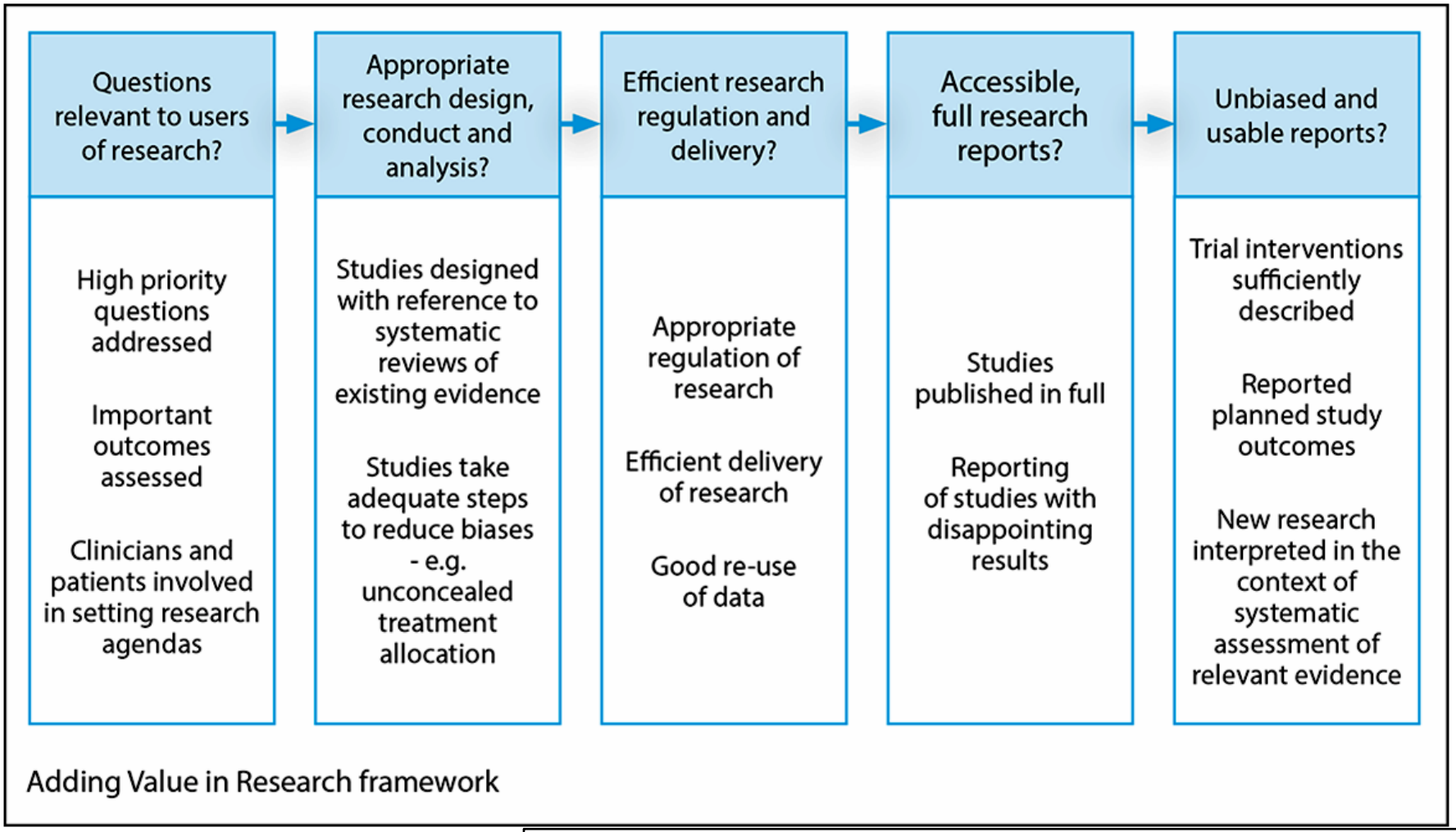
The problem of non-reported clinical trials

Why this matters

What can be done to fix it and to make trials useful and usable

- where have we got to?
- what challenges are coming?

85% research wasted, costing >$100bn/yr



Adding Value in Research framework

thebmj

Chalmers I, Glasziou P. Avoidable waste in the production and reporting of research evidence. Lancet 2014; 374: 86-9.

Let's increase research value and reduce harm

- review identified reporting bias in 40 indications comprising around 50 pharmacological, surgical (e.g. vacuum-assisted closure therapy), diagnostic (e.g. ultrasound), and preventive (e.g. cancer vaccines) interventions
- many cases involved withholding of study data by manufacturers and regulatory agencies or active attempts by manufacturers to suppress publication
- reporting bias can overestimate efficacy and underestimate safety risks of interventions

thebmj

McGauran, N, Wieseler, B, Kreis, J, Schüler, YB, Kölsch, H, and Kaiser, T. Reporting bias in medical research—a narrative review. Trials. 2010; **11**: 37

Underestimation of benefit can be wasteful too

- data from published systematic reviews and meta-analyses and unpublished FDA reviews were used in revised meta-analyses for 9 drugs approved by U.S. FDA in 2001-2

- summary estimate of the single harm outcome found greater harm after inclusion of unpublished FDA trial data

- addition of unpublished FDA trial data caused 46% (19/41) of summary estimates to show lower drug efficacy, 7% (3/41) to show identical efficacy, 46% (19/41) to show greater efficacy

thebmj

Hart B, Lundh A, Bero L. Effect of reporting bias on meta-analyses of drug trials: reanalysis of meta-analyses *BMJ* 2012; 344:d7202

Transparent reporting of clinical trials

Trial registry

But providing access to full protocols is voluntary and inconsistent

Zarin DA, Tse T. Medicine. Moving toward transparency of clinical trials. Science 2008 Mar 7;319(5868):1340-2.

‘Negative’ trials less likely to be published

From meta-analysis by Song et al: pooled odds ratios of publication of studies with positive results, compared to those without positive results (publication bias):

- 2.78 (95% CI: 2.10 to 3.69) in cohorts of studies from inception
- 5.00 (95% CI: 2.01 to 12.45) in trials submitted to regulatory authority
- 1.70 (95% CI: 1.44 to 2.02) in abstract cohorts,
- 1.06 (95% CI: 0.80 to 1.39) in cohorts of manuscripts.

It seems that publication bias occurs mainly before the presentation of findings at conferences and before the submission of manuscript to journals (the “file drawer problem”)

thebmj

Song F, Parekh-Bhurke S, Hooper L, Yoke YK, Ryder JJ, Sutton AJ, Hing CB, Harvey I. BMC Medical Research Methodology 2009, **9**:79 doi:10.1186/1471-2288-9-79

AllTrials campaign

THE PROBLEM

Thousands of clinical trials have not reported results.

Information on what was done and what was found in these trials could be lost forever to doctors and researchers, leading to bad treatment decisions, missed opportunities for good medicine, and trials needlessly repeated.

WHAT ARE PEOPLE DOING?

- nearly 90,000 individuals have signed the petition for all clinical trials to be registered and the results reported
- more than 700 organisations around the world have joined the campaign
- they write about the campaign, donate to it, and get involved in other ways via www.alltrials.net
- campaign has directly influenced policies on clinical trial transparency in UK, EU, Canada

Published trials are poorly reported

30 years ago some statisticians looked at 45 reports of comparative trials published in The BMJ, the Lancet, or New England Journal of Medicine and found:

- common failure to specify in advance the intended size of a trial or statistical stopping rules for interim analyses
- summaries or abstracts of trials tended to emphasize the more statistically significant end points
- overall, reporting of clinical trials appeared biased towards exaggerating treatment differences.

thebmj

Pocock SJ, Hughes MD, Lee RJ. Statistical problems in the reporting of clinical trials. A survey of three medical journals. N Engl J Med 1987;317:426–32.

COMPARE

TRACKING SWITCHED OUTCOMES IN CLINICAL TRIALS

METHODS | RESULTS | TEAM | BLOG | FAQ

67	9	354	357
TRIALS CHECKED	TRIALS WERE PERFECT	OUTCOMES NOT REPORTED	NEW OUTCOMES SILENTLY ADDED

On average, each trial reported just 58.2% of its specified outcomes. And on average, each trial silently added 5.3 new outcomes.

58	18	8	32
LETTERS SENT	LETTERS PUBLISHED	LETTERS UNPUBLISHED AFTER 4 WEEKS	LETTERS REJECTED BY EDITOR

From Oct 2015 - Jan 2016 the COMPare Trials Project team systematically checked every trial published in NEJM, JAMA, The Lancet, Annals of Internal Medicine, and The BMJ to see if they misreported their findings

http://compare-trials.org

Replication: desirable, but not always possible

- scientific evidence strengthened when important findings are replicated by multiple investigators using independent data, analytical methods, laboratories, and instruments
- replication is standard in basic sciences
- time and expense required for epidemiological studies means many are often not fully replicable, so policy decisions must be made with available evidence

Should at least ensure **reproducibility**

thebmj

Peng RD, Dominici F, Zeger SL. Reproducible Epidemiologic Research. Am J Epidemiol 2006;163: 783-9 doi:10.1093/aje/kwj093

Reproducibility: should always be possible

- reproducibility is an attainable minimum standard
- independent investigators subject the original data to their own analyses and interpretations
- reproducibility requires datasets and software to be available for:
  - verifying published findings
  - conducting alternative analyses of the same data
  - eliminating uninformed criticisms that do not stand up
  - expediting interchange of ideas among investigators

thebmj

Peng RD, Dominici F, Zeger SL. Reproducible Epidemiologic Research. Am J Epidemiol 2006;163: 783-9 doi:10.1093/aje/kwj093

What does reproducibility actually mean?

- methods reproducibility
- results reproducibility (via data, metadata, code)
- robustness and generalisability
- inferential reproducibility
  - hampered by selective reporting, data mining/dredging/torturing, p-hacking, HARKing (hypothesising after results known)

Goodman SN, Fanelli D, Ioannidis JPA. What does research reproducibility mean? Sci Trans Med 2016: 341PS12
http://stm.sciencemag.org/content/8/341/341ps12.full

thebmj

Sharing de-identified individual participant data (IPD) from clinical trials

To maximise fidelity of evidence base ...and:

- allow testing of secondary hypotheses
- aid design of future trials
- simplify data acquisition for IPD meta-analysis
- aid developing/evaluating novel statistical methods
- ensure analyses can be reproduced and checked
- provide incentive to ensure accuracy of dataset
- reduce deliberate misconduct
- whose data are they anyway?

thebmj

Vickers AJ. Whose data set is it anyway? Sharing raw data from randomized trials. *Trials* 2006, 7:15doi:10.1186/1745-6215-7-15

Data sharing: regulator and industry initiatives

- EMA is providing public access to clinical study reports

- July 2014 Principles for Responsible Clinical Trial Data Sharing from European Federation of Pharmaceutical Industries and Associations (EFPIA) and Pharmaceutical Research and Manufacturers of America (PhRMA)

- controlled access to online databases of de-identified individual patient data from trials through initiatives such as:
  - https://www.clinicalstudydatarequest.com
  - Yale University Open Data Access (YODA) Project http://yoda.yale.edu/
  - https://www.projectdatasphere.org for cancer trial data

Data sharing in clinical trials: U.S. Institute of Medicine (IOM) report 2015

"Stakeholders in clinical trials should foster a **culture in which data sharing is the expected norm**, and should commit to responsible strategies aimed at maximizing the benefits, minimizing the risks, and overcoming the challenges of sharing clinical trial data for all parties."

thebmj

IOM (Institute of Medicine). 2015. Sharing clinical trial data: Maximizing benefits, minimizing risk. Washington, DC: The National Academies Press. www/iom.edu/datasharing

Data sharing took off...or was it data dumping?

Patient-level de-identified data available for 3255 trials (44.3% Phase 3) via three platforms between 2013 and 31 Dec 2015: CDSR, YODA, SOAR (Duke)

234 proposals, 154 approved, 113 with data sharing agreements

- requests from 17 countries with 61/113 (54%) from U.S.
- **requests covered 505/3255 (15.5%) of all available trials**

5/113 analyses to validate study endpoint, others to analyse subgroups

- 50 on treatment effect
- 31 on disease state
- 32 for other purposes eg IPD meta-analysis

Navar AM, Pencina MJ, Rymer JA, Louzao DM, Peterson ED. Research letter: Use of open access platforms for clinical trials data. JAMA 2016; 315:1283-4

The BMJ's policy: data sharing on request

Mandatory since 2013 for any paper reporting main endpoints of an RCT of one or more drugs or medical devices in current use. Extending 2015 to all trials submitted to The BMJ

Oct 2016: survey in BMJ Open covering 21 clinical trials bound by The BMJ policy reported that only 5/21 had made data sets available.

One data set was freely available on Dryad, leaving 20 RCTs whose authors were emailed to request data:

- 13 did not respond
- 4 made the data available
- 3 declined, citing caveats about the survey

thebmj

Has open data arrived at the *British Medical Journal (BMJ)*? An observational study. Rowhani-Farid A, Barnett AG.
BMJ Open 2016;6:10 e011784 doi:10.1136/bmjopen-2016-011784

ICMJE: principles of data sharing 2016 draft

- data sharing plan to made public at time of clinical trial registration
- de-identified individual patient data to be shared (as planned)
  - no later than 6 months after publication of the article
  - data and metadata required to reproduce the results in the article, including tables, figures, appendices and supplementary materials
  - 1 year wash-in period after publication of full ICMJE policy
  - data deposition in a repository will not count as prior publication
  - data users must attest to use in accordance with data sharing plan, credit source using unique dataset ID commit to making results of their analyses public, report methods, consider collaboration with data generators
- RECs/IRBs should ensure patient informed consent covers all this
- journals may investigate breaches, express concern, retract

International Committee of Medical Journal Editors
http://www.icmje.org/news-and-editorials/M15-2928-PAP.pdf

ICMJE 2017 policy on sharing clinical trial data

As of 1 July 2018 manuscripts submitted to any ICMJE journal* that report the results of clinical trials must contain a data sharing statement.

Clinical trials that begin enrolling participants on or after 1 January 2019 must include a data sharing plan in the trial's registration. If the data sharing plan changes after registration this should be reflected in the statement submitted and published with the manuscript, and updated in the registry record.

www.icmje.org/recommendations/browse/publishing-and-editorial-issues/clinical-trial-registration.html

*International Committee of Medical Journal Editors (ICMJE) current journals: *Annals of Internal Medicine*, *The BMJ*, *Deutsches Ärzteblatt*, *Ethiopian Journal of Health Sciences*, *JAMA*, *Journal of Korean Medical Science*, *New England Journal of Medicine*, *New Zealand Medical Journal*, *PLOS Medicine*, *Lancet*, *Revista Médica de Chile*, *Ugeskrift for Laeger* (members in 2016-17)

thebmj

ICMJE 2017 policy on data sharing statements

Statements must indicate:

- whether individual deidentified participant data (incl. data dictionaries) will be shared
- what data will be shared
- whether related documents will be available (eg trial protocol, statistical analysis plan)
- when the data will become available and for how long
- access criteria (with whom, for what types of analyses, by what mechanism)

Editors may take data sharing statements into account when making decisions on peer review and acceptance of papers

"A couple of weak sticks and no carrots...since patients are taking the risks here, maybe they should start boycotting all clinical trials whose investigators refuse to share their data."

Adam Marcus,
Ivan Oransky
statnews.com
6 June 2017

thebmj

●●○○○ vodafone UK 17:25 99%
statnews.com

STAT Sections Topics Multimedia Popular STAT Plus Newsletters Log In / Subscribe

THE WATCHDOGS

New science data-sharing rules are two scoops of disappointment

By ADAM MARCUS @armarcus *and* IVAN ORANSKY @ivanoransky / JUNE 6, 2017

APSTOCK

After more than a year of deliberation, editors of some of the world's leading biomedical journals have come up with a declaration on data sharing destined to usher in a glorious world of transparency and rigor in science.

No, sorry, scratch that. What they've come up with is a policy with a couple of weak sticks and no carrots.

VS.

PQRST Index for rewarding research

"To change the tide, the criteria by which scientists and their teams are rewarded for their efforts by agencies that fund them and institutions that host them should be revisited, aligning criteria with the desired outcomes: research that is productive, high-quality, reproducible, shareable, and translatable [PQRST] …funding agencies, universities, research institutions, academies, professional societies, and prestigious award organizations may also have PQRST indices based on the research work they sponsor or perform and the scientists behind this work."

Journal integration with data repositories

Dryad http://datadryad.org/ eg authors can start the data deposition process while submitting their papers to BMJ Open or to PLOS journals. Dryad provides a DOI to aid citation and provide a permanent link to the data, along with a CC0 Creative Commons licence

Figshare https://figshare.com/ allows users to upload any file format so that any research output, from posters and presentations to datasets and code, can be disseminated

Patient perspective: more research needed

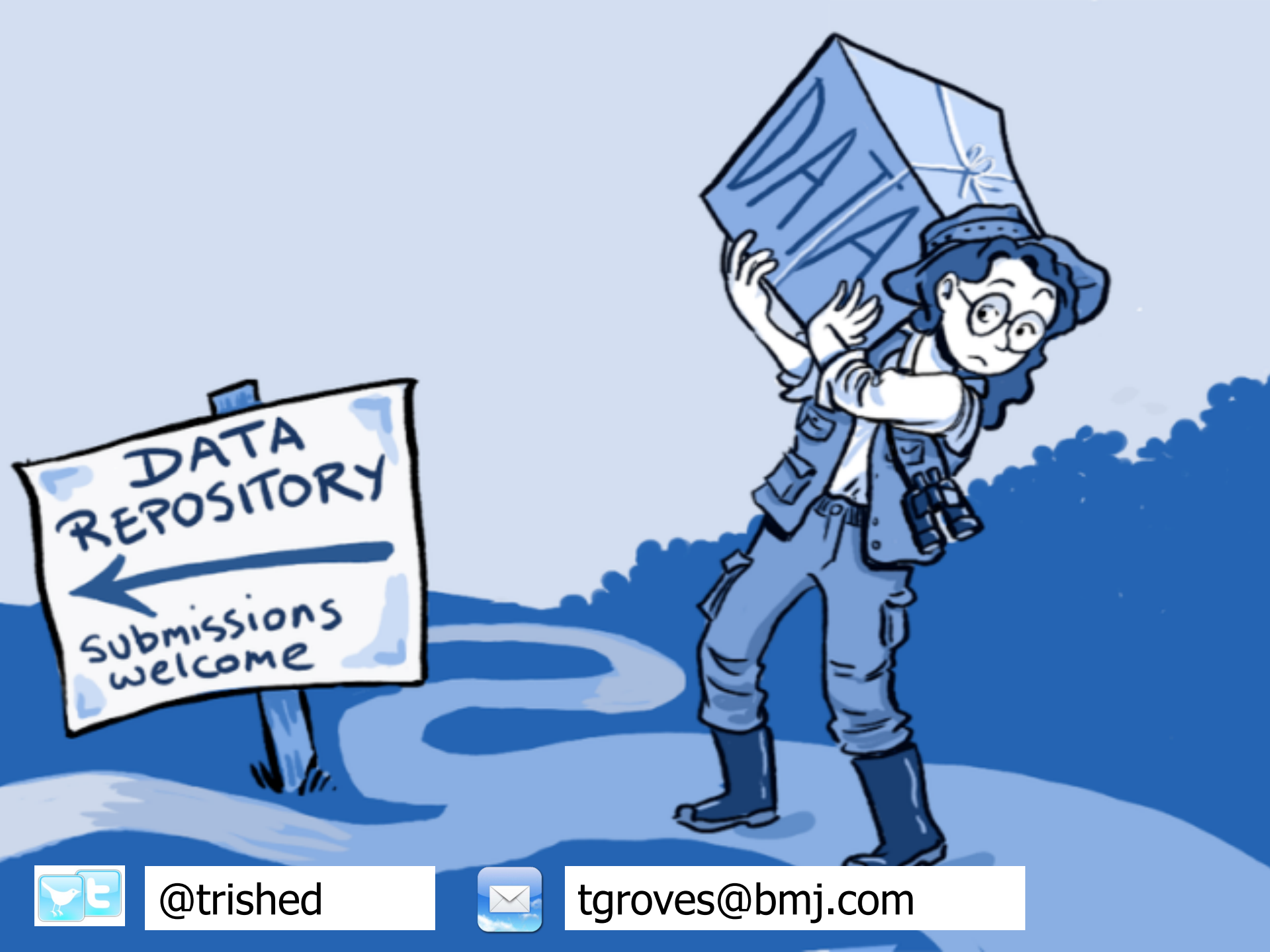
"Patients expect that health care professionals and researchers use patient data in the best possible way. That there is a fight over what the best way is perplexing and disappointing."

"If you have a life-threatening disease and need help, you do not care much about privacy. The question is also: How many people will die if we don't share data?

Patients cited in NEJM debate April 2017

thebmj

Haug CJ. Whose Data Are They Anyway? Can a Patient Perspective Advance the Data-Sharing Debate? NEJM 2017 DOI: 10.1056/NEJMp1704485



@trished

tgroves@bmj.com