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*

Florian Naudet,[1] Charlotte Sakarovitch,[2] Perrine Janiaud,[1] Ioana Cristea,[1,3] Daniele Fanelli,[1,4] David Moher,[1,5] John P A Ioannidis[1,6]

[1]Meta-Research Innovation Center at Stanford (METRICS), Stanford University, Stanford, California, USA

[2]Quantitative Sciences Unit, Division of Biomedical Informatics Research, Department of Medicine, Stanford University, Stanford, CA, USA

[3]Department of Clinical Psychology and Psychotherapy, Babes-Bolyai University, Romania

[4]Department of Methodology, London School of Economics and Political Science, UK

[5]Centre for Journalology, Clinical Epidemiology Program, Ottawa Hospital Research Institute, Ottawa, ON, Canada

[6]Departments of Medicine, of Health Research and Policy, of Biomedical Data Science, and of Statistics, Stanford University, Stanford, California, USA

Correspondence to:
J P A Ioannidis
jioannid@stanford.edu
Additional material is published online only. To view please visit the journal online.

ABSTRACT

OBJECTIVES
To explore the effectiveness of data sharing by randomized controlled trials (RCTs) in journals with a full data sharing policy and to describe potential difficulties encountered in the process of performing reanalyses of the primary outcomes.

DESIGN
Survey of published RCTs.

SETTING
PubMed/Medline.

ELIGIBILITY CRITERIA
RCTs that had been submitted and published by *The BMJ* and *PLOS Medicine* subsequent to the adoption of data sharing policies by these journals.

MAIN OUTCOME MEASURE
The primary outcome was data availability, defined as the eventual receipt of complete data with clear labelling. Primary outcomes were reanalyzed to assess to what extent studies were reproduced. Difficulties encountered were described.

RESULTS
37 RCTs (21 from *The BMJ* and 16 from *PLOS Medicine*) published between 2013 and 2016 met the eligibility criteria. 17/37 (46%, 95% confidence interval 30% to 62%) satisfied the definition of data availability and 14 of the 17 (82%, 59% to 94%) were fully reproduced on all their primary outcomes. Of the remaining RCTs, errors were identified in two but reached similar conclusions and one paper did not provide enough information in the Methods section to reproduce the analyses. Difficulties identified included problems in contacting corresponding authors and lack of resources on their behalf in preparing the datasets. In addition, there was a range of different data sharing practices across study groups.

CONCLUSIONS
Data availability was not optimal in two journals with a strong policy for data sharing. When investigators shared data, most reanalyses largely reproduced the original results. Data sharing practices need to become more widespread and streamlined to allow meaningful reanalyses and reuse of data.

TRIAL REGISTRATION
Open Science Framework osf.io/c4zke.

WHAT IS ALREADY KNOWN ON THIS TOPIC

The International Committee of Medical Journal Editors requires that a data sharing plan be included in each paper (and prespecified in study registration)

Two leading general medical journals, *The BMJ* and *PLOS Medicine*, already have a stronger policy, expressly requiring data sharing as a condition for publication of randomized controlled trials (RCTs)

Only a small number of reanalyses of RCTs has been published to date; of these, a minority was conducted by entirely independent authors

WHAT THIS STUDY ADDS

Data availability was not optimal in two journals with a strong policy for data sharing, but the 46% data sharing rate observed was higher than elsewhere in the biomedical literature

When reanalyses are possible, these mostly yield similar results to the original analysis; however, these reanalyses used data at a mature analytical stage

Problems in contacting corresponding authors, lack of resources in preparing the datasets, and important heterogeneity in data sharing practices are barriers to overcome

Introduction

Patients, medical practitioners, and health policy analysts are more confident when the results and conclusions of scientific studies can be verified. For a long time, however, verifying the results of clinical trials was not possible, because of the unavailability of the data on which the conclusions were based. Data sharing practices are expected to overcome this problem and to allow for optimal use of data collected in trials: the value of medical research that can inform clinical practice increases with greater transparency and the opportunity for external researchers to reanalyze, synthesize, or build on previous data.

In 2016 the International Committee of Medical Journal Editors (ICMJE) published an editorial[1] stating that "it is an ethical obligation to responsibly share data generated by interventional clinical trials because participants have put themselves at risk." The ICMJE proposed to require that deidentified individual patient data (IPD) are made publicly available no later than six months after publication of the trial results. This proposal triggered debate.[2-7] In June 2017, the ICMJE stepped back from its proposal. The new requirements do not mandate data sharing but only a data sharing plan to be included in each paper (and prespecified in study registration).[8]

Because of this trend toward a new norm where data sharing for randomized controlled trials (RCTs) becomes a standard, it seems important to assess how accessible the data are in journals with existing data sharing policies. Two leading general medical journals, *The BMJ*[9 10] and *PLOS Medicine*,[11] already have a policy expressly requiring data sharing as a condition for publication of clinical trials: data sharing became a requirement after January 2013 for RCTs on drugs and devices[9] and July 2015 for all therapeutics[10]

at *The BMJ*, and after March 2014 for all types of interventions at *PLOS Medicine*.

We explored the effectiveness of RCT data sharing in both journals in terms of data availability, feasibility, and accuracy of reanalyses and describe potential difficulties encountered in the process of performing reanalyses of the primary outcomes. We focused on RCTs because they are considered to represent high quality evidence and because availability of the data is crucial in the evaluation of health interventions. RCTs represent the most firmly codified methodology, which also allows data to be most easily analyzed. Moreover, RCTs have been the focus of transparency and data sharing initiatives,[12] owing to the importance of primary data availability in the evaluation of therapeutics (eg, for IPD meta-analyses).

Methods
The methods were specified in advance. They were documented in a protocol submitted for review on 12 November 2016 and subsequently registered with the Open Science Framework on 15 November 2016 (https://osf.io/u6hcv/register/565fb3678c5e4a66b5582f67).

Eligibility criteria
We surveyed publications of RCTs, including cluster trials and crossover studies, non-inferiority designs, and superiority designs, that had been submitted and published by *The BMJ* and *PLOS Medicine* subsequent to the adoption of data sharing policies by these journals.

Search strategy and study selection
We identified eligible studies from PubMed/Medline. For *The BMJ* we used the search strategy: "BMJ"[jour] AND ("2013/01/01"[PDAT]: "2017/01/01"[PDAT]) AND Randomized Controlled Trial[ptyp]. For *PLOS Medicine* we used: "PLoS Med"[jour]) AND ("2014/03/01"[PDAT]: "2017/01/01"[PDAT]) AND Randomized Controlled Trial[ptyp].

Two reviewers (FN and PJ) performed the eligibility assessment independently. Disagreements were resolved by consensus or in consultation with a third reviewer (JPAI or DM). More specifically, the eligibility assessment was based on the date of submission, not on the date of publication. When these dates were not available we contacted the journal editors for them.

Data extraction and datasets retrieval
A data extraction sheet was developed. For each included study we extracted information on study characteristics (country of corresponding author, design, sample size, medical specialty and disease, and funding), type of intervention (drug, device, other), and procedure to gather the data. Two authors (FN and PJ) independently extracted the data from the included studies. Disagreements were resolved by consensus or in consultation with a third reviewer (JPAI). One reviewer (FN) was in charge of retrieving the IPD for all included studies by following the instructions found in the data sharing statement of the included studies. More specifically, when data were available on request,

we sent a standardized email (https://osf.io/h9cas/). Initial emails were sent from a professional email address (fnaudet@stanford.edu), and three additional reminders were sent to each author two or three weeks apart, in case of non-response.

Data availability
Our primary outcome was data availability, defined as the eventual receipt of data presented with sufficient information to reproduce the analysis of the primary outcomes of the included RCTs (ie, complete data with clear labelling). Additional information was collected on type of data sharing (request by email, request using a specific website, request using a specific register, available on a public register, other), time for collecting the data (in days, time between first attempt to success of getting a database), deidentification of data (concerning name, birthdate, and address), type of data shared (from case report forms to directly analyzable datasets),[13] sharing of analysis code, and reasons for non-availability in case data were not shared.

Reproducibility
When data were available, a single researcher (FN) carried out a reanalysis of the trial. For each study, analyses were repeated exactly as described in the published report of the study. Whenever insufficient details about the analysis was provided in the study report, we sought clarifications from the trial investigators. We considered only analyses concerning the primary outcome (or outcomes, if multiple primary outcomes existed) of each trial. Any discrepancy between results obtained in the reanalysis and those reported in the publication was examined in consultation with a statistician (CS). This examination aimed to determine if, based on both quantitative (effect size, P values) and qualitative (clinical judgment) consideration, the discrepant results of the reanalysis entailed a different conclusion from the one reported in the original publication. Any disagreement or uncertainty over such conclusions was resolved by consulting a third coauthor with expertise in both clinical medicine and statistical methodology (JPAI). If, after this assessment process, it was determined that the results (and eventually conclusions) were still not reproduced, CS independently reanalyzed the data to confirm such a conclusion. Once the "not reproduced" status of a publication was confirmed, FN contacted the authors of the study to discuss the source of the discrepancy. After this assessment procedure, we classified studies into four categories: fully reproduced, not fully reproduced but same conclusion, not reproduced and different conclusion, and not reproduced (or partially reproduced) because of missing information.

Difficulties in getting and using data or code and performing reanalyses
We noted whether the sharing of data or analytical code, or both, required clarifications for which

additional queries had to be presented to the authors to obtain the relevant information, clarify labels or use, or both, and reproduce the original analysis of the primary outcomes. A catalogue of these queries was created and we grouped similar clarifications for descriptive purposes to generate a list of some common challenges and to help tackle these challenges pre-emptively in future published trials.

Statistical analyses

We computed percentages of data sharing or reproducibility with 95% confidence intervals based on binomial approximation or on Wilson score method without continuity correction if necessary.[14] For the purposes of registration, we hypothesized that if these data sharing policies were effective they would lead to more than 80% of studies sharing their data (ie, the lower boundary of the confidence interval had to be more than 80%). High rates of data sharing resulting from full data sharing policies should be expected. On the basis of experience of explicit data sharing policies in *Psychological Science*,[15] however, we knew that a rate of 100% was not realistic and judged that an 80% rate could be a desirable outcome.

When data were available, one researcher (FN) performed reanalyses using the open source statistical software R (R Development Core Team), and the senior statistician (CS) used SAS (SAS Institute). In addition, when authors shared their codes (in R, SAS, STATA (StataCorp 2017) or other), these were checked and used. Estimates of effect sizes, 95% confidence intervals, and P values were obtained for each reanalysis.

Changes from the initial protocol

Initially we planned to include all studies published after the data sharing policies were in place. Nevertheless, some authors who we contacted suggested that their studies were not eligible because the papers were submitted to the journal before the policy. We contacted editors who confirmed that policies applied for papers submitted (and not published) after the policy was adopted. In accordance, and to avoid any underestimation of data sharing rates, we changed our selection criteria to "studies submitted and published after the policy was adopted." For a few additional studies submitted before the policy, data were collected and reanalyzed but only described in the web appendix.

For the reanalysis we initially planned to consider non-reproducibility as a disagreement between reanalyzed results and results reported in the original publication by more than 2% in the point estimate or 95% confidence interval. After reanalysis of a couple of studies we believed that such a definition was sometimes meaningless. Interpreting a RCT involves clinical expertise and cannot be reduced to solely quantitative factors. Accordingly, we changed our definition and provided a detailed description of reanalyses and published results, mentioning the nature of effect size and the type of outcome considered.

Patient involvement

We had no established contacts with specific patient groups who might be involved in this project. No patients were involved in setting the research question or the outcome measures, nor were they involved in the design and implementation of the study. There are no plans to involve patients in the dissemination of results, nor will we disseminate results directly to patients.

Results

Characteristics of included studies

Figure 1 shows the study selection process. The searches done on 12 November 2016 resulted in 159 citations. Of these, 134 full texts were considered for eligibility. Thirty seven RCTs (21 from *The BMJ* and 16 from *PLOS Medicine*) published between 2013 and 2016 met our eligibility criteria. Table 1 presents
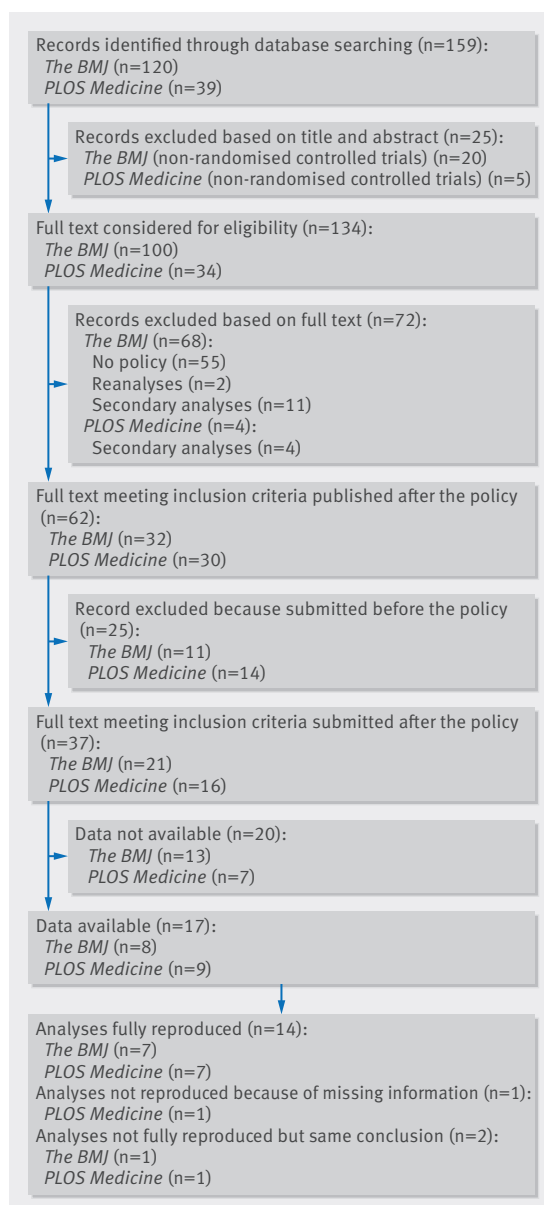

Fig 1 | Study flow diagram

Table 1 | Characteristics of included studies. Values are numbers (percentages) unless stated otherwise

Characteristics	All (37 studies)	The BMJ (21 studies)	PLOS Medicine (16 studies)
Geographical area of lead country:			
Europe	25 (67)	17 (80)	8 (50)
Australia and New Zealand	4 (11)	1 (5)	3 (19)
Northern America	3 (8)	1 (5)	2 (12.5)
Africa	3 (8)	1 (5)	2 (12.5)
East Asia	1 (3)	0 (0)	1 (6)
Middle East	1 (3)	1 (5)	0 (0)
Type of intervention:			
Drug	20 (54)	13 (62)	7 (44)
Device	8 (22)	8 (38)	0 (0)
Complex intervention	9 (24)	0 (0)	9 (56)
Medical specialty:			
Infectious disease	12 (33)	4 (19)	8 (50)
Rheumatology	5 (14)	5 (24)	0 (0)
Endocrinology/nutrition	4 (11)	1 (5)	3 (19)
Paediatrics	3 (8)	2 (9)	1 (6)
Mental health/addiction	2 (5)	1 (5)	1 (6)
Obstetrics	2 (5)	1 (5)	1 (6)
Emergency medicine	2 (5)	2 (9)	0 (0)
Geriatrics	2 (5)	0 (0)	2 (13)
Other	5 (14)	5 (24)	0 (0)
Designs:			
Superiority (head to head)	18 (49)	15 (71)	3 (19)
Superiority (factorial)	1 (3)	1 (5)	0 (0)
Superiority (clusters)	8 (21)	1 (5)	7 (43)
Non-inferiority+superiority (head to head)	4 (11)	1 (5)	3 (19)
Non-inferiority (head to head)	6 (16)	3 (14)	3 (19)
Median (interquartile range) sample size	432 (213-1070)*	221 (159-494)	1047 (433-2248)*
Private sponsorship:			
No	26 (70)	15 (71)	11 (69)
Provided device	1 (3)	1 (5)	0 (0)
Provided intervention	1 (3)	0 (0)	1 (6)
Provided drug	5 (13)	1 (5)	4 (25)
Provided drug and some financial support	2 (5)	2 (9)	0 (0)
Provided partial financial support	1 (3)	1 (5)	0 (0)
Provided total financial support	1 (3)	1 (5)	0 (0)
Statement of availability:			
Ask to contact by email	23 (62)	17 (81)	6 (38)
Explain how to retrieve data (eg, platform)	9 (24)	0 (0)	9 (56)
State "no additional data available"	2 (5)	2 (9)	0 (0)
Ask to contact by mail	1 (3)	0 (0)	1 (6)
Embargo	1 (3)	1 (5)	0 (0)
No statement	1 (3)	1 (5)	0 (0)

Rounded percentages add up to 100% for each variable.
*Exact sample size was not reported for one cluster trial in PLOS Medicine.

the characteristics of these studies. These RCTs had a median sample size of 432 participants (interquartile range 213-1070), had no industry funding in 26 cases (70%), and were led by teams from Europe in 25 cases (67%). Twenty RCTs (54%) evaluated pharmacological interventions, 9 (24%) complex interventions (eg, psychotherapeutic program), and 8 (22%) devices.

Data availability

We were able to access data for 19 out of 37 studies (51%). Among these 19 studies, the median number of days for collecting the data was 4 (range 0-191). Two of these studies, however, did not provide sufficient information within the dataset to enable direct reanalysis (eg, had unclear labels). Therefore 17 studies satisfied our definition of data availability. The rate of data availability was 46% (95% confidence interval 30% to 62%).

Data were in principle available for two additional studies not included in the previous count and both authored by the same research team. However, the authors asked us to cover the financial costs of preparing the data for sharing (£607; $857; €694). Since other teams shared the data for free, we considered that it would not have been fair to pay some and not others for similar work in the context of our project and so we classified these two studies as not sharing data. For a third study, the authors were in correspondence with us and discussing conditions for sharing data, but we did not receive that data by the time our data collection process was determined to be over (seven months). If these three studies were included, the proportion of data sharing would be 54% (95% confidence interval 38% to 70%).

For the remaining 15 studies classified as not sharing data, reasons for non-availability were: no answer to

the different emails (n=7), no answer after an initial agreement (n=2), and refusal to share data (n=6). Explanations for refusal to share data included lack of endorsement of the objectives of our study (n=1), personal reasons (eg, sick leave, n=2), restrictions owing to an embargo on data sharing (n=1), and no specific reason offered (n=2). The existence of possible privacy concerns was never put forward as a reason for not sharing data.

Among the 19 studies sharing some data (analyzable datasets and non-analyzable datasets), 16 (84%) datasets were totally deidentified. Birthdates were found in three datasets and geographical information (country and postcode) in one of these three. Most datasets were data ready for analysis (n=17), whereas two required some additional processing before the analysis could be repeated. In these two cases, such processing was difficult to implement (even with the code being available) and the authors were contacted to share analyzable data. Statistical analysis code was available for seven studies (including two that were obtained after a second specific request).

Reproducibility

Among the 17 studies[16-32] providing sufficient data for reanalysis of their primary outcomes, 14 (82%, 95% confidence interval 59% to 94%) studies were fully reproduced on all their primary outcomes. One of the 17 studies did not provide enough information in the Methods section for the analyses to be reproduced (specifically, the methods used for adjustment were unclear). We contacted the authors of this study to obtain clarifications but received no reply. Of the remaining 16 studies, we reanalyzed 47 different primary analyses. Two of these studies were considered not fully reproduced. For one study, we identified an error in the statistical code and for the other we found slightly different numerical values for the effect sizes measured as well as slight differences in numbers of patients included in the analyses (a difference of one patient in one of four analyses). Nevertheless, similar conclusions were reached in both cases (these two studies were categorized as not fully reproduced but reaching the same conclusion). Therefore, we found no results contradicting the initial publication, neither in terms of magnitude of the effect (table 2) nor in terms of statistical significance of the finding (fig 2).

We retrieved the data of three additional studies, published in *The BMJ* after its data sharing policy was in place (but submitted before the policy). Although these studies were ineligible for our main analysis, reanalyses were performed and also reached the same conclusions as the initial study (see supplementary e-Table 1).

Difficulties in getting and using data or code and performing reanalyses

Based on our correspondence with authors, we identified several difficulties in getting the data. A common concern pertained to the costs of the data sharing process—for example, the costs of preparing the data or translating the database from one language to another. Some authors wondered whether their team or our team should assume these costs. In addition, some of the authors balanced these additional costs with their perceived benefits of sharing data for the purpose of this study and seemed to rather value data sharing for the purpose of a meta-analysis than for reanalyses, possibly because of the risk acknowledged by one investigator we contacted about "naming and shaming" individual studies or investigators.

Getting prepared and preplanning for data sharing still seems to be a challenge for many trial groups; data sharing proved to be novel for some authors who were unsure how to proceed. Indeed, there was considerable heterogeneity between different procedures to share data: provided in an open repository (n=5), downloadable on a secured website (n=1) after registration, included as appendix of the published paper (n=3), or sent by email (n=10). On three occasions, we signed a data sharing request or agreement. In these agreements, the sponsor and recipient parties specified the terms that bound them in the data sharing process (eg, concerning the use of data, intellectual property, etc). In addition, typically there was no standard in what type of data were shared (at what level of cleaning and processing). In one case, authors mentioned explicitly that they followed standardized guidelines[33] to prepare the dataset.

Some analyses were complex and it was sometimes challenging to reanalyze data from specific designs or when unusual measures were used (eg, relative change in percentage). Obtaining more information about the analysis by contacting authors was necessary for 6 of 17 studies to replicate the findings. In one case, specific exploration of the code revealed that in a survival analysis, authors treated repeated events in the same patient (multiple events) as distinct observations. This was in disagreement with the methods section describing usual survival analysis (taking into account only the first event for each patient). However, alternative analyses did not contradict the published results.

Three databases did not provide sufficient information to reproduce the analyses. Missing data concerned variables used for adjustment, definition of the analysis population, and randomization groups. Communication with authors was therefore necessary and was fruitful in one of these three cases.

Discussion

In two prominent medical journals with a strong data sharing policy for randomized controlled trials (RCTs), we found that for 46% (95% confidence interval 30% to 62%) of published articles the original investigators shared their data with sufficient information to enable reanalyses. This rate was less than the 80% boundary that we prespecified as an acceptable threshold for papers submitted under a policy that makes data sharing an explicit condition for publication. However, despite being lower than might be desirable, a 46% data sharing rate is much higher than the average rate

Table 2 | Results of reanalyses

Study, treatment comparison; condition	Primary outcome measures	Effect size (type)	Effect size: published paper	P value	Effect size: reanalysis	P value
Luedtke 2015[27]*: transcranial direct current stimulation (TDCs) versus sham TDCs; chronic low back pain	Visual analogue scale, time 1	MD (99% CI)	1 (−8.69 to 6.30)	0.68	1.42 (−6.06 to 8.90)	0.62
	Visual analogue scale, time 2	MD (99% CI)	3 (−10.32 to 6.73)	0.58	5.19 (−3.62 to 14.00)	0.13
	Oswestry disability index, time 1	MD (99% CI)	1 (−1.73 to 1.98)	0.86	−0.12 (−1.99 to 1.75)	0.87
	Oswestry disability index, time 2	MD (99% CI)	0.01 (−2.45 to 2.62)	0.92	−0.49 (−3.15 to 2.18)	0.64
Cohen 2015[20]: epidural steroid injections versus gabapentin; lumbosacral radicular pain	Average leg pain, time 1	MD (95% CI)	0.4 (−0.3 to 1.2)	0.25	0.4 (−0.3 to 1.2)	0.26
	Average leg pain, time 2	MD (95% CI)	0.3 (−0.5 to 1.2)	0.43	0.3 (−0.5 to 1.2)	0.45
Hyttel-Sorensen 2015[24]: near infrared spectroscopy versus usual care; preterm infants	Burden of hypoxia and hyperoxia	RC% (95% CI)	−58 (−35 to −74)	<0.001	−58 (−36 to −72)	<0.001
Bousema 2016[18]†: hotspot targeted interventions versus usual interventions; malaria	nPCR parasite prevalence, time 1, zone 1	MD (95% CI)	10.2 (−1.3 to 21.7)	0.075; 0.024	10.2 (−2.5 to 22.9)	0.095; 0.047
	nPCR parasite prevalence, time 2, zone 1	MD (95% CI)	4.2 (−5.1 to 13.8)	0.326; 0.265	4.3 (−5.3 to 13.9)	0.328; 0.157
	nPCR parasite prevalence, time 1, zone 2	MD (95% CI)	3.6 (−2.6 to 9.7)	0.219; 0.216	3.6 (−3.3 to 0.1)	0.239; 0.339
	nPCR parasite prevalence, time 2, zone 2	MD (95% CI)	1.0 (−7.0 to 9.1)	0.775; 0.713	1.0 (−7.6 to 9.7)	0.778; 0.858
	nPCR parasite prevalence, time 1, zone 3	MD (95% CI)	3.8 (−2.4 to 10.0)	0.199; 0.187	4.0 (−2.1 to 10.2)	0.170; 0.253
	nPCR parasite prevalence, time 2, zone 3	MD (95% CI)	1.0 (−8.3 to 10.4)	0.804; 0.809	1.0 (−8.4 to 10.5)	0.805; 0.814
Gysin-Maillart 2016[21]‡: brief therapy versus usual care; suicide attempt	Suicide attempt	HR (95% CI)	0.17 (0.07 to 0.46)	<0.001	0.17 (0.06 to 0.49)	<0.001
Lombard 2016[26]: low intensity intervention versus general health intervention; obesity	Change in weight (unadjusted)	MD (95% CI)	−0.92 (−1.67 to −0.16)	.	−0.92 (−1.66 to −0.18)	.
	Change in weight (adjusted)	MD (95% CI)	−0.87 (−1.62 to −0.13)	.	−0.89 (−1.63 to −0.15)	.
Polyak 2016[30]: cotrimoxazole prophylaxis cessation versus continuation; HIV	Malaria, pneumonia, diarrhea, or mortality (ITT)	IRR (95% CI)	2.27 (1.52 to 3.38)	<0.001	2.27 (1.51 to 3.39)	<0.001
	Malaria, pneumonia, diarrhea, or mortality (PP)	IRR (95% CI)	2.27 (1.52 to 3.39)	<0.001	2.27 (1.52 to 3.40)	<0.001
Robinson 2015[31]: blood stage plus liver stage drugs or blood stage drugs only; malaria	*Plasmodium vivax* infection by qPCR, time 1	HR (95% CI)	0.18 (0.14 to 0.25)	<0.001	0.17 (0.13 to 0.24)	<0.001
	P vivax infection by qPCR, time 2	HR (95% CI)	0.17 (0.12 to 0.24)	<0.001	0.16 (0.11 to 0.22)	<0.001
Harris 2015[23]: nurse delivered walking intervention versus usual care; older adults	Daily step count	MD (95% CI)	1037 (513 to 1560)	<0.001	1037 (513 to 1560)	<0.001
Adrion 2016[16]: low dose betahistine versus placebo; Meniere's disease	Vertigo attack	RaR (95% CI)	1.036 (0.942 to 1.140)	0.759	1.033 (0.942 to 1.133)	0.777
	Vertigo attack	RaR (95% CI)	1.012 (0.919 to 1.114)	0.777	1.011 (0.921 to 1.109)	0.777
Selak 2014[32]: fixed dose combination treatment versus usual care; patients at high risk of cardiovascular disease in primary care	Use of antiplatelet, statin, ≥2 blood pressure lowering drugs	RR (95% CI)	1.75 (1.52 to 2.03)	<0.001	1.75 (1.52 to 2.03)	<0.001
	Change in systolic blood pressure	MD (95% CI)	−2.2 (−5.6 to 1.2)	0.21	−2.2 (−5.6 to 1.2)	0.21
	Change in diastolic blood pressure	MD (95% CI)	−1.2 (−3.2 to 0.8)	0.22	−1.2 (−3.2 to 0.8)	0.22
	Change in low density lipoprotein cholesterol level	MD (95% CI)	−0.05 (−0.17 to 0.08)	0.46	−0.05 (−0.17 to 0.08)	0.46
Chesterton 2013[19]: transcutaneous electrical nerve stimulation versus usual care; tennis elbow	Change in elbow pain intensity	MD (95% CI)	−0.33 (−0.96 to 0.31)	0.31	−0.33 (−1.01 to 0.34)	0.33
Atukunda 2014[17]: sublingual misoprostol versus intramuscular oxytocin; prevention of postpartum haemorrhage	Blood loss >500 mL	RR (95% CI)	1.64 (1.32 to 2.05)	<0.001	1.65 (1.32 to 2.05)	<0.001
Paul 2015[29]: trimethoprim-sulfamethoxazole versus vancomycin; severe infections caused by meticillin resistant *Staphylococcus aureus*	Treatment failure (ITT)	RR (95% CI)	1.38 (0.96 to 1.99)	0.08	1.38 (0.96 to 1.99)	0.08
	Treatment failure (PP)	RR (95% CI)	1.24 (0.82 to 1.89)	0.36	1.24 (0.82 to 1.89)	0.36
	All cause mortality (ITT)	RR (95% CI)	1.27 (0.65 to 2.45)	0.57	1.27 (0.65 to 2.45)	0.57
	All cause mortality (PP)	RR (95% CI)	1.05 (0.47 to 2.32)	1	1.05 (0.47 to 2.32)	1
Omerod 2015[28]: ciclosporin versus prednisolone; pyoderma gangrenosum	Speed of healing over six weeks (cm²/day)	MD (95% CI)	0.003 (−0.20 to 0.21)	0.97	0.003 (−0.20 to 0.21)	0.97
Hanson 2015[22]: home based counselling strategy versus usual care; neonatal care	Neonatal mortality	OR (95% CI)	1.0 (0.9 to 1.2)	0.339	1.0 (0.9 to 1.2)	0.339

nPCR=nested polymerase chain reaction; qPCR=quantitative real time PCR; MD=mean differences; RC%=relative change in percentage; HR=Hazard Ratio; RaR=rate ratio; RR=relative risk; OR=odds ratio; ITT=intention to treat; PP=per protocol.

*Study judged as not fully reproduced but reaching same conclusion because of differences in numerical estimates: for one of these analyses current authors found a difference in number of patients analyzed (one patient more in reanalysis).

†P values of adjusted analyses were computed to complement non-adjusted analyses.

‡Study judged as not fully reproduced but reaching same conclusion because initial analysis did not take into account repeated nature of data. This was corrected in reanalysis without major differences in estimates.
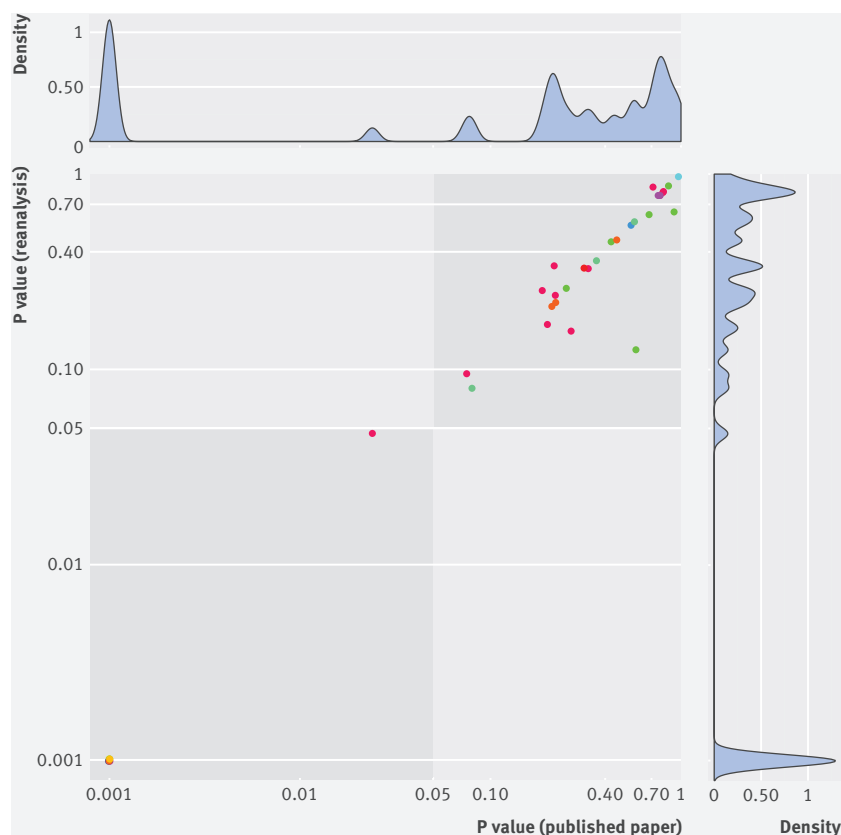
Fig 2 | P values in initial analyses and in reanalyses. Axes are on a log scale. Blue indicates identical conclusion between initial analysis and reanalysis. Dots of same colors indicate analyses from same study

of biomedical literature at large, in which data sharing is almost non-existent[34] (with few exceptions in some specific disciplines, such as genetics).[35] Moreover, our analyses focused on publications that were submitted directly after the implementation of new data sharing

policies, which might be expected to have practical and cultural barriers to their full implementation. Indeed, our correspondence with the authors helped identify several practical difficulties connected to data sharing, including difficulties in contacting corresponding authors, and lack of time and financial resources on their behalf in preparing the datasets for us. In addition, we found a wide variety of data sharing practices between study groups (ie, regarding the type of data that can be shared and the procedures that are necessary to follow to get the data). Data sharing practices could evolve in the future to deal with these barriers to data sharing (table 3).

For all results that we were able to reanalyze, we reached similar conclusions (despite occasional slight differences in the numerical estimations) to those reported in the original publication, and this result that at least the available data shared do correspond closely to the reported results is reassuring. Of course, there is a large amount of diversity on what exactly "raw data" mean and they can involve various transformations (from the case report forms to coded and analyzable data).[13] Here, we relied on late stage, coded, and cleaned data and therefore the potential for leading to a different conclusion was probably small. Data processing, coding, cleaning, and recategorization of events can have a substantial impact on the results in some trials. For example, SmithKline Beecham's Study 329 was a well known study on paroxetine in adolescent depression, presenting the drug as safe and effective,[36] whereas a reanalysis starting from the case report forms found a lack of efficacy and some serious safety issues.[37]

Strengths and weaknesses of this study

Some leading general medical journals—*New England Journal of Medicine*, *Lancet*, *JAMA*, and *JAMA Internal*

Table 3	Some identified challenges (and suggestions) for data sharing and reanalyses	
Identified problems	**Suggested solutions for various stakeholders**
Data sharing policies leaves responsibility and burden to researchers, to obtain all necessary resources (time, money, technical and organizational tools and services, ethical and legal compliance, etc)	**Patients and clinicians** should help to develop awareness of a common ownership of the data, intended as common responsibility also in providing all necessary resources to make data sharable for effective and ethical use **Researchers** should pre-emptively address and seek funding for data sharing **Funders** should allow investigators to use funds towards data sharing **Academic institutions** should reward data sharing activities through promotion and tenure
Getting prepared and preplanning for data sharing is still in progress in trials units. There is considerable heterogeneity between different procedures to share data and types of data that are shared	**All stakeholders** should adopt some clear and homogeneous rules (eg, guidelines) for best practices in data sharing **Clinical trials groups** should develop comprehensive educational outreach about data sharing **Institutional review boards** should systematically address data sharing issues
Data sharing on request leaves researchers exclusive discretion on decision about whether to share their own data to other research groups, for which objectives, and under which terms and conditions	**Editors** should adopt more binding policies than the current ICJME requirement When routine data deposition is not ethically feasible, **clinical trials groups** should prespecify criteria for data sharing by adopting effective and transparent systems to review requests
There could be some difficulties in contacting corresponding authors, limiting data sharing on request	**Researchers and editors** should favor data deposition when it is ethically possible
Some identifying information such as date of birth can be found in some databases	**Researchers** must ensure that databases are deidentified before sharing **Institutional review board** should provide guidance on the requested level of deidentification for each individual study. It included seeking consent for sharing IPD from trial participants
Shared databases need to be effectively sharable (eg, complete, homogeneous), including meta-data (eg, descriptive labels, description of pre-analysis processing tools, methods of analysis)	**Clinical trials groups** should develop comprehensive educational outreach about data sharing **Researchers** should provide details concerning the detailed labels used in the table and analytical code **Editors** who ask for data and code to be shared should ensure that this material is reviewed
Reproducible research practices are not limited to sharing data, materials, and code. Complete reporting of methods and statistical analyses are also relevant	**Clinical trials groups** should develop comprehensive educational outreach about data sharing **Researchers** should also share their detailed statistical analysis plan **Reporting guidelines** should emphasize best practices in data sharing about computational reproducibility

Medicine—have had no specific policy for data sharing in RCTs until recently. *Annals of Internal Medicine* has encouraged (but not demanded) data sharing since 2007.[38] *BMC Medicine* adopted a similar policy in 2015. *The BMJ* and *PLOS Medicine* have adopted stronger policies, beyond the ICMJE policy, that mandate data sharing for RCTs. Our survey of RCTs published in these two journals might therefore give a taste of the impact and caveats of such full policies. However, care should be taken to not generalize these results to other journals. First, we had a selected sample of studies. Our sample included studies (mostly from Europe) that are larger and less likely to be funded by the industry than the average published RCT in the medical literature.[39] Several RCTs, especially those published in *PLOS Medicine*, were cluster randomized studies and many explored infectious disease or important public health issues, characteristics that are not common in RCTs overall. Some public funders (or charities) involved in their funding have already open access policies, such as the Bill and Melinda Gates foundation or the UK National Institute for Health Research.

These reasons also explain why we did not compare journals. Qualitatively they do not publish the same kind of RCTs, and quantitatively *The BMJ* and *PLOS Medicine* publish few RCTs compared with other leading journals, such as the *New England Journal of Medicine* and *Lancet*. Any comparisons might have been subject to confounding. Similarly, we believed that before and after studies at *The BMJ* and *PLOS Medicine* might have been subject to historical bias and confounding factors since such policies might have changed the profile of submitting authors: researchers with a specific interest in open science and reproducibility might have been attracted and others might have opted for another journal. We think comparative studies will be easier to conduct when all journals adopt data sharing standards. Even with the current proposed ICJME requirements, *The BMJ* and *PLOS Medicine* will be journals with stronger policies. In addition, *The BMJ* and *PLOS Medicine* are both major journals with many resources, such as in-depth discussion of papers in editorial meetings and statistical peer review. Our findings concerning reproducibility might not apply to smaller journals with more limited resources. We cannot generalize this finding to studies that did not share their data: authors who are confident in their results might more readily agree to share their data and this may lead to overestimation of reproducibility. In addition, we might have missed a few references using the filter "Randomized Controlled Trial[ptyp]"; however, this is unlikely to have affected data sharing and reproducibility rates.

Finally, in our study the notion of data sharing was restricted to a request by a researcher group, whereas in theory other types of requestors (patients, clinicians, health and academic institutions, etc) may also be interested in the data. Moreover, we followed the strict procedure presented in the paper and without sending any correspondence (eg, rapid response) on the journals' website. In addition, our reanalyses were based on primary outcomes, whereas secondary and safety outcomes may be more problematic in their reproduction. These points need to be explored in further studies.

Strengths and weaknesses in relation to other studies

In a precedent survey of 160 randomly sampled research articles in *The BMJ* from 2009 to 2015, excluding meta-analyses and systematic reviews,[40] the authors found that only 5% shared their datasets. Nevertheless, this survey assessed data sharing among all studies with original raw data, whereas *The BMJ* data sharing policy specifically applied to clinical trial data. When considering clinical trials bound by *The BMJ* data sharing policy (n=21), the percentage shared was 24% (95% confidence interval 8% to 47%). Our study identified higher rates of shared datasets in accordance with an increase in the rate of "data shared" for every additional year between 2009 and 2015 found in the previous survey. It is not known whether it results directly from the policy implementation or from a slow and positive cultural change of trialists.

Lack of response from authors was also identified as a caveat in the previous evaluation, which suggested that the wording of *The BMJ* policy, such as availability on "reasonable request," might be interpreted in different ways.[40] Previous research across PLOS also suggests that data requests by contacting authors might be ineffective sometimes.[41] Despite writing a data sharing agreement in their paper, corresponding authors are still free to decline data requests. One could question whether sharing data for the purposes of our project constitutes a "reasonable request." One might consider that exploring the effectiveness of data sharing policies lies outside the purpose for which the initial trial was done and for which the participants gave consent. A survey found that authors are generally less willing to share their data for the purpose of reanalyses than, for instance, for individual patient data (IPD) meta-analyses.[42] Nevertheless, reproducibility checks based on independent reanalysis are perfectly aligned with the primary objective of clinical trials and indeed with patient interests. Conclusions that persist through substantial reanalyses are becoming more credible. To explain and support the interest of our study, we have registered a protocol and have transparently described our intentions in our email, but data sharing rates were still lower than we expected. Active auditing of data sharing policies by journal editors may facilitate the implementation of data sharing.

Concerning reproducibility of clinical trials, an empirical analysis suggests that only a small number of reanalyses of RCTs have been published to date; of these, only a minority was conducted by entirely independent authors. In a previous empirical evaluation of published reanalyses, 35% (13/37) of the reanalyses yielded changes in findings that implied conclusions different from those of the original article as to whether patients should be treated or not or about which patients should be treated.[43] In our assessment,

we found no differences in conclusions pertaining to treatment decisions. The difference between the previous empirical evaluation and the current one is probably due to many factors. It is unlikely that published reanalyses in the past would be have been published if they had found the same results and had reached the same conclusions as the original analysis. Therefore, the set of published reanalyses is enriched with discrepant results and conclusions. Moreover, published reanalyses addressed the same question on the same data but using typically different analytical methods. Conversely, we used the same analysis employed by the original paper. In addition, the good computational reproducibility (replication of analyses) we found is only one aspect of reproducibility[44] and should not be over-interpreted. For example, in psychology, numerous laboratories have volunteered to re-run experiments (not solely analyses), with the methods used by the original researchers. Overall, 39 out of 100 studies were considered as successfully replicated.[45] But attempts to replicate psychological studies are often easier to implement than attempts to replicate RCTs, which are often costly and difficult to run. This is especially true for large RCTs.

Meaning of the study: possible explanations and implications for clinicians and policy makers

The ICJME's requirements adopted in 2017 mandate that a data sharing plan will have to be included in each paper (and prespecified in study registration).[8] Because data sharing among two journals with stronger requirements was not optimal, our results suggest that this not likely to be sufficient to achieve high rates of data sharing. One can imagine that individual authors will agree to write such a statement in line with the promise of publication, but the time and costs involved into such data preparation might lead authors to be reluctant to answer data sharing requests. Interestingly the ICJME also mandates that clinical trials that begin enrolling participants on or after 1 January 2019 must include a data sharing plan in the trial's registration. An a priori data sharing plan might push more authors to pre-emptively deal with and find funding for sharing, but its eventual impact on sharing is unknown. Funders are well positioned to facilitate data sharing. Some, such as the Wellcome Trust, already allow investigators to use funds towards the charges for open access, which is typically a small fraction of awarded funding. Funders could extend their funding policy and also allow investigators to use a similar small fraction of funding towards enabling data sharing. In addition, patients can help promote a culture of data sharing for clinical trials.[5]

In addition, reproducible research does not solely imply sharing data but also reporting all steps of the statistical analysis. This is one core principle of the CONSORT statement "Describe statistical methods with enough detail to enable a knowledgeable reader with access to the original data to verify the reported results."[46] It is none the less sometimes difficult to provide such detailed information in a restricted number of lines (ie, a paper). We suggest that details of the statistical analysis plan have to be provided as well as the detailed labels used in the table, and efficient analytical code sharing is also essential.[47]

We suggest that if journals ask for data and code to be shared, they should ensure that this material is also reviewed by editorial staff (with or without peer reviewers) or at a minimum checked for completeness and basic usability.[47] This could also be positively translated in specific incentives for the paper; for example, many psychology journals use badges[15] as signs of good research practices (including data sharing) adopted by papers. And, beyond incentivizing authors, journals adopting such practices could also be incentivized by a gain in reputation and credibility.

Though ensuring patient privacy and lack of explicit consent for sharing are often cited as major barriers to sharing RCT data (and generally accepted as valid exemptions),[48] none of the investigators we approached mentioned this reason. This could suggest that technical constraints, lack of incentives and dedicated funding, and a general diffidence towards reanalyses might be more germane obstacles to be addressed.

Finally, data sharing practices differed from one team to another. There is thus a need for standardization and for drafting specific guidelines for best practices in data sharing. In the United Kingdom, the Medical Research Council Hubs for Trials Methodology Research has proposed a guidance to facilitate the sharing of IPD for publicly funded clinical trials.[49-51] Other groups might consider adopting similar guidelines.

Unanswered questions and future research

We recommend some prospective monitoring of data sharing practices to ensure that the new requirements of the ICJME are effective and useful. To this end, it should also be kept in mind that data availability is a surrogate of the expected benefit of having open data. Proving that data sharing rates impacts reuse[52] of the data is a further step. Proving that this reuse might translate into discovery that can change care without generating false positive findings (eg, in series of unreliable a posteriori subgroup analyses) is even more challenging.

tables) in the study and can take responsibility for the integrity of the data and the accuracy of the data analysis. FN is the guarantor.

1   Taichman DB, Backus J, Baethge C, et al. Sharing Clinical Trial Data: A Proposal from the International Committee of Medical Journal Editors. *PLoS Med* 2016;13:e1001950. doi:10.1371/journal. pmed.1001950.

2   Platt R, Ramsberg J. Challenges for Sharing Data from Embedded Research. *N Engl J Med* 2016;374:1897. doi:10.1056/ NEJMc1602016.

3   Kalager M, Adami HO, Bretthauer M. Recognizing Data Generation. *N Engl J Med* 2016;374:1898. doi:10.1056/NEJMc1603789.

4   Devereaux PJ, Guyatt G, Gerstein H, Connolly S, Yusuf S, International Consortium of Investigators for Fairness in Trial Data Sharing. Toward Fairness in Data Sharing. *N Engl J Med* 2016;375:405-7. doi:10.1056/NEJMp1605654.

5   Haug CJ. Whose Data Are They Anyway? Can a Patient Perspective Advance the Data-Sharing Debate? *N Engl J Med* 2017;376:2203-5. doi:10.1056/NEJMp1704485.

6   Rockhold F, Nisen P, Freeman A. Data Sharing at a Crossroads. *N Engl J Med* 2016;375:1115-7. doi:10.1056/NEJMp1608086.

7   Rosenbaum L. Bridging the Data-Sharing Divide - Seeing the Devil in the Details, Not the Other Camp. *N Engl J Med* 2017;376:2201-3. doi:10.1056/NEJMp1704482.

8   Taichman DB, Sahni P, Pinborg A, et al. Data sharing statements for clinical trials. *BMJ* 2017;357:j2372. doi:10.1136/bmj.j2372

9   Godlee F, Groves T. The new BMJ policy on sharing data from drug and device trials. *BMJ* 2012;345:e7888. doi:10.1136/bmj.e7888

10 Loder E, Groves T. The BMJ requires data sharing on request for all trials. *BMJ* 2015;350:h2373. doi:10.1136/bmj.h2373

11 Bloom T, Ganley E, Winker M. Data access for the open access literature: PLOS's data policy. *PLoS Biol* 2014;12:e1001797doi:10.1371/journal.pbio.1001797.

12 Mello MM, Francer JK, Wilenzick M, Teden P, Bierer BE, Barnes M. Preparing for responsible sharing of clinical trial data. *N Engl J Med* 2013;369:1651-8. doi:10.1056/NEJMhle1309073.

13 Zarin DA, Tse T. Sharing Individual Participant Data (IPD) within the Context of the Trial Reporting System (TRS). *PLoS Med* 2016;13:e1001946. doi:10.1371/journal.pmed.1001946.

14 Newcombe RG. Two-sided confidence intervals for the single proportion: comparison of seven methods[pii]. *Stat Med* 1998;17:857-72. doi:10.1002/(SICI)1097-0258(19980430)17: 8‹857::AID-SIM777›3.0.CO;2-E.

15 Kidwell MC, Lazarević LB, Baranski E, et al. Badges to Acknowledge Open Practices: A Simple, Low-Cost, Effective Method for Increasing Transparency. *PLoS Biol* 2016;14:e1002456. doi:10.1371/journal. pbio.1002456.

16 Adrion C, Fischer CS, Wagner J, Gürkov R, Mansmann U, Strupp M, BEMED Study Group. Efficacy and safety of betahistine treatment in patients with Meniere's disease: primary results of a long term, multicentre, double blind, randomised, placebo controlled, dose defining trial (BEMED trial). *BMJ* 2016;352:h6816. doi:10.1136/ bmj.h6816.

17 Atukunda EC, Siedner MJ, Obua C, Mugyenyi GR, Twagirumukiza M, Agaba AG. Sublingual misoprostol versus intramuscular oxytocin for prevention of postpartum hemorrhage in Uganda: a double-blind randomized non-inferiority trial. *PLoS Med* 2014;11:e1001752. doi:10.1371/journal.pmed.1001752.

18 Bousema T, Stresman G, Baidjoe AY, et al. The Impact of Hotspot-Targeted Interventions on Malaria Transmission in Rachuonyo South District in the Western Kenyan Highlands: A Cluster-Randomized Controlled Trial. *PLoS Med* 2016;13:e1001993. doi:10.1371/ journal.pmed.1001993.

19 Chesterton LS, Lewis AM, Sim J, et al. Transcutaneous electrical nerve stimulation as adjunct to primary care management for tennis elbow: pragmatic randomised controlled trial (TATE trial). *BMJ* 2013;347:f5160. doi:10.1136/bmj.f5160.

20 Cohen SP, Hanling S, Bicket MC, et al. Epidural steroid injections compared with gabapentin for lumbosacral radicular pain: multicenter randomized double blind comparative efficacy study. *BMJ* 2015;350:h1748. doi:10.1136/bmj.h1748.

21 Gysin-Maillart A, Schwab S, Soravia L, Megert M, Michel K. A Novel Brief Therapy for Patients Who Attempt Suicide: A 24-months Follow-Up Randomized Controlled Study of the Attempted Suicide Short Intervention Program (ASSIP). *PLoS Med* 2016;13:e1001968. doi:10.1371/journal.pmed.1001968.

22 Hanson C, Manzi F, Mkumbo E, et al. Effectiveness of a Home-Based Counselling Strategy on Neonatal Care and Survival: A Cluster-Randomised Trial in Six Districts of Rural Southern Tanzania. *PLoS Med* 2015;12:e1001881. doi:10.1371/journal.pmed.1001881.

23 Harris T, Kerry SM, Victor CR, et al. A primary care nurse-delivered walking intervention in older adults: PACE (pedometer accelerometer consultation evaluation)-Lift cluster randomised controlled trial. *PLoS Med* 2015;12:e1001783. doi:10.1371/journal.pmed.1001783.

24 Hyttel-Sorensen S, Pellicer A, Alderliesten T, et al. Cerebral near infrared spectroscopy oximetry in extremely preterm infants: phase II randomised clinical trial. *BMJ* 2015;350:g7635. doi:10.1136/bmj. g7635.

25 Liu X, Lewis JJ, Zhang H, et al. Effectiveness of Electronic Reminders to Improve Medication Adherence in Tuberculosis Patients: A Cluster-Randomised Trial. *PLoS Med* 2015;12:e1001876. doi:10.1371/ journal.pmed.1001876.

26 Lombard C, Harrison C, Kozica S, Zoungas S, Ranasinha S, Teede H. Preventing Weight Gain in Women in Rural Communities: A Cluster Randomised Controlled Trial. *PLoS Med* 2016;13:e1001941. doi:10.1371/journal.pmed.1001941.

27 Luedtke K, Rushton A, Wright C, et al. Effectiveness of transcranial direct current stimulation preceding cognitive behavioural management for chronic low back pain: sham controlled double blinded randomised controlled trial. *BMJ* 2015;350:h1640. doi:10.1136/bmj.h1640.

28 Ormerod AD, Thomas KS, Craig FE, et al, UK Dermatology Clinical Trials Network's STOP GAP Team. Comparison of the two most commonly used treatments for pyoderma gangrenosum: results of the STOP GAP randomised controlled trial. *BMJ* 2015;350:h2958. doi:10.1136/bmj.h2958.

29 Paul M, Bishara J, Yahav D, et al. Trimethoprim-sulfamethoxazole versus vancomycin for severe infections caused by meticillin resistant Staphylococcus aureus: randomised controlled trial. *BMJ* 2015;350:h2219. doi:10.1136/bmj.h2219.

30 Polyak CS, Yuhas K, Singa B, et al. Cotrimoxazole Prophylaxis Discontinuation among Antiretroviral-Treated HIV-1-Infected Adults in Kenya: A Randomized Non-inferiority Trial. *PLoS Med* 2016;13:e1001934. doi:10.1371/journal.pmed.1001934.

31 Robinson LJ, Wampfler R, Betuela I, et al. Strategies for understanding and reducing the Plasmodium vivax and Plasmodium ovale hypnozoite reservoir in Papua New Guinean children: a randomised placebo-controlled trial and mathematical model. *PLoS Med* 2015;12:e1001891. doi:10.1371/journal. pmed.1001891.

32 Selak V, Elley CR, Bullen C, et al. Effect of fixed dose combination treatment on adherence and risk factor control among patients at high risk of cardiovascular disease: randomised controlled trial in primary care. *BMJ* 2014;348:g3318. doi:10.1136/bmj.g3318.

33 Hrynaszkiewicz I, Norton ML, Vickers AJ, Altman DG. Preparing raw clinical data for publication: guidance for journal editors, authors, and peer reviewers. *BMJ* 2010;340:c181. doi:10.1136/bmj.c181

34 Iqbal SA, Wallach JD, Khoury MJ, Schully SD, Ioannidis JP. Reproducible Research Practices and Transparency across the Biomedical Literature. *PLoS Biol* 2016;14:e1002333. doi:10.1371/journal.pbio.1002333.

35 Mailman MD, Feolo M, Jin Y, et al. The NCBI dbGaP database of genotypes and phenotypes. *Nat Genet* 2007;39:1181-6. doi:10.1038/ng1007-1181.

36 Keller MB, Ryan ND, Strober M, et al. Efficacy of paroxetine in the treatment of adolescent major depression: a randomized, controlled trial. *J Am Acad Child Adolesc Psychiatry* 2001;40:762-72. doi:10.1097/00004583-200107000-00010.

37 Le Noury J, Nardo JM, Healy D, et al. Restoring Study 329: efficacy and harms of paroxetine and imipramine in treatment of major depression in adolescence. *BMJ* 2015;351:h4320. doi:10.1136/bmj.h4320

38 Laine C, Goodman SN, Griswold ME, Sox HC. Reproducible research: moving toward research the public can really trust. *Ann Intern Med* 2007;146:450-3. doi:10.7326/0003-4819-146-6-200703200-00154

39 Chan AW, Altman DG. Epidemiology and reporting of randomised trials published in PubMed journals. *Lancet* 2005;365:1159-62. doi:10.1016/S0140-6736(05)71879-1.

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

41 Savage CJ, Vickers AJ. Empirical study of data sharing by authors publishing in PLoS journals. *PLoS One* 2009;4:e7078. doi:10.1371/journal.pone.0007078.

42 Early Experiences With Journal Data Sharing Policies. A Survey of Published Clinical Trial Investigators. International Congress on Peer Review and Scientific Publication; 2017; Chicago.

43 Ebrahim S, Sohani ZN, Montoya L, et al. Reanalyses of randomized clinical trial data. *JAMA* 2014;312:1024-32. doi:10.1001/jama.2014.9646.

44 Goodman SN, Fanelli D, Ioannidis JP. What does research reproducibility mean? *Sci Transl Med* 2016;8:341ps12. doi:10.1126/scitranslmed.aaf5027.

45 Open Science Collaboration. PSYCHOLOGY. Estimating the reproducibility of psychological science. *Science* 2015;349:aac4716. doi:10.1126/science.aac4716.

46 Schulz KF, Altman DG, Moher D, CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMJ* 2010;340:c332. doi:10.1136/bmj.c332

47 Stodden V, McNutt M, Bailey DH, et al. Enhancing reproducibility for computational methods. *Science* 2016;354:1240-1. doi:10.1126/science.aah6168.

48 Lewandowsky S, Bishop D. Research integrity: Don't let transparency damage science. *Nature* 2016;529:459-61. doi:10.1038/529459a

49 Sydes MR, Johnson AL, Meredith SK, Rauchenberger M, South A, Parmar MK. Sharing data from clinical trials: the rationale for a controlled access approach. *Trials* 2015;16:104. doi:10.1186/s13063-015-0604-6.

50 Tudur Smith C, Hopkins C, Sydes MR, et al. How should individual participant data (IPD) from publicly funded clinical trials be shared? *BMC Med* 2015;13:298. doi:10.1186/s12916-015-0532-z.

51 Hopkins C, Sydes M, Murray G, 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. doi:10.1016/j.jclinepi.2015.07.002.

52 Navar AM, Pencina MJ, Rymer JA, Louzao DM, Peterson ED. Use of Open Access Platforms for Clinical Trial Data. *JAMA* 2016;315:1283-4. doi:10.1001/jama.2016.2374.

Supplementary information: e-Table 1: results of reanalyses for three ineligible studies published in The BMJ after the policy but submitted before