Making Protocols Available with the Article Improved Evaluation of Selective Outcome Reporting

Louis Calméjane, Agnès Dechartres, Viet Thi Tran, Philippe Ravaud

Making protocols available with the article improved evaluation of selective outcome reporting

Louis Calméjane, M.S., M.Sc.[1-3], Agnès Dechartres, M.D., Ph.D.[1-5], Viet Thi Tran[2-3], M.D, Ph.D. Philippe Ravaud, M.D., Ph.D. [1-4,6]

1 Faculté de Médecine, Université Paris Descartes, Sorbonne Paris Cité, Paris, France.

2 Centre d'Épidémiologie Clinique, Hôpital Hôtel Dieu, Assistance Publique–Hôpitaux de Paris (AP-HP), Paris, France

3 INSERM, U1153, Paris, France

4 Cochrane France, Paris, France

5 Sorbonne Université, INSERM U1136, Institut Pierre Louis d'Epidémiologie et de Santé Publique, Département de Biostatistique, Santé publique, Information Médicale, Hôpital Pitié Salpêtrière, AP-HP, Paris, France

6 Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, USA

Correspondence to: Agnès Dechartres, M.D., Ph.D.

Département Biostatistique, santé publique, information médicale -

Hôpital Pitié Salpêtrière,

47/83 boulevard de l'Hôpital

75013 Paris, France

Tel:+33 (0)1.42.16.05.99

E-mail:agnes.dechartres@aphp.fr

ABSTRACT (199 words)

Objective: To compare primary outcomes reported in publications, protocols and registries; and to evaluate the contribution of available protocols to assess selective outcome reporting (SOR) as compared with registration alone.

Study design and setting: We included all RCTs published in 2015 and 2016 in the five leading general medical journals. For each RCT, we evaluated whether the protocol was available and searched for registration. We extracted all primary outcomes reported in publications, registries and protocols. We evaluated whether SOR was suspected (ie, at least one discrepancy in primary outcomes), unclear or not suspected based on comparisons of publications and 1) trial registration alone or 2) protocols in addition to registration.

Results: SOR was suspected for 77/274 (28.1%), unclear for 30 (10.9%) and not suspected for 167 (60.9%) when comparing publications and trial registration alone. With protocols available, the classification changed for 38 RCTs (13.9%): 11 not suspected of SOR based on registration became suspected of SOR with protocols available, and 27 with unclear assessment based on registration became suspected of SOR (n=7) and not suspected of SOR (n=20) with protocols available.

Conclusions: Compared to registration alone, making protocols available allows for a more precise evaluation of selective outcome reporting.

Key-words: randomized controlled trial, selective outcome reporting, registries, protocols, publishing/standards, outcome, bias.

Running title: Availability of protocols for assessing selective outcome reporting

Word count: 2996

What is new?

Key findings

- Almost six in ten RCTs published in 2015 and 2016 in the leading general medical journals have a protocol available.
- Relying on only trial registration, selective outcome reporting was suspected for 77/274 RCTs (28.1%) and was unclear for 30 (10.9%) because of insufficient description of the primary outcomes in the registry entry.
- With protocols available, there were only 3 RCTs (1.1%) for which the risk of selective outcome reporting could not be assessed. In addition, we suspected additional cases of selective outcome reporting that were not identified with trial registration alone.

What this adds to what is known

- This study involves an original approach to evaluate the contribution of protocols to assessing selective outcome reporting as compared to trial registration alone in a large sample of recently published RCTs.

What is the implication, what should change now

- Making protocols available along with the article may improve evaluation of selective outcome reporting.

INTRODUCTION

Although randomized controlled trials (RCTs) are considered to have one of the highest levels of evidence[1], about 30% of them may be affected by undisclosed discrepancies between the initially planned outcomes and those reported in the final publications[2, 3]. These discrepancies can occur in various forms, including omission (non-reporting of outcomes), commission (changing definitions or measurements of outcomes) or over-reporting (reporting unplanned outcomes)[4, 5]. Such practices referred as selective outcome reporting, tend to favor positive findings[6] which could distort the body of evidence available to clinicians and patients.

To help reduce selective outcome reporting, the International Committee of Medical Journal Editors (ICMJE) required in 2005 all trials to be registered before the recruitment of the first patient on selected open-access trial registries as a condition for publication[7]. Overall, the number of trials registered increased after this statement[8], but several issues remain. Depending on the medical area, only 45% to 70% of trials are registered[9-11] and many are registered retrospectively, including after study completion[9, 12]. In addition, the quality of registration has been questioned, with a lack of precision in registry entries when reporting outcomes[5].

For these reasons, there is an increasing pressure to make protocols available[13, 14]. Since 2015, the Institute of Medicine[15] encourage authors to share clinical trial data, including initial, modified and final protocols, to increase transparency. Additionally, some general journals have recently started to require protocols of RCTs to be made available along with the article.

Due to these recent changes in journal policies, we aimed to 1) evaluate how many reports of RCTs published in the five leading general medical journals have their protocol available; 2) for these trials, compare primary outcomes reported in trial publications, registries and protocols; and

3) evaluate the contribution of available protocols to assess selective outcome reporting as compared with trial registration alone.

MATERIAL AND METHODS

Search for any requirement to make RCT protocols available in the five leading general medical journals

In February 2017, we systematically examined the "Instructions for authors" on the websites of the five leading general medical journals to assess whether there was any requirement regarding availability of RCT protocols and when this was implemented.

Search and selection of trials

We searched MEDLINE via PubMed for all RCTs published in 2015 and 2016 in these five journals by using the Cochrane Highly Sensitive Search Strategy for RCTs[16]. We manually screened all citations retrieved by the search and selected phase III or IV RCTs. We excluded pilot studies, phase I/II trials as well as commentaries, non-randomized studies, duplicate reports, follow-up studies, articles reporting results of several RCTs, factorial studies, and medico-economic studies. Reports were selected by a single reviewer (LC) with the help of a senior reviewer (AD) for any doubtful cases.

Evaluation of availability of protocols

For each RCT, we systematically evaluated whether a protocol was available along with the article or not. We considered that a protocol was available when it was provided as a supplementary appendix or via a functional Internet link in the publication. All trials without a protocol were further excluded.

Search for registration

For each RCT, we systematically searched for a registration number in the publication. At this step, we excluded trials registered after the primary completion date reported in the registry because selective outcome reporting cannot be assessed in this case. When the terms "currently recruiting" or "ongoing" were found in the registry, we looked for the primary completion date and included RCTs for which the primary completion date was before the publication date[11]. This was done to distinguish the truly ongoing trials from those where the authors simply forgot to update the registry.

Extraction of outcomes and general characteristics

For each RCT, we collected data from publications, protocols and registries by using three separate data extraction forms. To independently collect data from each source, we first collected all relevant information from the publications including appendices for all RCTs, then from the protocols and finally from registries.

Data collected from publications

- Outcomes: We recorded all primary and secondary outcomes reported in the methods or results sections or the abstract of the reports. If no primary outcome was clearly reported, we collected the outcomes used in sample size calculation. When sample size calculation was absent, we considered any primary objectives or analyses reported in the publication. If no primary outcomes were found at the end of this process, we excluded the article. We also systematically checked whether the authors gave reasons for any discrepancies in outcomes.

- General characteristics: We recorded the journal name and the date of publication. We also extracted the medical area, number of randomized patients, funding source, types of interventions (i.e., pharmacological, non-pharmacological, mixed) and controls (i.e., active treatment, placebo or sham or no treatment), recruitment period and the design of the trial (i.e., superiority, non-inferiority, equivalence).

Data collected from protocols

We collected the date of the available versions of the protocol. We then recorded all primary and secondary outcomes from the first available version of the protocol reporting primary outcomes. We used the same strategy used for the publications to identify primary outcomes.

Data collected from registries

We screened the registries to find the first version of the registration entry in which primary outcomes were reported. We recorded the date of the first registration of a primary outcome and all primary and secondary outcomes in this version. When the article provided several registration numbers, we extracted outcomes from all registration entries available.

Comparison of primary outcomes between publications, registries and protocols

Once all data had been extracted, we evaluated discrepancies in primary outcomes between the publication and

1) the first registry entry with primary outcomes reported. In a sensitivity analysis, we included only trials for which the registration of the primary outcomes was received before or no more than 3 months after the inclusion of the first participant, consistent with the definition proposed by Zarin et al.[17]. When several registration numbers were

reported, we compared each registry entry to the publication and evaluated whether at least one of the registry entry had a discrepancy.

2) the first available version of the protocol reporting primary outcomes. In a sensitivity analysis, we included only trials for which this version of the protocol preceded the inclusion of the first participant in the trial.

We classified each discrepancy as follows relying on a classification published in a Cochrane review by Dwan et al.[18]:

- Omission of a primary outcome in the publication
- Introduction of a new primary outcome in the publication
- Combined omission and introduction of a primary outcome in the publication (i.e., two completely different primary outcomes in both sources)
- Primary outcome in the protocol or registry reported as a secondary outcome in the publication
- Secondary outcome in the protocol or registry reported as a primary outcome in the publication
- Measurement methods (e.g., use of a subscale) or timeframes were changed. This also included changes in the components of a composite primary outcome.
- Other types of discrepancies such as defining a primary safety outcome in the protocol or registry but not reporting it in the publication as well as discrepancies not fitting any previous definitions

We considered that we could not compare primary outcomes when they were insufficiently described in the registry or protocol. They were not considered as discrepancies but were categorized as "insufficient description not allowing comparison".

We first pilot-tested 15% of trials identified at random in duplicate (LC, AD), then all three sources — publications, protocols and registries — were assessed by a single reviewer (LC) referring to a senior researcher (AD) for any doubtful case.

Contribution of protocols to assessing selective outcome reporting versus registration alone

For each trial, we evaluated whether selective outcome reporting was suspected (ie, at least one discrepancy in primary outcomes as defined above), unclear (ie, insufficient description of primary outcomes) or not suspected (ie, all primary outcomes were clearly described and matched). We first relied on trial registration only, which can be considered the usual approach and second, we considered protocols in addition to trial registration to evaluate the additional value of protocols.

Statistical analysis

We described qualitative variables with frequencies and percentages. Quantitative variables are reported as medians and interquartile range (IQR). All statistical analyses involved use of R v3.3.2 (http://www.R-project.org, the R Foundation for Statistical Computing, Vienna, Austria).

RESULTS

Evaluation of the requirement to make protocols available

From our screening of "Instructions to authors", three journals (*NEJM*, *JAMA* and *Annals of Internal Medicine*) clearly required the protocol to be available along with the article. This was implemented in April 2016 for *Annals of Internal Medicine*, as reported on its website and around January 2010 for the *NEJM* and September 2014 for *JAMA* according to a PubMed search. We did not retrieve any clear policy for the two remaining journals (*The Lancet* and *The BMJ*). Both journals encourage authors to make protocols available but did not require it as a condition for publication.

Trial selection and availability of protocols

Among 3821 references retrieved for 2015 and 2016 in the five leading journals, we identified 488 eligible RCTs; 279 had a protocol available (57.1%) (**Figure 1)**. A protocol was available for all trials published in the *NEJM* (n=166) and for 77/81 trials (95.1%) in *JAMA*. We retrieved few protocols for *The BMJ* (5 protocols for 37 trials [13.5%]), *The Lancet* (25 for 173 trials [14.5%]) and *Annals of Internal Medicine* (6 for 31 trials [19.4%], including 4 of 9 trials published after April 2016). Most protocols were available as a supplementary appendix (n=265 [95.3%]). All trials had at least one registration number reported in the publication.

We then excluded 2 ongoing trials and 3 trials registered after the primary completion date, for a final sample of 274 trials. Their characteristics are available in **Table 1**. Briefly, 142 (51.8%) were published in 2015 and 132 (48.2%) in 2016. The trials enrolled a median of 630 patients (IQR [318-1839]). Industry funding including mixed funding was reported in 144 articles (52.6%). A primary outcome was clearly reported as such in 256 publications (93.4%). For 157

trials (57.3%), the first version of the protocol available preceded the inclusion of the first patient and for 214 (78.1%), the primary completion date.

Comparison of publications and trial registration

A total of 77 trials (28.1%) had at least one discrepancy in primary outcomes between the trial registration and the publication and for 7 only, these modifications were justified in the publication. Among these 77 trials, 23 reports had two or more reasons for differences in primary outcomes, for a total of 100 discrepant outcomes. Most common discrepancies consisted of reporting different timeframes or methods of assessment (37 [13.5%]) and primary outcomes from the registries described as secondary in the publication (17 [6.2%]). In addition, five primary outcomes (1.8%) never registered were introduced in the final papers; an example is "achieved maternal 25-hydroxyvitamin D level of 30/ng/ml or higher at the third trimester sampling" (**Appendix 1)** in a trial for preventing asthma in babies[19].

In a sensitivity analysis considering only trials for which the registration of the primary outcomes preceded or occurred less than 3 months after the inclusion of the first patient, the proportion of discrepancies between registry and publication was 30.4% (69/227).

Comparison of publications and protocols

In total, 76 trials (27.7%) had at least one discrepancy in primary outcomes between the protocol and publication; 16 had two or more discrepancies, for a total of 93 discrepant outcomes (**Table 2)**. Only 10 of these 76 trials justified the discrepancies in publication. The most frequent discrepancies were the modification of timing or methodology of assessment (42/274 trials [15.3%]) and the switching of outcomes from primary in the protocol to secondary in the publication (16/274 trials [5.8%]). For example, in the protocol of a trial assessing the impact of

lifestyle-focused text messages on risk factors of coronary heart disease, the primary outcomes were "low-density lipoprotein cholesterol, BMI and systolic blood pressure," but BMI and systolic blood pressure were reported as secondary outcomes in the publication[20]. Other examples are provided in **Appendix 2.**

In a sensitivity analysis considering only trials for which the first version of the available protocol reporting primary outcomes preceded the inclusion of the first patient, the proportion of discrepancies between the protocol and publication was 28.0% (44/157).

Contribution of protocols to assessing selective outcome reporting versus trial registration alone (Figure 2)

Relying on trial registration alone, 77 RCTs (28.1%) were suspected of selective outcome reporting because of at least one discrepancy in primary outcomes as compared with the publication. Selective outcome reporting was not suspected for 167 trials (10.9%) and unclear for 30 (10.9%). Regarding these 30 trials, 23 (76.7%) were registered at ClinicalTrials.gov and the median year of registration was 2010 (minimum: 1999, maximum: 2014), with 2 trials registered before 2005. With protocols available, the assessment of selective outcome reporting changed for 38 RCTs (13.9%): 11 with no suspicion of selective outcome reporting based on trial registration became suspected of selective outcome reporting, and 27 with unclear assessment based on trial registration became suspected (n= 7) and not suspected (n= 20) of selective outcome reporting. Therefore, for the 30 RCTs for which the situation was unclear with trial registration alone, having the protocol available clarified the situation for 27 (90%). In the end, with both registration and protocols available, 95 RCTs were suspected of selective outcome reporting (33.6%), and selective outcome reporting remained unclear for 3 (1.1%).

DISCUSSION

In this study, we used an original approach to evaluate the contribution of protocols to assessing selective outcome reporting as compared to trial registration alone in a large sample of RCTs recently published in the five leading general medical journals. Relying on trial registration only, selective outcome reporting was suspected for 77 RCTs (28.1%) and was unclear for 30 (10.9%) because of insufficient description of primary outcomes in the trial registration. With protocols available, there were only three RCTs (1.1%) for which the risk of selective outcome reporting could not be assessed, and we suspected additional cases of selective outcome reporting as compared with relying on trial registration alone.

Despite the ICMJE statement in 2005 and initiatives such as the CONSORT statement[21], the proportion of trials affected by selective outcome reporting seems to remain high even in the five leading general medical journals. Our results are close to those reported in previous studies on the same topic. In three studies published between 2004 and 2008 comparing protocols and trials, 30% to 60% of study reports were affected by discrepancies in primary outcomes[22-24]. A study published in 2009 comparing primary outcomes in registries and publications reported 31% of discrepancies[11], and a recent systematic review found the same proportion of distortions[2]. Despite the recommendation of the CONSORT statement to report any changes to trial outcomes with reasons (item 6b)[21], many articles neither stated nor justified these discrepancies. Such little improvement over time even in the five leading medical journals highlights that selective outcome reporting may be a "deep-rooted cultural problem" in research[25].

Reducing the proportion of selective outcome reporting was one of the main goals of mandatory registration[7]. Nevertheless, we found that for 10.9% of trials, we could not assess the risk of selective outcome reporting based on trial registration alone because of insufficient description of the primary outcomes in the registry. In that situation, having access to the protocol was particularly useful, with 90% of trials showing imprecision in registries benefiting from available

protocols to evaluate the risk of selective outcome reporting. In addition, we identified some cases for which selective outcome reporting was not suspected based on trial registration alone but was suspected with available protocols. These results outline the importance of making protocols more available.

The availability of protocols was not systematic in the five general medical journals, as highlighted by the low number of protocols retrieved during our search of some journals. This low availability may be explained by relatively vague language in the "Instructions for authors" such as "accepted articles *should* include a link to the full study protocols". More precise instructions for authors and systematic checking of the availability of protocols by editorial staff may help increase the availability of protocols.

Nevertheless, making protocols available is not sufficient. The interest is to systematically check the consistency in primary outcomes with the publication and to ask authors for explanations in case of discrepancies. To improve and facilitate this assessment, during the peer-review process, editors could require authors to provide at least an original and final version of the protocol. It would also be helpful if the protocols were well-standardized with a clear description of the primary outcomes as recommended by the SPIRIT Statement[26, 27]. Systematic checking of the protocols including consistency in outcomes takes substantial time and resources. Large journals can devote some editorial staff to do this, but smaller journals may not[28]. One solution could be asking junior peer-reviewers to check protocols for consistency with the submitted manuscript, especially when registry entries are imprecise[29].

Another possible solution would be to directly involve the research community in checking primary outcomes and to confront authors and sponsors with undisclosed discrepancies, as highlighted by the COMPARE trials project. This collaborative project aims at detecting and

publicly reporting discrepancies in published reports, which induced some journals such as the *BMJ* to publish letters about outcome switching and issuing corrections[25].

Our study has some limitations. First, we focused on RCTs published in general medical journals with a high impact factor, which may be better reported than in other journals[30, 31]. This choice could lead to an underestimation of the proportion of discrepancies. We included trials only when the protocols were available as Internet links or supplementary files and did not evaluate published protocols. We evaluated the first version of the protocol or registry available so as to evaluate transparency in modifications to the primary outcome because amendments in the primary outcomes should be clearly reported and justified in the publication as recommended by the CONSORT Statement. We found a justification in publications corresponding to an amendment for 3 discrepancies between registries and publications and 6 discrepancies between protocols and publications. Evaluation of discrepancies may be subjective and it was evaluated in duplicate for only 15% of trials. Nevertheless, all doubtful cases were reviewed by a senior researcher.

In conclusion, we identified that about 30% of RCTs published in leading medical journals had discrepancies in primary outcomes between the published reports and trial registration or available protocols. As compared with trial registration alone, protocols allow for a more precise evaluation of selective outcome reporting.

REFERENCES

1. **The Oxford 2011 Levels of Evidence** [http://www.cebm.net/wp-content/uploads/2014/06/CEBM-Levels-of-Evidence-2.1.pdf]
2. Jones CW, Keil LG, Holland WC, Caughey MC, Platts-Mills TF: **Comparison of registered and published outcomes in randomized controlled trials: a systematic review**. *BMC Medicine* 2015, **13**(1):282.
3. Kirkham JJ, Dwan KM, Altman DG, Gamble C, Dodd S, Smyth R, Williamson PR: **The impact of outcome reporting bias in randomised controlled trials on a cohort of systematic reviews**. *BMJ* 2010, **340**:c365.
4. Raghav KP, Mahajan S, Yao JC, Hobbs BP, Berry DA, Pentz RD, Tam A, Hong WK, Ellis LM, Abbruzzese J *et al*: **From Protocols to Publications: A Study in Selective Reporting of Outcomes in Randomized Trials in Oncology**. *J Clin Oncol* 2015, **33**(31):3583-3590.
5. Zarin DA, Tse T, Williams RJ, Califf RM, Ide NC: **The ClinicalTrials.gov results database--update and key issues**. *N Engl J Med* 2011, **364**(9):852-860.
6. Dwan K, Gamble C, Williamson PR, Kirkham JJ: **Systematic review of the empirical evidence of study publication bias and outcome reporting bias - an updated review**. *PLoS One* 2013, **8**(7):e66844.
7. De Angelis C, Drazen JM, Frizelle FA, Haug C, Hoey J, Horton R, Kotzin S, Laine C, Marusic A, Overbeke AJ *et al*: **Clinical trial registration: a statement from the International Committee of Medical Journal Editors**. *Lancet* 2004, **364**(9438):911-912.
8. Zarin DA, Tse T, Ide NC: **Trial Registration at ClinicalTrials.gov between May and October 2005**. *New England Journal of Medicine* 2005, **353**(26):2779-2787.
9. Dechartres A, Ravaud P, Atal I, Riveros C, Boutron I: **Association between trial registration and treatment effect estimates: a meta-epidemiological study**. *BMC Med* 2016, **14**(1):100.
10. Farquhar CM, Showell MG, Showell EAE, Beetham P, Baak N, Mourad S, Jordan VMB: **Clinical trial registration was not an indicator for low risk of bias**. *J Clin Epidemiol* 2017, **84**:47-53.
11. Mathieu S, Boutron I, Moher D, Altman DG, Ravaud P: **Comparison of registered and published primary outcomes in randomized controlled trials**. *JAMA* 2009, **302**(9):977-984.
12. Harriman SL, Patel J: **When are clinical trials registered? An analysis of prospective versus retrospective registration**. *Trials* 2016, **17**:187.
13. Chan AW: **Bias, spin, and misreporting: time for full access to trial protocols and results**. *PLoS Med* 2008, **5**(11):e230.
14. Glasziou P, Altman DG, Bossuyt P, Boutron I, Clarke M, Julious S, Michie S, Moher D, Wager E: **Reducing waste from incomplete or unusable reports of biomedical research**. *Lancet* 2014, **383**(9913):267-276.
15. Medicine Io: **Sharing Clinical Trial Data: Maximizing Benefits, Minimizing Risk**. Washington, DC: The National Academies Press; 2015.
16. Higgins JPT, Green S, editors: ***Cochrane Handbook for Systematic Reviews of Interventions* Version 5.1.0 [updated March 2011]**. *The Cochrane Collaboration* 2011, **Available from www.cochrane-handbook.org**
17. Zarin DA, Tse T, Williams RJ, Rajakannan T: **Update on Trial Registration 11 Years after the ICMJE Policy Was Established**. *N Engl J Med* 2017, **376**(4):383-391.
18. Dwan K, Altman DG, Cresswell L, Blundell M, Gamble CL, Williamson PR: **Comparison of protocols and registry entries to published reports for randomised controlled trials**. *Cochrane Database Syst Rev* 2011(1):MR000031.

19. Litonjua AA, Carey VJ, Laranjo N, et al.: **Effect of prenatal supplementation with vitamin d on asthma or recurrent wheezing in offspring by age 3 years: The vdaart randomized clinical trial**. *JAMA* 2016, **315**(4):362-370.
20. Chow CK, Redfern J, Hillis GS, Thakkar J, Santo K, Hackett ML, Jan S, Graves N, de Keizer L, Barry T *et al*: **Effect of Lifestyle-Focused Text Messaging on Risk Factor Modification in Patients With Coronary Heart Disease: A Randomized Clinical Trial**. *JAMA* 2015, **314**(12):1255-1263.
21. Schulz KF, Altman DG, Moher D: **CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials**. *BMJ* 2010, **340**.
22. Al-Marzouki S, Roberts I, Evans S, Marshall T: **Selective reporting in clinical trials: analysis of trial protocols accepted by The Lancet**. *Lancet* 2008, **372**(9634):201.
23. Chan AW, Hrobjartsson A, Haahr MT, Gotzsche PC, Altman DG: **Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles**. *JAMA* 2004, **291**(20):2457-2465.
24. Chan AW, Krleza-Jeric K, Schmid I, Altman DG: **Outcome reporting bias in randomized trials funded by the Canadian Institutes of Health Research**. *CMAJ* 2004, **171**(7):735-740.
25. **The COMPare Trials Project** [www.COMPare-trials.org]
26. Chan A, Tetzlaff JM, Altman DG, et al.: **Spirit 2013 statement: Defining standard protocol items for clinical trials**. *Annals of Internal Medicine* 2013, **158**(3):200-207.
27. Ioannidis JPA, Greenland S, Hlatky MA, Khoury MJ, Macleod MR, Moher D, Schulz KF, Tibshirani R: **Increasing value and reducing waste in research design, conduct, and analysis**. *The Lancet*, **383**(9912):166-175.
28. Weber WE, Merino JG, Loder E: **Trial registration 10 years on**. *Bmj* 2015, **351**:h3572.
29. Moher D: **Reporting research results: A moral obligation for all researchers**. *Canadian Journal of Anesthesia* 2007, **54**(5):331.
30. Bala MM, Akl EA, Sun X, Bassler D, Mertz D, Mejza F, Vandvik PO, Malaga G, Johnston BC, Dahm P *et al*: **Randomized trials published in higher vs. lower impact journals differ in design, conduct, and analysis**. *Journal of Clinical Epidemiology* 2013, **66**(3):286-295.
31. Dechartres A, Trinquart L, Atal I, Moher D, Dickersin K, Boutron I, Perrodeau E, Altman DG, Ravaud P: **Evolution of poor reporting and inadequate methods over time in 20 920 randomised controlled trials included in Cochrane reviews: research on research study**. *Bmj* 2017, **357**:j2490.

COMPETING INTERESTS:

The authors declare that they have no competing interests in relation with this study.

FUNDING:

This study received no particular funding.

AUTHORS' CONTRIBUTIONS:

Louis Calméjane was involved in the study conception, selection of trials, data extraction, data analysis, interpretation of results and drafting the manuscript;

Agnès Dechartres was involved in the study conception, selection of trials, data extraction, data analysis, interpretation of results and drafting the manuscript;

Viet Thi Tran was involved in the study conception, interpretation of results and drafting the manuscript;

Philippe Ravaud was involved in the study conception, and drafting the manuscript.

ACNOWLEGMENTS:

We thank Elise Diard for help with the figure.

List of tables

List of figures

List of appendices

Table 1. Characteristics of included randomized controlled trials (N=274)

Characteristics	Number (%)
Journal	
NEJM	163 (59.5)
JAMA	76 (27.7)
The Lancet	24 (8.8)
Annals of Internal Medicine	6 (2.2)
The BMJ	5 (1.8)
Publication year	
2015	142 (51.8)
2016	132 (48.2)
Medical specialty	
Cardiovascular	48 (17.5)
Oncology	46 (16.8)
Infectious diseases	23 (8.4)
Rheumatology	21 (7.7)
Critical care	20 (7.3)
Pneumology	20 (7.3)
Neurology	16 (5.8)
Obstetrics and gynecology	16 (5.8)
Endocrinology	13 (4.7)
Pediatrics	11 (4)
Other specialties[a]	40 (14.7)
Number of patients randomized – median (IQR)	630 (318–1839)
Type of intervention	
Pharmacological	153 (55.8)
Non-pharmacological	87 (31.8)
Mixed	28 (10.2)
Others: screening, diagnosis	6 (2.2)
Type of control	
Active treatment (pharmacological, non-pharmacological, mixed)	181 (66)
Placebo and sham	78 (28.5)
Others: no treatment	15 (5.5)

Study design	
Superiority	204 (74.4)
Non-inferiority	39 (14.2)
Equivalence	4 (1.5)
Unclear	27 (9.9)
Funding source	
Public	129 (47)
Private	81 (29.6)
Mixed	63 (23)
Not reported	1 (0.4)
Primary outcome clearly stated as such in methods section	
Yes	256 (93.4)
No, primary objectives, hypotheses or analyses	8 (2.9)
No, use of sample size calculation	7 (2.6)
No, use of abstract	2 (0.7)
No, use of "main" outcome	1 (0.4)
Form of availability of the protocol	
Supplementary appendix	261 (95.3)
Link in the publication	13 (4.7)
Registry	
ClinicalTrials.gov	216 (78.8)
International Standard Randomized Controlled Trial Number registry	28 (10.2)
Others (National registries, Australia and New Zealand registry, EudraCT)	15 (5.5)
Double registration	13 (4.8)
Triple registration	2 (0.7)

[a] Other specialties included primary care, gastroenterology and hepatology, geriatrics, nephrology and urology, psychiatry, immunology, hematology, anesthesiology and ophthalmology

NEJM, New England Journal of Medicine; JAMA, Journal of the American Medical Association; The BMJ.

Table 2. Comparison of primary outcomes between publications and trial registration or protocol

Total number of trials: N=274	**Publication vs registration n (%)**	**Publication vs protocol n (%)**
Number of articles with at least one discrepancy in primary outcomes	77 (28.1)[a]	76 (27.7)[b]
Omission of a primary outcome	6 (2.2)	3 (1.1)
New primary outcome introduced in publication	5 (1.8)	3 (1.1)
Completely different primary outcomes	3 (1.1)	1 (0.4)
Primary outcome in the protocol or registry described as secondary in the publication	17 (6.2)	16 (5.8)
Secondary outcome in the protocol or registry described as primary in the publication	10 (3.6)	7 (2.6)
Different timing or method of assessment of primary outcome	37 (13.5)	42 (15.3)
Others, including Modification of safety outcomes	22 (8) 15 (5.5)	21 (7.7) 15 (5.5)
Number of articles with insufficient description in the protocol or registration not allowing comparison	30 (10.9)	14 (5.1)
Number of articles with no discrepancy	167 (60.9)	184 (67.2)

[a] 23 articles had 2 or more reasons for differences in primary outcome [b]16 articles had 2 or more reasons for differences in primary outcome..

Supplementary appendix 1: Examples of discrepancies in primary outcomes between trial registration and publications

We only present the necessary information to understand the discrepancies between outcomes, so outcome descriptions may have been shortened.

Author	Discrepancies	Outcome(s) reported in the registry	Outcome(s) reported in the publication
French (Lancet, 2016)	• **Omission of a primary outcome:** In this configuration, we decided to code this discrepancy as an omission of the primary outcomes in the publication. Our decision is based on the fact that each outcome in the registry had its own timeframe. Therefore, in the registry there were four separate primary outcomes. In the publication, only those with the 12-week timeframes remain, so two primary outcomes were omitted.	• **Primary outcomes:** **#1:** European Medicine Agency (EMA): Response rate [Week 6] **#2:** European Medicine Agency (EMA): Response rate [Week 12] **#3:** Food & Drug Administration (FDA): Percentage reduction in partial onset seizure frequency [Week 6] **#4:** Food & Drug Administration (FDA): Percentage reduction in partial onset seizure frequency [Week 12]	• **Primary efficacy endpoint** was changed from baseline in seizure frequency for each of the two everolimus Cmin ranges compared with placebo during the 12-week maintenance period of the core phase, expressed as response rate (reduction in seizure frequency) and median percentage reduction in seizure frequency.
Litonjua (JAMA, 2016)	• **Introduction of a new primary outcome:** The second primary outcome was never reported in the protocol.	• **Primary outcome:** Asthma or recurrent wheeze in the child. [Time Frame: 1 year and 3 years]	• **Primary outcomes:** **#1:** Parental report of physician diagnosis of asthma or occurrence of recurrent wheeze in the child's first 3 years of life ascertained from questionnaires administered every 3 months.

			#2: Achieved maternal 25-hydroxyvitamin D level of 30 ng/mL or higher at the third trimester sampling
Chow (JAMA, 2015)	• **Primary outcome in the registry reported as secondary in the publication:** Two primary outcomes in the first source were reported as secondary outcomes in the second source (systolic blood pressure and BMI). • Furthermore, one of these primary outcomes was split into two secondary outcomes (BMI and waist circumference).	• **Primary outcomes:** **#1:** Low-density lipoprotein cholesterol measured by fasting blood sample. Timepoint [1] Baseline and 6 months. **#2:** Systolic blood pressure Timepoint [2] Baseline and 6 months **#3:** Body mass index and waist circumference Timepoint [3] Baseline and 6 months	• **Primary outcome:** Low-density lipoprotein cholesterol (LDL-C) level at 6 months • **Secondary outcomes:** **#1:** Systolic blood pressure **#2:** BMI **#3:** Waist circumference

Elias (NEJM, 2016)	• **A secondary outcome in the registry was reported as a primary outcome in the publication** • **Measurement methods or timeframe were changed:** However, these two outcomes are not similar in terms of timeframe (3 months vs. approximately up to 12 months) • **Omission of a primary outcome:** The primary outcome used in the registry does not appear in the final publication	• **Primary outcome:** Severity of Device and Procedure related complications [Time Frame: At the time of ExAblate Transcranial thalamotomy procedure] • **Secondary outcome:** Effectiveness of of the ExAblate Transcranial MRgFUS treatment determined using the Clinical Rating Scale for Tremor (CRST) [Time Frame: Participants will be followed from the date of treatment until study completion, approximately up to 12 months]	• Change from baseline to 3 months in the tremor score for the hand derived from the CRST, Part A (three items: resting, postural, and action or intention components of hand tremor), and the CRST, Part B (five tasks involving handwriting, drawing, and pouring).
He (JAMA, 2015)	• **Measurement methods or timeframe were changed:** 6 months vs. 3 years	• Visual Acuity [Time Frame: 6 months]	• 3-year cumulative incidence rate of myopia. Myopia was defined as a spherical equivalent refractive error (sphere +½ cylinder) of at least −0.50 D
Jolly (NEJM, 2015)	• **Others:** We included in the category named "others" outcomes such as safety, pharmacokinetics and imaging outcomes that	• **Primary outcome:** Composite of cardiovascular death, recurrent myocardial infarction,	• **Primary outcomes:** **#1:** Death from cardiovascular causes, recurrent myocardial infarction,

	showed some discrepancies. We had two reasons for this choice: - These outcomes are often reported as "main" or "key" outcomes so it is difficult to assess whether they are primary or secondary outcomes. - To avoid having too many outcomes in their sample sizes, some authors prefer to create subcategories of outcomes so we used a separate category to take this into account. • **Of note, these outcomes were analysed in the category "others" in the main analysis but their numbers are reported separately in Table 2** • In this particular case, we can see that one secondary outcome became a "key safety outcome" and that one "key net-benefit outcome" that was not registered appeared in the publication	cardiogenic shock, or new or worsening New York Heart Association (NYHA) Class IV heart failure at 180 days • **Secondary outcome:** Stroke	cardiogenic shock, or new or worsening NYHA class IV heart failure within 180 days **#2: Key safety outcome** was stroke within 30 days **#3: Key net-benefit outcome** was the occurrence of the primary outcome or stroke within 180 days.
Mistry	• **Insufficient description in the protocol or**	• **Primary outcome (registered in the**	• **Primary outcome:** least-square mean

(JAMA, 2015)	**registry not allowing comparison:** This example illustrates the insufficient description of primary outcomes retrieved in some registry entries. It was not considered as a discrepancy but rather as a separate category	**corresponding section on ClinicalTrials.gov):** The primary objective is to compare the effects of Genz-112638 to placebo in patients with Type 1 Gaucher Disease [Time Frame: 39 weeks]	percentage change in spleen volume by [MRI] from baseline to 9 months

Supplementary appendix 2: Examples of discrepancies in primary outcomes between protocols and publications

We only present the necessary information to understand the discrepancies between outcomes, so outcome descriptions may have been shortened.

Author	Discrepancies	Outcome(s) reported in the protocol	Outcome(s) reported in the publication
Pathan (Lancet, 2016)	• **Omission of a primary outcome:** The protocol for this study had both a primary outcome and a primary endpoint that we considered as two separate primary outcomes. The so-called primary endpoint was omitted in the publication.	• **Primary outcome:** proportion of patients achieving "significant pain reduction". In this trial, we consider significant pain reduction as a drop in the initial pain score of 50% or more at 30 min after analgesia administration. • **Primary endpoint:** Meeting the optimal pain management expectation of the patient: being totally pain free, or improved a lot, no need for further pain medicine or a reduction of >50% on NRS from initial score recorded.	• The proportion of participants achieving significant pain reduction. We considered significant pain reduction as at least a 50% drop in the initial pain score at 30 min after analgesia administration
Kalra (Lancet, 2015)	• **Introduction of a new primary outcome:** The original protocol did not precise the co-primary outcome. Of note, in the results section of the publication, outcomes were reported separately, which convinced us that they	• **Primary outcome:** The clinical primary outcome measure is the incidence of PSP in the first 14 days after stroke onset or prior to discharge home if sooner.	• **Primary outcomes:** **#1:** Post-stroke pneumonia in the first 14 days, assessed with both a criteria-based, hierarchical algorithm and **#2:** Diagnosis of pneumonia made by the local treating physician was also

	were two different primary outcomes.		recorded as a co-primary outcome
Chow (JAMA, 2015)	• **Primary outcome in the protocol reported as secondary in the publication:** Two outcomes originally reported in the protocol as primary became secondary outcomes in the publication. • Of note, we had to use the sample size calculation to assess which outcomes were primary or not in the protocol.	• **Primary outcomes:** **#1**: Low density lipoprotein cholesterol **#2:** BMI **#3:** Systolic blood pressure	• **Primary outcome:** Low-density lipoprotein cholesterol level at 6 months • **Secondary outcomes:** **#1:** Systolic blood pressure **#2:** BMI
Kantarjian (NEJM, 2016)	• **A secondary outcome in the protocol was reported as a primary outcome in the publication**	• **Primary outcome:** Hematological remission defined as Complete response and complete response with incomplete count recovery. **Secondary outcome:** Overall survival	• **Primary outcomes:** **#1:** Complete remission (including complete remission with incomplete hematologic recovery) **#2:** Overall survival
Cherkin (JAMA, 2016)	• **Measurement methods or timeframe were changed:** Change of measurement methods in the sense that the outcomes were reported as continuous or	• Back-related dysfunction using the Modified Roland-Morris Disability Questionnaire (back-related dysfunction).	• Percentages of participants with clinically meaningful (≥30%) improvement from baseline in functional limitations (modified Roland Disability

	categorical in the protocol but the published outcomes have been dichotomized.		Questionnaire [RDQ]; range, 0-23) at 26 weeks
		• Bothersomeness of low back pain (0 to 10 scale).	• Percentages of participants with clinically meaningful (≥30%) improvement from baseline in self-reported back pain bothersomeness (scale, 0-10) at 26 weeks
Hernández (JAMA, 2016)	• **Measurement methods or timeframe were changed:** Possible change of measurement method because the Glasgow Coma Scale (GCS) has a maximum score of 15, so losing more than two points could be not equivalent to having a GCS score ≤ 8	• Reintubation within 72 hr after extubation. Predefined criteria for reintubation and postextubation respiratory failure including among others: Non-respiratory reasons, such as urgent surgery or a GCS ≤8 points not related to hypercapnia.	• Postextubation respiratory failure within 72 hr of extubation was defined as the presence and persistence of several criteria including: Decreased level of consciousness (GCS >1 point decrease)
Moseley (JAMA , 2015)	• **Measurement methods or timeframe were changed:** The protocol specified that the primary outcome was measured at 6 months and not more than 6 months. Since the primary time point in the publication is 3 months, we considered a	• Quality-adjusted life years: Utility will be measured at 6 months by the Assessment of Quality of Life (AQoL) instrument.	• Quality of life assessed using the AQoL at 1, 3 (primary time point), and 6 months.

	discrepancy in terms of timeframe.		
D'Cruz (NEJM, 2015)	• **Others:** We classified this discrepancy in the category "others" because we could not assess which outcome was indeed the primary outcome. Of note, the authors justify changing the outcome: "Overall Survival has always been the primary endpoint of the study as sample size calculations have been based on overall survival assumptions since trial initiation"	• "Loco-regional recurrence" clearly stated as primary outcome **BUT** "overall survival" used as the outcome of interest in the sample size calculation	• Overall survival, which was defined as the interval between the date of randomization and the date of death from any cause
Lee (JAMA, 2015)	• **Others:** Safety endpoints not clearly stated as primary or secondary outcomes but the publication reports principal safety outcomes. • We included in the category named "others" outcomes such as safety, pharmacokinetics and imaging outcomes that showed some discrepancies. We had two reasons for this choice: - These outcomes are often reported as "main" or "key" outcomes so it is	• Safety endpoints will consist of major bleeding and clinically non-major bleeding, overall mortality.	• Principal safety outcomes were major bleeding, clinically relevant non-major bleeding, and all-cause mortality.

	difficult to assess whether they are primary or secondary outcomes. - To avoid having too many outcomes in their sample sizes, some authors prefer to create subcategories of outcomes, so we used a separate category to take this into account.	