Commit ID: a51decd

2023-10-24

# Reporting quality has improved over time but remains sub optimal

## Evidence from observational studies

Many studies have compared quality of reporting before and after reporting guidelines were published and/or journals began asking authors to use them. For example, in 2023 Kilicoglu et al. [1] analysed 176,620 randomized trial reports published between 1966 and 2018. They used software to identify sentences pertaining to CONSORT methodology items. They found reporting quality in clinical trials has improved over time but remains sub optimal: articles published before 1990 reported 24% of CONSORT items, whereas articles published between 2010-2018 reported 48%. The fastest improvement occurred in the years immediately after CONSORT was published.

This result mirrors similar, manual, efforts to monitor reporting quality over time. For example, Dechartres et al. [2] looked at CONSORT items related to bias in 20,920 trial reports. They found information on sequence generation was missing from 69% of articles published before CONSORT (1986-1990), but only 31% of articles published after CONSORT (2011-2014). The proportion of articles missing information on allocation concealment fell from 70% to 45% over the same periods.

These two examples focus on clinical trials, but other research types have seen modest improvements too. De Jong et al. [3] reviewed 284 qualitative evidence syntheses. These syntheses appraised whether their primary studies had adhered to COREQ (a reporting guidelines for qualitative research). Studies published before COREQ reported 16 of COREQ’s 32 items, whereas studies published after reported 18. Similarly, when comparing the reporting quality of genetic association studies, Nedovic et al. [4] found reporting quality was better in generic association articles published after STREGA but only in journals that endorsed STREGA (63% vs 49% showing full adherence).

Not all studies have found relationships like these. For example, Howell et al. found no improvement in the reporting of quality improvement studies after the SQUIRE guidelines were published [5]. Similarly, Pouwels et al. found no improvement in observational epidemiology following STROBE’s publication [6] and another study [7] found that reporting quality improved *before* STROBE was published, but not *afterwards*.

Other studies have explored the effect of different journal policies. For example, Hopewell et al. [8] assessed the reporting of clinical trial abstracts before and after the publication of CONSORT for Abstracts and compared journals that a) had no reporting guideline policy b) endorsed the guideline and c) actively enforced guideline adherence. They only found an effect in the active enforcement group, where 5.41 items (out of 9) were reported, which was 50% higher than expected. No improvement was seen in journals that endorsed the guideline without enforcing it.

Some observational studies have focussed on a single time period. In 2018, before my PhD, I collaborated with EQUATOR and BMJ Open to compare manuscript versions before and after authors completed a reporting checklist [9]. Authors made few edits to their work. Of 20 included authors, three added information for a single reporting item, one added information for two items, and one added information for six items. The remaining 15 authors made no changes. On average, manuscripts described 57% of necessary reporting items before the author completed the checklist and 60% afterwards.

In 2018 the senior managing editor of the Journal of the National Cancer Institute asked 2000 submitting authors whether they had used a reporting guideline [10]. She then asked peer reviewers to rate manuscripts for their clarity and adherence to reporting guidelines. Declared guideline use was associated with better adherence to guidelines, but not associated with improved clarity nor acceptance rates.

## Evidence from intervention studies #Most are randomised but not Pandis

Some experimental studies have tried to isolate the effect of journal policies by randomising authors. In 2015 PLOS One randomly allocated 1689 incoming submissions reporting *in vivo* animal research manuscripts to either a) request completion the ARRIVE reporting checklist or b) usual practice [11]. No article achieved full compliance with ARRIVE, and only one sub item (details of animal husbandry) showed improvement between groups.

In another study [12], 197 authors submitting to 46 participating journals were randomly allocated to receive either a) access to an online tool (WebCONSORT) to generate customised reporting checklists and flow diagrams based on CONSORT and its extensions or b) a standard CONSORT flow diagram generator without a reporting checklist. There was no difference in reporting quality between groups: authors only reported half of required items.

In two earlier studies, Cobo et al. explored the roll of peer review. Simply providing peer reviewers with reporting guidelines had no effect [13]. Adding a reviewer whose task was to check for guideline adherence did lead to improved reporting quality [14], but “the observed effect was smaller than hypothesised and not definitively demonstrated”.

It would be unsurprising to find that the stricter the enforcement, the better the adherence. Pandis et al. [15] describe an enforcement strategy used in a small Dentistry journal where the associate editor checked manuscripts for adherence to CONSORT. The associate editor would complete a CONSORT checklist for each manuscript, making note of unclear or unreported items and suggesting ways to improve the manuscript. This would be sent back to the author. Resubmitted manusctripts were subject to the same process, and the manuscript would only be sent out for peer review once the reporting was deemed satisfactory. Over two years, 23 manuscripts were handled in this way. The policy was effective. All studies reported at least 33 of 37 CONSORT items (compared to 15 items before the policy was introduced). However, even with this heavy handed approach, “four items were still unreported in all trials: changes to methods (3b), changes to outcomes (6b), interim analysis (7b), and trial stopping (14b).”.

To summarise, reporting standards may have improved over the last two decades. There is some evidence that reporting guidelines may have contributed to this change, but if they have, their effect has only been modest and the bottom line is that most research still does not include the details these guidelines call for. There is huge room for improvement. In a systematic review of 124 studies assessing adherence to one of eight reporting guidelines, Jin et al [16] found 88% of studies reported suboptimal adherence to reporting guidelines. Similarly, Dal Santo et al [17] reviewed 148 studies of reporting quality from the previous few years and almost all described reporting quality as suboptimal.

## Studies exploring the efficacy of reporting guidelines are actually exploring the efficacy of the reporting guideline *system*

Because reporting guidelines do not exist in a vacuum, it is difficult to separate the guidelines themselves from the policies, people, websites, and tools involved in their implementation. For example, many before-and-after studies use the publication of the guideline as their defining time point. However, it is impossible for these studies to disentangle the effect of guidelines coming into existence with the effect of subsequent journal policies and editorial practices. Similarly, in experimental studies comparing the effect of asking authors to complete a checklist or use a resource, the intervention groups included changes to editorial workflows. These changes were external to the resource being tested, but could be equally important to its success: in the WebCONSORT study, editors’ inability to identify randomised trial reports was an important source of failure external to the tool being tested. Hence, in describing reporting guidelines as being part of a complex behaviour change intervention, I believe I am explicitly articulating a systems perspective that previous studies have hinted at, and I am exploring that system’s scope in more granularity.

1. Kilicoglu H, Jiang L, Hoang L, Mayo-Wilson E, Vinkers CH, Otte WM (2023) Methodology reporting improved over time in 176,469 randomized controlled trials. Journal of Clinical Epidemiology. <https://doi.org/10.1016/j.jclinepi.2023.08.004>

2. Dechartres A, Trinquart L, Atal I, Moher D, Dickersin K, Boutron I, Perrodeau E, Altman DG, Ravaud P (2017) [Evolution of poor reporting and inadequate methods over time in 20 920 randomised controlled trials included in Cochrane reviews: Research on research study](https://doi.org/10.1136/bmj.j2490). BMJ 357:j2490

3. de Jong Y, van der Willik EM, Milders J, Voorend CGN, Morton RL, Dekker FW, Meuleman Y, van Diepen M (2021) [A meta-review demonstrates improved reporting quality of qualitative reviews following the publication of COREQ- and ENTREQ-checklists, regardless of modest uptake](https://doi.org/10.1186/s12874-021-01363-1). BMC Medical Research Methodology 21:184

4. Nedovic D, Panic N, Pastorino R, Ricciardi W, Boccia S (2016) [Evaluation of the Endorsement of the STrengthening the REporting of Genetic Association Studies (STREGA) Statement on the Reporting Quality of Published Genetic Association Studies](https://doi.org/10.2188/jea.JE20150173). Journal of Epidemiology 26:399–404

5. Howell V, Schwartz AE, O’Leary JD, Donnell CM (2015) [The effect of the SQUIRE (Standards of QUality Improvement Reporting Excellence) guidelines on reporting standards in the quality improvement literature: A before-and-after study](https://doi.org/10.1136/bmjqs-2014-003737). BMJ Quality & Safety 24:400–406

6. Pouwels KB, Widyakusuma NN, Groenwold RHH, Hak E (2016) [Quality of reporting of confounding remained suboptimal after the STROBE guideline](https://doi.org/10.1016/j.jclinepi.2015.08.009). Journal of Clinical Epidemiology 69:217–224

7. Bastuji-Garin S, Sbidian E, Gaudy-Marqueste C, Ferrat E, Roujeau J-C, Richard M-A, Canoui-Poitrine F, European Dermatology Network (EDEN) (2013) [Impact of STROBE statement publication on quality of observational study reporting: Interrupted time series versus before-after analysis](https://doi.org/10.1371/journal.pone.0064733). PloS One 8:e64733

8. Hopewell S, Ravaud P, Baron G, Boutron I (2012) [Effect of editors’ implementation of CONSORT guidelines on the reporting of abstracts in high impact medical journals: Interrupted time series analysis](https://doi.org/10.1136/bmj.e4178). BMJ 344:e4178

9. Struthers C, Harwood J, de Beyer JA, Dhiman P, Logullo P, Schlüssel M (2021) [GoodReports: Developing a website to help health researchers find and use reporting guidelines](https://doi.org/10.1186/s12874-021-01402-x). BMC medical research methodology 21:217

10. Botos J (2018) [Reported use of reporting guidelines among JNCI: Journal of the National Cancer Institute authors, editorial outcomes, and reviewer ratings related to adherence to guidelines and clarity of presentation](https://doi.org/10.1186/s41073-018-0052-4). Research Integrity and Peer Review 3:7

11. Hair K, Macleod MR, Sena ES, et al (2019) [A randomised controlled trial of an Intervention to Improve Compliance with the ARRIVE guidelines (IICARus)](https://doi.org/10.1186/s41073-019-0069-3). Research Integrity and Peer Review 4:12

12. Hopewell S, Boutron I, Altman DG, et al (2016) [Impact of a web-based tool (WebCONSORT) to improve the reporting of randomised trials: Results of a randomised controlled trial](https://doi.org/10.1186/s12916-016-0736-x). BMC Medicine 14:199

13. Cobo E, Selva-O’Callagham A, Ribera J-M, Cardellach F, Dominguez R, Vilardell M (2007) [Statistical Reviewers Improve Reporting in Biomedical Articles: A Randomized Trial](https://doi.org/10.1371/journal.pone.0000332). PLoS ONE 2:e332

14. Cobo E, Cortés J, Ribera JM, et al (2011) [Effect of using reporting guidelines during peer review on quality of final manuscripts submitted to a biomedical journal: Masked randomised trial](https://doi.org/10.1136/bmj.d6783). BMJ 343:d6783

15. Pandis N, Shamseer L, Kokich VG, Fleming PS, Moher D (2014) [Active implementation strategy of CONSORT adherence by a dental specialty journal improved randomized clinical trial reporting](https://doi.org/10.1016/j.jclinepi.2014.04.001). Journal of Clinical Epidemiology 67:1044–1048

16. Jin Y, Sanger N, Shams I, et al (2018) [Does the medical literature remain inadequately described despite having reporting guidelines for 21 years? A systematic review of reviews: An update](https://doi.org/10.2147/JMDH.S155103). Journal of Multidisciplinary Healthcare 11:495–510

17. Santo TD, Rice DB, Amiri LSN, Tasleem A, Li K, Boruff JT, Geoffroy M-C, Benedetti A, Thombs BD (2023) [Methods and results of studies on reporting guideline adherence are poorly reported: A meta-research study](https://doi.org/10.1016/j.jclinepi.2023.05.017). Journal of Clinical Epidemiology 159:225–234