Non-existent ClinicalTrials.gov identifiers in abstracts indexed by PubMed

Benjamin Gregory Carlisle PhD¹

¹ QUEST (Quality | Ethics | Open Science | Translation), BIH Center for Transforming Biomedical Research. Anna-Louisa-Karsch-Str. 2, 10178 Berlin. Tel: +49 450 543 694; Email: benjamin.carlisle@bihealth.de.
Abstract

Prospective registration plays an important role in ensuring the transparency and reliability of clinical trials. Preregistration of clinical trials has been required by the ICMJE since 2005 and mandated by law for most clinical trial types since 2007. It is one of the roles of peer reviewers of a clinical trial publication to confirm that there is concordance between the registry entry and the submitted publication. On October 22, 2019, abstracts for all articles indexed by PubMed with publication type “Clinical Trial” and a publication date after January 1, 2003 were downloaded. Clinical trial registry identifiers were automatically extracted and tested for the existence of a corresponding entry on ClinicalTrials.gov. Among 38,001 published clinical trial registry numbers, 215 (0.6%) do not correspond to a legitimate clinical trial registry entry. While there is a small proportion of non-existent NCT numbers in our sample, even a single non-existent NCT number in a publication represents a failure on the part of journals who publish clinical trials to systematically ensure that reviewers always check clinical trial registry entries for concordance with the text submitted for publication. These results cast doubt on how frequently editors and reviewers evaluate clinical trial reports in light of their corresponding registry entries.
Background

Preregistration of clinical trials plays an important part in preventing selective reporting of outcomes, switching of primary and secondary endpoints, and certain kinds of publication bias.\textsuperscript{1,2} Prospective registration also allows for greater transparency in the conduct and reporting of clinical trials,\textsuperscript{3} which can bolster confidence in the proper functioning of the machinery of human research. Prospective clinical trial registration has been required by the ICMJE since 2005\textsuperscript{4} and has been legally mandated in the United States for most clinical trial types since 2007.\textsuperscript{5}

Methods

On October 22, 2019, all articles indexed by PubMed with publication type “Clinical Trial” and a publication date after January 1, 2003 were downloaded. The PubMed ID, date of publication, abstract and journal name were extracted from the PubMed XML metadata. Reported trial registration (“NCT”) numbers were automatically extracted from abstracts and checked for the existence of a corresponding record on ClinicalTrials.gov. NCT numbers that did not exist on ClinicalTrials.gov were verified manually to ensure that they had been extracted correctly from the abstract text and that the registry entry does not exist. The complete data set and code for extraction of NCT numbers and checking for a corresponding record on ClinicalTrials.gov are available online.\textsuperscript{6,7}

Results

Among 488,364 clinical trial publications indexed by PubMed, there were 38,001 published NCT numbers. Of these NCT numbers, 215 (0.6%) do not correspond to a clinical trial registry entry on ClinicalTrials.gov; see Figure 1.

The abstracts containing non-existent NCT numbers were published by 132 distinct journals which published an average of 577 clinical trials in our sample (range 7-7315). These journals reported an average of 134 NCT numbers in our sample (range 1-1,696). For 9 of the 215 non-existent NCT numbers (4%), there was a separate, secondary publication of that clinical trial in our sample that also referred to the same non-existent NCT number.

In a random sample of 22 (10%) of the 215 extracted NCT numbers, a manual search based on trial details (drug, indication, etc.) or other identifiers (e.g. non-NCT trial identifier) in the abstract or the full-text of the publication found that all 22 do have a valid corresponding record, either in ClinicalTrials.gov or another public registry. Seventeen (77%) of the extracted NCT numbers differed from the actual NCT number by only a single digit. In two cases, an identifier for another public registry was incorrectly reported as a ClinicalTrials.gov NCT number. Among the random sample of 22 NCT numbers, 8 of the non-existent NCT numbers were
repeated in the full text, 10 had no NCT number reported in the full text at all, and there were 4 cases where the non-existent NCT number appeared only in the abstract and a legitimate NCT number was presented in the full article text. See data set for details.

Discussion

While there is a small proportion of non-existent NCT numbers in clinical trial abstracts indexed by PubMed, publication of even a single non-existent NCT number represents a failure on the part of journals who publish clinical trials to systematically ensure that clinical trial registry entries are always checked for concordance with the trial report submitted for publication.

Because we were able to identify legitimate registry entries for the entire random sample that we checked manually, and in a small number of cases, the correct NCT number was reported in the full text publication, these errors appear to be simple data entry mistakes on the part of the authors, rather than a malicious attempt to hide a clinical trial registry from scrutiny. The journals implicated here are not predatory journals, but rather reputable, high-impact journals such as *The Lancet*,$^8–11$ *Blood*,$^{12–18}$ *PLOS One*,$^{19–23}$ *BMJ*,$^{24,25}$ *Circulation*,$^{26–29}$ and *Trials*. $^{30–34}$

In the 215 cases of non-existent NCT numbers that were identified, because the error was not corrected during the publication process, and because our results suggest that there are few cases in which the full publication text contains an accurate registry number when the number in the abstract does not exist, it is very likely that none of the editors or reviewers attempted to even access the clinical trial registry entry. If this is the case, peer review and editorial scrutiny provided no evaluation at all of the concordance between the details provided in the ClinicalTrials.gov record, such as primary and secondary outcomes, and those reported in the publication. This is consistent with what others have reported regarding peer review,$^{35}$ and suggests that editors and others involved in journal publication also do not review every clinical trial for concordance with the reported trial registry entry. The finding of even a single non-existent NCT number in a publication, to say nothing of serial publications of non-existent NCT numbers referring to the same clinical trial, suggests that there is no mechanism ensuring confirmation of clinical trial registry details by editors and reviewers of clinical trial publications, and casts doubt on how frequently editors and reviewers evaluate clinical trial reports in light of their corresponding registry entries.

A survey of 203 peer-reviewers found that verification of trial registration details was never clearly requested by journal editors.$^{36}$ The problem of published NCT numbers that do not correspond to a legitimate ClinicalTrials.gov record could be resolved by the introduction of a check-list to be used by reviewers and editors requiring them to confirm concordance of important clinical trial registry details with the publication text. It would even be possible to introduce an automated tool for reviewers that identifies numbers within a manuscript that are formatted as a clinical trial registry entry and retrieves details of a clinical trial registry entry from
the text of a paper or abstract, or in the case of non-existent ones, flags them for correction.

This study is limited in that for feasibility, this sample only includes registry numbers extracted from abstracts indexed by PubMed. Clinical trials may also report non-existent NCT numbers in the full text of the article.\(^{37}\) Hence, this analysis places a lower bound on the number of published NCT numbers that do not correspond to a legitimate ClinicalTrials.gov registry entry.

One of the main ways that clinical trial preregistration allows for accountability is at the point of trial publication. Even in cases where a clinical trial is pre-registered, inconsistencies such as outcome-switching between the registered trial and the publication are possible. It is the responsibility of peer reviewers of a clinical trial publication to confirm that there is concordance between the clinical trial registry entry data and the submitted text for publication. The role of clinical trial registration in holding investigators accountable is compromised if journal editors and peer reviewers do not refer to clinical trial registries in order to ensure that there is concordance between the registry record and the publication text.

END
Figures

Figure 1. Non-existent ClinicalTrials.gov registry numbers published in abstracts indexed by PubMed

Acknowledgements

Many thanks to Alex Bannach-Brown, Peter Grabitz and Jonathan Kimmelman for conversations that inspired this paper; and to Maia Salholz-Hillel for insightful feedback.

References


36. Chauvin, A., Ravaud, P., Baron, G., Barnes, C. & Boutron, I. The most important tasks for peer reviewers evaluating a randomized controlled trial are not congruent with the tasks most often requested by journal editors. BMC Med 13, 158 (2015).