A simple hint to improve Robinson and Dickersin’s highly sensitive PubMed search strategy for controlled clinical trials

From GIUSEPPE GL BIONDI-ZOCCAI,1 PIERFRANCESCO AGOSTONI,2 ANTONIO ABBATE,3 LUCA TESTA4 and FRANCESCO BURZOTTA1

Systematic literature searches are of paramount importance for investigators involved in clinical research as well as for those overviewing clinical evidence.1 While several electronic databases are available to search for controlled clinical studies, PubMed, the online version of the US National Library of Medicine MEDLINE database, is probably one of the most complete and user-friendly.2 Nonetheless, significant expertise is needed to tackle the vast amount of information available in PubMed. Among other useful tools, explicit search strategies using Boolean operators, Medical Subject Heading (MeSH) terms and keywords have been proposed to improve the sensitivity and specificity of PubMed searches for clinical trials.3 One of the most sensitive as well as specific of such systematic search algorithms is the strategy originally proposed by Robinson and Dickersin in 2002, which has been adopted by several Cochrane Collaboration reviewers.4 However, even this well-tested strategy may lack optimal specificity, as many initially retrieved citations, often in the order of hundreds, eventually end up being non-pertinent or duplicate hits. Unfortunately, it is well known that the number of initial reports is one of the key factors determining length, complexity, and economic burden of a systematic review. Moreover, readers and reviewers of biomedical literature always hope to reduce the number-needed-to-read in order to pursue their scholarly projects or continuing medical education.5

Our aim was to devise a few simple ameliorations to the Robinson and Dickersin’s highly sensitive PubMed search strategy for controlled trials. We thus simply excluded non-controlled and non-experimental studies, such as reviews, commentaries, practice guidelines, meta-analyses and editorials from the potential citations of interest, by using the ‘NOT’ Boolean operator. Letters or journal correspondence were instead considered as potential sources of randomized trial data and thus not explicitly excluded.6 Actually, we simply added to Robinson and Dickersin’s string the following: ‘NOT (comment[pt] OR editorial[pt] OR meta-analysis[pt] OR practice-guideline[pt] OR review[pt])’. The modified search strategy is thus the following (additions are in italics):


In order to test the performance of this simple modification of the Robinson and Dickersin’s algorithm, we compared the standard Robinson and Dickersin’s PubMed search strategy versus the modified algorithm hereby proposed in a pool of five systematic reviews already completed by our group.6–10 The median number (25–75% range) of initial citations was significantly lower using the modified search in comparison to Robinson and Dickersin’s standard PubMed algorithm, respectively 354 (55–1653) versus 393 (87–2041) (P = 0.043 at non-parametric Wilcoxon test). This translates into a median 19% (14–37) reduction in the number of initial citations. We also tested whether any pertinent trial, finally included in the meta-analyses published by our group, and originally retrieved by the standard Robinson and Dickersin’s search strategy, was missed by our more stringent modified algorithm. Actually, no pertinent study was excluded, implying unimpaired sensitivity.

In conclusion, we believe that a minor modification to the well-established Robinson and Dickersin’s PubMed search strategy for clinical trials may significantly increase search specificity without any reduction in sensitivity.

Acknowledgements

The present study is part of an ongoing training project of the Center for Overview, Meta-analysis, and Evidence-based medicine Training (COMET), based in Milan, Italy.

References

Author’s response
From KAREN A ROBINSON

It is always gratifying to know that colleagues have found your work useful and so it was with some interest that I read the article from Biondi-Zoccai, et al.1 I have a couple of comments regarding this work. First, it would be helpful if the authors would clarify the statement that the PubMed strategy retrieves many citations that ‘end up being non-pertinent or duplicate hits.’ For the former, I am assuming that they mean citations that are subsequently determined not to be reports of controlled trials. For the latter, however, I am puzzled as to how it is possible that the strategy would retrieve duplicate citations.

I would also like to clarify that the PubMed strategy evolved as a translation of the original Cochrane strategy developed for use in Ovid and SilverPlatter.2,3 If we had set out to develop a new strategy, such as through an analysis of the medical subject heading (MeSH) and text words in known controlled trials, it is possible that it would have been different. Having said that, however, both the revised and original Cochrane strategy avoided the use of limit statements. The addition of ‘NOT’ statements leaves a strategy vulnerable to indexing errors. The one limit statement included in both the original and PubMed strategies uses a phrasing that avoids this vulnerability (i.e. NOT (animal[mh] NOT human[mh])). There is also the possibility that articles indexed with one of the publication types in the Biondi-Zoccai limit statement may be a source of information concerning a controlled trial, such as is sometimes the case with a letter. I would suggest that those needing to complete comprehensive searches use caution in adding any limit statements, be they statements related to the specific topic of interest or the limit statement proposed by Biondi-Zoccai and colleagues.

References

Johns Hopkins University School of Medicine, 1830 E Monument Street, Baltimore, MD21205, USA. E-mail: krobin@jhmi.edu