

**Searching for observational studies: what does citation tracking add to Pub
Med? A case study in depression and coronary heart disease**

Hannah Kuper¹, Amanda Nicholson², Harry Hemingway²

¹*International Centre for Eye Health, Clinical Research Unit, London School of
Hygiene & Tropical Medicine, Keppel Street, London WC1E 7HT. E-mail:
hannah.kuper@lshtm.ac.uk

²International Centre for Health and Society, Department of Epidemiology and Public
Health, University College London Medical School, 1-19 Torrington Place, London
WC1E 6BT. E-mail: amandan@public-health.ucl.ac.uk, harryh@public-
health.ucl.ac.uk

* Corresponding author

Abstract

Background

Pub Med is the most widely used method for searches of the medical literature, but often fails to identify many relevant papers. Electronic citation tracking offers an alternative search method.

Methods

Papers investigating the role of depression in the aetiology and prognosis of coronary heart disease were sought through two methods: a) Pub Med, and b) citation tracking of index papers using Science Citation Index. The number and quality of eligible papers identified by the two methods were compared.

Results

52 studies met our inclusion criteria; 12 were identified through Science Citation Index alone, 9 through Pub Med alone, and 31 through both methods. Papers identified by Science Citation Index alone were published in higher impact factor journals, were published more recently and were less likely to show a positive association.

Conclusions

Science Citation Index identified more eligible papers than PubMed, and these differed qualitatively. Failing to use citation tracking in a systematic review of observational studies may result in bias.

Background

Highly sensitive methods have been developed for the identification of randomised trials, but there are few guidelines for searching for observational studies [1]. PubMed is the most widely used database for searches and is freely available. PubMed may miss many relevant papers [2] and this could influence the conclusions drawn [3]. Science Citation Index offers a potential complement to Pub Med, by allowing subsequent papers citing an index paper to be identified (so called forward citation tracking) as well as the bibliography of the index paper. No previous studies have assessed whether forward citation tracking improves upon a Pub Med Search. We sought therefore to compare the number and quality of papers investigating depression and coronary heart disease (CHD) that were identified through Science Citation Index and Pub Med.

Methods

Eligible papers were restricted to prospective studies of healthy populations (aetiological) or populations with defined CHD (prognostic), published before 2004. The exposure was depression and the outcome was defined as fatal CHD or incident non-fatal myocardial infarction (for aetiological studies) and all cause mortality or CHD mortality (for prognostic studies) [4].

For each of the 56 papers in the index review [4], we used Science Citation Index to identify all subsequent papers in which they were cited (*forward citation tracking*). Two independent reviewers went through the standard process of screening titles, abstracts and full text versions against the eligibility criteria, with recourse to a third reviewer in the event of a disagreement. All titles of papers in the bibliographies of the index papers were identified (*backward citation tracking*) and the selection procedure was repeated.

We devised a Pub Med search strategy of medical sub heading (MeSH) terms and text words, using the index review [4]. This identified all but six of the index papers (four of which did not include depression or depressive in the keywords, title or abstract and one did not mention heart disease; extending the search to allow identification of the sixth paper would have added more than 500 unique titles). The search was run on Pub Med until the end of 2003. The procedure for reviewing titles, abstracts and full text papers was repeated as above.

The publication year, journal impact factor and a summary of association was recorded for the eligible papers. Analysis of Variance (ANOVA) and Chi-square tests were carried out to test for differences in the characteristics of papers (i.e. year of publication, journal impact factor, type of study and strength of reported association) identified by Science Citation Index Alone, Pub Med or through both strategies.

Results

Science Citation Index identified more unique titles than Pub Med (2906 and 2153 respectively), as well as more abstracts (832 and 794) and papers for review (345 and 254). 52 studies met our inclusion criteria; 12 were identified through Science Citation Index alone (7 forward, 5 backward), 9 through Pub Med alone, and 31 through both methods.

Papers identified by Science Citation Index alone were significantly more likely to be published in higher impact factor journals and were least likely to show a positive association (Table 1). Papers identified by both Pub Med and Science Citation Index were published more recently.

Conclusion

In this case study, Science Citation Index identified more eligible papers than PubMed. These papers were published in higher impact journals and were more likely to show null results. However, starting the review with a smaller number of index papers may influence the relative efficiency and time taken of the two methods. It is not surprising that citation tracking improved upon PubMed. Forward citation tracking offers the integral of multiple searches carried out by different publishing research groups using different (unreported) search methods. Citation tracking is wholly independent of the need to specify search strategies or use MeSH headings, which are a recognised limitation of PubMed. Ten of the 12 papers that were identified through Science Citation Index alone were within the PubMed database, but were not detected by the search because they did not include depression or depressive in key words or MeSH headings (n=6) and/or did not include the relevant heart disease terms (n=5).

Failing to use citation tracking in a systematic review of observational studies may result in bias. It is well known that existing search methods fail to identify the complete set of eligible studies [2], and conclusions drawn from a systematic review may be influenced by the search strategy used [3]. Two systematic reviews of depression and CHD identified few eligible papers [5, 6]. The present case study does not prove that citation tracking improves upon PubMed in other observational settings, but we suspect that the chances of funding such bibliographic research are low. In the meantime, although Science Citation Index is only available by subscription, since citation tracking involves only a modest additional work load (in this case approximately two person weeks) and may offer an opportunity to reduce bias, we

propose that the onus should be on systematic review protocols to justify situations where citation tracking has *not* been used.

List of abbreviations

CHD	-	Coronary Heart Disease
MeSH	-	Medical Sub Heading
s.d.	-	Standard Deviation
ANOVA	-	Analysis of Variance

Competing interests

None.

Authors contributions

HK and AN were responsible for reading the titles, abstracts and papers and determining their eligibility, with recourse to HH in the event of a disagreement. HK was responsible for drafting the paper, and AN and HH were responsible for revising it critically for important intellectual content. All authors read and approved the final manuscript.

References

1. Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, Moher D, Becker BJ, Sipe TA, Thacker SB: **Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group.** *JAMA* 2000, **283**: 2008-12.
2. Dickersin K, Scherer R, Lefebvre C: **Identifying relevant studies for systematic reviews.** *BMJ* 1994, **309**: 1286-91.
3. Borsody MK, Yamada C: **Effects of the search technique on the measurement of the change in quality of randomized controlled trials over time in the field of brain injury.** *BMC Med Res Methodol* 2005 **5**: 7.
4. Kuper H, Marmot M, Hemingway, H: **Systematic review of prospective cohort studies of psychosocial factors in the etiology and prognosis of coronary heart disease.** *Sem Vasc Med* 2002, **2**: 267-314.
5. Rugulies R: **Depression as a predictor for coronary heart disease. a review and meta-analysis.** *Am J Prev Med* 2002, **23**: 51-61.
6. Wulsin LR, Singal BM: **Do depressive symptoms increase the risk for the onset of coronary disease? A systematic quantitative review.** *Psychosom Med* 2003, **65**: 201-10.

Table 1. Comparison of papers identified through Science Citation Index and Pub Med searches

Search method(s) that identified paper	Mean publication year (s.d.)	Mean impact factor (s.d.)*	Aetiologic or prognostic study	Summary of association**
Pub Med and Science Citation Index (n=31)	1999 (6.4)	4.2 (2.8)	8 aetiologic 23 prognostic	25 positive 6 null
Science Citation Index only (n=12)	1994 (9.3)	9.3 (9.4)	3 aetiologic 9 prognostic	5 positive 7 null
Pub Med only (n=9)	1994 (13.2)	3.8 (2.4)	2 aetiologic 7 prognostic	7 positive 2 null
p-value	0.08	0.05	0.98	0.04

* Four papers were published in journals with no impact factor (1 both methods, 1 Science Citation Index only, 2 Pub Med only).

** A positive study was one where the effect estimate indicated some association between depression and CHD outcome.