

Supplementary searches of PubMed to improve currency of MEDLINE (Ovid) searches for systematic reviews: time taken for a record to move to MEDLINE

Steven Duffy¹, Kate Misso¹, Shelley de Kock¹, Caro Noake¹, Janine Ross¹, and Lisa Stirk¹

¹*Kleijnen Systematic Reviews, York, United Kingdom*

Corresponding author: Steven Duffy, steven@systematic-reviews.com

Abstract

Introduction: When conducting literature searches for systematic reviews one bibliographic database is almost always included, MEDLINE. MEDLINE content can be searched via numerous different search interfaces, the majority of which are provided by fee-based subscription services (e.g. Ovid), although it can be accessed for free using PubMed. Many information specialists would choose to search using sophisticated interfaces (usually fee-based) that enable the design of complex search strategies. PubMed has more limited search capabilities, for example proximity searching is not possible. However, 2% of PubMed records are not found in MEDLINE, including newly published and ahead of print articles.

Objectives: Investigations of in-house Kleijnen Systematic Reviews Ltd (KSR) systematic reviews suggest that conducting a supplementary search of PubMed in addition to the main MEDLINE (Ovid) search is worthwhile. We wanted to see if our findings could be supported by following a randomly selected set of records from entry into PubMed to finally being indexed in MEDLINE.

Methods: A search of PubMed for ahead of print records was conducted. 'Ahead of print' records are those citations electronically submitted to PubMed in advance of the journal issue's release: following publication of the completed issue, the record then moves to the next stage of the indexing process. Using a random number generator, we selected a set of twenty records from the results of this search, and followed their progress through PubMed. A selection of ahead of print records from major medical journals were also followed: BMJ, JAMA, Lancet.

Results: We recorded movement through the PubMed indexing process to see if there was any consistency or disparity in the time it took for a record to proceed to the next stage of indexing, and whether certain journals took precedence.

Conclusions: The time it takes for a record to move from initial entry into PubMed, through to the In-Process stage, and finally to full Medical Subject Heading (MeSH) indexing and entry into MEDLINE, can influence whether a relevant study for inclusion in a systematic review is not retrieved by the literature search. The majority of records tracked during this investigation took between 4-11 months to migrate from PubMed to MEDLINE; however 4 records remained in PubMed for the duration of our investigation. Supplementary searches of PubMed can ensure that potentially relevant studies yet to reach MEDLINE are not missed by the literature searches.

Key words: Databases, Bibliographic; Information Storage and Retrieval; Medical Subject Headings; MEDLINE; PubMed; Review Literature as Topic.

Introduction

When conducting comprehensive literature searches for systematic reviews of health care interventions one bibliographic database is almost always included, MEDLINE. MEDLINE content can be searched via several search interfaces, the majority of which are provided by fee-based subscription services (e.g. Ovid, EBSCO or ProQuest), although it can be accessed for free using PubMed. When searching for systematic reviews PubMed has been shown to have a higher retrieval rate than fee-based versions of MEDLINE, such as Ovid.[1] PubMed enables users to search

MEDLINE, PubMed citations that will never be indexed on MEDLINE, as well as other sources, such as online books and reports. Despite this, and the fact that PubMed is freely available, many information specialists prefer to search using sophisticated interfaces (usually fee-based) that enable the design of complex search strategies.

Objective

Previously we investigated whether conducting a supplementary search of PubMed in addition to the main MEDLINE (Ovid) search would be worthwhile. We wanted to ascertain whether this supplementary PubMed search could be conducted quickly and if it retrieved unique, recently published and ahead of print studies that were subsequently considered for inclusion in the final systematic review. Searches of PubMed were conducted after MEDLINE (Ovid) and MEDLINE In-Process (Ovid) searches had been completed for six KSR reviews (and a review update). Search strategies were simplified and retained the original conceptual structure of the MEDLINE (Ovid) strategy. The searches were limited to records not on MEDLINE or MEDLINE In-Process (Ovid). Our findings indicated that supplementary PubMed searches identified additional unique records with potential studies for inclusion in the final systematic review. For KSR reviews, conducting supplementary PubMed searches for studies unavailable elsewhere was worthwhile, and improved the currency of the final review.[2] We wanted to see if our findings could be supported by following a randomly selected set of records from entry onto PubMed to finally being indexed for MEDLINE. We hoped that tracking a sample of PubMed records would provide a clearer picture of the time-lag preceding record migration to MEDLINE. We assumed this information would clarify at what time point a MEDLINE search would be sufficient to capture potentially relevant records available on PubMed.

Methods

A search of PubMed for ahead of print records was conducted. 'Ahead of print' records are those citations electronically submitted to PubMed in advance of the journal issue's release: following publication of the completed issue, the records can move to the next stage of the indexing process, in-process.

PubMed was searched using the following search strategy:

pubstatusaheadofprint AND 2014/05/16 [edat]

This search retrieved 3133 records that had entered PubMed on 16 May 2014 (EDAT: Entrez Date), all of which were downloaded into EndNote bibliographic software. Twenty records were then randomly chosen via an online random number generator (<http://www.randomizer.org/>). Each correspondingly numbered record was highlighted in EndNote, and added to a Group set. The PubMed record identifier (PMID) for each record in the group was then retrieved and combined to produce a saved search strategy in PubMed using 'My NCBI'. The saved strategy of 20 PMIDs was then combined with terms to help determine when records move to in-process or MEDLINE. An email alert was also created to notify whenever a record had moved to the next stage of the process.

However, there were a number of articles in the sample of 20 records from journals not usually included in systematic reviews of healthcare interventions (e.g. Journal of Heredity, Journal of Histochemistry and Cytochemistry, Inorganic Chemistry, Angewandte Chemie, European Journal of Human Genetics, Analytical Chemistry, etc.). A second, smaller, more representative sample was therefore created from the same 3133 records. This sample was pragmatically generated to include records from more prominent journals that we would expect to include in systematic reviews of

healthcare interventions (e.g. from the JAMA, BMJ and Lancet title collections). We also followed citation records of included studies that had been identified from our original investigations.

A record of any movement through the PubMed indexing process was kept to see if there was any consistency between the time it took for a citation to proceed to the next stage, and whether certain journals took precedence over others.

Results

For the random sample of 20 records there did not appear to be any discernible pattern in the time it took for citation records to move through the PubMed indexing process. Only 13 records of the 20 had been indexed for MEDLINE (MeSH indexed) by May 2016. It took those 13 records over 3 months on average to reach the in-process stage (5-300 days; 96 mean days; 90 median days; range 295 days), and a further almost 8 months (7 months, 3 weeks) to finally reach MEDLINE (50-432 days; 235 mean days; 266 median days; range 382 days).

However, even from this small sample there were 7 records that had not been indexed for MEDLINE by May 2016: one record had been removed entirely from PubMed; 2 records were still at the 'as supplied by publisher' stage; 2 records were at the in-process stage; and 2 records had reached in-process and subsequently moved into PubMed-not-MEDLINE.

The record that no longer exists [PMID: 24824232] was a citation for a 'correction' and the erratum details were added to the original citation record [PMID: 24727128]. No amendments were made to the 2 records still at the 'as supplied by publisher' stage since entering PubMed. Record citations only move to in-process once they have been published (i.e. receive volume/issue, page numbers), however both of these records were published in 2014 (PMID: 24829364 and 24828513). Both records should have proceeded to either in-process or PubMed-not-MEDLINE at the very least. It is not clear what has happened to these two records.

There appears to be no similarity between the 2 records that are currently at the in-process stage (May 2016). One record took almost a year to move to the in-process stage, (PMID: 24829153; 11 months, 13 days), whilst the other took 5 months to move to in-process (PMID: 24828620).

The 2 records that have moved to PubMed-not-MEDLINE were out of scope for MEDLINE.

The second sample of 12 records from more prominent journals appeared to move to MEDLINE a little quicker than those records in the random sample.

Certainly the prominent journals included in this sample moved their ahead of print articles fairly quickly to full publication (less than 2 months on average), thus enabling progression to in-process in PubMed (52-153 days; 49.3 mean days; 52 median days; range 101 days). The average time for a record in this sample to move through MeSH indexing and onto MEDLINE was just over another 2 months (53-81 days; 63.3 mean days; 62 median days; range 28 days). However, 2 of the records from this sample have still yet to be indexed for MEDLINE, and took a lot longer than 2 months to even reach the in-process stage: one record took a year and 9 months (648 days) to reach in-process (PMID: 24823693); whilst the other took 4 months (126 days) to move to in-process (PMID: 24825895).

Four records were followed after they had been identified by our supplementary PubMed searches in the original investigation of in-house KSR reviews [2]. Supplementary PubMed searches for a review about medicinal cannabis [3] identified one study [4] that was included in the final systematic review. After being published online on 13 January 2014 as 'Early View', the article was

eventually published in the European Journal of Pain in the August issue (which itself was published online on 8 July 2014). The article first appeared on PubMed on 15 January 2014, then moved to in-process on 22 July 2014, and was finally indexed for MEDLINE on 27 May 2015. Its first appearance in Ovid was on MEDLINE In-process on 23 July 2014, making it onto MEDLINE Ovid on 4 June 2015: MEDLINE 2015/May week 5 (Entry Date 20150527). The record took 497 days (1 year, 4 months, 12 days) to reach MEDLINE, and crucially over 6 months (188 days) of this time to move to in-process. For over a year this study would not have been retrieved by a search of MEDLINE, and for 6 months even a search of MEDLINE In-process would not have been successful.

Update searches for the cannabis review [5] also identified one included study [6] through the supplementary PubMed searches. This article was first cited on PubMed on 7 April 2015, but did not reach in-process until 29 June 2015 (83 days later), and then took another 277 days (over 9 months) to be indexed for MEDLINE (1 April 2016).

Our review about Procalcitonin testing [7] also included one study [8] identified by the supplementary PubMed searches. The study was added to PubMed on 10 July 2014 and the searches were undertaken on 14 July 2014. The record moved to in-process on 1 August 2014; and finally to MEDLINE status on 20 April 2015. It would not have been identified in searches of MEDLINE until April 2015, or in-process until August 2014. For our review about secondary hyperparathyroidism (SHPT), again, 1 included study was identified in PubMed. This study entered PubMed on 23 October 2014, with the PubMed search being conducted on 15 January 2015. In-process status was reached on 1 April 2015, five months after first entering PubMed, and almost three months after the search had been conducted. The study was finally indexed for MEDLINE on 16 February 2016, over a year and 3 months after entering PubMed.

Conclusions

Our investigations gave no definitive answers, despite following the sample records for 2 years. We could discern no pattern from our small samples: records were removed from PubMed, some records took over a year to move from one indexing stage to the next, whilst others took just a week. What was clear from our investigations was that it was impossible to give a precise time for how long it takes a record to move from one stage of the indexing process to the next in PubMed. For this reason alone we would recommend supplementary searches of PubMed to capture potentially useful records for inclusion in systematic reviews.

Update

At the end of April 2016 Ovid introduced a new segment to their MEDLINE suite, 'Epub Ahead of Print'. This segment contains the 'as supplied by publisher' records from PubMed. To all intents and purposes this new segment includes those records previously only available on PubMed that have been discussed throughout this investigation, and so negates the need to conduct supplementary searches of PubMed. However, it is unclear what exactly is being uploaded from PubMed to this new Ovid MEDLINE segment. It looks as though citations from the NLM PMC bookshelf are not being loaded, and that other PubMed-not-MEDLINE citations remain in PubMed alone. Updates by Ovid are currently weekly, whilst citations are loaded onto PubMed on a daily basis. However, Ovid are expecting as of June to be loading Epub records from the previous day. It is also unclear how differing search approaches in the different search interfaces are going to affect search results. This is a very welcome addition to Ovid MEDLINE, but further investigation is warranted. In the meantime we would suggest continuing to conduct supplementary searches of PubMed.

REFERENCES

1. Kaunelis D, Farrah K, Severn M, Canadian Agency for Drugs and Technologies in Health. The missing 2%: PubMed NOT MEDLINE. Canadian Health Libraries Association/Association des bibliothèques de la santé du Canada (CHLA/ABSC). Kingston, Ontario, 2010.
2. Stirk L, Duffy S, Misso K, Noake C, Ross J. Supplementary searches of PubMed to improve currency of MEDLINE and MEDLINE In-Process searches via Ovid: evidence from practice. Paper presented at 23rd Cochrane Colloquium; 3-7 Oct 2015; Vienna, Austria. 2015: Rapid Oral session 23.7.
3. Whiting P, Wolff R, Westwood M, Duffy S, Keurentjes C, Lang S, et al. Systematic review of cannabis for medical use. Commissioned by Swiss Federal Office of Public Health FOPH. York: Kleijnen Systematic Reviews Ltd, 2014
4. Serpell M, Ratcliffe S, Hovorka J, Schofield M, Taylor L, Lauder H, et al. A double-blind, randomized, placebo-controlled, parallel group study of THC/CBD spray in peripheral neuropathic pain treatment. *Eur J Pain* 2014;18(7):999-1012.
5. Whiting PF, Wolff RF, Despande S, Di Nisio M, Duffy S, Hernández AV, et al. Cannabinoids for medical use: a systematic review and meta-analysis. *JAMA* 2015;313(24):2456-73.
6. Wallace MS, Marcotte TD, Umlauf A, Gouaux B, Atkinson JH. Efficacy of inhaled cannabis on painful diabetic neuropathy. *Journal of Pain* 2015;16(7):616-27.
7. Westwood ME, Ramaekers BLT, Whiting P, Tomini F, Joore MA, Armstrong N, et al. Procalcitonin testing to guide antibiotic therapy for the treatment of sepsis in intensive care settings and for suspected bacterial infection in emergency department settings: a systematic review and cost-effectiveness analysis. *Health Technol Assess* 2015;19(96):1-236.
8. Schuetz P, Kutz A, Grolimund E, Haubitz S, Demann D, Vogeli A, et al. Excluding infection through procalcitonin testing improves outcomes of congestive heart failure patients presenting with acute respiratory symptoms: results from the randomized ProHOSP trial. *Int J Cardiol* 2014;175(3):464-72.