

Unraveling the threads of overlapping coverage among the three most commonly explored biomedical databases

Mangat G¹; Pilkhwal, N¹; Sharma, S²

¹Parexel International, Mohali, India;
²Parexel International, Chandigarh, India

Background

Healthcare researchers traditionally search for evidence in biomedical databases, such as Embase, PubMed, and Cochrane Library. Embase covers MEDLINE, so a search in Embase alone should cover every citation in both Embase and MEDLINE. PubMed is separately searched to identify “MEDLINE in-process” OR “ahead-of-print” citations. Cochrane Library comprises citations from the Cochrane Database of Systematic Reviews (CDSR) and Cochrane Central Register of Controlled Trials (CENTRAL) and is used for a comprehensive search for clinical trials.

Objective

There is a certain degree of overlap between these databases; however, explicit details on the overlap of these databases are not available. Therefore, to investigate this issue, we compared the databases to analyze the level of overlap and establish the best practice for using these databases.

Methods

We conducted two searches on mesothelioma and lymphoma as case examples and followed a step-by-step approach to verify citations manually across databases. We developed a search strategy for the Embase.com interface using industry-standard search filters and translated it for use in PubMed and Cochrane.

Results

PubMed vs. Embase

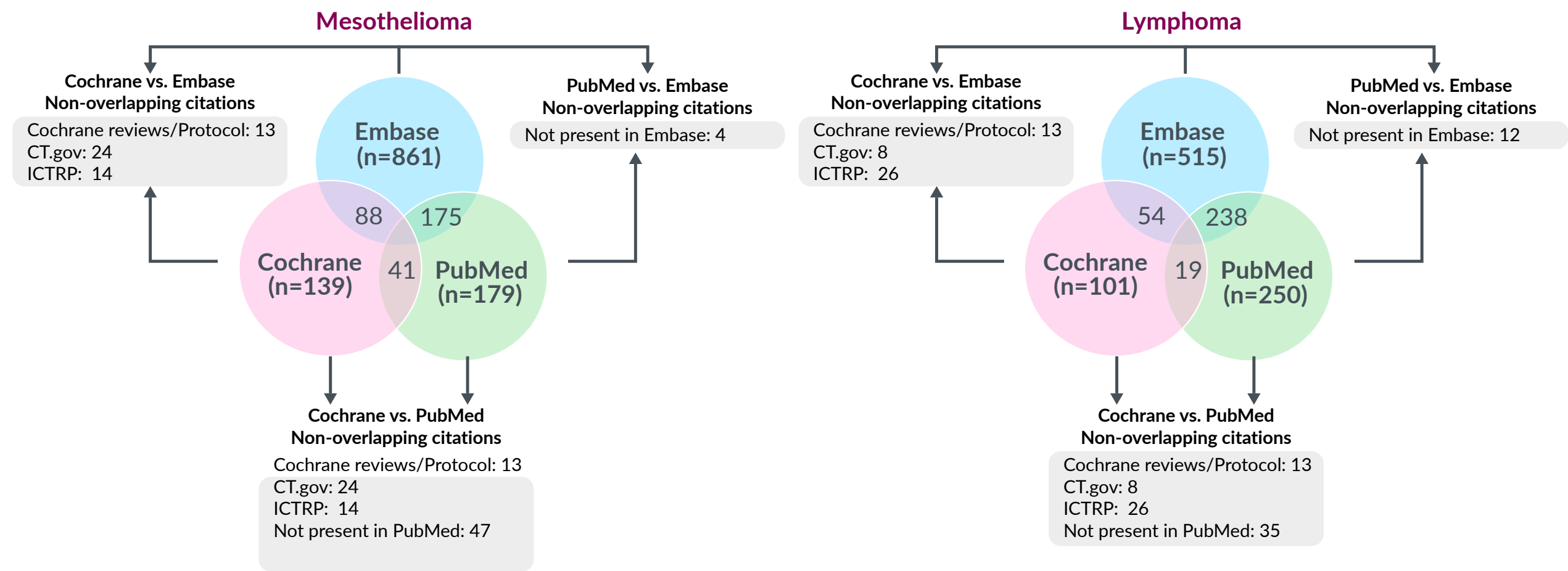
➤ We observed some critical variations. Of the 179 PubMed citations retrieved in mesothelioma, 4 were missing in Embase (both in search output and content archive). Similarly, among the 250 and 515 citations retrieved from PubMed and Embase for lymphoma, 12 MEDLINE citations were missing in Embase. This refutes the post-2010 theoretical consensus that all MEDLINE citations reside in Embase. We further noticed that all citations missing in Embase were published in the last 1-3 years, so potentially not missed due to time lag but due to coverage of different journals. This highlights the necessity to search PubMed separately for its overall content rather than considering Embase as a one-stop shop

Results

Cochrane Library vs. Embase and PubMed

➤ The Cochrane search identified 139 and 101 citations for mesothelioma and lymphoma, respectively. As Cochrane includes controlled trials (available in CENTRAL) retrieved from Embase and PubMed (either through autofeed or crowdsourcing), 100% of the controlled trial evidence (n=88 and 54 citations for mesothelioma and lymphoma, respectively) overlapped with Embase, while 47% and 35% with PubMed for mesothelioma and lymphoma, respectively. The difference in overlap % between Embase and PubMed could be attributed to their differential coverage (e.g., Embase.com has a unique set of citations unavailable in PubMed, no or less coverage of conference abstracts in PubMed, while Embase and Cochrane index conference abstracts). Both Embase and PubMed do not cover the output retrieved from clinical trials registries (i.e., clinicaltrials.gov and ICTRP) and CDSR (reviews and protocol)

Figure 1: Comparison of study coverage



Conclusion

The results suggest that researchers should view Embase and PubMed as independent sources to be searched without dependence on cross-database coverages. The relevance of Cochrane to cover scientific evidence beyond trials needs further testing.