No.
145752

A search filter to identify natural history studies for a systematic review: notes on structure, design and next steps

Maria Rapoport,[1] Chris Cooper,[2] Antoinette Randet[1]

[1] Wickenstones Ltd, Oxford, UK; [2] Independent researcher, London

WICKENSTONES
PARTNERS IN BUILDING VALUE

BACKGROUND

There are no existing search filters to identify natural history studies or data, despite these data being a core component of decision models and informing analyses in Health Technology Assessment.

In our work on the rare condition osteogenesis imperfecta (brittle bone disease), we designed a multi-stranded search filter to identify natural history studies and data. In this poster, we:

► Report the design and structure of this search filter

► Offer reflections on its use and possible changes

DEFINITIONS

Search filter: a prescribed (often validated) set of terms used to identify a specific type of study or data. A good example is the Cochrane RCT filter to identify studies reporting randomised or controlled studies

Multi-stranded search filter: a search filter that uses differing combinations (strands) of concepts to address complex research questions or information needs. For instance, within the same filter, searching for specific study designs, methods of data collection, time points or outcomes.

Natural history data: data collected from a studied group of people over time to develop an understanding of how a disease or condition develops

METHODS

We designed a multi-stranded search filter based on scoping searches and an informal review of studies and systematic reviews of natural history data (Table 1). The search filter took the following form:

Strand 1: terms for natural history or longitudinal data; OR
Strand 2: terms for study designs that include or would report data captured over time; OR
Strand 3: terms for data collection over time

We combined the strands with terms for the condition and a date limit of 2014–present.

Filter performance was compared to known eligible marker papers and drawn from the authors' experience undertaking this type of review.

RESULTS

We identified 1,447 studies after deduplication based on a search of MEDLINE and Embase (Ovid interface).

We included 66 studies in our review. Our filtered database search identified 58 studies (88%). Eight studies were identified by snowballing. No other eligible studies were identified from the suite of five reviews we are undertaking in parallel (e.g. HRQoL or economic evaluations) or following expert feedback.

We conclude, therefore, that the filter was effective in this topic area for our purpose.

NEXT STEPS

A number of animal studies were returned, which increased yield and screening burden. These could be removed using the following logic in MEDLINE Ovid interface: (search results NOT exp animals/not humans.sh.) Re-running the search in October 24, this reduced yield by 46 studies (3.6%). We would urge caution here. Since the advent of automated study indexing, we informally observe that human studies are incorrectly being indexed in MEDLINE. We would not limit like this, but it is possible.

Eight studies were identified by snowballing. Of these, one was not available for retrieval in MEDLINE/Embase, three were studies designs excluded in the protocol and four indicated potential adjustments to improve the sensitivity of our filter.

1. Adding the terms development or developmental to Line 29. This increases n to screen by 399 in MEDLINE alone but picked up two of the four studies. We would also consider adding the following indexing term in MEDLINE: Time Factors/ which adds n=23 to our search overall.
2. Focused versus unfocused indexing terms. Strand 2 focused on study designs likely to report natural history or longitudinal data. We focused the indexing terms (indicated by * preceding the indexing term) which limits retrieval to items where the primary focus of an item is the indexing term. Two of the eight studies identified in snowballing but not picked up by the search were indexed by study design, but study design was not reported in title or abstract (and there was no author indexed keywords). Unfocusing the indexing terms would lead to the two studies identified in the search, but yield increased by 10.71% (n=122).

CONCLUSIONS

► We present what we believe is the first filter for natural history studies.

► Our search filter is provided with a search narrative that we hope explains the decision making behind the filter, making it amendable for use by others or adaptation to specific projects.

► In future projects, we will add the terms development or developmental and the indexing term Time Factors/ We will also unfocus study design indexing terms.

► We will report this filter to the ISSG search filter resource for future use.

Table 1. Natural history study search filter

Database: MEDLINE (MEDALL)
Host: Ovid
Data parameters: 1946-current
Date of search: 24 March 2024

1 Osteogenesis imperfecta/ or (Osteogenesis imperfecta or (brittle adj3 bones)).ti,ab,kf. (6750)	Condition terms: The search opens at Line 1 with the controlled indexing (MeSH) term for osteogenesis imperfecta (the / indicates that this is an indexing term). We also utilise free-text terms. These are terms identified by the research team. They seek to cover deficiencies in indexing, to identify new or unindexed records and to search in content fields beyond indexing. As this was a targeted search, whilst we tested a broader free-text structure, the results exceeded our resources. This search uses terms for both osteogenesis imperfecta or brittle bones, as variation in naming is noted in registry reports (c.f. [1] or [2]). We do not distinguish between or define types of osteogenesis imperfecta (e.g. Type 1), as the underlying condition term does not change by type (c.f.[3-5]. that is, there is not a different sub-name by type, as is the case for some conditions). Free-text terms are searched in the following fields: ti = title ab = abstract kf = author generated keyword filed (i.e. terms chosen by authors to describe their papers) We use proximity markers, which (e.g. adj3) allow us to search between search terms to a gap of one space and with the terms in either order. So, we will identify 'brittle leg bone' or 'bones were brittle'. We checked the search strategies and studies included or excluded in the two Cochrane reviews focused on osteogenesis imperfecta.[6, 7] The aim was to critique our draft condition search terms to see if any terms had been overlooked. We also scoped the literature and studies. This process did not lead us to alter our terms.
2 Natural histor*.ti,ab,kf. (57155) 3 Osteogenesis Imperfecta/hi or History.ti. (101029) 4 ((patient* or client or after or provid* or family or families) adj3 histor*).ti,ab,kf. (169888) 5 2 or 3 or 4 [1. terms for natural history or historical data] (298105)	In this section of the search, we aim to identify natural history studies or studies that report longitudinal data. There are no existing search filters, nor is there guidance on searching specifically for natural history studies or data. So, in this aspect of the search, we search in three ways. 1. terms for natural history or historical data Lines 2-4 describe terms for natural history. We search for this phrase explicitly, but also using truncation to pick up plurals. The MeSH term is re-used from Line 1 (above) but focused down using a floating subheading to identify historical data specifically (Line 3). NB: we do not use the MeSH heading for natural history, since this is meant for the classification of natural objects.
6 *Retrospective Studies/ (561) 7 (retrospective* adj2 (stud* or review* or analysis or analysed)).ti,ab,kf. (733108) 8 chart review*.ti,ab,kf. (55973) 9 (Medical record* or patient chart*1 or patient record* or registry or registries or claims data* or patient survey*).ti,ab,kf. or *Registries/ or *"Surveys and Questionnaires"/ (435568) 10 *Prospective Studies/ (486) 11 (prospective* adj2 (stud* or review* or analysis or analysed)).ti,ab,kf. (434772) 12 *Case-Control Studies/ (1390) 13 case control*.ti,ab,kf. (166478) 14 *Longitudinal Studies/ (1963) 15 (longitudinal or longitudinally).ti,ab,kf. (368118) 16 *Cross-Sectional Studies/ (571) 17 (cross section or cross sectionally).ti,ab,kf. (35316) 18 *Interrupted Time Series Analysis/ (180) 19 time series.ti,ab,kf. (49515) 20 *Cohort Studies/ or *Birth Cohort/ (1615) 21 cohort*.ti,ab,kf. (928701) 22 Osteogenesis Imperfecta/ep [Epidemiology] (132) 23 (epidemiology or epidemiological).ti,ab,kf. (457897) 24 *Incidence/ or Incidence.ti,ab,kf. (950673) 25 *Prevalence/ or Prevalence.ti,ab,kf. (830899) 26 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 [2. study designs which would capture and may report data] (4083402)	2. study designs that would capture and may report data (even where not identified as natural history) We focus on studies that track data over time, as indicated in guidance produced by the FDA.[8] These are represented in Lines 6 to 25. We also include: • ITS designs, as these studies may track cohorts over time with the interruptions indicating data collection. • Cohort studies, where people known to have the condition or to support people with the condition have been tracked. • We search for incidence and prevalence. Whilst not strictly interested in these data for this review, the numbers returned by conclusion of this these search lines are low. It is possible that the studies in reporting data also capture data over time.
27 (Osteogenesis imperfecta and year*).ti. (69) 28 ((200* adj3 201*) or (200* adj3 202*) or (201* adj3 202*)).ab. (547965) 29 Disease progression/ or Progression.ti,ab,kf. (822816) 30 ("follow* up*" or "follow* on*" or "growth plane*" or "patient journey" or "clinical journey" or "over time" or "time course" or "life course" or "life span" or "natural course" or "end to* end" or "from* diagnosis" or "long term").ti,ab,kf. (2513777) 31 27 or 28 or 29 or 30 [3. terms for data collection over a time period] (3577457)	3. terms for data collection over a time period Line 27 focuses on the primary condition term and it will pick up any study that includes a year in the title (e.g. [9-11]). Line seeks to identify cohorts tracked over time, such as [12].
32 5 or 26 or 31 (6658156)	Line 32 combines the three sets of terms for NH.
33 (2004* or 2005* or 2006* or 2007* or 2008* or 2009* or 2010* or 2011* or 2012* or 2013* or 2014* or 2015* or 2016* or 2017* or 2018* or 2019* or 2020* or 2021* or 2022* or 2023* or 2024*).dt,dp,ed,ep,yr. (22827954) 34 1 and 32 and 33 (1217)	Line 33 reports the date limit 2004-2014. Line 37 completes the search by combining: • Terms for condition (Line 1) • Terms for NH (Line 32) • Date limit (Line 33)

REFERENCES

1 Ultragenyx Pharmaceutical Inc, 2023. https://trialsearch.who.int/Trial2.aspx?TrialID=NCT05768854
2 Mereo BioPharma 3 Ltd., 2017 https://trialsearch.who.int/Trial2.aspx?TrialID=EUCTR2016%E2%80%90005096%E2%80%9027%E2%80%90DK
3 Ultragenyx Pharmaceutical Inc, 2017 https://trialsearch.who.int/Trial2.aspx?TrialID=NCT03118570
4 Letocha AD, et al. 2005 10.1359/jbmr.050109
5 Rauch F et al. 2003 10.1210/jc.2002-021371
6 Dwan K, et al. 2016 10.1002/14651858.CD005088.pub4
7 Beecham E, et al. 2015 10.1002/14651858.CD010750.pub2
8 U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER), Center for Biologics Evaluation and Research (CBER), Office of Orphan Products Development (OOPD). 2019. https://www.fda.gov/media/122425/download
9 Liang X, et al. 2022 10.1111/pan.14454
10 Xi L, et al. 2021 10.3389/fgene.2021.622078
11 Zhang Z, et al. 2024 10.1097/MS9.0000000000001787
12 Joshi KK, et al. 2023 10.1002/jbm4.10782

ACKNOWLEDGEMENTS

The systematic review project was supported by Mereo BioPharma, UK.

Presented at: ISPOR Europe; Barcelona, Spain; 17-20 November 2024