Bringing it all together:
Interpreting literature reviews

Shaun Prentice, Ashley Blowes, Andrew Killcross

This article is part of a series of articles on research literacy.

Background
Literature reviews can summarise a range of evidence on a topic, making them a valuable resource. However, decisions made during different processes within literature reviews can affect the appropriateness of applying their findings to clinical encounters.

Objective
To support general practitioners (GPs) in understanding and critiquing literature reviews so they can incorporate evidence into their practice.

Discussion
This paper offers nine key questions to consider when reading literature reviews. These range from considering the sources researchers use to locate evidence, to the rigour of deciding which papers to include, to the way researchers interpret results. Consideration must also be given to the types of evidence included in the review and how applicable these are to your clinical situations. An example literature review is referred to throughout to provide specific examples of applying these questions to support you in discerning the utility of literature review findings.

WHEN FACED WITH A CLINICAL QUESTION, general practitioners (GPs) are encouraged to turn to the evidence.[1] Yet scouring a body of literature can be resource intensive and impractical within GPs' time constraints. Literature reviews can provide a solution by collectively summarising key messages from multiple literary sources into a single document. Literature reviews take many forms ranging from describing the literature (narrative reviews) to mapping the types of research (scoping reviews) to quantitatively (meta-analysis) or qualitatively (meta-synthesis) integrating findings.[2] Fundamentally, all reviews provide an overview of existing knowledge on a specific topic by summarising and integrating publications, efficiently offering authoritative answers to questions. However, applying literature reviews' findings to clinical settings can be difficult because of their breadth. This article is designed to help GPs interpret literature reviews, arming them with key questions to guide their reading. For illustrative purposes, we draw on a published systematic review and meta-analysis by Nguyen et al[3] examining the association between dietary patterns with bone density and fractures in adults. We encourage you to read that review alongside this paper to give you practical experience in critiquing reviews. Box 1 provides definitions of key terms used in reviews, with glossary terms made bold throughout the text.

Critically exploring the methods of a literature review

Where, how and when did the authors look for studies?

A **literature review** is only as good as its search; an incomplete search will miss relevant evidence. The first step in evaluating a literature review is examining the **search strategy** (ie how the authors searched for studies). To do this, ask yourself where the authors searched, how the search was done and when the search was conducted. Nguyen et al describe these details in their 'Literature Search' paragraph on page 394.

Regarding where the authors searched, researchers will use various academic **databases**. Each has its own focus, although some overlap. Searching multiple databases ensures maximum coverage of available evidence. Nugyen et al searched Medline, Embase and CENTRAL, covering international content, and searched ProQuest for theses, which other databases sometimes exclude. Specifically searching for 'grey literature' (ie research findings not published in a peer-reviewed journal such as theses or reports) can help uncover different perspectives, although might not be held to the same scrutiny as journal articles.

© The Royal Australian College of General Practitioners 2026

How researchers search is by selecting terms or phrases within academic databases. These need to cover all the topics in the research question. Nguyen et al's search terms align with the topics of interest (dietary patterns, bone mineral density and fractures), suggesting the search would be comprehensive.

Given the breakneck pace of publications, searches need to be up to date. Delays between searching and publishing could make the findings outdated. The impact of delays depends on how much research is being published in the area. Looking at the age of included studies suggests how active researchers are in the area, but updating searches is best practice.[4] Nguyen et al ran their search in 2017, but their review was published in 2021; a 4-year gap between searching and publication. However, they noted little literature was published over this time (page 400, paragraph 3).

What did the authors include?

Eligibility criteria define the types of studies the authors decided would answer their question. These criteria concern attributes such as study types (eg randomised controlled trials, qualitative research), participants (eg age, having particular diagnoses), treatments, and health outcomes assessed. These criteria are sometimes framed as inclusion and exclusion criteria. The eligibility criteria help you decide if the findings will help answer your question.

Nguyen et al clearly define their eligibility criteria (page 394, paragraph 5). They only included studies with adult participants who did not have bone metabolism diseases or the medications affecting bone metabolism. Although these criteria are helpful for understanding the 'pure' effect

Box 1. A glossary of terms commonly used in literature reviews.

Term	Definition
Abstract	A paragraph that summarises the key details of the research article, often used as part of screening in reviews.
Bias	A systematic error distorting the accuracy of results within a study. Biases can emerge from study design, implementation and analysis processes.[12]
Database	A large collection of academic articles. Common databases in medicine include Medline, PubMed, Embase and CINAHL.
Eligibility criteria	The characteristics of studies that are eligible to be included within a literature review. This often includes the types of participants that can be included (eg participant age, gender, profession), study designs (eg randomised controlled trials, longitudinal studies, interview studies), date of publication (eg within the last 10 years), and study setting (eg outpatient community clinics), among other characteristics.
Forest plot	Following a meta-analysis, the forest plot displays the results graphically. This represents the individual studies and their weighting in the meta-analysis.
Heterogeneity	A representation of the variability in results between studies. In meta-analysis, this can be represented statistically and graphically.
Literature review	A broad term for any research project that provides an overview of existing knowledge on a specific topic by summarising and integrating publications. There are multiple types of literature reviews, varying in the processes used and outcomes generated.[2]
Meta-analysis	A statistical technique to combine and analyse quantitative results from multiple studies of similar research topics.
Narrative synthesis	A process of synthesising findings from multiple studies into an overarching narrative.
Reporting quality assessment	A procedure to determine the degree to which important details about study processes have been adequately reported. This is used as a proxy to assess the quality of the study itself. Various assessment tools are available.
Screening	A process to review studies against the review's eligibility criteria to filter out irrelevant studies returned from database searches.
Search strategy	A collection of key words and phrases relevant to the topics of interest that are entered into databases to find relevant articles.
Subgroup analysis	A statistical technique used within meta-analyses to examine effects within subgroups (eg specific patient groups, different study types, data collected in different countries). This is one method of examining contributors to heterogeneity.
Synthesis	A process of combining and integrating information from multiple studies into an overarching set of findings.
Systematic review	A formal, replicable method of consolidating literature on a particular topic to answer a specified research question. These reviews involve exhaustive, comprehensive searches to identify all relevant literature.

© The Royal Australian College of General Practitioners 2026

of the interventions they studied, if your question relates to children, patients with comorbidities, or patients taking medications that do affect bone metabolism, the review's findings might have limited clinical relevance for your patients.

How did the authors decide what papers to use?

Database searching can be imprecise, often returning many irrelevant articles. **Screening** is the next process to filter out irrelevant articles. It involves two phases of examining articles against the eligibility criteria: (1) a quick scan of the title and **abstract** to see if the article could be relevant; and (2) then reading the full paper of retained articles for a more in-depth decision. Screening should be completed by two independent reviewers to minimise errors, especially excluding relevant papers.[5–7] Disagreements between reviewers are ideally resolved by a third reviewer's opinion, or the two reviewers reaching consensus through discussion.

Two researchers in Nguyen et al's team reviewed all articles (page 395, paragraph 2), showing a robust screening process. Disagreements between these reviewers were resolved through consensus (page 395, paragraph 2). Although this approach increases feasibility, it can lead to **biased** decisions (eg if one reviewer is more senior, the junior reviewer might feel pressured to defer to the senior researcher).

How did the authors assess the quality of evidence they found?

Researchers are bound by practical constraints, influencing the quality of evidence that their studies produce. Even the most rigorous literature review will yield misleading results if it only draws on low quality evidence. To minimise this, researchers undertake a **reporting quality assessment** of a study (ie how much detail do authors provide about their study) as a proxy for the study's actual quality. Assessment checklists are available for different study types, allowing authors to focus on relevant details. Ideally, two reviewers assess reporting quality to minimise bias and errors. These assessments help authors understand the limitations of studies they include, or even inform which studies they include.

Nguyen et al assessed reporting quality using a tool specifically for musculoskeletal observational studies (page 395, paragraph 2). Using a topic-specific tool makes the assessment more meaningful by focusing on issues relevant to the subject (eg the rigour of dietary pattern assessments). The authors defined studies with a reporting quality score of 60% or over as being 'high quality'. Although in this case all studies were deemed 'high quality', when some studies do not meet this threshold, researchers can distinguish all studies' findings from high quality studies' findings. This provides nuance in understanding how much trust to place in the review's findings.

How did the authors make sense of the results?

Once researchers have found relevant studies, they use a process termed '**synthesis**' to combine all the information from different studies into an overarching set of findings. Synthesis can take many forms. One approach is **narrative synthesis**, where researchers tell the story of the findings. This ranges from a simple description of each study's findings to a rich narrative that groups ideas. **Meta-synthesis** applies qualitative techniques to the findings or raw data in individual qualitative studies, such as comparing and contrasting findings or generating themes.[8] **Meta-analysis** statistically combines data from individual quantitative studies, allowing researchers to robustly test research hypotheses with large datasets. To do this, studies must be similar in how they collect and handle data. In synthesising findings, reviews are taking findings out of their original context and aggregating them in novel ways. Consequently, reviewers must robustly recognise limitations of any overarching findings.

Nguyen et al used multiple synthesis approaches. They noted high variability in the data collected across studies (page 395, paragraph 4), so narratively synthesised findings from 21 studies, which are detailed in the results section and Table 2. They considered the similarity in dietary patterns reported within the available studies and concluded that only the 'healthy' dietary pattern was consistent enough across studies to meta-analyse.

Interpreting the results

Who was included?

The eligibility criteria tell you what you could find, but the results section tells us what the authors did find. Understanding details about what was found is critical for applying the review's findings. For example, Nguyen et al required study participants to be 18 years and older (page 395, paragraph 1). Table 1 in their study shows that just over half of their included studies comprised adults aged 50 years and over, suggesting the findings will largely represent this age bracket. If the patient you have in mind is aged 30 years, this might limit the findings' applicability. Similar considerations might be worthwhile for other contextual factors, such as included studies' settings (eg hospital vs community could affect treatment feasibility), countries (eg Western vs non-Western samples might have cultural implications) or gender composition (eg predominantly female vs more balanced).

How were outcomes measured?

Reviews sometimes require studies to measure outcomes in specific ways (eg only measuring distress via the Kessler psychological distress scale [K10]). Sometimes reviews allow different types of measures to be included; although this can broaden available findings, it also makes it difficult to determine if studies' findings differ because of real-world differences or variability in the instruments, questionnaires or clinical tests used. Understanding how consistent measurement is between the studies helps understand this nuance.

The differences are illustrated when comparing Nguyen et al's outcomes. Fractures were assessed either by self-report or medical records. The results of the meta-analysis showed that the findings were more consistent (ie less **heterogeneity**) when restricted to studies where fractures were determined only via medical records (page 396, paragraph 3).

What do the results say?

Sometimes a review's most clinically useful feature is quickly summarising relevant studies rather than providing an overarching answer. Some studies might be more relevant to your situation than others, meaning it is more appropriate

© The Royal Australian College of General Practitioners 2026

to focus on these studies' results instead of the overarching synthesis. Reviews often provide a table or appendix that summarises each study's characteristics and findings (ideally alongside its reporting quality). It is also worthwhile examining how consistent the results are; this is often graphically displayed in meta-analyses (Box 2).[9]

Nguyen et al summarised individual studies' findings in Table 2, categorising these by different outcomes of interest. However, they did not list the results and characteristics of each study, making it difficult for readers to evaluate the usefulness of each study for their context. For reference, Lim, Smith and Usherwood[10] provide such a table in their **systematic review** of barriers to spirometry.

How consistent are the results?

Studies are designed with different objectives in mind under different constraints and conditions. Collating these can be fraught, as these differences in how studies are designed produce variability in results. Differences also arise from real differences, which we are most interested in; these show where we can expect the effect of a treatment to differ between patients. Meta-analyses allow us to measure this variability (termed 'heterogeneity'). Broadly speaking, if authors report a large amount of heterogeneity, this suggests the effect of a treatment could vary a lot between patients on the basis of their characteristics.[13] Nguyen et al report high heterogeneity (page 396, paragraph 3). This heterogeneity is represented in their forest plot (Box 2; Figure 2).

What is it the results do not say?

As authoritative as reviews can be, they nonetheless have limitations. Many of the above questions help you consider the scope that the results can be applied to. These considerations can be expanded by reviewing the authors' own reflections on the review's limitations, both the available studies and the review itself. Likewise, it is worthwhile reviewing any declarations by authors of conflicts of interest, as these might bias results.

Nguyen et al highlight the limitations of their review in the central column of page 400. They recognise that they could only find cross-sectional data, which prevented them from assessing causal relationships. They also noted how their search was somewhat outdated, raising the possibility that newer evidence was not considered. The authors also made several declarations about personal fees and funding for the project, providing transparency about potential influences.

Conclusion

Literature reviews can authoritatively summarise evidence to inform clinical decisions, representing a go-to evidence source for busy clinicians. When reading a

Box 2. How consistent are the results? (How to interpret heterogeneity)

Study or Subgroup	Highest level Events	Total	Lowest level Events	Total	Weight	Risk ratio M-H, Random, 95% CI
Zeng 2013	132	326	266	459	24.5%	0.70 (0.60, 0.81)
Dai 2014	240	12631	438	12631	24.4%	0.55 (0.47, 0.64)
Fung 2015	480	22569	433	22569	25.1%	1.11 (0.97, 1.26)
Warensjo 2017	1002	14184	1501	14184	26.0%	0.67 (0.62, 0.72)
Total (95% CI)		**49710**		**49843**	**100.0%**	**0.73 (0.56, 0.96)**
Total events	1854		2638			

Heterogeneity: Tau² = 0.07; Chi² = 59.03; df = 3 (P < 0.00001); I² = 95%
Test for overall effect: Z = 2.26 (P = 0.02)

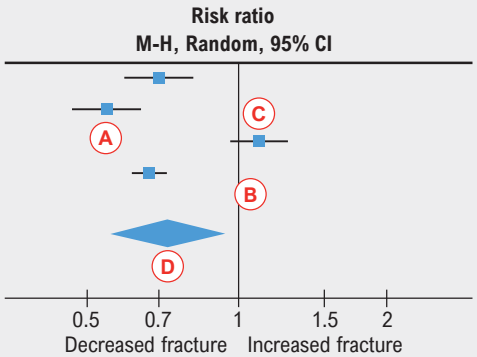


Figure 2. Forest plot of the association between healthy dietary pattern and hip fracture.

Figure 2 is taken from Nguyen HH, Wu F, Makin JK et al. Associations of dietary patterns with bone density and fractures in adults: A systematic review and meta-analysis. Aust J Gen Pract 2021;50(6):394–401.

Here we see that three studies found decreased risk of fractures for people following a healthy diet (ie squares **(A)** and to the left of the centre line **(B)**), but one study suggested the opposite **(C)**. The authors explored this heterogeneity using a 'subgroup analysis' with only studies that assessed fractures from medical records. From that, they found lower heterogeneity (ie more consistent findings).

The diamond **(D)** represents the collective findings; its centre corresponds to the estimated effect, and its width shows the range of effects one could expect from treatment of a healthy diet (ie graphically representing the total risk ratio, **(E)**). The range spans from a risk ratio of 0.56 (ie nearly halving the risk) to 0.96 (which corresponds to a 4% drop in risk of hip fracture (100% − 96 = 4)). The diamond does not cross the centre line, which indicates that the pooled result is statistically significant at the *P* <0.05 level.

Adapted from The Royal Australian College of General Practitioners, Associations of dietary patterns with bone density and fractures in adults: A systematic review and meta-analysis (www1.racgp.org.au/ajgp/2021/june/associations-of-dietary-patterns-with-bone-density), with permission.

© The Royal Australian College of General Practitioners 2026

review, be mindful of its processes to ensure the findings' relevance to your situation. More detailed information about critically appraising systematic reviews is available (Page et al).[11] For clinical applications, Box 3 summarises key questions to pose. Answering these questions will help you make an informed judgement about any literature review's merits and applications to your circumstances.

Box 3. Questions to ask when reading a literature review

Methods

- Where, how and when did the authors look for studies?
- What did the authors include?
- How did the authors decide what papers to use?
- How did the authors assess the quality of the evidence they found?
- How did the authors make sense of the results?

Results

- Who was included?
- How were outcomes measured?
- What do the results say?
- How consistent are the results?
- What is it the results do not say?

Key points

- Literature reviews summarise information from multiple studies on a specific topic and can take different forms.
- Literature reviews can support clinical decision making by efficiently and thoroughly describing current evidence.
- When interpreting literature reviews, clinicians should pay attention to the decisions made by researchers to ensure the clinical applicability of findings to individual patients.
- Close examination of results, particularly the consistency of findings, is warranted to consider the utility of findings.
- Sometimes it is more clinically appropriate to use reviews to identify individual studies relevant to particular clinical contexts rather than rely on overarching findings.

Authors

Shaun Prentice BPsychSc (Hons), MPsych (Clin), PhD MAPS, Research Officer, General Practice Training Research Team, The Royal Australian College of General Practitioners, Adelaide, SA; Clinical Associate Lecturer, School of Psychology, Adelaide University, Adelaide, SA

Ashley Blowes BND (Hons), MPH, Research Assistant, General Practice Training Research Team, The Royal Australian College of General Practitioners, Sydney, NSW

Andrew Killcross MBChB, FRACGP, Medical Educator (Practice Experience Program Specialist), The Royal Australian College of General Practitioners, Adelaide, SA; Honorary Clinical Senior Lecturer, Adelaide University Medical School, Adelaide, SA

Competing interests: None.

Funding: None.

Provenance and peer review: Commissioned, externally peer reviewed.

AI declaration: The authors confirm that there was no use of artificial intelligence (AI)-assisted technology for assisting in the writing or editing of the manuscript and no images were manipulated using AI.

Correspondence to:
shaun.prentice@racgp.org.au

References

References are available online only.

correspondence **ajgp@racgp.org.au**

© The Royal Australian College of General Practitioners 2026

References

1. The Royal Australian College of General Practitioners (RACGP). Standards for general practice. 5th edn. RACGP, 2017. Available at www.racgp.org.au/getattachment/ece472a7-9a15-4441-b8e5-be892d4ffd77/Standards-for-general-practices-5th-edition.aspx [Accessed 24 September 2025].
2. Grant MJ, Booth A. A typology of reviews: An analysis of 14 review types and associated methodologies. Health Info Libr J 2009;26(2):91–108. doi: 10.1111/j.1471-1842.2009.00848.x.
3. Nguyen H, Wu F, Makin J, et al. Associations of dietary patterns with bone density and fractures in adults: A systematic review and meta-analysis. Aust J Gen Pract 2021;50(6):394–401. doi: 10.31128/AJGP-02-20-5245.
4. Aromataris E, Lockwood C, Porritt K, Pilla B, Jordan Z, editors. JBI Manual for evidence synthesis – 2024 edition. JBI, 2024. Available at https://synthesismanual.jbi.global [Accessed 15 August 2025].
5. Porritt K, Gomersall J, Lockwood C. JBI's systematic reviews: Study selection and critical appraisal. Am J Nurs 2014;114(6):47–52. doi: 10.1097/01.NAJ.0000450430.97383.64.
6. Higgins J, Green SE. Cochrane handbook for systematic reviews of interventions version 5.1.0. The Cochrane Collaboration, 2011. Available at www.cochrane.org/authors/handbooks-and-manuals/handbook/current [Accessed 27 April 2026].
7. Stoll CRT, Izadi S, Fowler S, Green P, Suls J, Colditz GA. The value of a second reviewer for study selection in systematic reviews. Res Synth Methods 2019;10(4):539–45. doi: 10.1002/jrsm.1369.
8. Walsh D, Downe S. Meta-synthesis method for qualitative research: A literature review. J Adv Nurs 2005;50(2):204–11. doi: 10.1111/j.1365-2648.2005.03380.x.
9. Ried K. Interpreting and understanding meta-analysis graphs – A practical guide. Aust Fam Physician 2006;35(8):635–38.
10. Lim R, Smith T, Usherwood T. Barriers to spirometry in Australian general practice: A systematic review. Aust J Gen Pract 2023;52(9):585–93. doi: 10.31128/AJGP-09-22-6568.
11. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. BMJ 2021;372:n71. doi: 10.1136/bmj.n71.
12. Pannucci CJ, Wilkins EG. Identifying and avoiding bias in research. Plast Reconstr Surg 2010;126(2):619–25. doi: 10.1097/PRS.0b013e3181de24bc.
13. Borenstein M, Hedges LV, Higgins JP, Rothstein HR. Introduction to meta-analysis. John Wiley & Sons, Ltd, 2009. doi: 10.1002/9780470743386.

© The Royal Australian College of General Practitioners 2026