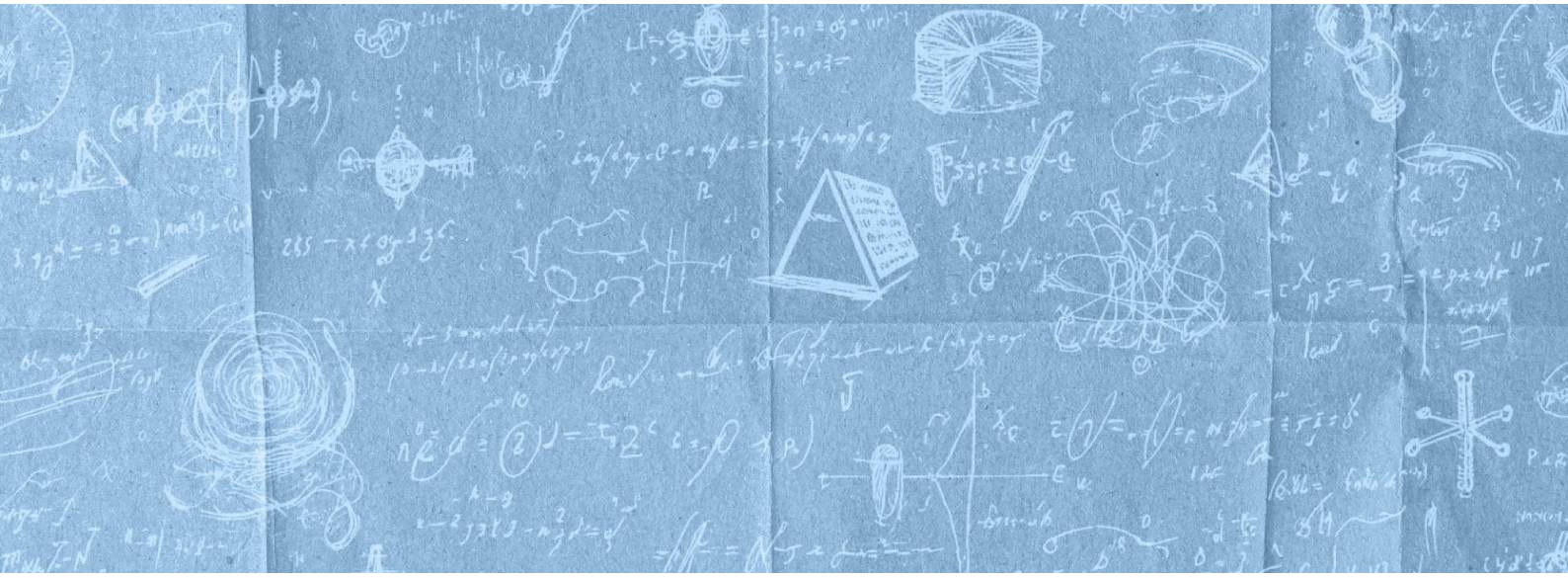
Consensus



CLASSROOM RESOURCES FOR FACULTY

Thinking Critically with AI

Practical guides for teaching students to use AI research tools with judgement — starting with the question they type in.

INSIDE

01 **Using AI to Build Judgment, Not Replace It**
Case study · Megan Heuer, Southern Methodist University Libraries

02 **Thinking Critically with AI in Research**
Using AI responsibly without outsourcing your judgement

03 **Framing Unbiased Research Prompts**
How the way you phrase a question shapes the evidence you find

04 **From Framework to Prompt**
PICO, SPIDER, PEO, ECLIPSE and CoCoPop, turned into prompts

05 **Fact-checking your Social Media Feed**
Practical guide · applying all of the above to a live claim

Guides written by **Dr. Benita Olivier**, Professor of Rehabilitation, Oxford Brookes University

CASE STUDY · 01

Using AI to Build Judgment, Not Replace It

How one university library rebuilt its research teaching around AI — and what it changed about the skills students actually leave with.

FEATURING

Megan Heuer

Director of Educational Initiatives, Southern Methodist University Libraries

WHY THIS COMES FIRST

The guides that follow are practical: how to frame a question, choose a framework, and test a claim. This case study is here to answer the question underneath them — why any of it is worth class time. It describes a librarian who reached the same conclusion from the other direction, by watching what students could and could not do. Read it as the argument for the rest of the pack.

Megan Heuer has spent her career teaching students to find and evaluate information. She is the Director of Educational Initiatives at Southern Methodist University's libraries, the liaison to its political science and public policy programs, co-designer of the university's Information Literacy Certificate, and author of a forthcoming research paper on how students read while they search.

On the topic of AI in the classroom, accuracy matters to her. But she sees a bigger problem, one that predates AI: research tools were already failing students before it showed up.

What AI Search Is Actually For

Heuer can walk into any library database and get what she needs, but only because she has spent a career learning to think the way the machine thinks. Most students never will. "We could teach until the end of time," she says, "and students will still not know how to find things." Instruction cannot scale to every freshman, and the students who struggle most are the ones drowning in results from searches they did not know how to run.

That reframe is why she adopted Consensus in her teaching. Her image for it is a student throwing a dart at a dartboard: rather than train them to aim from a distance, she moves the target closer.

> **"I just brought the dartboard closer, so they can think a little more about what's actually relevant, because they're less overwhelmed by everything else."**
>
> MEGAN HEUER

The payoff is not less teaching but different teaching. Cutting the time students spend fighting search syntax frees class time for the work she considers central: teaching students how to evaluate sources, judge relevance, refine their topics, and understand how a body of research fits together. This is especially important in the early stages of the research process where discovery is more fluid and open-ended, and you can still move to traditional databases when you require more specificity and control.

The Assignment: From Broad Topic to Concept Map

Her core assignment asks students to develop a research question on a broad topic, say gerrymandering in the United States. Before writing any question, the student has to learn the shape of the existing research, because, as she puts it, "ideas don't come out of a vacuum. They can't come up with some great focused research question out of nothing."

The deliverable is a concept map: a diagram the student draws by hand, with the field's subtopics as nodes and lines showing how they connect. Building one forces students to categorize what they find and to see the field as a structure rather than a list.

The problem was that novices rarely got far enough to start. Inferring a field's structure from pages of unsorted database results is an expert skill, and first-year students would stall on it, spending the session fighting the interface and scrolling lists of results instead of thinking about the field.

So the assignment now opens with a sequence of prompts in Consensus. She starts students broad, then makes them iterate: ask a narrower question, then narrower again, until they are working inside a subtopic of a subtopic.

> **Map the field.**
> `"What are the areas of research within gerrymandering in the United States?"`

> **Drill for specificity.**
> `"How does gerrymandering affect voter turnout?"`
> `"Does gerrymandering reduce turnout among young voters specifically, and do studies disagree?"`

Then come two lessons a traditional database cannot demonstrate. The first is disagreement. Students take a claim and prompt for the studies that contradict it. "That, particularly in databases, is impossible," she says, and it solves a problem she sees constantly: the student who cannot find counterevidence for their argument.

> **Hunt the counterargument.**
> `"Show me studies that disagree with this."`

The second is time. She prompts for how research on a topic has changed, which teaches the concept the ACRL Framework for Information Literacy calls scholarship as conversation: knowledge as an ongoing exchange rather than a stack of isolated facts. Colleagues who teach first-year writing use the citation graph to put that conversation on screen to show how papers connect and reference each other.

> **Trace the conversation.**
> `"How has research on this topic changed over time?"`

Then choose a paper and show the citation graph in Consensus, to demonstrate how previous papers are connected to new ones.

> **"It's like seeing the visualization of the conversation. I use that as a teaching point."**
>
> MEGAN HEUER

Preventing AI from becoming a crutch

Heuer will not blame AI for the reading problem. Students' reading skills were declining before these tools arrived, driven by phone use, internet reading habits, and what they carried up from K to 12. What AI adds is a way to fake the reading.

> **"AI becomes a crutch for students not reading."**
>
> MEGAN HEUER

Students who once skipped course readings can now pretend to have done them by dumping the text into a chatbot. Her answer is not a ban but a drill she runs in class: read the abstract, decide whether it is relevant, defend the decision. "No matter what the AI does," she tells students, "you still have to be able to read the summary and judge it." An abstract has always demanded the same evaluative reading an AI summary does now; the underlying skill has not changed.

Her larger argument is the one faculty can build a syllabus around: evaluation cannot be delegated to a machine by someone who has not developed it, because using AI well consists of evaluating what it returns.

> **"You ultimately don't know anything. You have to develop those evaluative skills to begin with."**
>
> MEGAN HEUER

Consensus fits this classroom because it works in the same direction as her teaching. It does not write the paper; it searches peer-reviewed literature and returns targeted summaries of what the research says, with every claim linked to its source. For a student, that summary is not a way around the reading but a way into it: a short, readable account of the state of the research that points them to the specific papers worth their attention. It strengthens the judgment it depends on; it does not replace it.

Designing the Assignment So It Holds

Heuer also runs a faculty cohort program, with a $1,000 stipend, to redesign research assignments that chatbots cannot fake.

Its most transferable finding is about assessment: if you want students to read, the reading has to be graded and verified.

> **"It has to be in the assessment, and you have to check it."**
>
> MEGAN HEUER

In practice, the cohort's faculty require students to turn in the annotated source itself, a PDF marked up with their own comments; they require page numbers; and they require a direct connection between every claim and the passage that supports it. The design does not fight AI with detection. It makes faking engagement more work than doing it.

A tool that helps students find the right peer-reviewed source is doing what libraries have always wanted, and several faculty at SMU now write Consensus into their assignments as a sanctioned way to use AI in research.

It is also how use spread on campus: faculty training other faculty and librarians teaching it in class took SMU from near zero to about 2,000 Consensus users in the tool's first year on campus.

Heuer's ambition is not that students search better. It is that they read more, argue more carefully, and learn to weigh evidence, skills that were under strain long before AI arrived. The tool she reaches for is the one that clears the busywork and then gets out of the way of the judgment and real learning.

GUIDE · 02

Thinking Critically with AI in Research

A practical guide to using AI responsibly in academic research, with principles for maintaining intellectual rigor, verifying evidence, and protecting research integrity.

WRITTEN BY

Dr. Benita Olivier

Professor of Rehabilitation, Oxford Brookes University

IN THIS GUIDE

01 Why this matters

02 The core principle: augmentation, not substitution

03 AI outputs are not evidence

04 Appropriate and inappropriate use

05 Systematic reviews and evidence synthesis

06 Qualitative research: specific considerations

07 Practical rules for responsible use

08 Data, transparency, and institutional alignment

09 Decision-making checklist

01

Why this matters

AI is now part of the research landscape. The challenge is not whether to use it, but how to use it well. Used thoughtfully, AI can support learning and sharpen thinking. Used uncritically, it can undermine the very skills researchers are trying to develop.

Most researchers reach a point where AI produces something genuinely impressive — a summary, a structure, a first draft — in the time it takes to make a cup of tea. The reaction is often mixed: relief, and then a flicker of unease. If it can do that, what is my role here?

This guide does not sidestep that question. It is built around it. Because how you use AI — and what you allow it to do on your behalf — will shape the kind of researcher you become.

> DISCLAIMER
>
> Read this guide in the context of your own institution, discipline, and programme.

02

The core principle: augmentation, not substitution

There is a distinction worth noting as you read this guide: the difference between doing research and directing it.

Doing is the execution layer — the searching, drafting, summarising, and formatting. AI is already competent here, and increasingly so. Directing is something else: deciding which question deserves your attention, recognising which direction within a project is most promising, knowing when a result is genuinely good rather than merely plausible. Directing requires a deep familiarity with your field — a feel for the terrain that tells you where to go next.

This is important for both ethical and practical reasons. AI is most powerful in the hands of researchers who can direct it well — who know what to ask for, how to evaluate what comes back, and what it means for the work they are actually trying to do. That directing capacity does not come from having access to the tools. It comes from having developed the judgement to use them. And judgement develops through engagement with the substance of research — exactly the engagement that AI makes it easiest to skip.

PAUSE AND REFLECT

What are you developing?

Think about the research tasks you have used AI for recently — or are tempted to. For each one, ask: is this a task where the doing itself builds something in me? Or is it genuinely routine?

The goal is not to avoid AI. It is to be honest about when the engagement matters — and when it does not. Some tasks are worth protecting. Others not.

AI should support your thinking, not replace it.

03

AI outputs are not evidence

Evidence-based practice depends on a hierarchy of evidence — systematic reviews and meta-analyses at the top, expert opinion and anecdote at the bottom. AI outputs occupy no recognised place in this hierarchy. They are not a source of evidence. They are a tool for processing, organising, and communicating evidence that must originate elsewhere.

In practice, this means:

- Any claim that originates from or has been shaped by AI must be traced back to a verified primary source before it can be treated as a research finding.
- AI-generated summaries of the literature reflect patterns in training data, not a systematic or methodologically sound synthesis. They cannot substitute for a proper literature review.

04

Appropriate and inappropriate use

The distinction often comes down to whether you remain cognitively active — interpreting, questioning, and deciding. This applies across all phases of research, not just the writing-up stage: data collection, instrument design, qualitative analysis, and knowledge translation into practice are equally governed by these principles.

✓ Appropriate use	✕ Not appropriate
Reading a paper yourself, then using AI to challenge or refine your understanding	Asking AI to summarise a paper and relying on the summary without reading it

✓ Appropriate use	✕ Not appropriate
Using AI to create summaries during initial scoping — before deciding whether a paper is relevant	Citing a study based on the findings in the summary without reading the paper
Using AI to suggest structure when the argument is already yours	Asking AI to write or construct the argument on your behalf
Asking AI to improve the clarity of a sentence you have written	Accepting AI-generated text containing reasoning or claims you did not generate
Using AI to help draft interview questions, then reviewing them against your methodology	Using AI to generate themes from qualitative data without engaging directly with participants' words
Asking AI to explain a difficult concept, then testing your own understanding	Asking AI to appraise a paper's methodology on your behalf and accepting that appraisal unchecked
Using Consensus to discover relevant papers, then reading and critically appraising them	Treating Consensus results as a synthesis of evidence without reading the primary sources

A useful test: your name is on the work. The expectation is that the ideas, reasoning, and interpretations are your intellectual contribution. If AI is constructing the argument or speaking in your place, it is no longer supporting your work — it is replacing it.

PAUSE AND REFLECT

Are you still in the driver's seat?

After your next AI-assisted session, try this test: set the output aside and write — without referring to it — one paragraph explaining the argument, finding, or idea in your own words.

If you can do it fluently, you were cognitively present. If you struggle, that is useful information: not a reason for guilt, but a signal that the engagement needs to go deeper before you move on.

05

Systematic reviews and evidence synthesis

Systematic reviews are the gold standard of evidence synthesis in health research. They operate under strict methodological protocols — including PRISMA reporting standards, PROSPERO registration, structured database searching, and formal risk of bias assessment. AI is increasingly being used at various stages of this process, and the risks are specific and significant.

What AI cannot do in a systematic review

- **Replace structured database searching.** Clinical databases — MEDLINE, CINAHL, Cochrane, EMBASE — remain the methodological standard for evidence retrieval. AI search tools do not currently meet the requirements for reproducibility, transparency, and comprehensiveness in systematic review methodology.
- **Screen abstracts or extract data without validation.** If AI is used to support these tasks, the process must be validated against human screening, and the method must be explicitly reported.
- **Conduct risk of bias assessment.** This requires methodological expertise, contextual judgement, and accountability that AI cannot provide.

> **DECLARATION IN METHODS**
>
> Any AI involvement in a systematic review — including abstract screening, data extraction, or synthesis support — must be declared explicitly in the methods section, not simply noted in a general disclosure statement. Reviewers and readers need to be able to evaluate the methodological implications.
>
> The methodological landscape is evolving, and a small but growing number of published reviews have incorporated AI with explicit validation and reporting. If you are considering AI involvement, seek current field-specific guidance — but the principle of transparency is non-negotiable.

06

Qualitative research: specific considerations

The interpretive, reflexive engagement with participants' words and experiences is not a preliminary step in qualitative research — it is the research. Using AI to generate themes from interview data, summarise field notes, or suggest codes risks introducing researcher distance from the data and undermining the methodological rigour that gives qualitative work its validity and trustworthiness.

This does not mean AI has no role in qualitative research. It may legitimately support literature searching, help structure a methodology section you have already drafted, or improve the clarity of your writing. But the analytic process itself — the sustained, reflexive engagement with data — must remain yours.

DATA PROTECTION IN QUALITATIVE WORK

Uploading participant interview transcripts or field notes to a public AI system may breach your ethical approval, your institution's data protection policy, and GDPR — even when participants are not named. Contextual details in qualitative data can be identifying. When in doubt, do not upload.

07

Practical rules for responsible use

1

Stay cognitively active

Question AI outputs, compare them to your own understanding, and refine them. Sit with uncertainty rather than using AI to generate premature answers — tolerance of ambiguity is a genuine research skill, and one that AI can quietly erode.

2

Verify everything — and know where to look

AI can fabricate citations, misattribute ideas, and inaccurately summarise research — generating authoritative-sounding text with no basis in reality.

WHAT A HALLUCINATED CITATION LOOKS LIKE

AI might generate: '`Smith, J. (2019). Methodological pluralism in qualitative health research. Journal of Nursing Studies, 45(3), 112–128.`' — complete with volume and page numbers. That paper may not exist. Always locate sources before citing them.

3

Use AI after thinking, not before

Form your own view and attempt the task before involving AI. This preserves the learning value and ensures that what AI supports is genuinely yours.

4

Protect your voice

If a revised passage no longer sounds like you — or contains reasoning you did not generate — revise it until it does.

5

Know your tools

There is an important difference between retrieval-augmented tools and generative AI. Consensus searches indexed academic databases and returns real, citable papers — making it significantly more reliable for source discovery than a general-purpose language model. But retrieving real papers is not the same as synthesising evidence. Consensus can show you what exists; it cannot tell you whether those studies are methodologically sound, whether their findings are applicable to your context, or what they collectively mean. That judgement is yours.

General-purpose large language models (such as ChatGPT or Claude) generate text based on their training data and can fabricate sources with confidence. They are useful for drafting, explanation, and exploration — not for evidence retrieval. Using them to identify or summarise research without verification is methodologically indefensible.

Use the right tool for the right purpose. And with every tool, read the original.

6

Keep a reflective log

Answer these questions in your reflective log: what did I use AI for, what did I accept or modify, what was my reasoning? This is good reflective practice, and it provides a defensible record should your methods ever be questioned.

7

Check guidance on grant writing and ethics applications

AI use in grant applications and research ethics submissions is an area where institutional policies are often unclear or still developing. Check with your institution before using AI to draft these documents, as some funders and ethics committees have specific requirements or restrictions.

08

Data, transparency, and institutional alignment

Do not input unpublished data, participant information, interview transcripts, or any confidential material into public AI systems. Doing so may breach your ethical approval, your institution's data protection policy, and GDPR.

Disclose AI use as required — but note that disclosure does not make a use acceptable. Policies vary across institutions, programmes, and journals, and are still evolving. Disciplinary norms also vary: what is standard in computational research may be inappropriate in qualitative health research. Seek field-specific guidance alongside this framework.

Journal and publisher policies on AI-assisted writing differ, including on which tools must be declared. Check these before submission.

Before inputting anything that touches participant data into any AI tool, ask three questions:

1. Does my institution have an approved agreement with this provider?
2. Does my ethics approval permit this use?
3. Could any of this data identify a participant, even indirectly?

09

Decision-making checklist

When using AI, or including AI-supported work in your research:

#	When using AI — ask yourself	Response
1	Am I using AI to support my thinking, or to replace it?	Yes / No / Unsure
2	Do I understand the output — could I explain it confidently in my own words?	Yes / No / Unsure
3	Have I verified all claims and references against a reliable, verified source?	Yes / No / Unsure
4	Does this reflect my own voice and reasoning?	Yes / No / Unsure
5	Is this aligned with my institution's, programme's, or journal's current guidance?	Yes / No / Unsure
6	Have I checked this against my ethical approval where data is involved?	Yes / No / Unsure
7	If this involves evidence synthesis, have I used appropriate clinical databases?	Yes / No / Unsure
8	Have I declared AI use in my methods where required — not just in a general statement?	Yes / No / Unsure
9	Could I defend this work as genuinely my own intellectual and professional contribution?	Yes / No / Unsure

If any answer is 'no' or uncertain, pause and re-engage with the task.

PAUSE AND REFLECT

The researcher you are becoming

Looking at your research over the past few months — not the outputs, but the process — where do you feel your judgement and expertise are genuinely developing? Where do you feel it is being quietly outsourced?

Researchers who stay intellectually in charge of AI are not those who use it least — they are those who remain curious about what they are becoming in the process.

A FINAL THOUGHT

The researchers who benefit most from AI are not those who use it most — but those who use it most wisely, keeping their own intellectual development and professional accountability at the centre.

You are the researcher. AI is the tool. Keep it that way.

GUIDE · 03

Framing Unbiased Research Prompts

How the way you phrase a research question shapes the evidence you find, and how neutral questions help surface more balanced scientific results.

WRITTEN BY

Dr. Benita Olivier

Professor of Rehabilitation, Oxford Brookes University

IN THIS GUIDE

01 The tool reflects the question

02 Leading vs neutral: see the difference

03 Common pitfalls to avoid

04 A simple formula

05 Your pre-search checklist

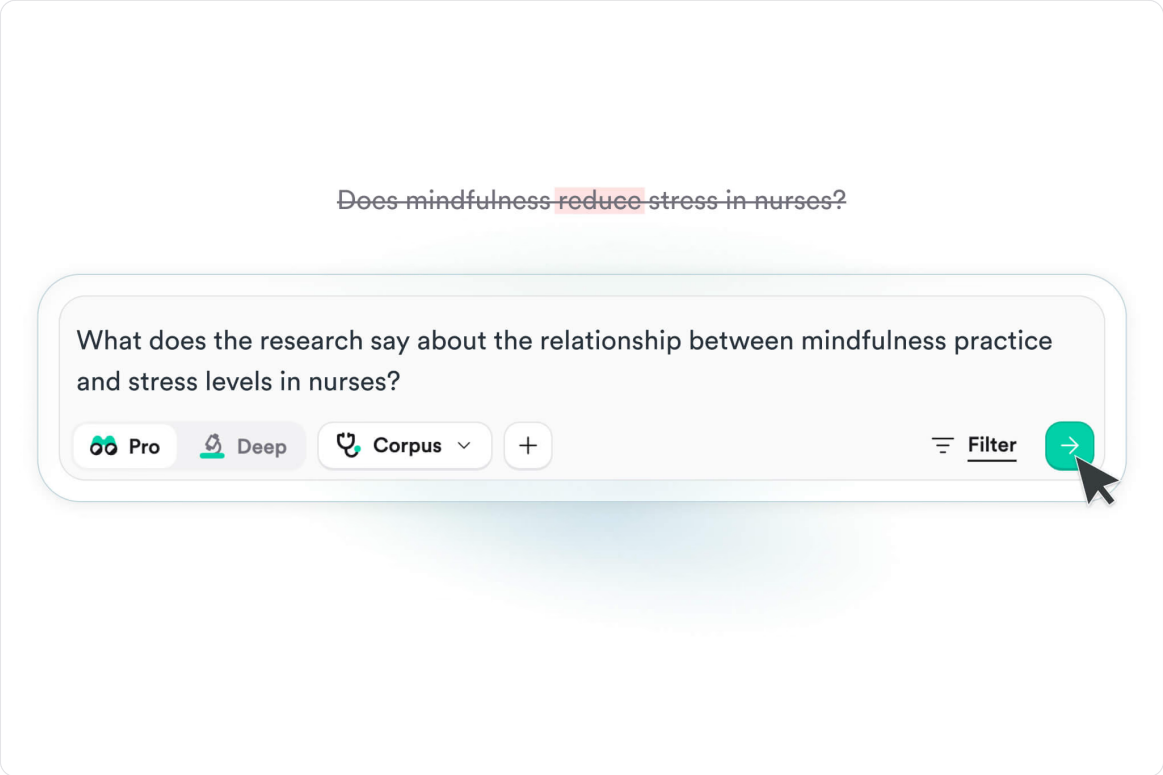
06 What comes next

01

The tool reflects the question

The way you phrase a question shapes the evidence you find. Ask a leading question — one that assumes a direction of effect — and you will tend to surface confirmatory evidence. Ask a neutral question — one that opens the topic up rather than closing it down — and you will get a fuller, more balanced picture.

This is important because AI tools do not correct for the direction of your question. When you type a question into an AI research tool, you are not just searching — you are directing it. The words you choose, and the assumptions baked into them, shape what comes back.



The same topic, asked neutrally. The rewrite drops the assumed direction of effect and lets the evidence answer.

02

Leading vs neutral: see the difference

The examples below show the same topic asked two ways. Notice how the leading version assumes an outcome before any evidence has been reviewed.

✕ Leading question	✓ Neutral question
Does social media **cause** depression in teenagers?	What is the relationship between social media use and depression in teenagers?
Does mindfulness **reduce** stress in nurses?	What does the research say about the relationship between mindfulness practice and stress levels in nurses?
Is online learning **worse** than face-to-face teaching?	How does online learning compare with face-to-face teaching in terms of student outcomes?
Does poor sleep **harm** academic performance?	What is the evidence on the relationship between sleep duration and academic performance in students?

THE WORDS THAT GIVE IT AWAY

Cause, **reduce**, **worse**, **harm** — each one assumes an answer before any evidence has been considered. The neutral version replaces that assumption with an open question: what does the evidence actually say?

03

Common pitfalls to avoid

Leading questions are the most common problem, but they are not the only one. Here are five pitfalls that can undermine your search before it begins.

Pitfall	What goes wrong	Quick fix
Leading questions	Assumes the answer in the question itself	Does x help y? → What is the evidence on the relationship between x and y?
Built-in assumptions	Treats an unproven claim as fact	Why does stress cause burnout? → What is the relationship between stress and burnout in healthcare workers?
Vague phrasing	Too broad to surface useful evidence	Is technology bad for you? → What does the evidence say about the effects of screen time on sleep quality in adolescents?

Pitfall	What goes wrong	Quick fix
Double-barrelled questions	Two questions hiding in one	`Does exercise improve mood and concentration?` → ask separately: `…the relationship between exercise and mood in adults?` and `…between exercise and concentration in adults?`
Jargon and abbreviations	Misread by AI; limits results	`Does CBT reduce GAD?` → `What is the effectiveness of cognitive behavioural therapy for generalised anxiety disorder symptoms in adults?`

04

A simple formula

If you are unsure whether your question is neutral, try this reframe:

INSTEAD OF	“Does X help Y?”
TRY	**“What is the evidence on the relationship between X and Y?”**

You can make it more specific by adding context: who the population is, what the setting is, or what outcome you care about. That specificity is what takes a decent question and makes it a good one.

Putting it into practice: a worked example

Here is how you can move from a first instinct to a well-formed question.

1 FIRST INSTINCT

`“Does social media cause anxiety in university students?”`

2 WHAT’S WRONG

The word “cause” assumes a causal relationship that the evidence may not support. “Social media” is also vague — it could mean passive scrolling, active posting, or comparison behaviour, all of which the literature treats differently.

3 APPLY THE FORMULA

`“What is the evidence on the relationship between social media use and anxiety symptoms in university students?”`

4

ADD SPECIFICITY

“What does the evidence say about the relationship between passive social media use and anxiety symptoms in undergraduate students aged 18–25?”

WHY THE FINAL VERSION WORKS

It is open, specific, and free of assumptions, so it will surface a wider and more useful range of evidence than the original. Studies that fit your exact question may not exist; Consensus will give you the closest available research, because it now knows what you are looking for.

05

Your pre-search checklist

Before you search, ask yourself:

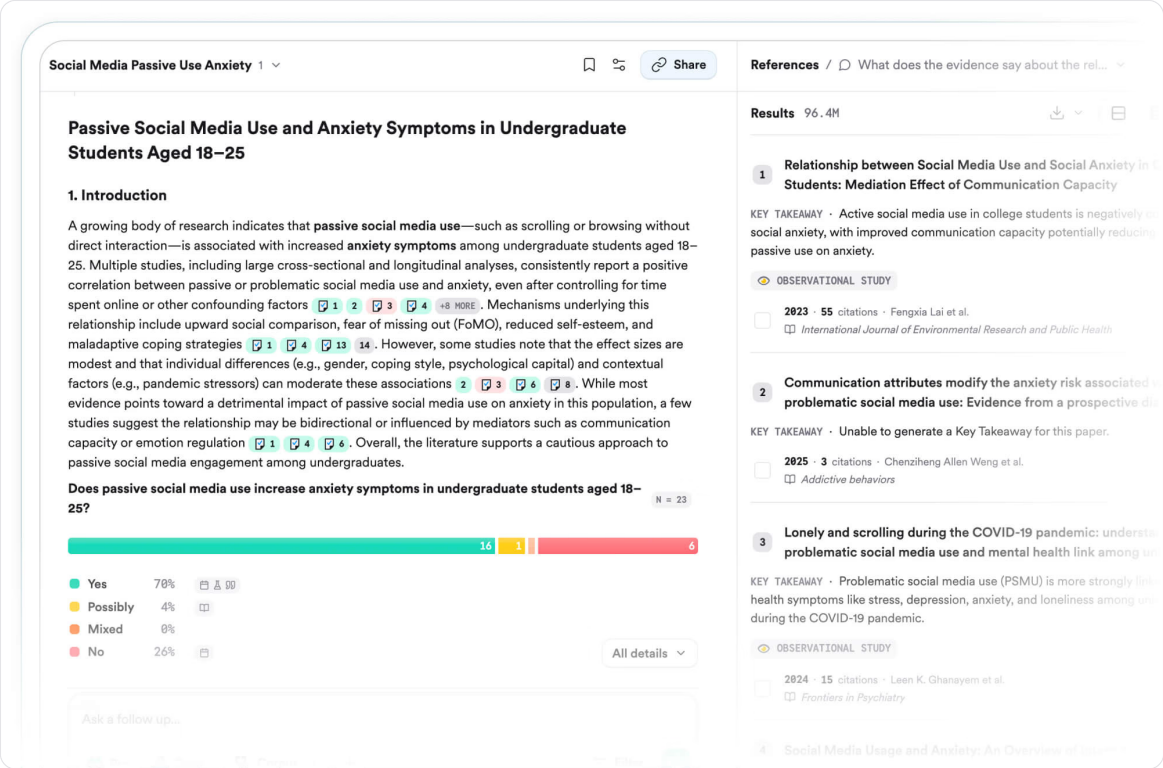
- [ ] Does my question assume a direction of effect — help, harm, improve, worsen?
- [ ] Does it contain a hidden assumption I haven’t tested?
- [ ] Is the population specific enough — age, setting, condition?
- [ ] Is it one question, or two questions in disguise?
- [ ] Have I used plain language rather than jargon or abbreviations?
- [ ] Would a different phrasing of this question lead to a different answer?

Try it with your own topic

Write your first-instinct question, then rewrite it using the formula from section 04.

06

What comes next



What a neutral, specific question returns: a synthesis with citations, a spread of findings across the literature, and the individual papers behind it.

Once you can ask a neutral, well-framed question, you are ready to go further. The companion guide **From Framework to Prompt** introduces structured frameworks — PICO, SPIDER, PEO, and others — that help you build on the skills practised here. You can find it at `consensus.app/home/resources/prompting-for-consensus`.

The ability to ask a good question is not a small thing. It is where all good research begins.

A FINAL THOUGHT

AI does not think for you — it starts where you point it. The quality of your question is the quality of your research.

GUIDE · 04

From Framework to Prompt

Learn how to write clear, unbiased research questions for Consensus using evidence-based frameworks like PICO, SPIDER, PEO, ECLIPSE, and CoCoPop.

WRITTEN BY

Dr. Benita Olivier

Professor of Rehabilitation, Oxford Brookes University

IN THIS GUIDE

– Why your question is important

– AI and bias: what you need to know

– Choose the right framework for your question

01 PICO — for intervention and effectiveness questions

02 SPIDER — for qualitative and mixed-methods questions

03 PEO — for etiologic, risk, and cause-and-effect questions

04 ECLIPSE — for health service and policy questions

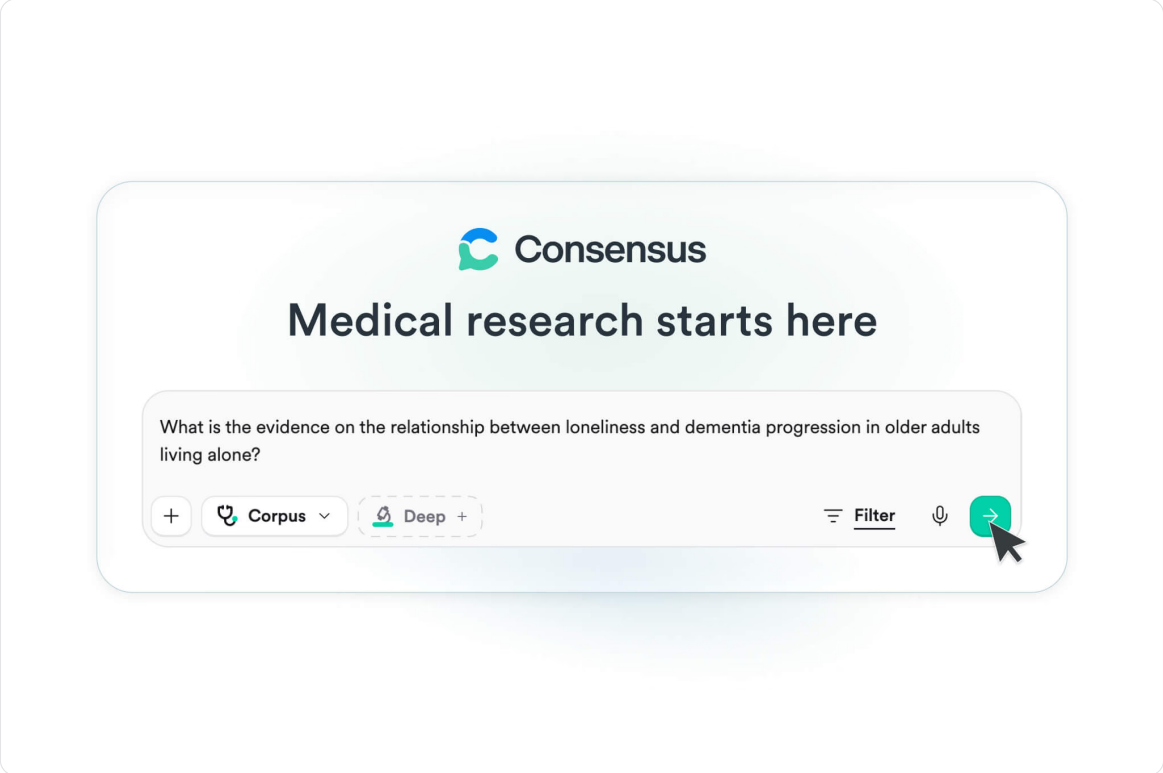
05 CoCoPop — for incidence and prevalence questions

– Your prompting checklist

Why your question is important

When you type a question into Consensus, you are instructing an AI system to search, filter, and synthesise evidence from over 200 million peer-reviewed papers on your behalf. What you put in shapes what you get out.

Consensus is not a chatbot. It searches real scientific literature and uses AI to summarise findings — with citations. It will not fabricate sources. But it can retrieve irrelevant results if your question is vague, and it can misinterpret a paper if your question misleads it. The quality of the evidence you find begins with you.



A question that names its population, its exposure and its outcome, and assumes nothing about the direction of the result.

AI and bias: what you need to know

Every AI tool reflects the world it was trained on — and that world is imperfect. There are three types of bias you need to be aware of:

- **Bias in the published literature.** Research has historically under-represented women, older people, ethnic minorities, and populations in low-income countries. AI searches what exists — it cannot correct for what was never published.
- **Bias in your question.** A leading question (“Does exercise help anxiety?”) will surface confirmatory evidence. A neutral one (“What is the evidence on the relationship between exercise and anxiety?”) will surface the full picture.

- **Publication bias.** Positive results are more likely to be published than neutral or negative ones. Strong-looking evidence may reflect a skewed literature rather than clinical truth. This is relevant whether you use AI or traditional databases to find research.

> **THE PRINCIPLE TO REMEMBER**
>
> AI reflects the biases of the world it learned from. Your job is to ask questions that are neutral in language, specific in scope, and aware of whose evidence may be missing.

Choose the right framework for your question

Before you prompt, identify what type of question you are asking, then select the appropriate framework.

- Effectiveness of a treatment or intervention — use **PICO** or **PICOS**
- Patient or service user experiences, perceptions, or attitudes — use **SPIDER**
- A naturally occurring exposure, risk factor, or association — use **PEO**
- Health service evaluation, policy, or management — use **ECLIPSE**
- Incidence or prevalence of a condition within a population — use **CoCoPop**

The frameworks

Hosseini, M-S., Jahanshahlou, F., Akbarzadeh, M.A., Zarei, M., & Vaez-Gharamaleki, Y. (2024). Formulating research questions for evidence-based studies. Journal of Medicine, Surgery, and Public Health, 2, 100046. https://doi.org/10.1016/j.glmedi.2023.100046

01

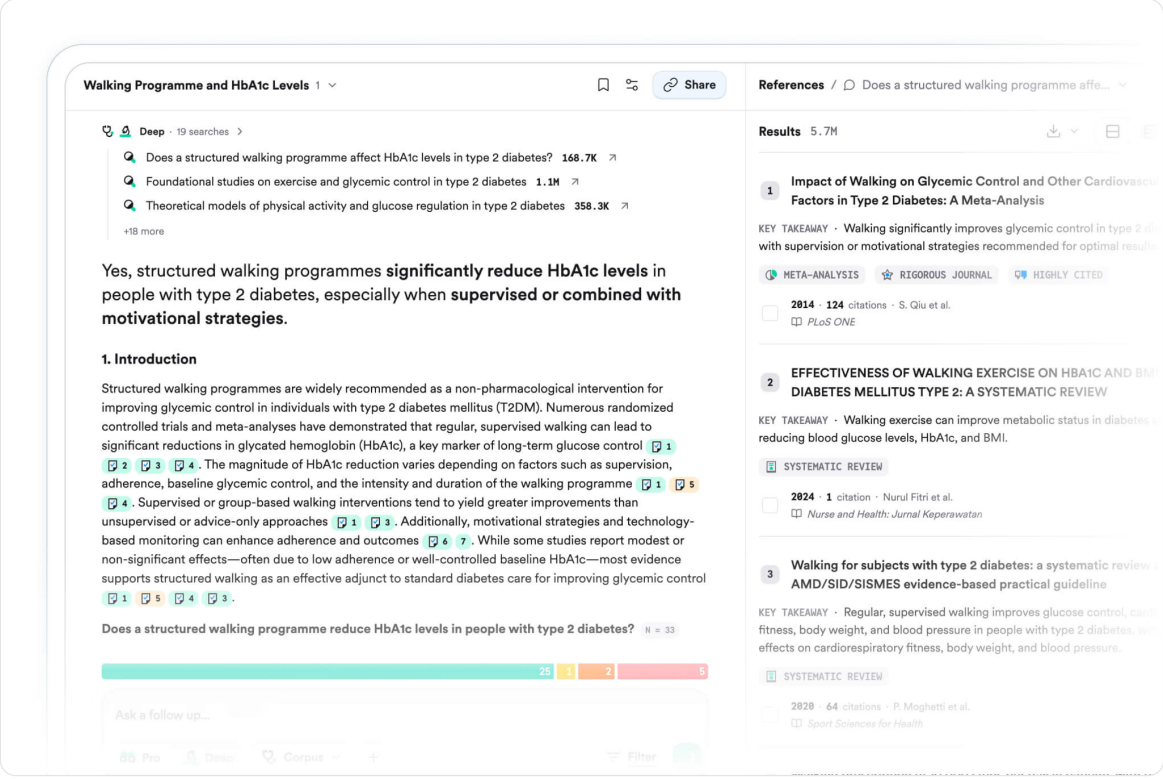
PICO — for intervention and effectiveness questions

Use PICO when you want to evaluate the effectiveness of a treatment, intervention, or clinical approach.

- **P — Patient / Population:** Who are you studying? Be specific — include age group, diagnosis, setting, and relevant characteristics, for example, “registered nurses in intensive care units”, rather than just “healthcare workers”.
- **I — Intervention:** What is the main intervention or treatment being considered? Use generic clinical terms, rather than brand names.
- **C — Comparison:** What is it being compared to? This might be standard care, a placebo, or an alternative intervention. If there is no comparator, state this explicitly.
- **O — Outcome:** What are you trying to measure or achieve? State this neutrally — “effect on wound healing rates” rather than “improvement in wound healing”.

- **PICOS adds S — Study design:** Specify the type of study you are looking for (e.g., randomised controlled trial, systematic review) when evidence level matters.

> **EXAMPLE**
>
> In adult patients with type 2 diabetes (P), does a structured walking programme performed five days per week (I), compared with no structured exercise (C), affect HbA1c levels over 12 weeks (O)?



The PICO example run in Consensus.

To note: Consensus may simplify the question above to: “Does a structured walking programme affect HbA1c levels in type 2 diabetes?” and show you the Consensus meter. The output will still be relevant to your search query, and the act of you specifying each component in the question means you know what to focus on when reviewing the output.

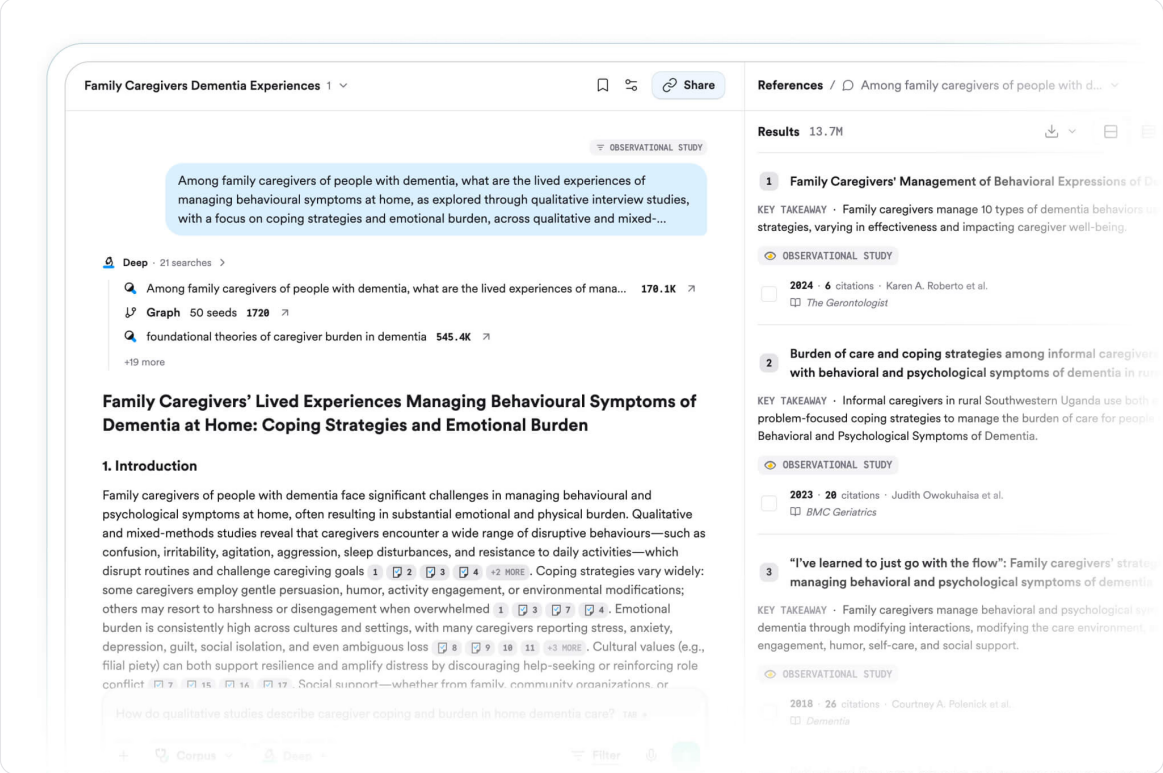
02

SPIDER — for qualitative and mixed-methods questions

Use SPIDER when you are exploring experiences, perceptions, attitudes, or behaviours, and your evidence will come from qualitative or mixed-methods studies.

- **S — Sample:** Who is being studied? In qualitative research, “sample” is preferred over “population”, as participants are typically purposively selected rather than randomly sampled.
- **PI — Phenomenon of Interest:** What experience, behaviour, or situation is being explored? Use neutral, descriptive language and avoid making assumptions about what the experience involves.
- **D — Design:** What study design are you looking for? For example: phenomenology, grounded theory, ethnography, qualitative interview study.
- **E — Evaluation:** What is being evaluated or measured? This might include experiences, perceptions, attitudes, behaviours, or processes.
- **R — Research type:** Are you looking for qualitative, quantitative, or mixed-methods studies?

> **EXAMPLE**
>
> Among family caregivers of people with dementia (S), what are the lived experiences of managing behavioural symptoms at home (PI), as explored through qualitative interview studies (D), with a focus on coping strategies and emotional burden (E), across qualitative and mixed-methods designs (R)?



The SPIDER example run in Consensus.

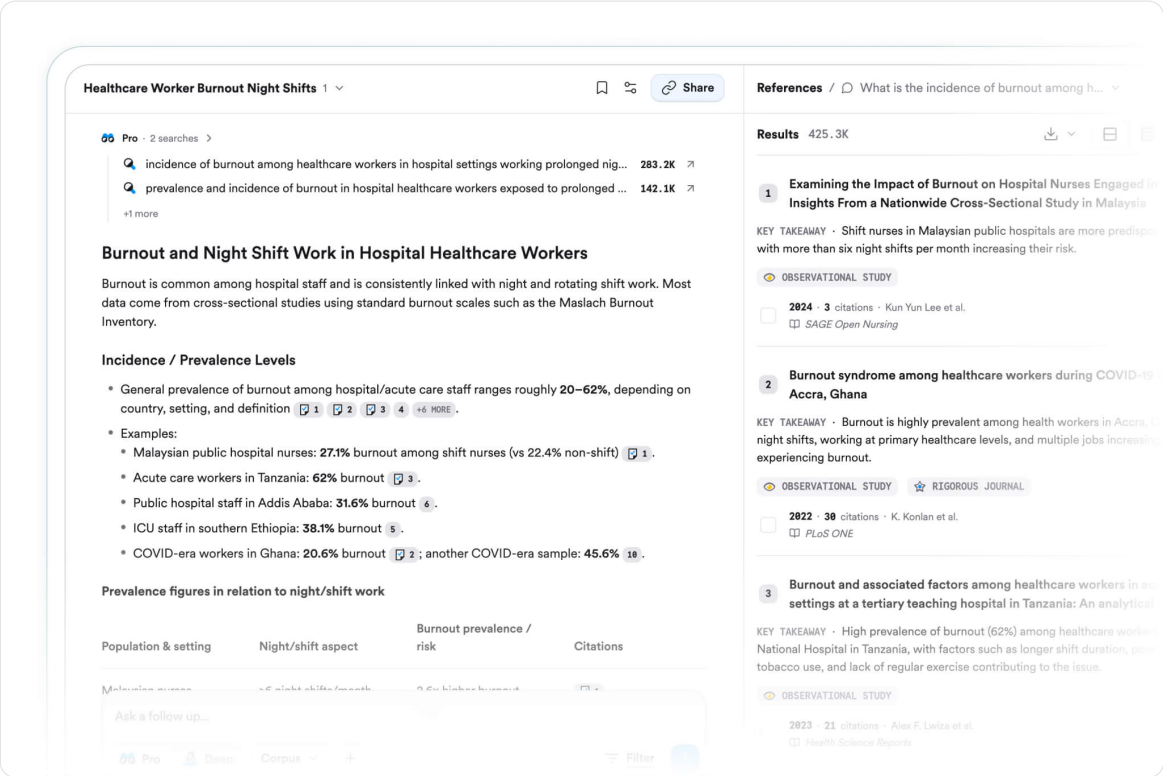
03

PEO — for etiologic, risk, and cause-and-effect questions

Use PEO when you are investigating the relationship between an exposure or environmental factor and a health outcome. The exposure is something participants experience naturally — not a deliberate intervention.

- **P — Population:** What groups of population or patients with common characteristics are targeted?
- **E — Exposure / Environment:** What is the primary exposure or environmental factor of interest? Examples include shift work, air pollution, socioeconomic deprivation, or occupational hazard.
- **O — Outcome:** What outcomes are anticipated to be assessed, measured, or improved?

> **EXAMPLE**
>
> `What is the incidence of burnout (O) among healthcare workers in hospital settings (P) working prolonged night shift work (E)?`



The PEO example run in Consensus.

04

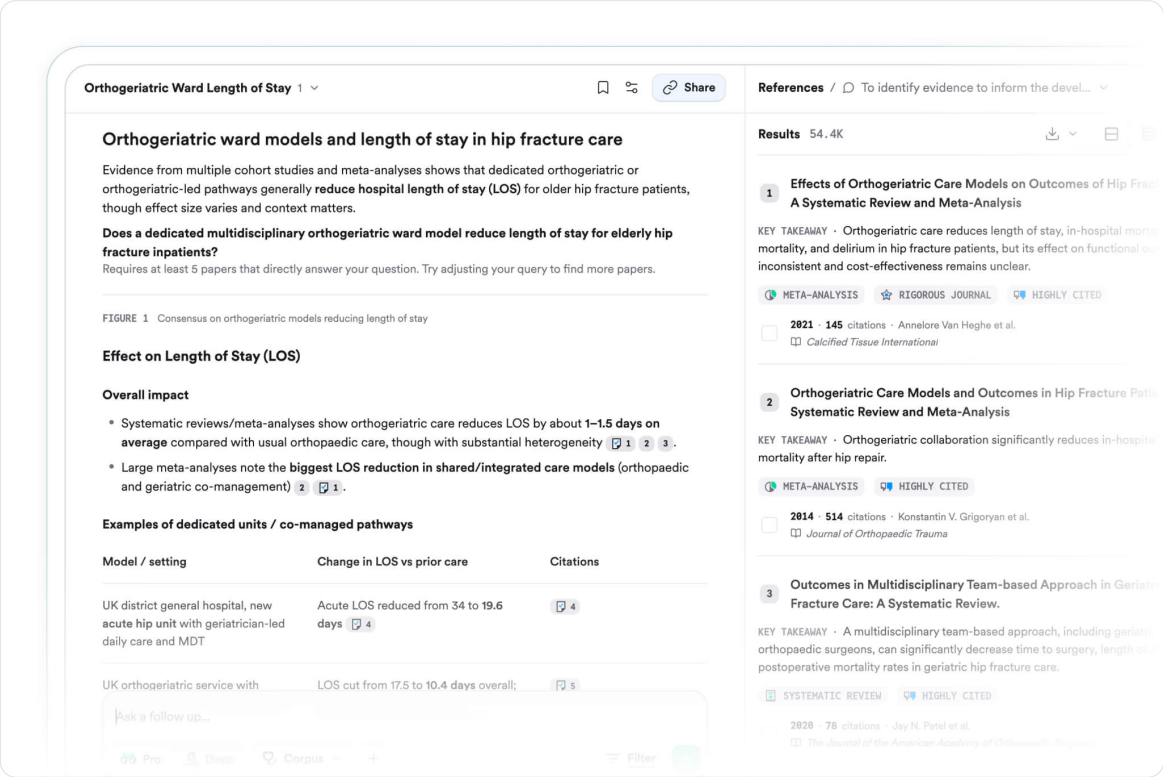
ECLIPSE — for health service and policy questions

Use ECLIPSE when you are evaluating a service, informing a policy decision, or exploring how a health service is organised or delivered.

- **E — Expectation:** What does the person or organisation want from the information? Are they seeking improvement, justification, or information?
- **C — Client group:** Who is the service or intervention aimed at?
- **L — Location:** Where is the service provided? Include setting, geographic context, or system level.
- **I — Impact:** How is change or effectiveness being measured? What outcomes or impact indicators are of interest?
- **P — Professionals:** Who delivers or is involved in the service?
- **S — Service:** What type of service or programme is being examined?

EXAMPLE

To identify evidence to inform the development of inpatient care pathways (E) for elderly patients with hip fractures (C) in a district general hospital in the United Kingdom (L), what does the evidence show about length of stay (I) when a dedicated orthogeriatric ward model (S) is delivered by a multidisciplinary orthogeriatric team (P)?



The ECLIPSE example run in Consensus.

05

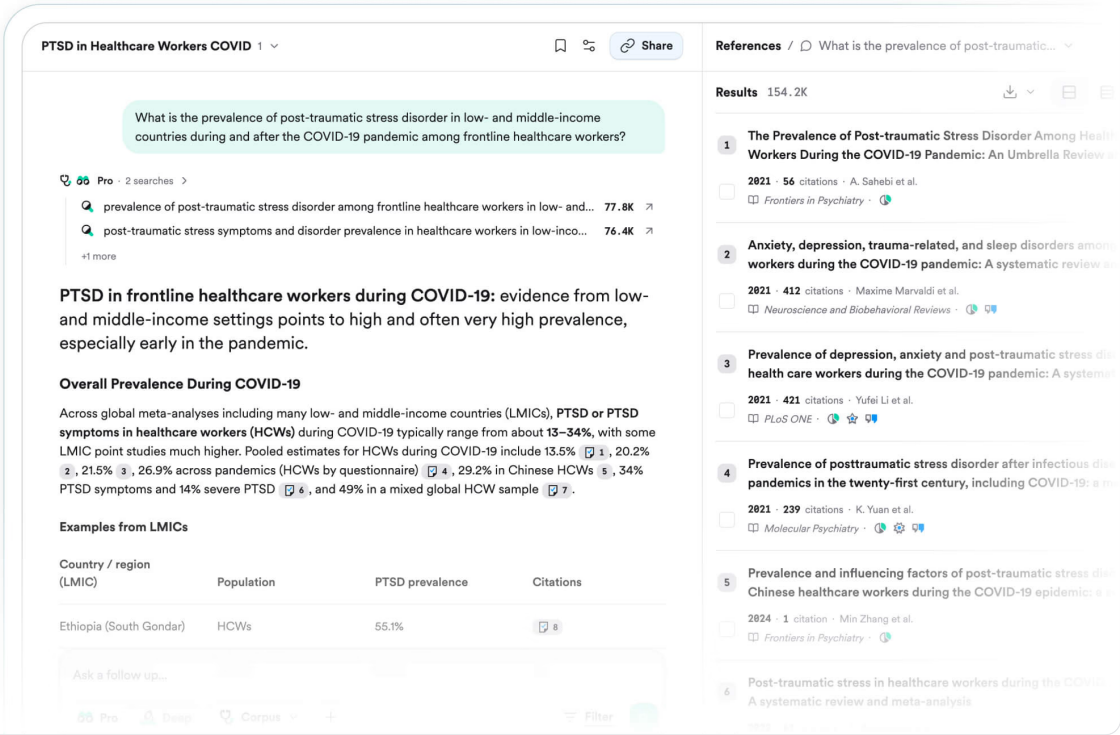
CoCoPop — for incidence and prevalence questions

Use CoCoPop when you want to understand how common a condition or event is within a specific population and context. This framework is particularly suited to epidemiological and public health questions.

- **Co — Condition:** What condition, disease, or event is the focus of the study?
- **Co — Context:** In what setting, geographic region, or background circumstances should the condition be studied?
- **Pop — Population:** What group of participants or patients is being studied?

EXAMPLE

What is the prevalence of post-traumatic stress disorder (Co - condition) in low- and middle-income countries during and after the COVID-19 pandemic (Co - context) among frontline healthcare workers (Pop)?



The CoCoPop example run in Consensus.

A NOTE ON CHOOSING YOUR FRAMEWORK

These five frameworks cover the most common question types you are likely to encounter. If your question does not fit neatly into one, the Hosseini et al. (2024) paper provides a comprehensive overview of additional frameworks for diagnostic accuracy, prognostic, economic, and other specialist question types.

Your prompting checklist

Before you search

- [ ] I know what type of question I am asking and have chosen the right framework.
- [ ] My population is defined specifically — setting, condition, demographic context.

- [ ] My language is neutral and does not assume a direction of effect.
- [ ] I have included relevant context (country, setting, professional group).

After you search

- [ ] I have read the original papers, not just the AI summaries.
- [ ] I have appraised the quality and relevance of the evidence.
- [ ] I have considered whether the populations studied match my clinical context.
- [ ] I have noted evidence gaps — absence of evidence is itself a finding.
- [ ] I have disclosed my use of Consensus in accordance with institutional guidelines.

A FINAL THOUGHT

The best researchers don’t just use AI tools, they use them wisely.

Consensus is a powerful research companion, but it is your question, your critical thinking, and your integrity that determine whether the evidence you find is genuinely useful.

PRACTICAL GUIDE · 05

Fact-checking your Social Media Feed

Everything from the guides so far, applied to a live claim: frame the question, choose the framework, read the evidence, and decide what it actually supports.

WRITTEN BY

Dr. Benita Olivier

Professor of Rehabilitation, Oxford Brookes University

IN THIS GUIDE

01 Why check claims

02 What cherry-picking looks like

03 The principle of aggregation

04 How Consensus helps

05 A guided exercise: testing a viral claim

06 Red flags in social media health claims

07 Sharing what you find

01

Why check claims

Health and wellness claims travel fast on social media. A confident creator with good lighting and a punchy hook can persuade thousands of viewers that seed oils are toxic, that a specific supplement transforms cognition, or that one particular diet will fix everything from low energy to autoimmune disease. These claims often sound scientific. Some cite a single study. A few even tag a peer-reviewed paper in the description.

The problem is not that the science is being mentioned. The problem is that a single study, stripped of context, is being presented as settled evidence. That is rarely how science works. This guide shows you how to use Consensus to test claims you encounter in your feed: to look beyond the headline and see what the wider body of evidence actually says. The skill it builds is learning to think in terms of evidence rather than anecdotes.

02

What cherry-picking looks like

Cherry-picking is the practice of selecting evidence that supports a particular conclusion while ignoring evidence that contradicts it. It is not always deliberate. Sometimes, a creator genuinely believes their claim because they have only encountered the studies that support it. Sometimes the motivation is more cynical. Either way, the effect on the viewer is the same: a partial picture presented as a complete one.

Here is what cherry-picking typically looks like in practice:

- One small study showing a possible link is cited as proof of causation.
- A finding from an animal model is presented as if it applies to humans.
- A result from a highly specific population (athletes, people with a particular condition) is generalised to everyone.
- Limitations, conflicts of interest, and study design issues are not mentioned.
- The dozens or hundreds of studies pointing in a different direction are simply absent from the conversation.

> **A USEFUL RULE OF THUMB**
>
> Any claim that depends on a single study to make its case is, by definition, not yet supported by the body of evidence. It may eventually be. It may not be. You cannot tell from one study.

03

The principle of aggregation

In evidence-based practice, the strength of a claim is not determined by how compelling a single study sounds. It is determined by what the body of literature looks like when you take a step back.

This is sometimes called the principle of aggregation, which means to know what the evidence actually says, you need to look at multiple studies, ideally including systematic reviews and meta-analyses, and ask:

- How many studies have investigated this question?
- Do they point in the same direction, or do they conflict?
- What are the strongest study designs, and what do they show?
- What are the limitations, and where are the gaps?

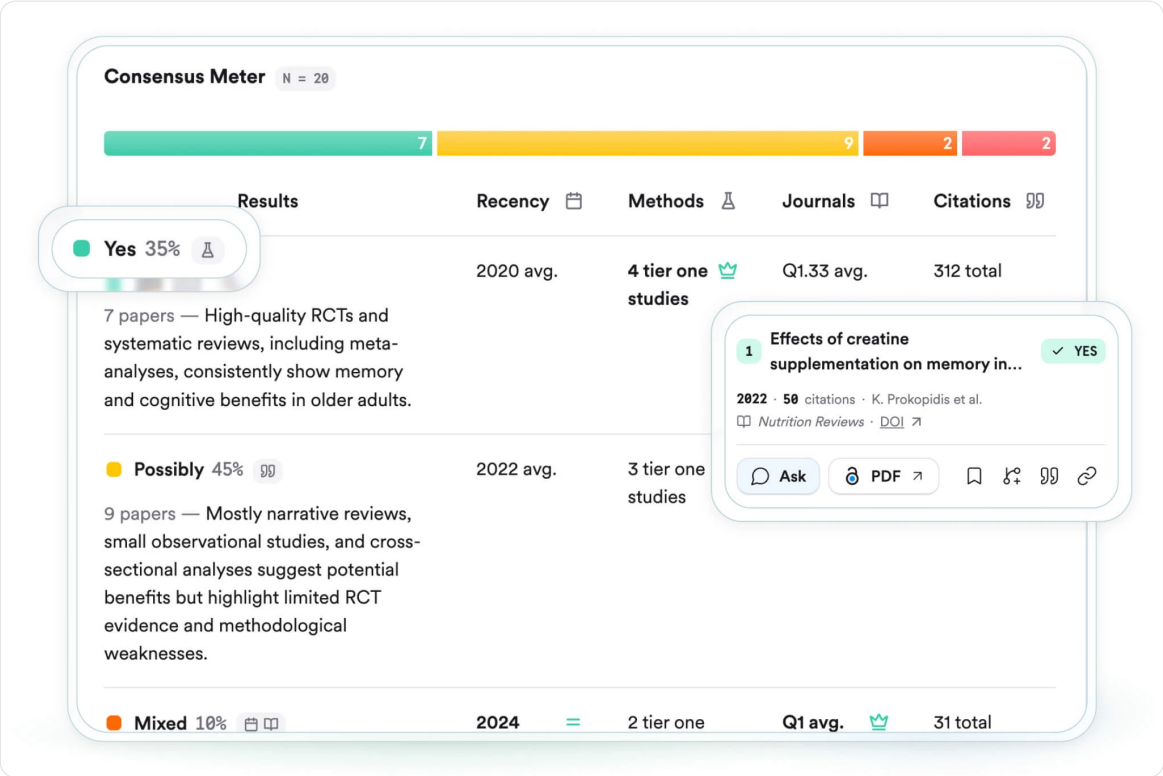
A claim supported by a single study is interesting. A claim supported by a body of consistent evidence is robust. Learning to tell the difference is the heart of this skill.

> A **systematic review** is a structured summary of all the studies that address a particular question, found and assessed using explicit, reproducible methods. Because it sets out to capture the whole literature rather than a convenient slice of it, it sits near the top of the evidence hierarchy.
>
> A **meta-analysis** goes one step further: it combines the numerical results of multiple studies into a single pooled estimate, giving a statistical answer drawn from many studies at once.

04

How Consensus helps

When you ask Consensus a yes/no research question, it does not just retrieve papers. It shows you what the aggregate of those papers says using the Consensus Meter, which displays the proportion of relevant papers that support, are mixed on, or do not support the claim.



The Consensus Meter: the proportion of relevant papers supporting, mixed on, or not supporting the claim, broken down by recency, method, journal and citation count.

This is exactly the view that social media tends to obscure. Where a viral post will show you one study, Consensus shows you many. And where a creator will pick the most striking finding, Consensus shows you the spread.

A few caveats are worth holding alongside this:

- The Consensus Meter reflects how individual papers' findings align with your question. It is a useful starting point, not a replacement for a robust systematic review with meta-analyses or a substitute for reading the studies themselves.
- A meter showing strong support does not mean a claim is settled. It means the available evidence in Consensus's index leans in that direction. Quality and methodology still matter.
- A meter showing mixed or contradictory findings is not a failure of evidence. It often reflects the genuine state of the science, and that is valuable information in itself.

You will still need to read the original papers, especially for any claim you intend to act on, share, or cite. But Consensus gives you something a creator's post cannot: a view of the wider literature.

05

A guided exercise: testing a viral claim

Let's work through an example using a claim that has circulated widely on social media: that seed oils are inflammatory and damaging to health.

1

Put the claim to the test

Turn the claim into a question and add that question to Consensus. The Consensus Meter requires at least five papers that directly answer your question in order for it to show up.

For example, "Do common dietary seed oils raise systemic inflammatory markers such as CRP, IL-6, or TNF-α in humans?" may not show the Consensus Meter, because there may only be three studies that investigated this. But a broader question, such as "Are seed oils inflammatory and damaging to health?" will show the Consensus Meter.

Have a look at the Consensus Meter, and read the summary. This will give you a good overview of what the evidence says about the claim. But don't stop here. Delve deeper.

2

Make the claim more specific

A vague claim such as "seed oils are bad for you" is too imprecise to test. What is actually being claimed?

Usually one of:

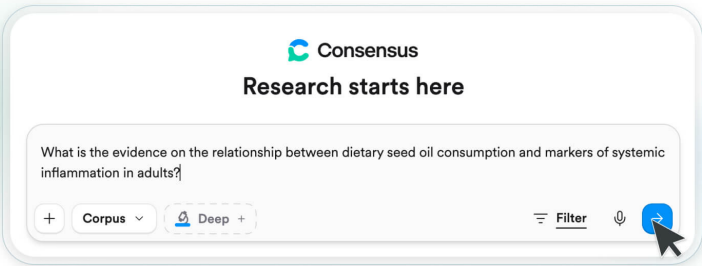
- Seed oils raise markers of systemic inflammation.
- Seed oils contribute to chronic disease (heart disease, obesity, metabolic syndrome).
- Seed oils are worse for health than alternative fats (butter, tallow, olive oil).

Each of these is a distinct empirical question. Pick one to test.

For example, "What is the evidence on the relationship between dietary seed oil consumption and markers of systemic inflammation in adults?" (For more on this, see the companion guide **Framing Unbiased Research Prompts**, guide 01.)

3 Run the search

Type the question into Consensus.



4 Read the papers

Click through to the papers themselves and look for:

- Study designs (randomised controlled trials, observational studies, or animal studies) – some study designs carry more weight than others.
- Any systematic reviews or meta-analyses, which carry more weight than single studies.
- Populations studied (healthy adults compared with people who have specific conditions), and whether these correspond with what the claim promotes.
- Outcomes measured (inflammation markers, clinical outcomes, or surrogate markers) and how these fit with the claim.

5 Form a considered view

You may conclude that the claim is unsupported, partially supported, contested, or simply that the evidence is too limited to say. Any of those is a legitimate, defensible conclusion. What you no longer have to do is take the original creator's word for it.

06

Red flags in social media health claims

Use this table to recognise when a claim deserves a second look.

Red flag	What it suggests	What to do
Cites one study with great confidence	Likely cherry-picked from a wider, more mixed literature	Search the topic in Consensus and read the meter
Uses absolute language ("toxic", "destroys", "miracle")	Overstates findings that are almost always more nuanced	Reframe as a neutral question and check the evidence

Red flag	What it suggests	What to do
Generalises from animal or in vitro studies to humans	Effects in cells or animals do not always translate to people	Look for human studies on the same question
Presents a finding from a specific population as universal	Results in athletes, patients, or particular groups may not generalise	Check whether the study population matches the population being advised
Recommends a product the creator sells or promotes	Conflict of interest	Read the underlying study independently
Does not mention conflicting evidence	The body of literature is being ignored	Search Consensus for the whole picture
Frames complex science as a "truth they don't want you to know"	A persuasive narrative, not an evidence claim	Treat the framing as opinion, then check the actual evidence

07

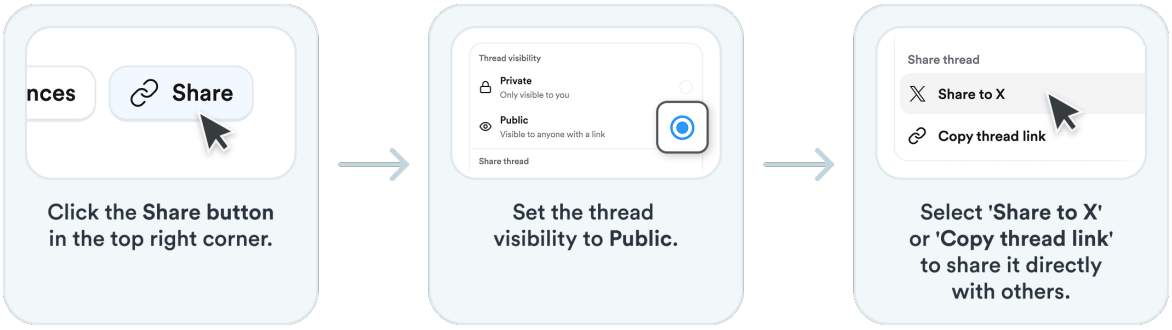
Sharing what you find

Once you have tested a claim, you may want to respond to the original post or share what you found with others. This is worth doing, and worth doing well. A response that is generous and grounded in evidence does far more good than a combative correction, which tends to make people defensive rather than curious.

- Lead with the evidence. Showing people the body of literature is more persuasive than telling them they are wrong.
- Stay curious rather than contemptuous. The aim is to widen the conversation, not to win it.
- Be honest about uncertainty. If the evidence is genuinely mixed, say so. That honesty is what sets your response apart from the original claim.
- Hold your own conclusion lightly. Now and then, the evidence will surprise you, and you will be the one updating your view.

Consensus makes the evidence easy to share directly. Once you have run a search and read the underlying papers, you can share the thread so that others can see the same body of evidence you did.

How to share a Consensus thread



A FINAL THOUGHT

The ability to test a claim against the wider evidence is one of the most useful research skills you can develop. It applies to your studies, to your clinical practice, and to the conversations you have with friends, family, and patients about what the science actually says.

Social media rewards confidence. However, the evidence is the more reliable guide.

Consensus

Got an idea for a resource? We'd love to hear it.

These guides are shaped by what faculty tell us they need. If there is a topic you would like covered — or feedback on this one — get in touch.

daniel@consensus.app