

The Study Screening Workflow

How to take thousands of search results down to your included studies without bias or wasted effort, in two clean stages with the numbers a PRISMA flow diagram needs. A workflow from PhD methodologists.



Written by PhD methodologists researchgold.org

Screening is where a search becomes a set of studies

A comprehensive search returns far more records than belong in your review, often thousands. Screening is the disciplined process of removing what does not fit while keeping everything that does. Done in two clean stages, against criteria fixed in advance, it is fast and defensible. Done ad hoc, it is the step where relevant studies quietly vanish.

PRINCIPLE

Screen against eligibility criteria written before you start, never against a sense of what feels relevant. The criteria, set in your protocol, decide what stays and what goes. When the rule lives in your head rather than on paper, two reviewers will screen differently and the process cannot be reproduced.

The two stages of screening

1 Deduplicate first

The same study appears across databases. Remove duplicates before screening, so you do not assess the same record twice and your counts stay honest.

2 Stage one: title and abstract

Screen each record on its title and abstract alone, keeping anything that might be eligible. This stage is deliberately generous; when in doubt, carry it forward to the next stage.

3 Stage two: full text

Read the full text of everything that passed stage one, and apply the eligibility criteria strictly. Record a reason for every exclusion at this stage, because the flow diagram requires it.

4 Reconcile between reviewers

Two reviewers screen independently and resolve disagreements by discussion, bringing in a third when needed. Independent dual screening is what protects the review from one person bias.

Track the counts as you go

Every record has to be accounted for, from the raw search to the final included set. Keep this tally live; reconstructing it afterwards is how reproducibility breaks.

Screening tally (fills the PRISMA flow diagram)

Records identified from databases:	[number]
Records identified from other sources:	[number]
Duplicates removed:	[number]
Records screened on title and abstract:	[number]
Records excluded at this stage:	[number]
Full texts sought and assessed:	[number]
Full texts excluded, with reasons:	[number]
Reason: [wrong population]	[number]
Reason: [wrong intervention]	[number]
Reason: [wrong outcome]	[number]
Reason: [wrong study design]	[number]
Studies included in the review:	[number]

WATCH OUT FOR

At full-text screening, the exclusion numbers and their reasons must add up exactly. The full texts assessed, minus those excluded for each recorded reason, must equal the studies included. A flow diagram whose numbers do not reconcile is the first thing a careful peer reviewer checks, and a mismatch undermines confidence in everything that follows.

Make the process fast without cutting corners

Good practice and good tooling let you screen thousands of records quickly while keeping every decision documented.

PRACTICES THAT KEEP SCREENING FAST AND CLEAN

Practice	Why it helps
Pilot the criteria on a sample first	Surfaces ambiguous criteria before they cause inconsistency
Use dedicated screening software	Tracks decisions, hides conflicts until reconciliation, logs reasons
Screen independently, then compare	Prevents one reviewer bias from shaping the included set
Record full-text exclusion reasons live	Builds the flow diagram as you go, not from memory
Decision-support, with humans deciding	Can prioritise likely-relevant records; never replaces human judgement

PRINCIPLE

Any tool that ranks or filters records is decision support, not a decision. A human makes every inclusion and exclusion call, and no automated screen should be described as fully accurate. The role of software is to make a careful reviewer faster, not to replace the judgement that makes the review trustworthy.

WATCH OUT FOR

Pilot the criteria on a shared sample, then measure how far the two reviewers actually agreed before screening the full set. Raw percent agreement flatters you, because most records are obvious excludes that both reviewers reject for free. Report Cohen's kappa instead, which corrects for that chance agreement: it is the standard the reporting guidelines expect, and a low value on the pilot is a signal to redraft an ambiguous criterion, not to push on and reconcile thousands of conflicts later. A kappa around 0.6 or below at the pilot stage usually means a criterion is doing two jobs at once or is worded so it can be read two ways. Fix the wording, re-pilot, and only then screen at scale.

The screening readiness checklist

If every box is ticked, your screening will be consistent, fast, and fully documented.

- ☒ Your eligibility criteria are written down before screening begins.
- ☒ Duplicates are removed before the first screening stage.
- ☒ You piloted the criteria on a shared sample and checked inter-reviewer agreement with Cohen's kappa, not raw percent agreement, before screening at scale.
- ☒ Title and abstract screening is generous, carrying doubtful records forward.
- ☒ Full-text screening applies the criteria strictly, with a reason for each exclusion.
- ☒ Two reviewers screen independently and reconcile disagreements.
- ☒ The counts are tracked live and reconcile at every stage.
- ☒ Any screening software is treated as decision support, with humans deciding.

WHEN THE SCREENING HAS TO BE DEFENSIBLE

Have screening run cleanly and documented for you

If you want dual, independent screening run against clear criteria, with every exclusion reason logged and a reconciled flow diagram at the end, our PhD methodologists carry it out with you, then on through extraction and synthesis. You stay the author and make every judgement call.

Get a quote for your review

researchgold.org . Written by PhD methodologists . We work in writing, on your schedule.