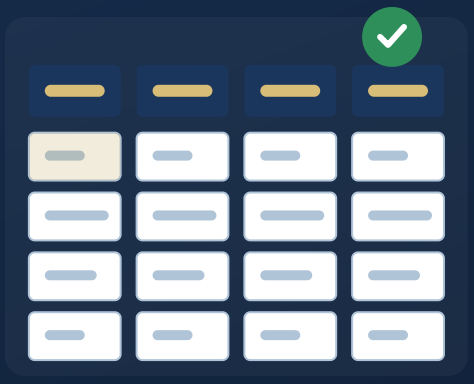
Research Gold

RESEARCH GOLD FIELD GUIDE

The Data Extraction Form Kit

The fields to capture, the form to build, and the piloting routine that stops a systematic review from stalling at its most error-prone step. Templates and a checklist from PhD methodologists.



Written by PhD methodologists researchgold.org

START HERE

Extraction is where careful reviews quietly go wrong

Searching and screening get the attention, but data extraction is where the most consequential errors enter a systematic review. A field you forgot to capture means re-reading every paper. A number transcribed wrongly flows straight into the synthesis. The fix is a form designed before you open the first study, and piloted before you trust it.

PRINCIPLE

A good extraction form is designed once, piloted on a handful of studies, and then applied identically to every paper. The goal is that two people extracting the same study independently produce the same numbers. If the form leaves room for interpretation, the data will not be reproducible.

The principles behind a form that holds up

1. **Capture more than you think you need**
   It is far cheaper to record a field and not use it than to discover a missing field after extracting forty studies. Build generously at the design stage.

2. **One field, one meaning**
   Each field should have a single, written definition, so that anyone extracting interprets it the same way. Ambiguous fields produce data that cannot be pooled or trusted.

3. **Pilot before you commit**
   Extract a small set of studies independently in duplicate, compare, and refine the form. Almost every form needs revision after the pilot, which is exactly what the pilot is for.

4. **Record the source, not just the value**
   Note where in the paper each number came from, and flag anything you derived or estimated, so every value can be checked later.

THE FIELDS

What every extraction form should capture

Group your fields so the form follows the logic of the study. These groups cover the needs of most reviews; add domain-specific fields on top.

THE CORE FIELD GROUPS

Group	Fields to capture	Why it matters
Identification	Study identifier, first author, year, country, funding source	Lets you trace, cite, and check for conflicts and duplication
Design	Study design, setting, recruitment, unit of analysis	Determines eligibility and how the study can be combined
Participants	Sample size per group, key characteristics, inclusion criteria	Defines who the result applies to and the weight the study carries
Intervention or exposure	What was done, dose or intensity, comparator	The heart of the comparison; vague capture here is fatal to synthesis
Outcomes	Outcome definition, timepoint, how measured	Outcomes measured differently cannot be pooled without an explicit rule
Results	Effect estimate, variability, group sizes, probability value	The numbers the synthesis actually uses
Quality	Per-domain risk of bias on the right tool (RoB 2 for trials, ROBINS-I for non-randomised, QUADAS-2 for diagnostic accuracy)	Carries into interpretation and sensitivity analysis; the tool must match the design

WATCH OUT FOR

For any continuous outcome, capture the variability term, a standard deviation, standard error, or confidence interval, alongside every mean. A mean with no measure of variability and no way to recover one cannot enter a meta-analysis, and this single omission strands more studies than any other.

THE TEMPLATE

A starting extraction sheet you can copy

Build this as one row per study in a spreadsheet, with these as the columns. Add or remove fields to fit your question, then pilot it.

Extraction sheet columns (one row per study)

```
Study ID  |  Author  |  Year  |  Country  |  Design  |  Setting

Participants:  N (group 1)  |  N (group 2)  |  Key characteristics  |  Inclusion criteria
Intervention:  What  |  Dose or intensity  |  Duration
Comparator:    What  |  Dose or intensity

Outcome:       Definition  |  Timepoint  |  Measurement method
Results (continuous):  Mean (g1)  SD (g1)  |  Mean (g2)  SD (g2)
Results (binary):      Events (g1)  Total (g1)  |  Events (g2)  Total (g2)
Effect as reported:    Value  |  95% confidence interval  |  Probability value

Risk of bias:  Domain ratings
Source:        Page or table the data came from
Notes:         Anything derived, estimated, or unusual (and how)
```

PRINCIPLE

Decide your effect measure before you build the results columns. If you are pooling continuous outcomes, you need means, standard deviations, and group sizes. If you are pooling binary outcomes, you need events and totals. Building the wrong results columns means re-reading every paper.

THE ROUTINE

Pilot the form before you trust it

A short pilot turns a form that looks complete into one that actually is. Do this before extracting the full set.

1. **Choose a small, varied set**
   Pick a handful of included studies that differ in design and reporting, so the pilot stresses the form rather than flattering it.

2. **Extract in duplicate, independently**
   Have two people extract the same studies without conferring, then compare field by field. Disagreements reveal exactly where the form is ambiguous.

3. **Resolve and refine**
   For each disagreement, sharpen the field definition or add a missing field. Record the rule you agreed, so it is applied consistently from then on.

4. **Lock the form**
   Once the pilot extractions agree, freeze the form and apply it identically to every remaining study.

WHEN THE PAPER DOES NOT REPORT WHAT YOU NEED

Recover the statistic before you strand the study

Studies rarely report data in the exact form a synthesis needs. Discarding them is the wrong reflex: most missing statistics can be reconstructed from what is reported, using accepted formulae, and the recovery is recorded as a derived value.

WHAT IS REPORTED, AND HOW TO RECOVER WHAT YOU NEED

You have	You need	The accepted route
Standard error, or a 95% confidence interval	Standard deviation	SD = SE times the square root of n; or recover SD from the CI width and the t multiplier (Cochrane Handbook formulae)
Median, interquartile range, or range	Mean and SD	Use the Wan (2014) or Luo (2018) estimators, which outperform the older Hozo range method, especially for skewed data
Only a p-value and group sizes	Standard error of the difference	Back-calculate the SE from the exact p and the t or z it implies
Change-from-baseline without its SD	SD of the change score	Impute using a correlation coefficient from a study that reports both, and test it in sensitivity analysis
A graph but no numbers	Point estimates	Digitise the figure (for example WebPlotDigitizer), and flag every value as estimated

WATCH OUT FOR

Two cautions that keep recovery honest. First, the median-to-mean estimators assume an approximately symmetric distribution; when the data are clearly skewed, recovering a mean to force the study into a mean-difference meta-analysis can mislead, and the better answer may be to synthesise on a different scale or narratively. Second, every imputed correlation or digitised value is an assumption, not a measurement, so each must be flagged in the extraction sheet and varied in a sensitivity analysis. The fallback before any of this is to email the corresponding author for the raw statistic, which a careful review documents having attempted.

BEFORE YOU EXTRACT AT SCALE

The extraction readiness checklist

If every box is ticked, your extraction will be reproducible and your synthesis will not stall for want of a number.

- [ ] Your effect measure is decided, and the results columns capture exactly what it needs.
- [ ] Every continuous outcome has a variability term recorded alongside the mean.
- [ ] Both group sizes are captured separately, not just the combined total.
- [ ] Each field has a single written definition, so two extractors would agree.
- [ ] Outcome timepoints are recorded, with a rule for pooling across different times.
- [ ] Every derived or estimated value is flagged with the method used.
- [ ] Missing statistics were recovered with an accepted formula, not the study discarded, where recovery was sound.
- [ ] Risk of bias was rated with the tool that matches each design, not a generic domain list.
- [ ] The form was piloted in duplicate and refined before full extraction began.
- [ ] A second reviewer checks extraction, at least for the results fields.

WHEN THE REVIEW HAS TO HOLD UP

Have the extraction and synthesis done to a standard that survives review

If you want the screening, extraction, risk of bias assessment, and synthesis handled by people who do this every day, our PhD methodologists run the review with you, with duplicate extraction and a reproducible record behind every number. 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.