BASIRA

CANCER METABOLISM
RESEARCH & TRAINING

بصيرة

TRAINING CURRICULUM

CHAPTER SIX

Build Your Own

Now go and find a question of your own

Everything so far has been preparation. This chapter is not read — it is done. You will leave it holding a question that is yours, sharpened until it could fail, designed so that it could be answered, and honest about what its answer would and would not mean.

Founder & Programme Lead	Aroob Gomosani
Prepared & delivered by	Haneen Marghalani
Version	2.0 · second cohort
Date	2026

Gomosani, A., & Marghalani, H. (2026). Basira Curriculum. Zenodo.
https://doi.org/10.5281/zenodo.21791751

6·0

How this chapter works

This is a workbook, not a chapter. Each stage ends with something you produce — a filled worksheet, a written sentence, a decision — and you bring it to the session. Reading ahead is fine; skipping the writing is not, because the writing is where the thinking happens.

Two things to accept before you start. First, your first question will be too broad. Everyone's is. The funnel in §6.2 exists because narrowing is a skill, not a failure. Second, you may reach the end and conclude your question is not worth pursuing — and if you reach that conclusion honestly, with evidence, you have done the assignment perfectly. Chapter 3 listed four directions this group abandoned. Killing an idea well is the work.

WHERE THIS SITS

	Chapter	What it answers
	1 · Thinking	How does science reason, and how do I avoid fooling myself?
	2 · Practice	How do we behave — integrity, collaboration, communication.
	3 · The Science	What is cancer, and how does a cell power itself?
	4 · Data & Tools	What data exists, and how do we handle it without breaking it?
	5 · The Toolkit	The statistics and code, as a reference you return to.
▶️	6 · Build Your Own	Now go find a question of your own and run it.

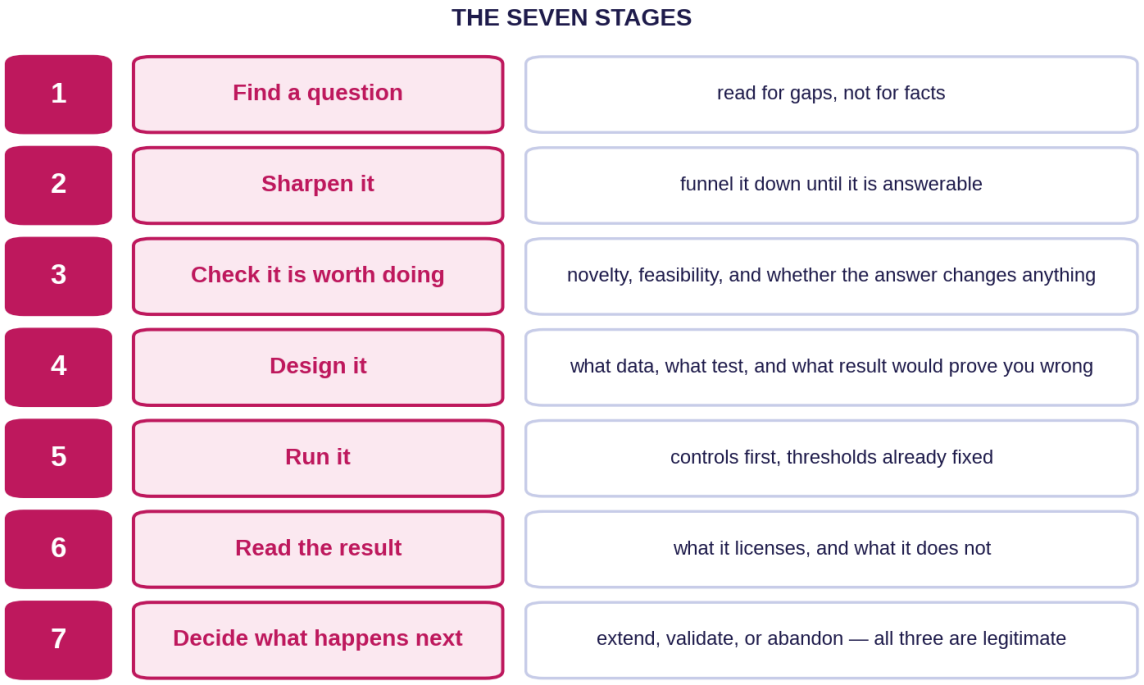


Figure 6.1 The shape of the whole chapter. Each stage produces something.

6.1
Stage one — find a question

You are not looking for a topic. You have a topic already: cancer metabolism, or whatever part of Chapter 3 caught you. You are looking for the place where the field's knowledge runs out.

Chapter 3 gave you the four territories — saturated, active, assumed and untouched — and said the third is the richest. So read differently from how you read for an exam. Do not read for what is known. Read for sentences that are asserted rather than demonstrated, for "it is widely accepted that", for claims that every paper repeats and cites back to one paper from 2009 that measured something slightly different.

> ◆ **KEY IDEA — The reading habit to build this week**
>
> Keep a running note. Every time you hit a sentence you cannot immediately trace to evidence, write it down with its citation.
>
> After two weeks you will have a list. Most entries will turn out to be fine. One or two will not — and those are candidate questions that cost you nothing to find.

WORKSHEET · Stage one — candidate questions

Three claims I could not trace to direct evidence	
Where I found each (paper, section)	
Which one interests me most, and why	

6.2
Stage two — sharpen it

A question arrives too broad. Narrowing it is not a compromise; it is the work that converts an interest into something answerable.

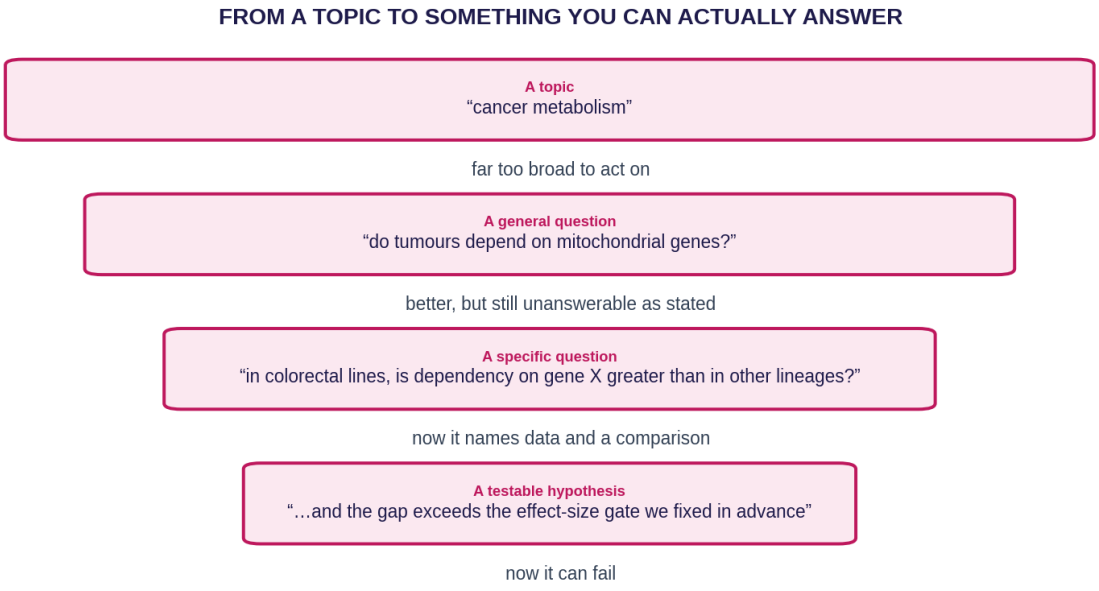


Figure 6.2 Four rungs. Most people stop one rung too early and wonder why the project will not start.

The structure below forces the specificity the funnel is aiming at. If you cannot fill in all four boxes, you do not yet have a question.

A STRUCTURE THAT FORCES YOU TO BE SPECIFIC

P	Population	what samples, cells or patients?
E	Exposure	what is the thing whose effect you are asking about?
C	Comparator	compared against what? A question with no comparator is not a question.
O	Outcome	what will you measure, and how?

If you cannot fill all four boxes, you do not yet have a question — you have an interest.

Figure 6.3 Four slots. The comparator is the one people forget, and a question without one cannot be tested.

WORKSHEET · Stage two — PECO

P — Population	
E — Exposure or factor	
C — Comparator	
O — Outcome and how measured	
My question, in one sentence	

6.3
Stage three — check it is worth doing

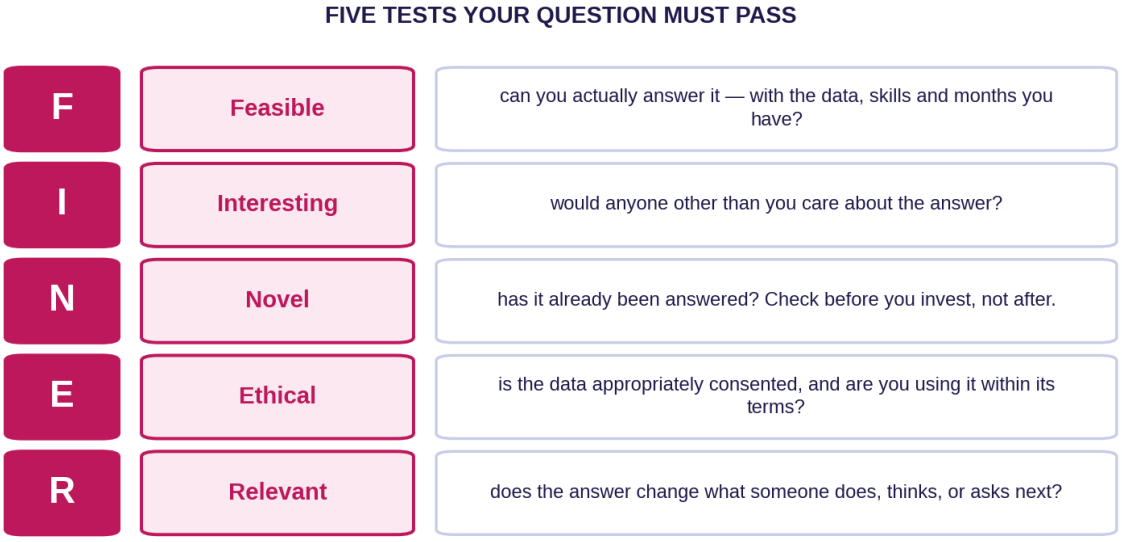


Figure 6.4 Five tests. Be honest about F — it is the one enthusiasm ignores.

Novelty deserves its own process, because it is the check most often skipped and the one whose omission is most expensive.

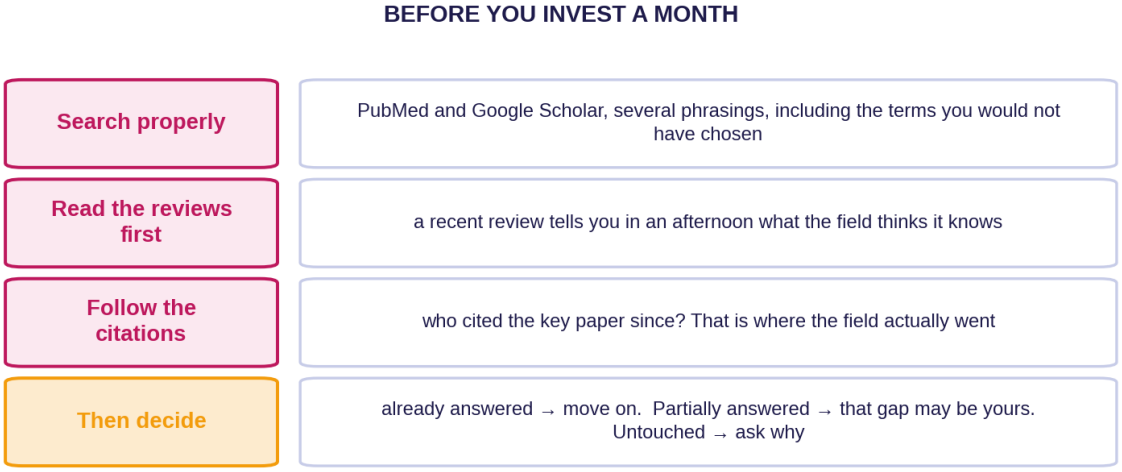


Figure 6.5 An afternoon here saves a month later.

▲ CAUTION — **The honest possibility you must leave open**

This stage might end your project, and that is a legitimate outcome. If a good search shows the question has been answered, you have saved yourself weeks — report it, say what you found, and go back to stage one with a sharper sense of the field.

What you must not do is avoid searching properly because you are afraid of what you will find. That is not optimism; it is the motivated reasoning of Chapter 2, wearing a hopeful face.

WORKSHEET · Stage three — FINER and novelty

Feasible? What data, what skills, how long	
Interesting? Who besides me cares	
Novel? What my search found, and where the gap is	
Ethical? Data source and its terms of use	
Relevant? What changes if the answer is yes. And if it is no	
Verdict: proceed / narrow further / abandon	

6.4
Stage four — design it

Now you decide, in advance and in writing, exactly what you are going to do. Everything Chapters 1, 2 and 5 argued for converges here, because this is the stage where honesty is cheap and after it honesty becomes expensive.

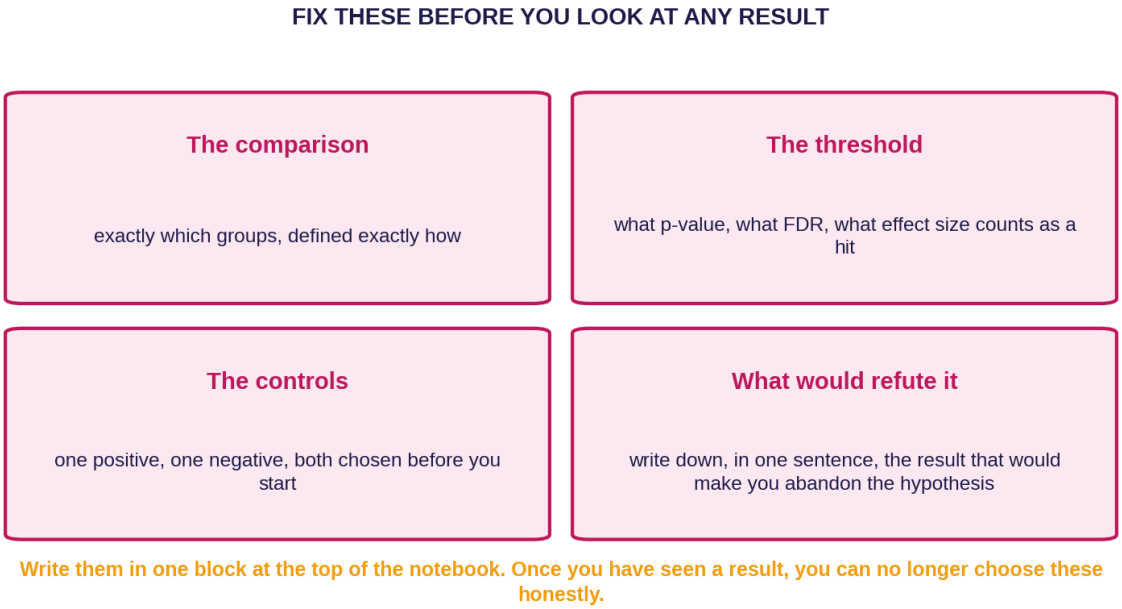


Figure 6.6 Four things fixed before you look. The fourth is the one that makes it science.

Use Figure 5.1 to choose your method. Write down which test, on which data, with which threshold — and write down what you expect a positive control to show and a negative control to show. If you cannot name a control, you do not yet have a design.

WORKSHEET · Stage four — the pre-registration

The exact comparison	
The data, and its version	
The method (§ in Chapter 5)	
Thresholds: p / FDR / effect size	
Positive control, and what it should show	

Negative control, and what it should show	
The result that would refute my hypothesis	

◆ KEY IDEA — **Date it and hand it in before you run anything**

This worksheet is your pre-registration. Once it is submitted, you cannot quietly adjust a threshold because the result was almost significant — and knowing that in advance is precisely what makes the eventual result mean something.

6.5
Stage five — run it

The mechanics are Chapter 5's. Two things about the order matter more than any of the statistics.

- **Controls before results.** Run the positive and negative controls first, and look at them first. If the positive control does not fire, your method is broken and nothing downstream means anything. Discovering this after you have interpreted a result is a very bad afternoon.
- **Then run it once.** The analysis you pre-specified, with the thresholds you fixed. If you find yourself wanting to try a variation, write down what you want to try and why — and label anything that comes out of it exploratory, because it is.

▲ CAUTION — **When something breaks — and it will**

Code fails, files are malformed, joins return nothing. That is normal and is not a sign you are bad at this.

What matters is the Chapter 2 rule: fix it in the open. Say what broke, what you changed, and whether the change touched anything that could affect the result. A silent fix is how an honest error becomes a dishonest one.

6.6
Stage six — read the result

Chapter 1 gave you the discipline for this and it applies now in full: a result is a lead, not a conclusion, and the result you are most excited about is the one that most needs scrutiny. Before you interpret anything, name which rung you are standing on.

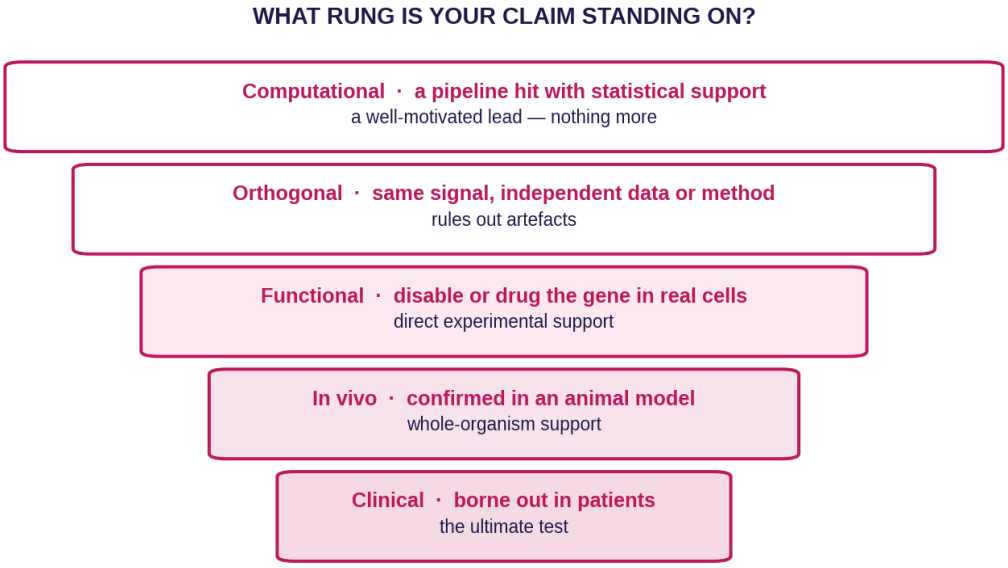


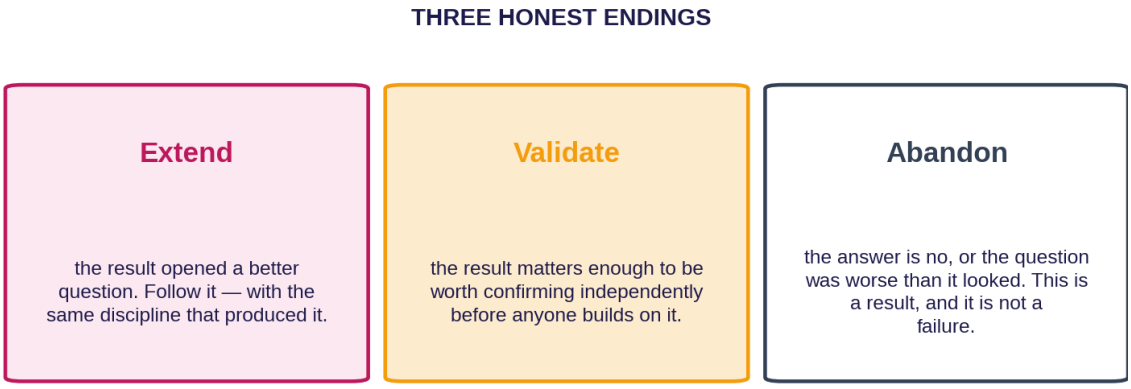
Figure 6.7 Everything produced in this programme starts at the bottom. Say so plainly.

Then write your finding twice. Once in the strongest language you would like to use, and once in the language the evidence actually supports — using the certainty ladder from Chapter 1. The gap between those two sentences is the exact size of the temptation you are resisting.

WORKSHEET · Stage six — the result, honestly

What I found, in one sentence	
The rung of evidence it stands on	
Every caveat attached to it	
The three rival explanations I considered	
Which of them I ruled out, and how	
Which remain open	

6.7
Stage seven — decide what happens next



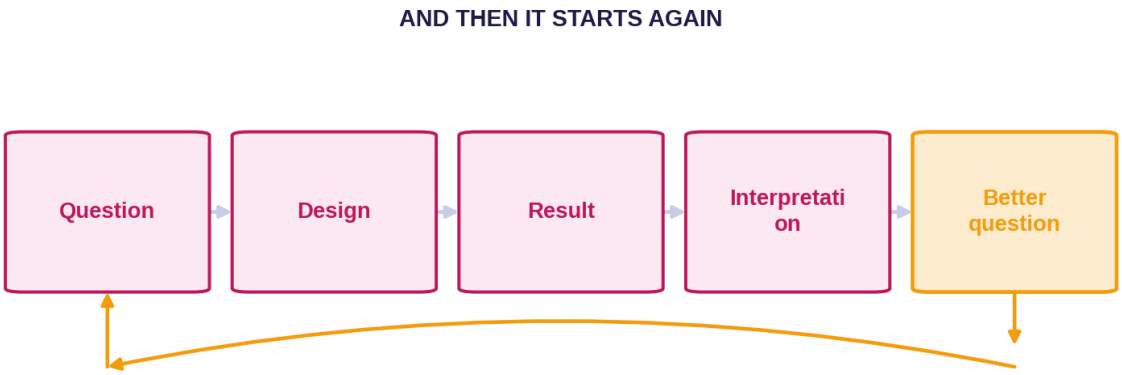
The third is the one people find hardest, and it is the one that separates a scientist from an enthusiast.

A question you killed properly is worth more than one you kept alive by not looking too closely.

Figure 6.8 Three endings, all of them legitimate. Only one of them is comfortable.

If you extend, extend with the same discipline: state the new question, fix new thresholds, run new controls. An extension bolted onto a finished analysis without any of that is how the garden of forking paths gets planted — many optional analyses, quietly recreating the multiple-comparisons problem across an entire project.

If you abandon, write down why in enough detail that the next student does not repeat it. A well-documented dead end is a genuine contribution to the group, and it is one you will be thanked for.



Research is not a line that ends in an answer. It is this loop, run again and again, asking a little better each time.

Your first question will not be your best one. It is not supposed to be.

Figure 6.9 Whichever ending you chose, you arrive back at the beginning holding a better question.

? CHECKPOINT — Final checkpoint

You are ready to present when you can answer all of these without notes.

- What is your question, in one sentence, with all four PECO slots filled?
- How do you know it has not already been answered?
- What did you fix in advance, and where is that written down?
- What did your controls show, and what would you have concluded if the positive control had not fired?
- What did you find, and what rung of evidence does it stand on?
- What is the strongest rival explanation still standing, and what would settle it?
- What happens next — extend, validate, or abandon — and why?

6.8
What you take with you

If the whole curriculum reduced to one sentence, it would be the one from Chapter 1: research is disciplined truth-seeking under uncertainty. Every chapter since has been an attempt to make each of those three words concrete — the discipline in Chapter 2's processes and Chapter 5's controls, the truth-seeking in Chapter 3's biology and Chapter 4's careful loading, and the uncertainty in every checkpoint that asked what else could explain your result.

You will forget most of the specifics. That is fine and expected. What should survive is a set of reflexes: to ask what else could explain this, to check the units before trusting the number, to say exactly what the evidence supports and no more, and to be most suspicious of the result you most want to be true.

◆ KEY IDEA — **The standard, one last time**

Every number is computed; nothing is typed by hand. Thresholds are fixed before results are seen. Controls run before claims. What could not be tested is reported as untested, not as negative.

Do work that is worth building on — so that the next result, yours or someone else's, stands on solid ground.

6.9
Go deeper

[Read] How to Write a Great Research Question — NCBI Bookshelf / PubMed Central — Search “research question FINER” for worked treatments from several fields.

[Read] **Center for Open Science — preregistration** — Templates and explanations for exactly what stage four asks you to do.

[Watch] **Crash Course Scientific Thinking — full series** — Worth rewatching now. It lands very differently once you have tried to do this yourself.

[Source] **Platt (1964) — Strong Inference** — Read it again at the end of the programme. It is a different paper once you have run something.

6.10
Glossary

Abduction Reasoning from an observation to the explanation that would best account for it.

Comparator The group or condition against which your exposure of interest is compared. A question without one cannot be tested.

Feasibility Whether a question can actually be answered with the data, skills and time genuinely available to you.

FINER Feasible, Interesting, Novel, Ethical, Relevant — five tests a research question should pass.

Forking paths The many optional analytic choices available in a project, which together recreate the multiple-comparisons problem.

Hierarchy of evidence The ordering from computational prediction through orthogonal replication and functional validation to clinical confirmation.

PECO Population, Exposure, Comparator, Outcome — a structure that forces a question to be specific.

Pre-registration Writing down the design, thresholds and refuting result before running anything.

Research cycle The loop from question to design to result to interpretation to a sharper question.

Saturation The state of a research area in which so much work exists that a marginal contribution adds little.