POS Lab: Biotech Research

Dr. Yves J. Hilpisch[1]

December 11, 2025

Purpose and Connection

This lab-style document invites you to apply ideas from *Introduction to the Philosophy of Science: Concepts, Practice, and Case Studies* [1]—about measurement, causality, statistics, complexity, and reflexivity—to a concrete domain: biotechnology. It is written as a guided case study rather than a formal chapter. You can work through it on your own or in a small group, using it as a bridge between conceptual themes in the book and real research or product questions in life sciences and biotech.

[1]Get in touch: `https://linktr.ee/dyjh`. Web page `https://hilpisch.com`. Research, structuring, drafting, and visualizations were assisted by GPT 5.1 as a co-writing tool under human direction.

Contents

1 Scenario: A Biomarker-Guided Therapy

Imagine a biotech company developing a new targeted therapy for a subtype of cancer. The therapy is expected to work primarily in patients whose tumours carry a specific molecular signature (a biomarker panel based on gene expression and protein levels).

High-Level Narrative

- **Scientific hope:** if the biomarker panel is well chosen, treatment effects will be strong in biomarker-positive patients and weak in others.
- **Business hope:** a companion diagnostic test will allow the company to sell the drug in a focused segment at a premium price.
- **Regulatory perspective:** agencies will require solid evidence that the biomarker is measured reliably and that treatment effects are genuinely linked to it.

Your task in this lab is to design and critique a research programme around this therapy using the tools from the book.

From Lab to Life

Keep concrete examples in mind as you work:

- a phase II trial exploring dose and early signals of efficacy,
- a phase III randomised trial with biomarker stratification,
- real-world post-market data from hospital registries.

At each step, ask: what are the hypotheses, how are they tested, and where can things go wrong?

2 Measurement and Error: Defining the Biomarker

Relate this section to Chapter 5 (*Measurement and Error*) and Appendix A.

Step 1: Concept to Measurement

- Start by writing, in words, what biological property the biomarker is meant to capture (for example "pathway activation level" or "DNA repair deficiency").
- List at least two different operationalisations of that concept (for example RNA expression panel, protein staining, functional assay).
- For each operationalisation, classify likely sources of systematic error (bias) and random error (noise).

Guiding Questions

- Which measurement choices would make cross-trial comparison easier later?
- Are there latent variables (for example tumour microenvironment) that your biomarker only indirectly reflects?
- How might pre-analytic factors (sample handling, storage time) enter as hidden auxiliaries in your measurement model?

Step 2: Visualising Measurement Error

Using the ideas behind the measurement-error figure in the main book [1, Chapter 5], sketch—on paper or in code—two distributions:

- a "true" underlying biomarker value for a patient group,
- the measured distribution after assay bias and noise.

Mark how much of the spread you expect from random error versus biological heterogeneity.

3 Causal Structure and Trial Design

Relate this section to Chapters 6 and 4.

Step 3: Drawing a Causal Diagram

Draw a small causal diagram for the trial:

- nodes for treatment T, biomarker status B, outcome Y (for example tumour response), and key covariates (age, stage, co-morbidities),
- arrows showing how you think these variables influence one another.

Then, drawing on the causal-inference discussion in the main text [1, Chapter 6]:

- Decide whether you are comfortable assuming that randomisation of T breaks back-door paths from covariates to outcome.
- Mark any remaining open paths (for example post-treatment variables or selection effects) that could bias estimates of the treatment effect in biomarker-positive patients.

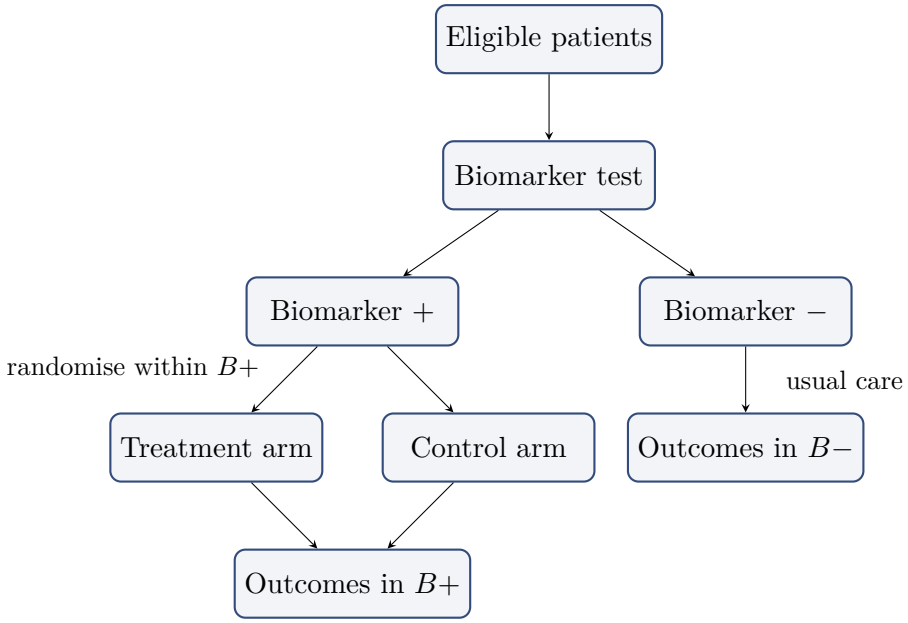

Figure 1: Schematic flow for a biomarker-guided trial: starting from an eligible population, patients are tested, stratified by biomarker status, randomised within the biomarker-positive group, and followed for outcomes.

Design Variants

Compare at least two designs:

- All-comers trial with pre-specified biomarker subgroup analysis.
- Enriched trial recruiting mainly biomarker-positive patients, with a smaller biomarker-negative cohort.

For each, sketch benefits and risks in terms of power, generalisability, and susceptibility to post-hoc story-telling.

For a bird's-eye view, Figure 1 summarises an idealised path from an eligible population through biomarker testing to treatment arms and outcome comparison. Use it as a prompt when checking that the causal diagram you drew really matches the practical flow of patients, samples, and decisions.

4 Evidence, Updating, and Decision Rules

Relate this section to Chapter 2 and the Bayesian update example in Appendix A of the main book [1].

Step 4: Prior and Posterior for Treatment Effect

Pretend you are part of a data monitoring committee.

- Write down a prior for the treatment effect in biomarker-positive patients (for example a distribution over hazard ratios or response-rate differences).
- Describe what kind of interim data (number of events, effect estimate, uncertainty) would lead you to update strongly in favour of efficacy.

- Describe what pattern (for example modest effect with large variance) would push you toward trial modification or early stopping.

You do not need full formulas, but you should be clear about what would count, in your view, as strong vs. weak evidence.

Step 5: Multiple Testing and Research Degrees of Freedom

List at least five analytic choices that create a "garden of forking paths":

- choice of primary endpoint (response vs. progression-free survival vs. overall survival),
- biomarker cut-off definition,
- subgroup selections (for example prior lines of therapy),
- covariates included in the model,
- handling of missing data.

For each, say how you would preregister or constrain decisions to avoid p-hacking while keeping room for legitimate scientific exploration.

5 Complexity, Reflexivity, and Real-World Rollout

Relate this section to Chapters 9, 12, and 13.

Step 6: From Trial to Health System

Once the therapy and biomarker test are on the market:

- Hospitals and payers adopt guidelines.
- Clinicians decide when to order the test.
- Patients and advocacy groups react to reported success rates.

Describe at least two feedback loops:

- one that could make the biomarker look better than it is (for example preferential testing in healthier patients),
- one that could hide benefits in certain subgroups (for example lack of access in underserved populations).

Reflexive Effects

Ask how your own models and decision rules enter the system:

- If hospitals adopt a risk score derived from your trial to allocate intensive care, how might that change future data?
- If payers reimburse only biomarker-positive cases, how does that affect who gets tested and which outcomes you ever observe?

This mirrors the reflexivity discussion in quantitative finance: model-based rules reshape the world that later data describe.

6 Putting It Together: A Mini Research Plan

As a final exercise, write a one-page research plan for the biomarker-guided therapy that:

- states the core scientific and clinical hypotheses,
- justifies the measurement strategy and addresses key sources of error,
- explains the causal identification strategy for treatment effects,
- outlines the statistical analysis plan with attention to multiple testing,
- notes major computational or data limitations,
- anticipates at least two reflexive feedbacks in real-world deployment.

Try in 60 Seconds

If time is short:

- Circle one design choice you would *never* compromise on for this therapy (for example randomisation, blinded outcome assessment, pre-specified biomarker cut-off) and explain why.
- Name one assumption you would happily treat instrumentally—useful for calculation and coordination, even if you doubt it is literally true.

References

[1] Y. J. Hilpisch. *Introduction to the Philosophy of Science: Concepts, Practice, and Case Studies*. 2025. Available at `https://hilpisch.com/philosophy.pdf`.

[2] U.S. Food and Drug Administration. *Guidance for Industry: Enrichment Strategies for Clinical Trials to Support Approval of Human Drugs and Biological Products*. 2019. Available at `https://www.fda.gov/media/121320/download`.

[3] M. A. Hernán and J. M. Robins. *Causal Inference: What If*. Chapman & Hall/CRC, 2020.

[4] J. P. A. Ioannidis. Why most published research findings are false. *PLoS Medicine*, 2(8):e124, 2005. Available at `https://doi.org/10.1371/journal.pmed.0020124`.