How to Detect Questionable Research Practices in Clinical Trial Protocols

Younes Boulaguiem, PhD
Postdoctoral Researcher in Statistics

University of Geneva – Faculty of Medicine
Clinical Research Center, HUG

An 8min overview of QRPs, why they matter, and what they look like in a real clinical trial protocol.

Delivered October 2025

Questionable Research Practises (QRPs)

What are QRPs ?

- Behaviors that **fall short of misconduct** but **undermine credibility**.
- The constitute the gray areas between good science and misconduct .
- Blur the line between **clarification** and **redefinition**, often retrospectively.

Why they matter?

- They **shape** evidence before data collection.

Where do they flourish?

- In environments where **ambiguity**, **incentive pressure**, and **lack of transparency** intersect:
  - ▷ Ambiguity ⇒ HARKing, selective reporting, redefinition, etc.
  - ▷ Incentive pressure ⇒ underpowered studies, unjustified alpha adj., etc.
  - ▷ Opacity ⇒ poor blinding/randomization, etc.

What are their impact on science?

→ "**Most published research findings are false** because of bias and flexibility in design and analysis." — *Ioannidis, 2005*

→ "QRPs are the everyday behaviors that collectively **erode reproducibility**." — *Munafò et al., 2017*

Methodological Roadmap to Identify Redflags

The idea is to transform a study protocol into a clear audit trail of where transparency is strong and where gray zones remain:

1. **Context & Timing:** registration, amendments
2. **Design Integrity:** randomization, inclusion criteria, power
3. **Statistical Transparency:** endpoints, analysis plan, multiplicity
4. **Governance & Oversight**: data sharing, SPIRIT traceability

Many methods exist to help assess these steps:

- NLP/LLM flagging of vague language (e.g., "as appropriate...").
- Rule-based scoring based on SPIRIT checklist (DeVito et al., 2020).
- Check unrealistic size effects, power analysis, fragility analysis, etc.

Case Study: The Prevent-TAHA8 Trial

Good practice:

- Web-based randomization, ethical approval, defined clinical objective...

Potential red flags:

- Retrospective registration $\Rightarrow$ unclear pre-specification.
  First patient: 2021, registration: 2022.
- Subjective inclusion $\Rightarrow$ selection bias.
  "Successful PPCI", "poor echo window".
- Fixed block size, single-blind $\Rightarrow$ performance bias.
  Clinicians know who gets the cells; adjudication not transparent: reviewer "excluded from the group".
- Undefined primary endpoint, no SAP $\Rightarrow$ analytical flexibility.
  Heart failure (HF) never defined.
- Large effect size $\Rightarrow$ feasibility bias.
  HF: 12% $\rightarrow$ 3.9%; protocol: 390, registry: 420.
- SPIRIT checklist inconsistencies $\Rightarrow$ traceability gaps.
  Page references don't match; marked 'no' to plans for trial results communications to the public.

Conclusions

- Most bias originates in the planning phase, not in data analysis...
  $\Rightarrow$ ...and no amount of good statistics can rescue a poorly planned study.
- QRPs thrive in gray zones where definitions, timing, or analysis plans stay flexible.
  $\Rightarrow$ They rarely break the rules, but a good amount can bend the truth.
- A structured, reproducibility-driven audit can help expose those blind spots before trials start.
  $\Rightarrow$ By combining statistical rigor with meta-scientific thinking, we can build reproducibility by design.

References

An-Wen Chan, Jennifer M. Tetzlaff, Douglas G. Altman, Andreas Laupacis, Peter C. Gøtzsche, Karmela Krleža-Jerić, Asbjorn Hrobjartsson, Humphrey Mann, Kay Dickersin, Jesse A. Berlin, et al. Spirit 2013 statement: Defining standard protocol items for clinical trials. *Annals of Internal Medicine*, 158(3): 200–207, 2013.

Kerry Dwan, Carrol Gamble, Paula R. Williamson, and Jamie J. Kirkham. Systematic review of the empirical evidence of study publication bias and outcome reporting bias. *PLoS One*, 9(7):e100011, 2014.

John P. A. Ioannidis. Why most published research findings are false. *PLoS Medicine*, 2(8):e124, 2005.

Marcus R. Munafò, Brian A. Nosek, Dorothy V. M. Bishop, Katherine S. Button, Christopher D. Chambers, Nathalie Percie du Sert, Uri Simonsohn, Eric-Jan Wagenmakers, and John P. A. Ioannidis. A manifesto for reproducible science. *Nature Human Behaviour*, 1:0021, 2017.

Joseph P. Simmons, Leif D. Nelson, and Uri Simonsohn. False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, 22(11):1359–1366, 2011.