Avoid p-Hacking and Bias

By Dr Pedram Rashidi and Emma Cooney, University of Queensland Centre for Policy Futures

P-hacking is generally defined as '*choosing to report a subset of multiple dependent variables or adding observations until the effect of interest is significant*'. [1] There are many ways one can engage in p-hacking. [2] Typically, we think of p-hacking as the practice of using trial and error to fish for statistically significant relationships (a low p-value) with no regard to the theoretical underpinnings of the research. Selective reporting is a broader concept and often involves failing to publish studies that do not produce statistically significant results. The following example illustrates the practice of p-hacking:

> A researcher studies whether a new diet improves weight loss compared to a control group.
>
> 1. They collect data and initially find no statistically significant effect (p = 0.12).
> 2. Instead of reporting this null result, they:
>   - Remove a few 'outliers' (participants who gained weight).
>   - Re-run the analysis -> p = 0.07 (still not significant).
>   - Then:
>   - Try different statistical models
>   - Test only a subgroup (e.g., participants under age 40)
>   - Adjust which variables are controlled for
> 3. Eventually, one analysis shows p = 0.04.
> 4. The researcher reports only this significant result and omits all other analyses.

Why is it relevant?

P-hacking often arises from the academic 'publish or perish' culture. The problem is exacerbated by scientific journals' tendency to favour statistically significant or novel results. In this environment, researchers may, often unintentionally, make biased

analytical choices. P-hacking is also sustained by a lack of clear education and sometimes even institutional mandates. [3]

However, it is a detrimental practice that contributes to the proliferation of false-positive results and the inability to replicate prior studies. This has undermined confidence in some scientific results. [4] It can also lead to devastating consequences when data that has been manipulated is relied upon, especially in sensitive fields such as medicine. [5]

Practical Steps

You may inadvertently engage in outcome fishing as there are many ways in which one can analyse a dataset, and each of these decisions may contain implicit bias. [3] There are some practical steps to avoid this:

1. *Acknowledge degrees of freedom.* Be aware, and explicitly report on, the degrees of freedom within the design, analysis, and reporting of statistical analysis which can be used to generate statistically significant results. [3]
2. *Justify all choices in data preparation and analysis*. Ensure that all choices to exclude outliers, transform variables, analyse a certain range or with certain dependent variables are well-justified.
3. *Consider implementing preregistration*. Preregistration requires you to provide a specific, precise and exhaustive plan of the study. Each step within the plan needs to be detailed, providing only one way in which it can be interpreted or implemented. [6] This avoids outcome fishing.
4. *Minimise bias in analysis.* Whenever possible, perform analyses without knowing which data points belong to which groups, making it difficult to intentionally manipulate results toward a specific outcome.
5. *Report results transparently.* Fully report all measures, conditions, and analyses including insignificant results.

Case Study

The head of Cornell University's Food and Brand Lab, Brian Wansink, has become notorious for instructing his research students to dig through data sets to discover relationships that were statistically significant. In an email to a student, he described the processes of data analysis as "squeez[ing] some blood out of this rock." Following

these revelations, Cornell opened an investigation into research misconduct and many of these studies were retracted. [7]

References

1. Bruns, S.B. and Ioannidis, J.P. (2016) 'P-curve and p-hacking in observational research', *PLOS ONE*, 11(2), p. 2.
2. Reis, D. and Friese, M. (2022) 'The myriad forms of p-hacking', in *Avoiding questionable research practices in applied psychology*. Available at: https://doi.org/10.1007/978-3-031-04968-2_5
3. Stefan, A. *et al.* (2023) 'Big little lies: a compendium and simulation of p-hacking strategies', *Royal Society Open Science*. Available at: https://royalsocietypublishing.org/doi/10.1098/rsos.220346.
4. Banks, G.C., et al. (2016) 'Editorial: evidence on questionable research practices: the good, the bad, and the ugly', Journal of Business and Psychology, 31. Available at: https://link.springer.com/article/10.1007/s10869-016-9456-7
5. DataCamp (2025) 'P-hacking: how to (not) manipulate the p-value'. Available at: https://www.datacamp.com/tutorial/p-hacking
6. Wicherts, J.M., *et al.* (2016) 'Degrees of freedom in planning, running, analyzing, and reporting psychological studies: a checklist to avoid p-hacking', *Frontiers in Psychology*, 7. Available at: https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2016.01832/full
7. *The New York Times* (2018) 'More evidence that nutrition studies don't always add up', 29 September. Available at: https://www.nytimes.com/2018/09/29/sunday-review/cornell-food-scientist-wansink-misconduct.html.