

---

☰ Menu

---

1 WEEK AGO    EDIT

# Questionable Research Practices in Nursing Science

## Questionable Research Practices in Nursing Science: .05 Shades of Grey

Richard Gray

**NURSE AUTHOR & EDITOR, 2019, 29(2), 5**

Many editorial colleagues will have had to investigate cases of alleged misconduct and impose sanctions that can include retraction of a manuscript. Cases of research misconduct can garner considerable attention. The website Retraction Watch ([www.retractionwatch.com](http://www.retractionwatch.com)) is an important repository of scientific misdeeds and has witnessed exponential growth in the number of manuscripts retracted over the past decade.

Seemingly, comparatively few nursing science papers get retracted (Al-Ghareeb et al., 2018). Why is this? Is it because because nurse researchers are exceptionally

well behaved or because we fail to scrutinize colleagues' work with sufficient rigor? If we are not spotting research misconduct, how well do we handle more subtle and dubious research practices? In other disciplines, there is a growing interest in these dark arts, collectively termed “questionable research practices” (John, Loewenstein, & Prelec, 2012).

Simmons (2011) demonstrated in a series of experimental studies how questionable research practices can profoundly inflate the chance of finding support for a hypothesis that is, in fact, false. There are several studies where authors have attempted to estimate how common questionable research practices are in different disciplines (though not nursing). For example, John et al., 2012 surveyed 2,000 psychologists and reported that a surprisingly high proportion—about one-third—of participants indicated that they had engaged in questionable research practices. That said, it is worth noting that authors have criticized this work for overestimating the prevalence of these activities (Fiedler & Schwarz, 2016).

## **EXAMPLES OF QUESTIONABLE RESEARCH PRACTICES**

Questionable research practices can be considered under two umbrella headings: HARKing (Hypothesizing After the Results are Known) and P-hacking (Probability hacking). Other practices that may surpass the questionable threshold include salami slicing and non-random allocation of participants in randomized controlled trials.

### **HARKing**

HARKing is defined as presenting a post hoc hypothesis (i.e., one based on or informed by one's results) in a paper as if it were, in fact, an a priori hypotheses (Kerr, 1998). It is not hard to imagine how HARKing might emerge when a group of researchers meet to review the study findings: results are presented and

seemingly show that the exciting new treatment has no significant effect against the primary outcome. A debate ensues about “What hypothesis were we actually testing?” And that “Come to think of it, surely we would expect to see an effect on [insert outcome] which seems to be what the data are telling us.” Naïvely the group talk themselves into misrepresenting the presentation of their work.

## P-hacking

P-hacking refers to the conscious—or perhaps unconscious—manipulation of data until they become significant (Head, Holman, Lanfear, Kahn, & Jennions, 2015). Examples of P-hacking, include stopping data collection as soon as results reach  $p < .05$ , analyzing multiple measures but only reporting those that are significant and removing (or adding) covariates to get a  $p < .05$ .

### *Stopping data collection when $p < .05$*

Researchers naturally want to review their data as a study progresses to see if significant findings emerge. This is a practice that needs to be resisted as researchers may be tempted to stop data collection and write the paper if they spot an interesting (likely significant) observation in their data. Stopping a study when observations are significant runs the risk of reporting a false positive (Type 1 error) because significance can appear and disappear as more data are collected. However tempting, data analysis should only be undertaken once recruitment has been completed, ideally following a pre-specified analysis plan. The only caveat or possible exception is in a clinical trial that requires independent safety monitoring. One of the tasks of the safety committee undertakes is to consider stopping a trial if it seems the treatment is spectacularly effective or conversely, spectacularly dangerous.

### *Many measures but only reporting significant findings*

Authors of observational studies have a habit of administering multiple measures and not specifying the hypotheses they are seeking to test. When writing up their findings, authors then selectively report only the associations that were significant.

### *Add and remove covariates (confounders)*

Confounding is an alternative explanation for the observed association. Many researchers have experienced the heart sink moment when initially significant observations disappear once adjustments are made for confounding. What if the confounder that does all the damage is removed from the analysis? What if the researcher pretends the variable wasn't measured it in the first place? The significant finding is back and with a bit of luck reviewers and editors (biased towards publishing significant observations) won't spot the ruse.

## **OTHER QUESTIONABLE RESEARCH PRACTICES**

Other questionable research practices include adding/removing outlying participants, rounding a p-value just above .05 (e.g. .051) and claiming it as significant, and not reporting all arms of a trial (e.g., dropping a third arm). Perhaps these may not be so common in nursing science but there are some candidate practices that seem to me to be common in nursing science and should be discussed.

### Salami Slicing

Salami slicing refers to the reporting in multiple papers, a study that could (or should) be reported in one (Gray & Baker, 2016). For example, a researcher completes a cross-sectional survey of a few hundred nurses. The questionnaire used in the study has a number of subscales, each with a separate focus (educational needs, attitudes, competence, etc.), and each reported in a separate paper. It might be argued that the authors are giving each individual subscale

justice, after all, there are many large observational studies—the Whitehall or Framingham cohorts spring obviously to mind—that have spawned many dozens of papers. A simple survey of nurses, multiple papers, seriously? The question to ask, could the study be reasonably reported in a single article? If the answer is yes, then that is what should be done.

### Randomization Procedures

Many nursing science trials—undoubtedly many more than might be expected by chance— have exactly equal group sizes (check the next time you edit or review a trial). This happens when researchers, to balance the groups, force group allocation. For example, in an experiment with a target sample size of 20, the researchers have so far recruited 15 participants. Ten have been randomized (using simple randomization procedures) to the experimental and five to the comparator condition. To balance the groups, the researcher decides to allocate the final five participants to the control condition. This is an issue because the moment a researcher *chooses* which group to allocate participants, unpredictability—core to the randomized controlled trial—is lost and bias is inevitably introduced. Further, if authors do not report in the paper that this is what they did, the reader will be misled about the conduct of the trial influencing how the observations are interpreted.

## QUESTIONABLE RESEARCH PRACTICES IN QUALITATIVE RESEARCH

Authors have written extensively about questionable research practices in experimental or observation designs. What about in qualitative methodology, perhaps the dominant paradigm in nursing science? There are many points along the qualitative research process when a researcher may potentially manipulate data. It is hard to believe that researchers have not subtly edited a quote to exemplify a theme, or attributed words from one participant to another to ensure

more are represented in the data. There have, seemingly, been few attempts to probe how common these practices are in qualitative paradigms and if they impact on the reported findings in any meaningful way.

## REFLECTIONS

There is unquestionable pressure on aspiring academics to find significance in their data. This may induce researchers to engage in questionable research practices. Researchers may not have been adequately trained that such conduct is unacceptable; innocently believing that they are merely presenting their work in the most positive way. Self-evidently there is a task to raise awareness among nurse researchers about these practices. Many questionable research practices can be brought to a grinding halt by pre-registration and it is positive to see a number of nursing science journals, including the *Journal of Clinical Nursing*, introducing registered reports that focus on the reporting rigorously done rather than statistically significant research (Smith et al., 2018).

## CONCLUSION

Questionable research practices exist in a grey zone some way after good practice and before misconduct. This puts journal editors and reviewers in an awkward position when they spot such conduct. Their response needs to be considered and proportionate. What cannot be allowed to happen is that they are swept under the carpet and ignored.

## REFERENCES

1. Al-Ghareeb, A., Hillel, S., McKenna, L., Cleary, L., Vistentin, D., Jones, M., ... Gray, R. (2018). Retraction of publications in nursing and midwifery research: A systematic review. *International Journal of Nursing Studies*, 81, 8-13.

2. Fiedler, K., & Schwarz, N. (2016). Questionable research practices revisited. *Social Psychological and Personality Science*, 7(1), 45–52.  
<https://doi.org/10.1177/1948550615612150>
3. Gray, R., & Baker, C. (2016). Salami slicing. *Journal of Psychiatric and Mental Health Nursing*. <https://doi.org/10.1111/jpm.12290>
4. Head, M. L., Holman, L., Lanfear, R., Kahn, A. T., & Jennions, M. D. (2015). The extent and consequences of P-Hacking in science. *PLOS Biology*, 13(3), e1002106. <https://doi.org/10.1371/journal.pbio.1002106>
5. John, L. K., Loewenstein, G., & Prelec, D. (2012). Measuring the prevalence of questionable research practices with incentives for truth telling. *Psychological Science*, 23(5), 524–532. <https://doi.org/10.1177/0956797611430953>
6. Kerr, N. L. (1998). HARKing: Hypothesizing After the Results are Known. *Personality and Social Psychology Review*, 2(3), 196–217.  
[https://doi.org/10.1207/s15327957pspr0203\\_4](https://doi.org/10.1207/s15327957pspr0203_4)
7. Smith, G. D., Gelling, L., Haigh, C., Barnason, S., Allan, H., Penny, K., & Jackson, D. (2018). Transparency in the reporting of nursing research. *Journal of Clinical Nursing*, 27(3–4), 475–477. <https://doi.org/10.1111/jocn.14212>

## ABOUT THE AUTHOR

Richard Gray, RN, PhD is Professor of Clinical Nursing Practice, School of Nursing and Midwifery, La Trobe University, Melbourne, VIC 3086, Australia. You can contact him via email at: [r.gray@latrobe.edu.au](mailto:r.gray@latrobe.edu.au).

[2019 29 2 5 Gray](#)

Copyright 2019: The Author. May not be reproduced without permission.

Journal Complication Copyright 2019: John Wiley and Son Ltd.