

Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:

Story, DA

Title:

Feasibility and pilot studies: dropping the fig leaf

Date:

2020-02-01

Citation:

Story, D. A. (2020). Feasibility and pilot studies: dropping the fig leaf. *Anaesthesia*, 75 (2), pp.152-154. <https://doi.org/10.1111/anae.14865>.

Persistent Link:

<https://hdl.handle.net/11343/286447>

Editorial

Feasibility and pilot studies: dropping the fig leaf

D.A. Story

Professor of Anaesthesia, Centre for Integrated Critical Care, The University of Melbourne, Victoria, Australia

Email: dastory@unimelb.edu.au

Keywords: feasibility; pilot study; research, study design.

Feasibility and pilot studies should ask the question: “*can we do this research?*” [1-4].

Unfortunately, many clinical researchers have used ‘pilot’ or ‘feasibility’ in their manuscript as a fig leaf to hide the problem that their research is, in reality, a small, underpowered, single centre study; or other poorly designed or conducted research. Some view pilot and feasibility as interchangeable terms, while another view is that pilot studies are a formal dress rehearsal of a definitive study and are therefore a subgroup of feasibility studies asking: “*can we do this research exactly like this?*” [1,5]. In this issue of *Anaesthesia*, Charlesworth et al. [6] examine studies published in leading anaesthesia journals over the last 10 years that claim to be feasibility or pilot studies.

The researchers [6] turned to recent definitions of pilot studies: “...*small-scale versions of a definitive randomised controlled trial or a prospective study examining the feasibility of a definitive study*”, and use the words pilot and feasibility interchangeably. They also considered the recent Consolidated Standards of Reporting Trials (CONSORT) pilot and feasibility trials extension [2], which is a useful blueprint for both reporting and conducting pilot trials. Charlesworth et al.’s [6] finding that only 34 (13%) of the 266 studies

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the [Version of Record](#). Please cite this article as [doi: 10.1111/ANAE.14865](https://doi.org/10.1111/ANAE.14865)

This article is protected by copyright. All rights reserved

claiming to be pilot or feasibility studies fulfilled the authors' criteria suggests that many researchers have used the 'pilot study' fig leaf, and many editors and reviewers don't know that fig leaf when they see it.

Outcomes

As the CONSORT extension emphasises [2], an important feature of reporting pilot studies is using metrics of feasibility such as ability to recruit patients and protocol adherence to test areas of research uncertainty, rather than more traditional clinical trial outcomes [3, 4, 6]. Analysing pilot studies often requires mixed methods using quantitative and qualitative analyses to tease out specific areas of strength and weakness in the protocol for systems, staff, or patients [2-4]. Feasibility analysis both in process and metrics has a lot of similarity to evolving implementation science which is at the far end of the research translation pipeline [3,7]. In addition to determining feasibility, pilot study analysis identifies potential beneficial protocol or trial management changes [3,4]. Beyond ensuring trial success, studies that iron-out feasibility barriers are likely to enhance subsequent wider implementation the intervention if it is shown to be beneficial; helping to bridge the so-called translation valley-of-death.

In the Charlesworth et al. study [6], a notable finding is that even when pilot studies were deemed to be consistent with their definitions, fewer than half had led to identifiable definitive studies. There are several possible reasons for this including: the time taken to fund, conduct and report a large study; in addition to publication bias [8]. The research may not have been done, may not yet have been completed, or may not have been published. Charlesworth et al. [6] speculate that some may have chosen not to continue on to further research when their pilot study had a fortuitous finding of $p < 0.05$ supporting the intervention. That is, the researchers had mistakenly assumed they had proved their case beyond reasonable doubt. Such reliance on isolated p values from small studies as the determinant of truth and signifying a job-well-done is unfortunate [1]. One outcome not discussed [6] is where a feasibility study concludes that the research design is not feasible, that is: "*Don't do what we did*". It is unclear whether Charlesworth et al. excluded negative pilot studies when examining the rate of subsequent publication of larger definitive studies. Publishing a negative pilot trial is important because it explains why the researchers did not proceed to further and helps prevent others from attempting research that is unlikely to succeed. Futile

research is a waste of time, money and effort for researchers, patient participants, and the community.

Not feasible

One group recently reported a negative feasibility study of web-based cognitive prehabilitation at home before major surgery to reduce postoperative delirium [9]. A lot of hot topics in one study! Their primary hypothesis was that ‘... *home-based cognitive prehabilitation will be feasible for older surgical patients.*’ This is a good example of a hypothesis for a feasibility study. They tested feasibility metrics using mixed methods. I quote their sample size statement verbatim because it captures the essence of the exercise: “*Given the feasibility nature of this study, the sample size represents a convenience sample, and power calculations were deferred. By having a relatively small group of participants, detailed analysis can be conducted for determining digital literacy, barriers to enrolment, and barriers to training program completion.*” The researchers [9] randomised sixty patients (a good size for a pilot study) to a commercial, computer-based cognitive training program for about 20 minute per day for seven days before surgery or usual care. They found that they could enrol only about 50% of patients screened; suggesting limited feasibility and acceptability. Of patients allocated to the intervention only 5 of 30 completed the per-protocol training; very poor protocol fidelity. Qualitative work with the patients led the authors to conclude “*enrolment in such a program over a short time frame may worsen preoperative anxiety and apprehension.*”. Their intervention distressed some patients. The authors proposed that future work look at supervised training, with breaks, over a longer period to optimise exposure and minimise distress. Suggested changes for future studies can be very useful.

Many argue that pilot studies should not be used to obtain preliminary answers to planned primary research questions [1,4]. The CONSORT pilot trial extension suggests only using descriptive statistics of intervention outcomes and confidence intervals to demonstrate degrees of uncertainty. The previously quoted pilot study [9] demonstrates this area of controversy. Despite a good feasibility primary aim, their second aim was to test their hypothesis that “... *such training will mitigate delirium.*” From the results, this is probably not what they found; but who knows? Unfortunately, part two of this good feasibility study analysis [9], is a more traditional analysis of a very small randomised trial that has a sample

size based on mixed methods for a very different question. Of 23 patients in the computer prehabilitation group, 6 (26%) suffered delirium vs. 5 of 29 (17%) in the usual care group. The authors [9] report an unsurprising and unhelpful $p=0.507$, but did not report the difference or the 95% CI. When one calculates these, the prehabilitation group are 9% worse and the 95% CI ranges from 30% worse to 13% better. No firm conclusions can be drawn from this. Further, such imprecise results are of limited use in calculating sample sizes for later research; not to mention the point estimate is in the wrong direction [1,3,4]. To be fair to the authors, both reviewers and editors facilitated, and perhaps even demanded, reporting these unhelpful secondary results.

Value

But what about value? There is limited evidence. Morgan et al. [10] examined the financial returns from 89 feasibility studies with total funding of £19.5 million from the National Institute for Health Research in the UK. About 60% of the researchers had published their results and most of the others were planning to publish. Average funding for the feasibility studies was £220,000. Of 57 of 89 (60%) judged feasible, a quarter had been funded for further research at an average of £1.2M; a five-fold increase on pilot funding. Another 17 of the 57 were still seeking funding. Of the 20 (22%) deemed not feasible none had sought further funding. The authors estimate if the not-feasible studies had been funded, up to £20M of funder money could have been wasted. But, even if none of this research had been funded, there are the considerable costs of preparing and assessing grant applications.

Both the Morgan study of financial waste [10] and the Charlesworth study [6] of reporting pilot research are good examples of meta-research: research about research [11]. One risk with meta-research is to be overly dogmatic. In judging the work of others, and reporting our own, we should be careful in the certainty of our statements. Rather than using words like ‘correct’ [6] I think it is often better to use less certain phrases like ‘consistent with’. Avoiding a black and white, true vs false, approach is part of embracing uncertainty including escaping the scientific tyranny of isolated p values [12]. We should be particularly careful applying definitive words such as correct to areas of evolving thinking such as pilot and feasibility studies.

Researchers, particularly emerging researchers, should move away from unhelpful attempts at small, dodgy clinical trials. Instead, they should consider feasibility and pilot research that is useful, publishable, and quicker to conduct than definitive trials [10]. This

research increasingly forms a fundamental component of successful grant applications and conduct of large trials. For funders and the broader community there is greater certainty that subsequent research projects can succeed. Researchers should no longer try to hide small studies behind the fig leaf of adding ‘pilot’ to their title, and editors and reviewers should recognise this fig leaf when they see it. All should embrace, fund, conduct, report, and publish high-quality feasibility research using valuable tools such as the CONSORT extension [2].

Pilot and feasibility studies are important research. It is time that journal editors, particularly in specialty journals, add a pilot and feasibility section to their journals and select reviewers with both content expertise and an appropriate understanding of feasibility methodology.

Acknowledgements

No conflict of interest declared.

References

1. Kistin C, Silverstein M. Pilot studies: a critical but potentially misused component of interventional research. *Journal of the American Medical Association* 2015; **314**: 1561-2.
2. Eldridge SM, Chan CL, Campbell MJ, et al. CONSORT 2010 statement: extension to randomised pilot and feasibility trials. *British Medical Journal* 2016; **355**: i5239.
3. Story D, Leslie K, French C. Feasibility and pilot studies: small steps before giant leaps. *Anaesthesia and Intensive Care* 2018; **46**.
4. Thabane L, Ma J, Chu R, et al. A tutorial on pilot studies: the what, why and how. *BMC Medicine Research Methodology* 2010; **10**: 1.
5. Eldridge SM, Lancaster GA, Campbell MJ, et al. Defining feasibility and pilot studies in preparation for randomised controlled trials: development of a conceptual framework. *PLoS One* 2016; **11**: e0150205.
6. Charlesworth M, Klein A, White S. A ten-year bibliometric analysis of the conversion and reporting of pilot studies published in six anaesthesia journals. *Anaesthesia* 2019 (in this issue of the journal).

7. Proctor E, Silmere H, Raghavan R, et al. Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. *Administration and Policy in Mental Health* 2011; **38**: 65-76.
8. Chong SW, Collins NF, Wu CY, Liskaser GM, Peyton PJ. The relationship between study findings and publication outcome in anesthesia research: a retrospective observational study examining publication bias. *Canadian Journal of Anesthesia* 2016; **63**: 682-90.
9. Vlisides PE, Das AR, Thompson AM, et al. Home-based cognitive prehabilitation in older surgical patients: a feasibility study. *Journal of Neurosurgical Anesthesiology* 2019; **31**: 212-7.
10. Morgan B, Hejdenberg J, Hinrichs-Krapels S, Armstrong D. Do feasibility studies contribute to, or avoid, waste in research? *PLoS One* 2018; **13**: e0195951.
11. Ioannidis JPA. Meta-research: Why research on research matters. *PLoS Biology* 2018; **16**: e2005468.
12. Ioannidis JPA. The importance of predefined rules and prespecified statistical analyses: do not abandon significance. *Journal of the American Medical Association* 2019; **321**: 2067-8.