

Article

Not peer-reviewed version

Simulation as Inquiry: Reframing Methodological Knowledge Through Boundary Mapping

[Sifiso Vilakati](#) *

Posted Date: 27 February 2026

doi: 10.20944/preprints202602.1967.v1

Keywords: non-superiority; boundary mapping; simulation studies; epistemic status; ADEMP framework; registered reports



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Article

Simulation as Inquiry: Reframing Methodological Knowledge Through Boundary Mapping

Sifiso Vilakati 

Biostatistics, University of the Free State; sifemman@gmail.com

Abstract

Methodological statistics has developed under the pervasive belief that the value of a simulation study lies in its ability to demonstrate the superiority of a proposed method. This belief has distorted the epistemic role of simulation research by encouraging selective reporting, narrow design spaces and the suppression of results that fail to show improvement. Drawing on philosophical accounts and on meta scientific critiques by several studies, this paper argues that non superiority is not a failure but an essential source of methodological knowledge. It marks the boundaries at which a method performs adequately, where it begins to degrade and where it breaks. Using empirical patterns evident in contemporary methodological literature and drawing on the widely cited case in which machine learning methods often failed to outperform logistic regression in clinical prediction, this paper demonstrates how non-superiority clarifies expectations and guides methodological refinement. The paper proposes a structural remedy that pairs the Registered Reports model with principled simulation design following the ADEMP framework of Morris et al. [1]. This alignment protects the visibility of non-superiority and restores the integrity of simulation research. Boundary mapping is therefore not a modest contribution but a scientific and ethical imperative for methodological transparency and applied reliability.

Keywords: non-superiority; boundary mapping; simulation studies; epistemic status; ADEMP framework; registered reports

1. Introduction

A longstanding assumption in methodological statistics is that scientific value is demonstrated primarily through improvement. Within this ideal of perfection, a simulation study that fails to show superiority is often viewed as a scientific disappointment rather than as an epistemic contribution. Yet the philosophical foundations of inquiry consistently reject such a narrow interpretation of progress. Popper [2] argued that knowledge advances when bold conjectures are subjected to severe testing rather than when they accumulate favourable confirmations. Kuhn and Hacking [3] described anomalies as the pressure points through which paradigms evolve and eventually break. Firestein [4] explained that ignorance and error are not liabilities but assets because they provoke the questions that drive science forward.

Contemporary metascience has amplified these classical insights. Ioannidis [5] demonstrated that in environments characterised by limited power, multiple testing and analytical flexibility, a substantial proportion of statistically significant results are likely to be false. Sterling [6] showed decades earlier that when non-significant studies remain unpublished, chance alone eventually produces significant findings that then distort the scientific archive. Loken and Gelman [7] later showed that noisy measurement exacerbates this distortion because the studies that pass significance thresholds under high noise tend to exaggerate effect sizes. Together, these arguments indicate that suppressing non-superiority corrupts the scientific record. These concerns apply acutely to simulation studies, which are uniquely positioned to test statistical tools across diverse data generating conditions. When the literature selectively rewards superiority, simulations become demonstrations rather than explorations.

This paper argues that their fundamental purpose should be boundary mapping, the identification of regions in which methods work, where they begin to deteriorate and where they fail.

2. Methods

2.1. *Locating Simulation Research in the Logic of Inquiry*

Simulation occupies a distinct position within the logic of scientific inquiry. It is neither deductive mathematics, which derives analytic truths from axioms, nor empirical fieldwork, which observes naturally occurring phenomena. Instead, simulation is a logic of consequences instrument: it asks what follows from a set of statistical assumptions when those assumptions are instantiated repeatedly in controlled hypothetical worlds. This characterisation aligns with the view that knowledge advances not through accumulation of confirmations but through structured confrontation with assumptions, as emphasised in classical accounts of scientific method by Popper [2] and Kuhn and Hacking [3]. Within this framework, simulation generates conditional knowledge statements of the form: given this data generating mechanism, these patterns of behaviour arise. The commitment to severe testing implicit in this stance, echoed in Firestein [4] account of the epistemic value of failure, requires that simulations explore challenging parameter regions as part of their core logic rather than as optional embellishments.

2.2. *The Role of Abstraction and Model Worlds in Method Evaluation*

Simulation studies operate through the deliberate creation of model worlds, abstract constructs designed to isolate the structural features most relevant to methodological performance. These model worlds are not intended as replicas of specific datasets but as analytic environments that make visible the mechanisms, noise levels, dependence structures, sparsity, nonlinearity that shape estimator behaviour. This perspective is consistent with the treatment of the Data Generating Process (DGP) as the “laboratory” in which theoretical constructs are made observable, a view articulated in methodological guidance such as the ADEMP framework of Morris et al. [1]. Abstraction, in this setting, is not a simplification but a philosophical stance: it foregrounds the conditions under which performance claims hold. This accords with the boundary mapping intuition advanced throughout this paper, that methodological insight arises not merely from documenting expected strengths but from exposing the limits and breakpoints of methods under diverse, well justified abstractions.

2.3. *Methodological Integrity as Design Discipline*

Under the boundary mapping philosophy, the credibility of knowledge from simulation does not derive primarily from the numerical outcomes but from the discipline of design that produced them. This design first perspective places epistemic responsibility on transparent declaration of aims, explicit articulation of the simulation space, careful justification of estimands, and principled selection of performance measures. Such an orientation resonates with the argument that selective visibility where only superior or favourable results are published distorts the methodological archive, a concern raised by Ioannidis [5], Sterling [6], and Loken and Gelman [7] in their analyses of publication bias and noise driven inflation of effects. It also aligns with structural reforms that shift scientific value from outcomes to process, most notably the Registered Reports model introduced by Chambers [8] and later developed with collaborators, which binds publication decisions to methodological clarity rather than result direction. This paper adopts that ethos: simulation is treated as epistemically sound only when its design is explicit, its assumptions defensible, and its limitations openly incorporated into the structure of inquiry.

2.4. *Simulation Studies as the Mapping of Methodological Terrain*

Simulation studies function as controlled experiments in which researchers can observe how methods behave across a variety of conditions. When conducted with transparency and rigour, they reveal the contours of methodological performance. Morris et al. [1] emphasised that principled simulation design requires pre-specification of aims, data generating mechanisms, estimands, methods and

performance measures, a structure captured in their ADEMP framework. This framework transforms simulations into scientifically disciplined experiments that resist selective reporting.

When simulations are designed and reported within this structure, the landscape becomes visible. Regions of clear superiority appear as peaks of the terrain, while areas of non-superiority appear as plateaus that indicate where established methods remain competitive. Sudden deterioration appears as cliffs where variance increases or convergence fails. Knowledge of this full landscape is essential for methodological understanding. A literature that depicts only peaks is not a map at all but a curated catalogue of successes that obscures the truth of the terrain.

From a philosophical perspective, this mapping aligns with the view that scientific knowledge grows through confrontation with the limits of theories. A simulation space that excludes difficult or unfavourable regions undermines the logic of severe testing and reduces the contribution of simulation studies to methodological understanding.

2.5. *The Epistemic Consequences of Suppressing Non-Superiority*

The scarcity of non superiority results in the methodological literature is a documented systemic artifact rather than a reflection of genuine methodological performance. As Sterling [6] noted decades ago, scientific publication systems selectively retain positive findings, and this logic persists in contemporary methodological statistics, where simulation studies are often expected to demonstrate improvement rather than to map methodological limits. Evidence from issues of major theoretical statistics journals consistently shows that published simulation studies overwhelmingly emphasise performance gains, with neutral or mixed outcomes appearing only rarely. These norms are not explicitly stated in editorial policies, but they emerge from a long standing cultural incentive structure: papers that fail to show superiority struggle to be framed as “contributions,” and editors and reviewers frequently interpret non superiority as insufficient novelty rather than as valuable boundary evidence. As a result, the scientific archive disproportionately reflects peaks in the methodological landscape while obscuring the plateaus and cliffs that are essential for a complete epistemic map.

While such norms undoubtedly accelerate the refinement of high performing procedures, they simultaneously suppress the publication of essential boundary information: the plateaus where classical methods remain competitive, the cliffs where new techniques deteriorate, and the unstable regions where estimation error or algorithmic failure emerges. Without systematic visibility into these boundaries, the literature presents an illusory terrain composed entirely of peaks, an epistemically distorted landscape that undermines severe testing, exaggerates methodological progress, and obscures the very conditions under which methods genuinely differ.

The epistemic consequences of this selective visibility are substantial. First, the suppression of non-superiority inflates the apparent performance of new methods, producing a landscape of exaggerated advances. This aligns with Ioannidis [5] finding that positive predictive value collapses under selectivity. Second, because noise inflates the apparent success of studies that pass significance thresholds, the literature becomes biased toward overstated claims, a process described by Loken and Gelman [7]. Third, the absence of reported failures prevents researchers from diagnosing why methods break, which parameters induce instability and which conditions limit performance.

The ethical consequences are equally serious. Applied researchers depend on methodological results when selecting tools. A literature that conceals valleys and cliffs misleads practitioners and produces decisions that are not well aligned with the limitations of methods. Methodological research therefore bears a responsibility to describe the full terrain.

2.6. *A Case Study in Boundary Mapping*

The comparison between logistic regression and machine learning models for clinical prediction offers a compelling example of the value of non-superiority. For many years, it was widely assumed that machine learning methods would outperform logistic regression due to their flexibility. However, Christodoulou et al. [9] reviewed dozens of real-world applications and found no consistent performance benefit for machine learning when logistic regression was properly validated and speci-

fied in Christodoulou et al. [9]. This non superiority result clarified that logistic regression remains competitive for structured tabular data with modest signal to noise ratios and limited non-linearities.

This example demonstrates that non superiority does not indicate methodological failure. Rather, it reveals the boundaries within which simpler models are optimal and the contexts in which more complex models yield genuine improvements. This is precisely the type of information simulation research is meant to uncover.

2.7. *The Role of Process First Publication Structures*

Transforming the simulation literature requires structural reform. The Registered Reports model introduced by Chambers [8] relocates peer review to the design stage, ensuring that publication decisions are based on the importance of the research question and the quality of the methodology rather than on the eventual results. Chambers [8] later elaborated how this format realigns publication incentives with scientific integrity by preventing authors from selectively reporting favourable results Chambers et al. [10]. Nosek and Lakens [11] argued that the Registered Reports model increases credibility by binding publication to methodological clarity.

Applied to simulation studies, Registered Reports operate synergistically with the ADEMP framework. Authors pre-commit to the simulation space they will investigate, including the regions where performance might deteriorate. Reviewers can strengthen the design before simulations are conducted, and the acceptance decision becomes independent of outcome direction. This structure legitimises the publication of non- superiority and protects the integrity of the methodological record.

2.8. *Failure as Methodological Insight*

Failure is often the most informative scientific event. When a simulation reveals that a new estimator offers no improvement, the result becomes a diagnostic clue. If a penalized estimator performs poorly in small samples, it suggests that penalty paths or tuning mechanisms may require revision. If a classifier overfits sparse data, the insight concerns regularisation and validation. If an estimator becomes unstable under heavy tailed distributions, the opportunity arises to explore robust loss functions. Each of these improvements begins with the honest documentation of non-superiority.

The neutral comparison tradition has long argued for this orientation, insisting that scientific progress depends on transparency in performance evaluation. When the literature embraces non-superiority, it embraces the raw material of methodological refinement.

3. Conclusions

Simulation studies are central to methodological science. They serve not merely as demonstrations of performance but as instruments for mapping the conditions under which methods succeed, under which they weaken and under which they fail. A literature that records only peaks presents a distorted scientific landscape. The suppression of non-superiority produces epistemic exaggeration, impedes methodological refinement, and undermines applied decision making.

A process first approach that integrates Registered Reports with principled simulation design restores the integrity of simulation studies. It shifts the focus from outcome to design and permits the honest publication of non-superiority. Boundary mapping is thereby recognised as a scientific and ethical imperative. In a community committed to truth and accountability, non-superiority is not an embarrassment. It is a contribution.

Author Contributions: Conceptualization, S.V.; methodology, S.V.; software, S.V.; validation, S.V.; writing—original draft preparation, S.V.; writing—review and editing, S.V. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The author declares no conflicts of interest.

References

1. Morris, T.P.; White, I.R.; Crowther, M.J. Using simulation studies to evaluate statistical methods. *Statistics in medicine* **2019**, *38*, 2074–2102.
2. Popper, K. *The logic of scientific discovery*; Routledge, 2005.
3. Kuhn, T.S.; Hacking, I. *The structure of scientific revolutions*; Vol. 2, University of Chicago press Chicago, 1970.
4. Firestein, S. *Failure: Why science is so successful*; Oxford University Press, 2015.
5. Ioannidis, J.P. Why most published research findings are false. *PLoS medicine* **2005**, *2*, e124.
6. Sterling, T.D. Publication decisions and their possible effects on inferences drawn from tests of significance—or vice versa. *Journal of the American statistical association* **1959**, *54*, 30–34.
7. Loken, E.; Gelman, A. Measurement error and the replication crisis. *Science* **2017**, *355*, 584–585.
8. Chambers, C.D. Registered reports: a new publishing initiative at Cortex. *Cortex* **2013**, *49*, 609–610.
9. Christodoulou, E.; Ma, J.; Collins, G.S.; Steyerberg, E.W.; Verbakel, J.Y.; Van Calster, B. A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. *Journal of clinical epidemiology* **2019**, *110*, 12–22.
10. Chambers, C.D.; Dienes, Z.; McIntosh, R.D.; Rotshtein, P.; Willmes, K. Registered reports: Realigning incentives in scientific publishing. *Cortex* **2015**, *66*, A1–A2.
11. Nosek, B.A.; Lakens, D. A method to increase the credibility of published results. *Social Psychology* **2014**, *45*, 137–141.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.