

Concept Paper

Not peer-reviewed version

From Irreproducible Science to Venture-Fundable Fiction: How Incentives, Paper Mills, and Capital Markets Enable a New Class of Biotech Startups

[Andrey Samsonov](#) *

Posted Date: 27 January 2026

doi: 10.20944/preprints202601.2034.v1

Keywords: reproducibility crisis; scientific misconduct; paper mills; biomedical research; venture capital; biotech startups; research integrity; translational medicine; clinical trial risk; AI drug discovery; incentive structures; grant funding



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.

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.

Concept Paper

From Irreproducible Science to Venture-Fundable Fiction: How Incentives, Paper Mills, and Capital Markets Enable a New Class of Biotech Startups

Running Title: Scientific Irreproducibility as a Startup Risk Factor

Andrey Samsonov

Samdolite Pharmaceuticals Inc., 112 Christine Cay Ct, Saint Charles, MO 63304, USA;
andrey.samsonov@samdolite.com

Abstract

The reproducibility crisis has progressed from a methodological concern into a systemic threat to biomedical innovation. While irreproducible findings have long been recognized as a barrier to scientific progress, recent evidence suggests the emergence of a more consequential downstream phenomenon: venture-backed biotech and pharmaceutical startups built on compromised scientific foundations, including irreproducible and, in some cases, fabricated results. This commentary argues that the modern incentive structure—where publication volume and journal prestige strongly influence grant success—creates conditions that can reward narrative certainty over empirical truth. These incentives may promote the growth of paper mills, selective reporting, and other forms of research misconduct, while simultaneously increasing the probability that venture capital funds companies whose core claims are not biologically valid. Such “venture-fundable fiction” can crowd out rigorously grounded programs, distort capital allocation, and erode trust in biomedical R&D. The consequences extend beyond economics: a culture of performative productivity may select for dishonesty in early-career researchers, accelerating talent loss from academia and reinforcing unethical norms within startups. Finally, this commentary highlights the risk that artificial intelligence (AI)-driven drug discovery may amplify these failures when trained on contaminated literature and datasets. Addressing this trajectory will require coordinated reforms in funding incentives, publication standards, reproducibility verification, and investment diligence practices to protect scientific integrity and accelerate genuine medical progress.

Keywords: reproducibility crisis; scientific misconduct; paper mills; biomedical research; venture capital; biotech startups; research integrity; translational medicine; clinical trial risk; AI drug discovery; incentive structures; grant funding

1. Introduction: The Reproducibility Crisis as a System-Level Risk

The reproducibility crisis is no longer a peripheral methodological debate; it has become a structural vulnerability in biomedical research and development. Over two decades, concerns have intensified that a substantial portion of published findings—particularly in preclinical and biomedical domains—cannot be reliably replicated (Ioannidis 2005, Begley and Ioannidis 2015). Recent analyses further emphasize that the entities enabling fraud at scale are resilient and expanding, raising the possibility that industrialized mechanisms now exist to sustain scientific misconduct beyond isolated cases (Richardson, Hong et al. 2025).

Although irreproducibility has multiple causes, including statistical underpowering, selective reporting, and poor experimental design, its consequences are increasingly visible across the innovation pipeline. The downstream effects are particularly severe in biotech and pharma, where

early stage claims often determine the direction of years of development and tens of millions of dollars in capital deployment.

2. Incentives That Reward Publishable Narratives Over Verifiable Truth

A central driver of irreproducibility is incentive design. Funding is limited, while competition is intense. In practice, laboratories with higher publication output - especially in high-impact journals - often gain advantage in securing NIH/NSF and other public or private grant support. This creates an incentive gradient where speed, volume, and perceived novelty can dominate methodological rigor.

Under such conditions, questionable research practices may become rational behaviors within the system. Occam's razor principle (https://en.wikipedia.org/wiki/Occam%27s_razor) offers a useful framing: rather than assuming widespread incompetence, the simplest explanation for widespread irreproducibility is that incentives reward "publishable outcomes" more than reproducible truth. The pressure is not distributed evenly; early-career scientists, trainees, and applicants to competitive programs may face especially acute incentives to produce publications rapidly, regardless of quality.

3. From Irreproducible Science to Venture-Fundable Startups

While grant incentives are widely discussed, a second layer of risk has become increasingly consequential: irreproducible or fraudulent science can become venture fundable.

Startup failure is common and expected; most startups fail, and many venture-backed companies do not return investor capital. Failure alone does not imply fraud. However, as irreproducible research increases, the probability rises that some venture-backed biotech startups are founded on scientific claims that were never valid in the first place. These companies may secure substantial financing not because their biology is correct, but because their materials—data packages, decks, patents, and narratives—are optimized for fundraising.

This creates a new class of startup distinct from "high-risk innovation": the startup whose primary risk is not scientific uncertainty, but the possibility that its underlying scientific premise is non-reproducible by design.

4. Why Fraudulent Startups Can Pass Diligence

Venture capital screening is often rigorous, yet biomedical diligence has unique limitations. Many biological claims are expensive and time-consuming to verify independently. As a result, diligence processes may overweight surface indicators of legitimacy such as:

- polished investor materials prepared by professional consultants,
- credible-seeming preclinical data packages,
- patent filings that signal defensibility,
- well-known advisors or institutional affiliations, and
- warm introductions through trusted networks.

In this context, startups built on compromised science can appear "high quality" because they are engineered to pass evaluation filters. Meanwhile, slower-moving companies committed to rigorous validation may be disadvantaged. The systemic result is not only investor loss, but opportunity cost: real science is crowded out by better-packaged fiction.

5. The Challenge of Accountability in Biotech

Biotech provides natural cover for misconduct because failure is expected. When a program fails in Phase 2 or Phase 3, leadership can plausibly attribute the outcome to disease complexity, patient heterogeneity, or trial design. Public statements often emphasize disappointment for patients and families—an appropriate sentiment that can also function as a shield against scrutiny of the program's foundational integrity.

Theranos Inc. remains a rare example where misconduct was prosecuted decisively ([https://en.wikipedia.org/wiki/Theranos#:~:text=Theranos%20Inc.%20\(%2F%CB%88%CE%B8%C9%9Br.%C9%99n.o%CA%8As%2F\)%20was%20an&text=o%CA%8As%2F\)%20was%20an%20American%20privately,in%202003%20by%20then%2019%2Dyear%2Dold](https://en.wikipedia.org/wiki/Theranos#:~:text=Theranos%20Inc.%20(%2F%CB%88%CE%B8%C9%9Br.%C9%99n.o%CA%8As%2F)%20was%20an&text=o%CA%8As%2F)%20was%20an%20American%20privately,in%202003%20by%20then%2019%2Dyear%2Dold)). Most cases in biomedical innovation do not resemble overt criminal fraud; instead, they hide within statistical noise, selective reporting, outsourced CRO workflows, and the inherent uncertainty of translational development.

6. The Most Concerning Consequence: Selection for Dishonesty Among Young Scientists

Perhaps the most damaging long-term outcome is cultural selection. When publication volume becomes a dominant currency, early-career researchers may feel compelled to engage with paper mills or other unethical pathways to remain competitive. Producing careful, reproducible science is slower and more difficult than producing attractive but unreliable results. Over time, honest scientists may lose opportunities and leave academia, weakening the long-term capacity of the research ecosystem.

The same dynamic can appear within startups. In organizations that operate under “first fake it, then make it,” scientists whose results contradict desired narratives may be pressured to repeat experiments until results align with expectations. When refusal is punished, scientific inquiry becomes narrative manufacturing. This threatens not only ethics but innovative efficiency, as decision-making becomes decoupled from reality.

7. AI-Driven Drug Discovery: Solution or Amplifier?

AI-based drug discovery is often positioned as a way to overcome biological complexity and accelerate target identification and lead optimization. However, AI systems inherit the quality of their training data. If the literature and datasets contain high levels of irreproducible or fabricated findings, AI models may learn false associations and generate biologically meaningless hypotheses at scale. In this scenario, AI does not solve the reproducibility crisis—it industrializes it.

Therefore, AI may improve productivity while simultaneously increasing the volume of false-positive programs, unless paired with disciplined experimental validation and robust reproducibility filtering.

8. Conclusions

The reproducibility crisis is now a translational and economic problem, not merely an academic one. When compromised science becomes fundable, it forms a pipeline that converts irreproducible publications into grant support, venture capital, and years of downstream development. The cost is paid not only by investors and institutions, but by patients—through delayed progress, misallocated resources, and diminished trust.

Addressing this trajectory will require reforms that realign incentives toward reproducibility, strengthen safeguards against industrialized publication fraud, and improve verification standards across funding and investment processes. The future of biomedical innovation depends not only on faster discovery, but on restoring the reliability of the scientific substrate on which discovery is built.

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

Ethics Statement: The study did not require ethical approval.

References

- Begley, C. G. and J. P. Ioannidis (2015). “Reproducibility in science: improving the standard for basic and preclinical research.” *Circ Res* **116**(1): 116-126.
https://en.wikipedia.org/wiki/Occam%27s_razor. “Occam’s razor.”

[https://en.wikipedia.org/wiki/Theranos#:~:text=Theranos%20Inc.%20\(%2F%CB%88%CE%B8%C9%9Br.%C9%99n.o%CA%8As%2F\)%20was%20an&text=o%CA%8As%2F\)%20was%20an%20American%20privately,in%202003%20by%20then%2019%2Dyear%2Dold.](https://en.wikipedia.org/wiki/Theranos#:~:text=Theranos%20Inc.%20(%2F%CB%88%CE%B8%C9%9Br.%C9%99n.o%CA%8As%2F)%20was%20an&text=o%CA%8As%2F)%20was%20an%20American%20privately,in%202003%20by%20then%2019%2Dyear%2Dold.) “Theranos Inc.”.

Ioannidis, J. P. (2005). “Why most published research findings are false.” PLoS Med **2**(8): e124.

Richardson, R. A. K., S. S. Hong, J. A. Byrne, T. Stoeger and L. A. N. Amaral (2025). “The entities enabling scientific fraud at scale are large, resilient, and growing rapidly.” Proc Natl Acad Sci U S A **122**(32): e2420092122.

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.