IRREPLICABLE PATENTS

Janet Freilich[*]

[*] Associate Professor, Fordham Law School. I thank Wendy Gordon, Yotam Kaplan, Joseph Kupferman, Laura Pedraza-Fariña, Nicholson Price, David Schwartz, Jacob Sherkow, Polk Wagner, and participants at the Junior Faculty Forum for Law and STEM, Regulation and Innovation in the Bioscience Workshop, and Roundtable on Empirical Methods in Intellectual Property. For excellent research assistance, I thank Jennifer Lei, Adam Mengler, Meredith Miller, and Allison O'Hara.

INTRODUCTION

The underlying theories, goals, and doctrines of patent law rely on the assumption that the inventions described in patents work and are replicable by others.[1] Similarly, patent law scholarship assumes that, when the patentee has actually made and tested the invention, the invention works and is replicable.[2] But, in contrast to this sensible conventional wisdom, this Article argues that most patented inventions – in the life sciences and likely elsewhere – probably do not work and are not replicable. I provide the first empirical evidence that, in the life sciences, even patents that disclose extensive experimentation to verify the utility and functionality of the invention still often do not work and cannot be replicated.[3]

My argument draws on the scientific literature on replicability. For the past decade, there has been widespread attention to the "replicability crisis" in science.[4] Studies attempting to replicate pre-clinical experiments have found that a shocking 90% of experiments published in well-respected, peer-reviewed journals are not replicable.[5] The cost of irreplicability is enormous – economists estimate that a conservative 50% irreplicability rate in pre-clinical research in the United States alone would cost $28 billion per year.[6] Irreplicability is also blamed for an increasing inability to translate promising pre-clinical research into effective human treatments, which delays bringing lifesaving drugs to market.[7] The popular press has dubbed irreplicability a "crisis" and has reported on it extensively.[8] The replicability crisis has similarly been a high priority for institutions such as the NIH[9] and NSF,[10] and hundreds

---

[1] Section I.B.1, *infra*. These requirements are formalized in 35 U.S.C. §§ 101, 112.

[2] *See, e.g.*, Christopher A. Cotropia, *The Folly of Early Filing in Patent Law*, 61 HASTINGS L.J. 1, 64 (2014); Mark A. Lemley, *Ready for Patenting*, 96 B.U. L. REV. 1171, 1198 (2016); Lisa Larrimore Ouellette, *Pierson, Peer Review, and Patent Law*, 69 VAND. L. REV. 1825, 1830 (2016); Sean B. Seymore, *Heightened Enablement in the Unpredictable Arts*, 56 UCLA L. REV. 127, 145 (2008).

[3] The only scholar who discusses the possibility that patents with experimental evidence might be wrong is Jacob Sherkow, however, his work is not empirical and does not estimate the scope of the problem. *Patent Law's Reproducibility Paradox*, 66 DUKE. L.J. 845, 846 (2017).

[4] Section I.A., *infra*.

[5] C. Glenn Begley and Lee M. Ellis, *Drug Development: Raise Standards for Preclinical Cancer Research*, 483 NATURE 531, 532 (2012).

[6] Leonard P. Freedman, Iain M. Cockburn, Timothy S. Simcoe, *The Economics of Reproducibility in Preclinical Research*, 13 PLOS BIOLOGY e1002165, 1 (2015).

[7] Jack W. Scannell and Jim Bosley, *When Quality Beats Quantity: Decision Theory, Drug Discovery, and the Reproducibility Crisis*, 11 PLOS ONE e0147215, 2 (2016).

[8] *E.g.*, Joel Achenbach, *Many Scientific Studies Can't Be Replicated. That's a Problem*, WASH. POST (Aug. 27, 2015) Aaron E. Carroll, *Science Needs a Solution for the Temptation of Positive Results*, N.Y. TIMES (May 29, 2017); Richard Harris, *The Breakdown in Biomedical Research*, WALL ST. J. (Apr. 7, 2017).

[9] Francis S. Collins & Lawrence A. Tabak, *NIH Plans to Enhance Reproducibility*, 505 NATURE 612, 612 (2014).

[10] Subcommittee on Replicability in Science Advisory Committee to the National Science Foundation Directorate, *Social, Behavioral, and Economic Sciences Perspectives on Robust and Reliable Science*, NAT'L. SCI. FOUND. (May 2015).

of prominent scientific journals have devised formal policies to combat irreplicability.[11] Though these policies have not been effective,[12] their prevalence underscores the gravity of the crisis.

Here, I show empirically that the replicability crisis in the scientific literature extends to patents and that the irreplicability rates of experiments described in patents are similarly high. I assess the replicability of experiments in patents by measuring their methodological quality, scoring methodological quality using a checklist developed by the journal *Nature*.[13] Methodological quality has been validated as a proxy for replicability in the scientific literature.[14] The rationale behind the proxy is that experiments that omit basic techniques to ensure reliability such as randomization or statistical analysis are less likely to be replicable.

Measuring methodological quality does not translate into a specific estimate of replicability rates. To overcome this issue, I compare the methodological quality scores of experiments in patents and experiments reported in the scientific literature. Because experiments from the scientific literature are known to be frequently irreplicable, if experiments in patents have comparably low methodology quality scores, then experiments in patents are likely also frequently irreplicable.

I hand-coded a random sample of 500 pre-clinical experiments from granted patents and applications and scored their methodological quality. I found that these experiments have very poor methodological quality. Only 62% of experiments in patents in my sample disclosed sample size, 12% were randomized, 4% were blinded, 2% conducted replicate studies, and 63% had statistical analysis of any kind.[15] This is worse than the methodological quality in scientific papers, where more than 70% of experiments disclosed their sample size, approximately 15% were randomized, approximately 20% were blinded, and over 90% included statistical analysis.[16] These methodological quality numbers in scientific papers are frequently used to support the existence of a replicability crisis,[17] so the lower numbers in patents suggest that a crisis exists there too.

I looked specifically at life sciences patents. This is because life sciences patents are the most likely to contain experiments[18] and because the replicability crisis in the scientific

---

[11] Section I.A.4, *infra*.

[12] *Id*.

[13] Nature Publishing Group, *Reporting Life Sciences Research* (April 2015), https://www.nature.com/authors/policies/reporting.pdf.

[14] Section II.A, *infra*.

[15] Figure 1, *infra*.

[16] These numbers are approximate because many studies have used this method to assess methodological quality in scientific papers, and results differ somewhat between the studies. None of the studies assessed whether replicates were disclosed. *See* Table 2 and accompanying footnotes, *infra*.

[17] Section II.A, *infra*.

[18] Janet Freilich, *Prophetic Patents* (2018), *available at* https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3202493.

literature is most often discussed in the context of life sciences research.[19] However, though I focused on the life sciences here, replicability is likely a problem across all industries – although the reasons may be different in different industries.[20]

I additionally validate my methodology for use in patents. Poor methodological quality may reflect poor experimental design, or it may reflect poor experimental reporting. For instance, an experiment may be well-designed, but the drafting patent attorney may omit methodological detail from the patent. Though the connection between quality of experimental design and quality of experimental reporting has been confirmed in scientific articles, patents are written using different conventions and for different purposes, so the link may not be present in patents.

I use several approaches to connect quality of methodological reporting with quality of methodological design. First, I show that patents covering FDA-approved drugs have better methodological quality than their non-commercialized counterparts.[21] Because these patents are commercialized and are – hopefully[22] – replicable, this suggests that better reported methodology correlates with genuine quality of experiment. Second, I show that methodological quality is correlated with the scientific institution filing the patent but is not correlated with the law firm drafting the patent.[23] This suggests that the methodology seen in patents reflects the design of the experiment, not the preference of the drafting attorney. Finally, I compare experiments in patents and papers with the same inventors/authors. The methodological quality scores between patents and papers are very similar.[24]

My findings demonstrate a serious mismatch between patent theory and doctrine and the way that patents function in practice. Patent theory and doctrine rely on the assumption that, when a patent is filed, it has been "reduced to practice" – meaning that the invention works.[25] The reality, however, is that most inventions do not work.

This mismatch generates several problems. First, because a patent's scope is generally broader than the experiments that support the patent, an irreplicable experiment can be used

---

[19] Section I.A.1, *infra*. However, there are also replicability crises in other disciplines, most notably psychology. *E.g.*, Monya Baker, *Over Half of Psychology Studies Fail Reproducibility Test*, NATURE NEWS (Aug. 27, 2015).

[20] Lisa Ouellette, *Who Reads Patents?* 35 NATURE BIOTECHNOLOGY 421, Supplementary Fig. 4 (2017) (when researchers where asked "Do you think you could recreate the invention described in the most recent patent you read in your field?" fewer than 45% of researchers in any field answered affirmatively).

[21] Section II.B.2(b), *infra*.

[22] Sherkow, *supra* note 3, at 846 (arguing that some Orange Book listed patents are not replicable, to the detriment of human health).

[23] Section II.B.2(a), *infra*.

[24] Section II.B.2(c), *infra*.

[25] *See, e.g.*, Wiesner v. Weigert, 666 F.2d 588, 594 (CCPA 1981) ("[the] invention is not reduced to practice until its practicability or utility is demonstrated pursuant to its intended purpose.").

to obtain a patent that covers technology that *does* work – creating the potential for harm.[26] For example, imagine that inventor A finds that a new drug treats cancer in mice and, on the strength of that finding, obtains a patent. The patent will cover any use of drug A – to treat cancer, to treat any other disease, or even for a non-medical use such as shoe polish. It is later found that inventor A is entirely wrong and the experiment is irreplicable – the drug does nothing to treat cancer. However, inventor B then discovers that the drug does treat HIV.[27] If inventor B wants to use or sell the drug for purposes of treating HIV, she must obtain a license from inventor A – even though inventor A was wrong and inventor B was right. Although inventor A's patent may not be valid,[28] it is time-consuming and expensive to prove invalidity in court, and so an inoperable patent can still be used to collect rents from innovators developing operable technology.[29]

Irreplicable experiments can cause additional harms. Patents on technologies that do not work overload the patent system, burdening examiners, creating patent thickets, and providing fodder for patent trolls.[30] They also simply fail to implement the goal of the patent system: incentivizing the development of useful technologies – since an irreplicable experiment is not useful. Moreover, they create considerable waste, since the labor and materials that went into conducting the experiment are squandered.

These problems arise because the patent system evaluates experiments in a way that makes little sense. Patents are filed early in the life cycle of an invention[31] and many of the experiments reported in patents are preliminary investigations into the functionality of the invention. Preliminary experiments are, by their nature, somewhat speculative and will often be proven wrong by later, more intensive, experimentation.[32] Yet these preliminary experiments are used to satisfy the requirements of patentability – they are the grounds upon which patents are granted.[33]

Although the experiments that provide evidence for patentability are tentative, the rights

---

[26] Section I.B.3(b), *infra*.

[27] This is loosely based on the real story of the development of azidothymidine (AZT). Alice Park, *The Story Behind the First AIDS Drug*, TIME (March 19, 2017).

[28] Patents must be operative to be valid. 35 U.S.C. § 101. Inoperative patents, or patents supported by many experiments that do not work, may also be invalid for failure to satisfy the enablement requirement. 35 U.S.C. § 112. However, not every patent with an irreplicable experiment will be invalid. *See* Section I.B.1, *infra*.

[29] Leveraging the cost of litigation to extract rents is a common "patent troll" strategy. Doug Lichtman & Mark A. Lemley, *Rethinking Patent Law's Presumption of Validity*, 60 STAN. L. REV. 45, 48 (2007).

[30] *E.g.*, James Bessen and Michael Meurer, PATENT FAILURE: HOW JUDGES, BUREAUCRATS, AND LAWYERS PUT INNOVATORS AT RISK, 67 (2008).

[31] Sichelman, *supra* note **Error! Bookmark not defined.**, at 343.

[32] Section III.A, *infra*.

[33] Specifically, experiments can be used to satisfy the utility, enablement, and written description requirements. 35 U.S.C. 101, 112. The relationship between experiments and patentability is discussed further in Section I.B, *infra*.

that attach to the patent are not. On the basis of early-stage experiments that are often incorrect, the patentee gets the powerful legal right to exclude others from making or using the invention. These rights are practically permanent for the life of the patent – a granted patent is presumed valid and is therefore difficult to challenge in litigation.[34] Further, because the patentee both need never update the experiment and can prevent others from repeating the experiment and therefore from testing its validity,[35] the preliminary science described in the patent stands as truth unless the patentee chooses otherwise. The replicability literature – and the findings of this Article – teaches us that experiments of the sort reported in patents are not reliable enough to merit this level of control and influence.

Here, I emphasize a different relationship between functionality and patenting that better reflects the actuality of how science progresses and how patents are filed. We should recognize that patents are not filed after an invention works; rather, they are early-stage inventions that may or may not work. I argue that it would be prohibitively expensive to delay patenting until we are quite confident that inventions work,[36] therefore the better solution is to reconceptualize patent law to adapt to the reality that we do not know if patented inventions are functional.

To this end, we should make it easier to update experimental disclosure and to identify and invalidate patents based on irreplicable experiments. I recommend clarifying the experimental use exception to make plain that replication attempts are not infringement.[37] I further propose creating a system to collect data obtained after patent filing.[38] Thus, if a patentee later finds that the experiment is irreplicable, that finding will be attached to the patent. Finally, I suggest reducing the cost and time needed to invalidate irreplicable patents, specifically removing the presumption of operability and considering procedures outside the courtroom for determining questions of enablement and utility.[39]

The Article proceeds as follows: Part I provides background on the irreplicability crisis in the scientific literature and then discusses the role of experiments in patents, why they might be irreplicable, and the harm irreplicable experiments in patents could cause. Part II describes the empirical study, including methodology and results. Part III explores the implications of the empirical findings and suggests policy reform.

I. UNDERSTANDING IRREPLICABILITY

[34] 35 U.S.C. 282.

[35] Making or using the patented invention is patent infringement under 35 USC 271(a). Although there is an experimental exception defense that may apply to replication attempts, the scope of this defense has been unclear in recent years. *See* Section I.B.3(d), *infra*.

[36] Section III.A, *infra*.

[37] Section III.B.1, *infra*.

[38] Section III.B.3, *infra*.

[39] Section III.B.2, *infra*.

Before filing a patent, inventors must show that their invention works.[40] To be sure, the inventor need not fully test the invention,[41] nor create a commercially viable version[42] – in fact, the inventor need not even create a physical model of the invention[43] – but the inventor must have some evidence that the invention is functional. A mere hunch is insufficient, as are "crude and imperfect experiments."[44]

It is particularly important than an invention work before patenting because the patent system relies on various aspects of an invention's functionality in order to properly determine patent scope and inventorship. The scope of the patent should correspond to the scope of the invention, so we limit the scope of a patent to the aspects of the invention that the patentee could make work.[45] In addition, patents should be granted to the inventor of the claimed invention, and the inventor is generally thought of as the person who makes the invention work.[46] Further, patents are supposed to disclose useful information about how to make and use new technologies[47] and instructions on how to make and use a product that does not work are not helpful. In short, if an invention does not work, the patent system as applied to that invention does not work. The functionality of inventions is a foundational assumption upon which the patent system stands.

The problem is that a lot of inventions probably don't work.

---

[40] *E.g.*, Mycogen Plant Sci., Inc. v. Monsanto Co., 243 F.3d 1316, 1332 (Fed. Cir. 2001) ("To establish an actual reduction to practice, it is necessary to show that…[the invention] was shown or known to work for its intended purpose."); Loral Fairchild Corp. v. Matsushita Elec., 366 F.3d 1358, 1362 (Fed. Cir. 2001) ("Once the invention has been shown to work for its intended purpose, reduction to practice is complete."); DSL Dynamics Scis. Ltd. V. Union Switch & Signal, Inc., 928 F.2d 1122, 1125 (Fed. Cir. 1991) (requiring patentees to show that "the embodiment relied upon as evidence of priority actually worked for its intended purpose.").

[41] Taskett v. Dentlinger, 344 F.3d 1337, 1342 (Fed. Cir. 2003) ("That Dentlinger did not test this step of the counter under conditions of actual use does not mean that he did not reduce it to practice. His test was sufficient to determine that the invention would work for its intended purpose.").

[42] *Id*. ("To hold otherwise would be to require an inventor to have created a viable commercial embodiment before the Board or a court could find reduction to practice. This the law does not require.").

[43] *See, e.g.*, The Telephone Cases, 126 U.S. 1, 535-536 (1888) (upholding a patent granted to Alexander Graham Bell even though Bell had not created a working version of the telephone before filing the patent application. The Court noted that Bell had written a set of instructions on how to make a telephone – instructions that were accurate – and that this was sufficient to enable his patent). *See also* Pfaff v. Wells Elecs., Inc., 525 U.S. 55, 56 (1998) (applying the rule from *The Telephone Cases*).

[44] Seymour v. Osborne, 78 U.S. 516, 517 (1870).

[45] *E.g.*, Plant Genetic Systems, N.V. v. DeKalb Genetics Corp., 315 F.3d 1335, 1339-40 (Fed. Cir. 2003); In re Vaeck, 947 F.2d 488, 495 (Fed. Cir. 1991) ("the first paragraph of §112 requires that the scope of the claims must bear a reasonable correlation to the scope of enablement provided by the specification.").

[46] *E.g.*, Newkirk v. Lulejian, 825 F.2d 1581, 1582 (Fed. Cir. 1987) ("proof of actual reduction to practice requires demonstration that the embodiment relied upon as evidence of priority actually worked for its intended purpose); Wiesner v. Weigert, 666 F.2d 588, 594 (CCPA 1981) ("invention is not reduced to practice until its practicability or utility is demonstrated pursuant to its intended purpose.").

[47] 35 U.S.C. § 112(a).

A. The Replicability Crisis in Science

The replicability crisis in science is fundamentally about the discovery that many inventions that we thought worked actually do not work – even when the inventors are reputable scientists, even when the invention has been thoroughly peer reviewed, and even when the invention is published in a prominent, respected journal.

Replicability is the ability of scientists to re-do an experiment.[48] For example, a study might find that a drug shrinks tumors in mice. Scientists attempting to replicate the experiment will try to test the drug in a new set of mice, following the protocol of the original experiment as closely as possible. Irreplicability occurs when an experiment is re-done and the original results cannot be repeated.[49] In the example above, the replicators might find that when done again, the drug has no effect on the size of tumors in mice. Irreplicability means that the experiment does not work. Irreplicable experiments tell us something about the world that is not true.

Irreplicability is not just about failure to replicate precise results. Rather, the irreplicability crisis has garnered so much attention because the big ideas from studies could not be repeated.[50] Even well-regarded studies that had been cited hundreds of times could not be replicated.[51] Human trials – carefully reviewed by the FDA – were based on pre-clinical studies that were later found to be irreplicable.[52] Irreplicability is therefore about more than just failure of a study to work when tried again – it is about a multitude of spectacular, impactful failures that have thrown the scientific world into crisis. Irreplicability impedes our ability to make scientific progress, to innovate, and ultimately to produce lifesaving technologies.

1. Overview of the Replicability Crisis

In 2005, John Ioannidis, a Professor of Medicine and Health Research at Stanford University, published a paper titled "Why Most Published Research Findings Are False."[53] The article argued that scientific researchers were influenced by certain incentives and design constraints that would inevitably lead to the publication of irreplicable results.[54]

---

[48] This is distinct from reproducibility, which is the ability to re-run an analysis from the same set of data. Steven N. Goodman, Daniele Fanelli, and John P. A. Ioannidis, *What Does Research Reproducibility Mean?*, 8 SCI. TRANSLATIONAL MEDICINE 341, 341 (2016).

[49] There are many different measures of replicability, and no clear agreement on how close one must be to the original to be considered replicable. Stefan Schmidt, *Shall We Really Do It Again? The Powerful Concept of Replication is Neglected in the Social Sciences*, 13 REV. GEN. PSYCHOLOGY 90, 91 (2009).

[50] C. Glenn Begley & John P.A. Ioannidis, *Reproducibility in Science: Improving the Standard for Basic and Preclinical Research*, 2015 CIRCULATION RES., 116, 117.

[51] *Id*.

[52] *Id*.

[53] John P.A. Ionnidis, *Why Most Published Research Findings Are False*, 2 PLOS MEDICINE e124 (2005).

[54] *Id*.

Ioannidis' theory has been influential – as of 2018, Ioannidis' paper has been cited 5,567 times.[55]

Over the next several years, other scholars attempted to test Ioannidis' theoretical claim by actually trying to replicate experiments. The results were staggering. One well-known attempt at experimentally replicating previous work found that an astonishing 89% of pre-clinical experiments in the fields of hematology and oncology were irreplicable.[56] The replicability team specifically chose "landmark" studies, so the irreplicable papers were not obscure, poorly regarded works, but were quite the opposite.[57] Another attempt at experimentally replicating pre-clinical studies in oncology, women's health, and cardiovascular disease found that only 20-25% of the studies were replicable.[58] An attempt at replicating mouse trials of drugs to treat amyotrophic lateral sclerosis (ALS) found that a striking 0% of the drugs showed any beneficial effect when the experiments were replicated.[59] Venture capital firms, which often assess the viability of early-stage data have an "unspoken rule" that "at least 50% of published studies, even those in the top-tier academic journals, can't be repeated with the same conclusions by an industrial lab."[60] A survey of researchers by the journal *Nature* found that over 70% of researchers failed to replicate a reported experiment (and over 50% failed to replicate their own experiments).[61]

2. Causes of Irreplicability

The causes of irreplicability are varied. Some lack of replicability is caused by improper statistical analysis – including low statistical power and *P*-hacking.[62] Incentives are also a problem.[63] Scientists are rewarded – through grant money, promotion, and otherwise – for publishing frequently and in well-regarded journals.[64] This is more likely to occur if the

---

[55] According to a citation count by Google Scholar on March 2, 2018.

[56] C. Glenn Begley and Lee M. Ellis, *Drug Development: Raise Standards for Preclinical Cancer Research*, 483 NATURE 531, 532 (2012). This study is called "best-known" by editors of the journal *Nature*. Monya Baker, *Is There a Reproducibility Crisis?* 533 NATURE 543, 543 (2016).

[57] *Id*.

[58] Florian Prinz, Thomas Schlange, and Khusru Asadullah, *Believe It Or Not: How Much Can We Rely on Published Data on Potential Drug Targets?*, 10 NATURE REVS. DRUG DISCOVERY 712, 712-13 (2011).

[59] Steve Perrin, *Preclinical Research: Make mouse studies work*, 507 NATURE 423, 424 (2014).

[60] Lev Osherovich, *Hedging Against Academic Risk*, 4 SCIENCE-BUSINESS EXCHANGE 1, 1 (2011).

[61] Monya Baker, *Is There a Reproducibility Crisis?* 533 NATURE 543, 543 (2016).

[62] ML Head, L Holman, R Lanfear, AT Kahn, and MD Jennions, *The Extent and Consequences of P-Hacking in Science*, 13 PLOS BIOLOGY e1002106, 1 (2015) (explaining that p-hacking "occurs when researchers try out several statistical analyses and/or data eligibility specifications and then selectively report those that produce significant results.").

[63] Marcus R. Munafo, Brian A. Nosek, Dorothy VM Bishop, et. al., *A Manifesto for Reproducible Science*, 1 NATURE HUMAN BEHAVIOR 21, 22 (2017).

[64] *E.g.*, C. Glenn Begley, Alastair M. Buchan, and Ulrich Dirnagl, *Institutions Must Do Their Part for Reproducibility*, 525 NATURE 25, 25-26 (2015); Francis S. Collins and Lawrence A. Tabak, *NIH Plans to Enhance Reproducibility*, 505 NATURE 612, 613 (2014); ("Perhaps the most vexed issue is the academic incentive system."); Elie Dolgin, *Drug Discoverers Chart Path to Tackling Data Irreproducibility*, 13 NATURE

scientist publishes a result that is positive, novel, and exciting. Researchers studying irreplicability believe that these incentives push scientists to use methodology that increases the likelihood of positive, novel, and exciting results – but decreases the likelihood of replicability.[65] For instance, there is no incentive to conduct studies on large numbers of samples, or to repeat one's study to ensure that it is correct before publishing.[66] Further, there is no incentive to publish negative results.[67]

Yet another cause of irreplicability is poor reliability of materials – scientists may be inadvertently working with impure samples or using the wrong cell line.[68] A final cause is poor methodological reporting.[69] Many studies are reported with insufficient details for another team to replicate their procedure.[70] If the replicating team must guess at details, it is not surprising that replication attempt will often fail.

3. Costs of Irreplicability

The costs of irreplicability are enormous. Economists estimate that a 50% irreplicability rate in pre-clinical research in the life sciences alone would cost $28 billion.[71] One major cost comes from waste – materials, time, and effort spent conducting an experiment that produces misleading results. The need to check whether experiments are replicable is also costly. Pharmaceutical companies try to replicate previous research before beginning a new project, and these attempts typically take up to two years and cost up to $2 million each.[72] Venture capital companies may also attempt to replicate experiments before fully investing in a company. Atlas Venture reports that it invests between $50,000 to $500,000 to validate

---

REVIEWS DRUG DISCOVERY 875, 875 (2014).

[65] AD Higginson and M Munafo, *Current Incentives for Scientists Lead to Underpowered Studies with Erroneous Conclusions,* 14 PLOS BIOLOGY e2000995, 1 (2016).

[66] *Id*.

[67] B. Nosek, J.R. Spies, M. Motyl, *Scientific Utopia: Restructuring Incentives and Practices to Promote Truth Over Publishability*, 7 PERSPECTIVES IN PSYCHOLOGICAL SCIENCE 615, 616 (2012). Failure to publish negative results can lead to the appearance of a positive result that is really due to chance. For instance, let us say that twenty scientists set out to do an experiment. One, by random luck, gets a positive result, while the other nineteen get negative results. The one successful researcher will publish the result, while the other studies are relegated to file drawers – and the technique will appear successful, even though it is clearly not (using a p-value of 0.05, one in twenty positive results will be due to chance).

[68] Begley, Buchan, and Dirnagl, *supra* note 64 at 26. HeLa cells, derived from a cervical cancer sample taken unknowingly from Henrietta Lacks (REBECCA SKLOOT, THE IMMORTAL LIFE OF HENRIETTA LACKS, 76 (2010)) are one of the most common contaminants in other cell lines. Jill Neimark, *The Dirty Little Secret of Cancer Research*, DISCOVER MAGAZINE (Oct. 2, 2014).

[69] Story C. Landis, Susan G. Amara, Khusuru Asadullah, Crhis P. Austin, et. al., *A Call for Transparent Reporting to Optimize the Predictive Value of Preclinical Research*, 490 NATURE 187, 187 (2012).

[70] *Id*.

[71] *Id*.

[72] Pharmaceutical Research and Manufacturers of America, 2013 BIOPHARMACEUTICAL INDUSTRY PROFILE, 78 (2013).

the data of an early stage company before making more substantial investments.[73]

There are also non-monetary costs to irreplicability. The first is lack of trust. It is demoralizing to scientists themselves, and also affects the public, as the public's trust in scientific research diminishes. There are additionally ethical and human costs. To illustrate, in the 1980s, scientists studying the use of high-dose chemotherapy with autologous bone marrow transplant (HDC-ABMT) to treat breast cancer produced promising preliminary results.[74] These results were so encouraging that patients pushed their doctors to provide the treatment, patient groups lobbied the FDA to allow access, and one woman won a $89 million in punitive damages against her insurance company for their failure to cover the treatment.[75] Over the next decade, 41,000 patients underwent the treatment.[76] In the late 90s, further study showed conclusively that the treatment did not work.[77]

4. Efforts to Fix Irreplicability

Many organizations have attempted to fix these problems. For example, *PLoS* (the Public Library of Science) is publishing findings even if the authors were not the first to discover the phenomenon, which provides an incentive for replication attempts.[78] *Nature* has implemented a checklist for methodological completeness that must be completed along with article submission.[79] The British Psychological Society has launched an initiative to allow authors to pre-register plans for experiments.[80] *Cell* is providing unlimited space online for a supplemental methodology section.[81] To improve study design and prevent publication of studies with inadequate methodological descriptions, over 300 groups have published guidelines or checklists on best practices in methodological design and reporting.[82] Several journals and grant-giving agencies now require submission of a checklist

[73] Osherovich, *supra* note 60, at 1.

[74] Michelle Mello and Troyen Brennan, *The Controversy Over High-Dose Chemotherapy With Autologous Bone Marrow Transplant for Breast Cancer*, 20 HEALTH AFFAIRS 101, 103 (2001).

[75] *Id*. at 106-07.

[76] *Id.* at 101.

[77] *Id*. at 107.

[78] PLOS Biology Staff Editors, *The Importance of Being Second*, 16 PLOS BIOLOGY e2005203, 1 (2018) ("we are formalizing a policy whereby manuscripts that confirm or extend a recently published study ('scooped' manuscripts, also referred to as complementary) are eligible for consideration at PLOS Biology…This new policy…addresses the current concern regarding the reproducibility, or lack thereof, of scientific findings.").

[79] Editorial Staff, *Towards Greater Reproducibility*, 546 NATURE 8 (2017); *See also* Editorial Staff, *Checklist Checked*, 556 NATURE 273-74 (2018).

[80] *Working with Wiley to Improve the Replicability and Transparency of Research*, BRITISH PSYCHOL. SOC'Y (2017), https://www.bps.org.uk/news-and-policy/we-are-working-wiley-improve-replicability-and-transparency-research.

[81] Emilie Marcus, *Scientific Credibility and Reproducibility*, 159 CELL 965, 965 (2014).

[82] Munafo, *supra* note 63, at 4. These guidelines are aggregated by The EQUATOR Network (Enhancing the QUAlity and Transparency Of Health Research) (2018), *Your One-Stop-Shop for Writing and Publishing High-Impact Health Research*, http://www.equator-network.org/.

with each article, or encourage their peer reviewers to use the checklists.

Unfortunately, most evidence shows that these measures are not working.[83] For example, in 2004, *Nature* instituted a policy requiring that papers published in the journal comply with certain guidelines on reporting statistics. Yet a follow-on paper in 2012 found that "it is still common" for articles published in *Nature* to not comply with the guidelines.[84] Further, even efforts that do work are limited to individual journals or groups of journals. Though the NIH – which has the power to affect a far greater audience than individual journals – has instituted some measures, the NIH does not control publication of results, and so, in the words of Francis Collins, the director of the NIH, irreplicability "is not a problem the NIH can tackle alone."[85]

B. Replicability and the Patent System

As mentioned above, the patent system is predicated on an assumption that the inventions described in patents actually work. However, many patents are based on scientific experiments – and the dominant conversation for the past decade about scientific experiments has been about the frequency with which they do *not* work. The natural question, then, is 'do experiments in patents work'?

Part 1, below, provides background on the role of experiments in patents. It discusses when and why patentees include experiments and how experiments are used to satisfy various doctrines of patentability. Part 2 asks whether we might expect experiments in patents to be replicable. It explores patentee incentives to avoid irreplicability and whether the PTO assess replicability. Part 2 concludes that it is as least plausible that irreplicable experiments are common in patents. Given this conclusion, Part 3 surveys potential consequences of irreplicability.

---

[83] For example, the Animal Research Reporting of *In Vivo* Experiments (ARRIVE) guidelines have been endorsed by over 300 journals, all major UK funding agencies, and the US National Research Council Institute for Laboratory Animal Research. However, a study of the guidelines two years after their implementation found that there was widespread failure to comply with the guidelines among artless published in the very journals that had endorsed the guidelines. David Baker, Katie Lidster, Ana Sottomayor, & Sandra Amor, *Two Years Later: Journals Are Not Yet Enforcing the ARRIVE Guidelines on Reporting Standards for Pre-Clinical Animal Studies*, 12 PLOS BIOLOGY e1001756, 1 (2014). Similarly, in 2003, the National Academies of Science, Engineering, and Medicine drafted guidelines about what information should be included in publications to support replicability. A study done in 2012 found that, while the number of journals implementing these guidelines was increasing, only a small minority had policies implementing all guidelines. Victoria Stodden, Peixuan Guo, and Zhaokun Ma, *Towards Reproducible Computational Research: An Empirical Analysis of Data and Code Policy Adoption by Journals*, 8 PLOS ONE e67111, 1 (2013).

[84] David L. Vaux, *Know When Your Numbers Are Significant*, 492 NATURE 180, 180 (2012).

[85] Francis S. Collins and Lawrence A. Tabak, *Policy: NIH Plans to Enhance Reproducibility*, 505 NATURE 612, 613 (2014).

1. The Role of Experiments in Patents

The inability of researchers to replicate experiments in the peer reviewed scientific literature is relevant to patents because experiments are a central component of the patent document. The Constitution empowers Congress to grant patents to "promote the Progress of Science."[86] Patents do this in two ways. They give inventors the exclusive right to make and use their invention, which allows inventors to profit from their inventions and thereby incentivizes the creation of those inventions.[87] In addition, patents disclose information about new technologies to the public so that the public can then build on and further develop those technologies.[88] These are part of the basic bargain of the patent system: the patentee gets an exclusive right and in return the public gets the creation and disclosure of new technology.

Experiments have two roles in this basic bargain. First, they help prove that patent applicants have in fact invented something that will benefit the public and second, they facilitate disclosure of information about the technology. Experiments are therefore tightly linked to the goals of the patent system.

a. Why Patentees Use Experiments

Patents do not need to contain experiments. There is no legal requirement that the patented invention be described or supported using experimental evidence.[89] However, patents, particularly in chemistry and the life sciences, commonly do contain experiments.[90] The purpose of describing experiments in the patent document is to satisfy the disclosure doctrines – utility, enablement, and written description.[91] Below, I describe each doctrine in turn and explain how experiments can help comply with the doctrine.

Utility: In order to satisfy the utility doctrine, patents must state specifically why the invention is useful.[92] For example, a patent on a particular molecule might explain that the molecule can be used to treat a disease.[93] The requirement prevents patenting of inventions

---

[86] U.S. CONST. art I, §8, cl. 8.

[87] *E.g.*, Dan L. Burk & Mark A. Lemley, *Policy Levers in Patent Law*, 89 VA. L. REV. 1575, 1597 (2003); Matthew Hershkowtiz, *Patently Insane for Patents*, 28 FORDHAM INTELL. PROP. MEDIA & ENT. L.J. 109, 116 (2017).

[88] *E.g.*, Jeanne Fromer, *Patent Disclosure*, 94 IOWA L. REV. 539, 548-50 (2008).

[89] MPEP 2164.02 ("Compliance with the enablement requirement…does not turn on whether an example is disclosed…lack of working examples or lack of evidence that the claimed invention works as described should never be the sole reason for rejecting the claimed invention on the grounds of lack of enablement.").

[90] Janet Freilich, *Prophetic Patents* (2018), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3202493 (finding that approximately 50% of chemistry and life sciences patents contain experiments, and noting that the true number of patents containing experiments is likely higher than this figure).

[91] 35 U.S.C. §§ 101, 112(a)-(b).

[92] 35 U.S.C. § 101.

[93] MPEP § 2107 ("Such statements will usually explain the purpose of or how the invention may be used

that have no purpose, do not work,[94] or where the inventor does not know the utility of his invention.[95] This ensures that the public receives a benefit – an invention that does something useful – in exchange for giving the patentee a monopoly.

The statement of utility must be credible, and the patent examiner assesses the credibility of the statement on the basis of the logic and facts that underlie the statement.[96] Showing utility is a low bar, and examiner will not ordinarily dispute an applicant's statement of utility.[97] However, an applicant's case for utility is stronger if the applicant can show experimental evidence that the invention is useful for the stated purpose.[98] The PTO has rejected applications where the patent provides only general statements of utility without experiments to back them up.[99] Thus, a mere statement that a molecule can treat a particular disease might not be seen as credible by an examiner whereas an experiment showing that a molecule effectively targets a disease is clear evidence of utility. For example, the experiment below demonstrates utility because it shows that the patented compound is useful for the purpose of reducing tumor size:[100]

> One hundred and twenty Balb/c mice were divided into…groups of 40. One group was immunized…with sterile saline [as a control. And another group] was immunized with [the patented invention]…Each mouse was administered $5x10^5$ tumor cells…The only group showing any protection…was the animals which received [the patented invention]…after 25 days, 6 of the animals showed no detectable tumor growth…In contrast, all the mice in other groups have tumors between 1.5 and 3.0cm.[101]

Enablement: Patents must contain a sufficient explanation of the invention that

(e.g. a compound I believed to be useful in the treatment of a particular disorder).").

[94] An invention that is totally inoperable is not useful because it is not functional. *E.g.*, In re Swartz, 232 F.3d 862, 863 (Fed. Cir. 2000).

[95] Brenner v. Manson, 383 US 519, 536 (1966) ("a patent is not a hunting license. It is not a reward for the search, but compensation for its successful conclusion.").

[96] MPEP § 2107.

[97] In re Langer, 503 F.2d 1380, 1391 (CCPA 1974) ("a specification which contains a disclosure of utility which corresponds in scope to the subject matter sought to be patented must be taken as sufficient to satisfy the utility requirement of § 101…unless there is a reason for one skilled in the art to question the objective truth of the statement of utility or its scope.") (emphasis in original).

[98] *See, e.g.*, In re Wright, 999 F.2d 1557, 1560 (1993) (finding that, in the absence of experimental evidence that a vaccine against RNA viruses would work, there was no evidence that the applicant's invention would be useful to create such vaccines.). *See also* Bratislav Stankovic, *The Use of Examples in Patent Applications*, 18 INTELL. PROP. & TECH. L.J. 9, 10 (2006).

[99] Ex Parte Sudilovsky, 1991 WL 332566, *6 (BPAI 1991).

[100] Ultimately, the goal is likely treatment of cancer in humans. However, the Patent Office will accept mouse models (or *in vitro* models) as evidence of utility in humans as long as there is a correlation between the model and use in humans. In re Brana, 51 F.3d 1560, 1566 (Fed. Cir. 1995); MPEP § 2164.02 .

[101] U.S. Patent No. 6,565,852, Example 2 (issued May 20, 2003).

another scientist could make and use the invention without "undue experimentation."[102] The enablement requirement pushes the Constitutional mandate that patents "promote the progress of science" by ensuring that patents disclose enough information that others can build on the invention.[103]

Experiments are also used to satisfy the enablement doctrine. As with the utility requirement, experiments are not strictly necessary to satisfy enablement.[104] However, experiments are a convenient and effective way to fulfill the enablement requirement and the number of experiments present in the patent is a formal factor in the enablement analysis.[105] An experiment that describes the step-by-step process of making or using the invention is essentially an instruction manual to others who want to make or use the invention, and so can enable a patent.

Written Description: Patents must describe the invention in terms that are sufficiently complete to demonstrate that the inventor was in "possession" of the invention.[106] Like enablement, the written description doctrine is linked to the Constitutional *quid pro quo* because it ensures that the public is given a meaningful invention in exchange for the patent.[107] Though the requirement is somewhat amorphous – the Federal Circuit has acknowledged that "the term 'possession'…has never been very enlightening" – the doctrine requires that scientists be able to read the patent and understand that the patentee actually invented the claimed invention.[108]

As with the other requirements of patentability, experiments are not necessary to satisfy the written description requirement, but they can help. For example, the court in *Wyeth v. Abbott* invalidated a patent for lack of written description because the inventors claimed to have invented a method of treating a condition by administering rapamycin rectally or transdermally, but the inventors had not tried these methods of administration, nor did they know if rectal or transdermal administration would work, and rapamycin had never been

---

[102] In re Wands, 858 F.2d 731, 737 (Fed. Cir. 1988).

[103] *See, e.g.*, AK Steel Corp. v. Sollac and Ugine, 344 F.3d 1234, 1244 (Fed. Cir. 2003) ("as part of the *quid pro quo* of the patent bargain, the applicant's specification must enable one of ordinary skill in the art to practice the full scope of the claimed invention.").

[104] MPEP § 2164.02 ("Compliance with the enablement requirement…does not turn on whether an example is disclosed.").

[105] Wands, 858 F.2d at 737.

[106] Capon v. Eshhar, 418 F.3d 1349, 1357 (Fed. Cir. 2005).

[107] Rochester v. GD Searle & Co., Inc., 358 F. 3d 916, 1922 (Fed. Cir. 2004). *See also*, Ariad Pharmaceuticals, Inc. v. Eli Lilly and Co., 598 F.3d 1336, 1345 (Fed. Cir. 2010) ("a separate requirement to describe one's invention is basic to patent law. Every patent must describe an invention. It is part of the *quid pro quo* of a patent…").

[108] Ariad Pharmaceuticals, Inc. v. Eli Lilly and Co., 598 F.3d 1336, 1351-52 (Fed. Cir. 2010); Hologic, Inc. v. Smith & Nephew, Inc. __ F.3d __ (Fed. Cir. 2018). Note that this does not mean that the inventor needs to understand every mechanism underlying the claimed invention.

administered by those routes by anyone else when the patent was filed.[109] The court specifically cited the lack of experiments as a reason for finding inadequate written description.[110]

Once a patent is granted, it is presumed valid, meaning that it is presumed useful, enabled, and adequately described.[111] As a practical matter, this means that any experiments used to prove utility, enablement, and written description are also presumed to work. This presumption can be challenged in litigation, but only at great expense in both money and time.

b. Relationship Between Experiments and Patent Scope

Though experiments help satisfy the requirements for patentability as described above, patents can be far broader than just the material described in the experiment.[112] The experiment demonstrates one way in which the invention can be used, but through this experiment, an inventor can get a patent covering *all* ways in which an invention can be used.[113] For example, a scientist who discovers that drug X treats disease Y will obtain a patent covering not only use of drug X to treat disease Y, but also the use of drug X to treat other conditions, the combination of drug X with drug Z treat yet another condition, or any other use of drug X for any purpose – irrespective of whether the patentee was aware of that purpose.[114] The relationship between experiment and patent claim can best be envisioned as the experiment as a core embodiment near the center of the broader claim.[115] The experiment represents the inventor's current thinking about how the invention might work, but the claim is intended to provide enough patent coverage to protect against downstream developments

---

[109] Wyeth v. Abbott Labs., 2012 WL 175023, at *8 (D.N.J. 2012), *aff'd* Wyeth & Cordis Corp. v. Abbott Labs., 720 F.3d 1380 (Fed. Cir. 2013).

[110] *Id*. at *9 ("Here, the specification contains no data, examples, or other disclosures sufficient to demonstrate that the inventors were in possession of the full scope of their invention.").

[111] 35 U.S.C. 282.

[112] *E.g.*, Cont'l Paper Bag Co. v. E. Paper Bag Co., 210 U.S. 405, 418-19 (1908) (explaining that patents cover not only the embodiment created by the inventor, but also the "principle" of the invention.). *See also* Janet Freilich, *The Uninformed Topography of Patent Scope*, 19 STAN. L. REV. 150, 152 (2015).

[113] This is because the utility requirement is satisfied by disclosure of one way in which the invention is useful, and the enablement requirement is satisfied by disclosure of one way in which the invention can be made or used. Thus, one experiment can enable a far broader claim. *See, e.g.*, MPEP § 2164.01(b) ("As long as the specification discloses at least one method for making and using the claimed invention that bears a reasonable correlation to the entire scope of the claim, then the enablement requirement of 35 U.S.C. § 112 is satisfied."). *See also* CFMT, Inc. v. Yieldup Int'l. Corp., 349 F.3d 1333, 1338-39 (Fed. Cir. 2003); Johns Hopkins Univ. v. CellPro, Inc., 152 F.3d 1342, 1359-61 (Fed. Cir. 1998).

[114] The patentee does not need to be aware of how the invention works. *See, e.g.*, Abbott Labs. V. Geneva Pharms., Inc., 182 F.3d 1315, 1318 (Fed. Cir. 1999).

[115] Patents contain formal "claims" which set out the boundaries of the patent's scope. *See, e.g.*, Warner-Jenkinson Co. v. Hilton Davis Chem. Co., 520 U.S. 17, 29 (1997). *See also*, Dov Hirsch, *The Riddle of the Mysterious Patent Dance Wrapped in an Enigma*, 27 FORDHAM INTELL. PROP. MEDIA & ENT. L.J. 654, 652 (2017).

and discoveries, and variations by competitors.

To illustrate, the patent on sildenafil (Viagra®) describes experiments showing the drug's efficacy at treating hypertension.[116] These experiments were sufficient to support the validity of a claim to sildenafil generally – i.e. the patent covered any use of sildenafil.[117] This broad claim became a goldmine when doctors realized that sildenafil's most profitable use was not the treatment of hypertension but the treatment of erectile dysfunction.[118]

c. Relationship Between Experiments and Patent Validity

If an experiment in a patent does not work, is the patent invalid? The answer depends on the other content of the patent. First, the requirements of patentability apply to the patent claim as a whole, rather than to individual experiments. Thus, if an experiment teaching how to make some aspect of the claimed invention is irreplicable but a scientist could still figure out how to make the invention as a whole without undue experimentation, the claim is enabled.[119] It is well established in case law that a patent that claims some totally inoperative variations on the invention can still be enabled.[120] Thus, merely because one experiment in a patent is irreplicable, it does not mean that the claimed invention as a whole will be deemed nonenabled.[121] Similarly, if a claim is enabled by just one embodiment in the specification and that embodiment is inoperable, the claim is invalid, but if the specification contains multiple possible embodiments then one inoperative embodiment will not render the claim invalid, as long as the other embodiment is enabled.[122] For example, if a patent discloses two cell lines that can be used to produce the claimed antibodies, but the antibodies can only be

[116] U.S. Patent No. 5,250,534, col. 6, ll. 35-59 (issued Oct. 5, 1993). Note that the experiments in question are prophetic.

[117] *Id*. at claim 1.

[118] Pfizer, *How Does Viagra Work?* (2016), https://www.viagra.com/learning/how-does-viagra-work.

[119] Application of Cook, 439 F.2d 730, 735 (1971) ("…many patented claims read on vast numbers of inoperative embodiments…There is nothing wrong with this so long as it would be obvious to one of ordinary skill in the relevant art how to include those factors [that are omitted in the description of the embodiments] in such manner as to make the embodiment operative rather than inoperative."). There is no clear line as to what precisely constitutes undue experimentation and it varies with context, but case law suggests that quite a bit of experimentation can be permitted. For example, in one case the Federal Circuit held that three years of experimentation was not undue. Transocean Offshore Deepwater Drilling, Inc. v. Maersk Drilling USA, Inc., 699 F.3d 1340, 1356 (Fed. Cir. 2012). However, in another case the Federal Circuit found that experimentation of eighteen months to two years was undue. Cephalon, Inc. v. Watson Pharmaceuticals, Inc., 707 F.3d 1330, 1336 (Fed. Cir. 2013). For further discussion of the link between reproducibility and enablement, see Dmitry Karshtedt, *Limits on Hard-to-Reproduce Inventions: Process Elements and Biotechnology's Compliance with the Enablement Requirement*, 3 HASTINGS SCI. & TECH. L.J. 109, 111-12 (2011).

[120] Atlas Powder Co. v. E.I. du Pont de Nemours & Co., 750 F.2d 1569, 1576 (Fed. Cir. 1984) ("Even if some of the claimed combinations were inoperative, the claims are not necessarily invalid."). *See also*, Warner Lambert Co. v. Teva Pharmaceuticals USA, Inc., 2007 WL 4233015, *12 (D.N.J. 2007); Application of Myers, 410 F.2d 420, 426 (CCPA 1969).

[121] However, if many or most embodiments of the claim are inoperative, the claim may be invalid. Graver Tank & Mfg. Co. v. Linde Air Products Co., 336 U.S. 271, 276 (1949).

[122] *E.g.*, Johns Hopkins University v. CellPro, Inc., 152 F.3d 1342, 1361 (Fed. Cir. 1998).

produced from one of those cell lines, the patent is still valid.[123]

2. Incentives and Disincentives for Replicability

One driver of the irreplicability crisis in the scientific literature is that academic scientists do not fully internalize the cost of irreplicable results, incentivizing poor experimental technique.[124] A second cause of irreplicability in science is the inability of peer reviewers to put sufficient time and energy into the process to catch irreplicability.[125] These two factors could look very different in patents. In this section, I discuss the incentives and abilities of patentees and patent examiners to avoid and catch irreplicable experiments.

a. Patentee Incentives

Patentees pay tens to hundreds of thousands of dollars to obtain a patent.[126] If the patented technology does not work, the patentee may not recuperate that cost. Companies filing patents therefore presumably do internalize the cost of irreplicability, suggesting that there is less incentive for patentees to file patents without designing experiments to ensure that the invention works.

In practice, however, patentee incentives may be closer to academic incentives. First, many academic scientists are also patentees.[127] But even industry patentees have incentives for irreplicability. Companies increasingly reward scientists whose work translates into a patent. The reward is for the patent itself, not for the underlying science. A survey by the American Intellectual Property Law Association reported that 60% of biotechnology and pharmaceutical companies pay bonuses for employee inventors when a patent application is filed.[128] This incentivizes the filing of many patents, but not necessarily the filing of high-

[123] *Id.* (finding that patent was valid even though defendant alleged that "no one ever succeeded in making CD34 antibodies using…purified My-10+ cells" because the defendant did not allege that antibodies could not be made using the KG-1/KG-1a cell lines, which were also disclosed in the patent).

[124] Notes 62-67 and accompanying text, *supra*.

[125] *E.g.*, Luke Oakden-Rayner, Andrew Beam, & Lyle Palmer, *Medical Journals Should Embrace Preprints to Address the Reproducibility Crisis*, 2018 INT'L J. EPIDEMIOLOGY 1, 2 (2018); Roger Peng, *The Reproducibility Crisis in Science: A Statistical Counterattack*, 12 SIGNIFICANCE 30, 30 (2015).

[126] Filing a patent in the United States can cost between $10,000 and $20,000. Gene Quinn, *The Cost of Obtaining a Patent in the US*, IP WATCHDOG (2015), http://www.ipwatchdog.com/2015/04/04/the-cost-of-obtaining-a-patent-in-the-us/id=56485; Mark Lemley, *Rational Ignorance at the Patent Office*, 95 NW. U. L. REV. 1, 5 (2001). Filing a patent internationally can cost many hundreds of thousands of dollars, depending on the number of countries in which patent protection is desired. Anthony de Andrade and Venkatesh Viswanath, *Estimating the Cost for Filing, Obtaining and Maintaining Patents Across the Globe*, IP WATCHDOG (2016), https://www.ipwatchdog.com/2016/08/28/cost-filing-obtaining-maintaining-patents/id=72336/.

[127] Bhaven N. Sampat, David C. Mowery, and Arvids A. Ziedonis, *Changes in University Patent Quality after the Bayh-Dole Act: A Re-Examination*, 21 INT'L J. INDUSTRIAL ORG. 1371, 1372 (2003).

[128] Soonhee Jang, *Inventor Compensation in the U.S.,* AIPLA, 9 (2013), http://www.aipla.org/committees/committee_pages/IP-Practice-in-Japan/Committee%20Documents/2013%20Japan%20Delegation/Jang%20Inventor%20compensation%20in

quality patents.

If individual inventors do not have an incentive to file replicable patents, surely the companies paying for patent filing do? Perhaps not. There is an increasing body of literature suggesting that pure numbers of patents are valuable, even if the contents of those patents are not useful.[129] Though this literature relates predominantly to patents in the high-tech fields, the concepts apply to some extent in biomedical patenting as well.[130]

Ultimately, industry patentees may have more incentive to ensure replicability than their academic counterparts. However, the incentive story is complex, and it is an empirical question that I investigate in Part II.

b. Examiner Incentives

It is hard to ask peer reviewers – who are uncompensated and busy – to carefully investigate the likely replicability of a scientific article. By contrast, examiners – though also busy – are paid to carefully review patents.[131] Thus, even if patentees are not incentivized to care about replicability, examiners could create such an incentive by rejecting patents containing irreplicable experiments.

The evidence on whether examiners assess the quality of experiments is mixed. First, it is well established that the applicant does not have to show utility "as a matter of statistical certainty."[132] Further, as a practical matter, the PTO does not need evaluate statistics at all, since examiners reject applications for lack of credible utility only when the invention could

---

%20the%20US.ppt.

[129] Colleen Chien, *From Arms Race to Marketplace: The Complex Patent Ecosystem and Its Implications for the Patent System*, 62 HASTINGS L.J. 297, 321 (2010); Stuart J.H. Graham and Ted Sichelman, *Why Do Start-Ups Patent?*, 23 BERKELEY TECH. L.J. 1064, 1082 (2008); David H. Hsu and Rosemarie H. Ziedonis, *Patents as Quality Signals for Entrepreneurial Ventures*, 2006 ACAD. MANAGEMENT BEST PAPER PROC. 1, 2, http://www.management.wharton.upenn.edu/hsu/inc/doc/2015/11.pdf; Clarisa Long, *Patent Signals*, 69 U. CHI. L. REV. 625, 626 (2002); Gideon Parchomovsky & R. Polk Wagner, *Patent Portfolios*, 154 U. PENN. L. REV. 1, 4 (2005); David L. Schwartz, *On Mass Patent Aggregators*, 114 COLUM. L. REV. SIDEBAR 51, 56 (2014); R. Polk Wagner, *Understanding Patent-Quality Mechanisms*, 157 U. PENN. L. REV. 2136, 2157 (2009).

[130] One way in which large numbers of patents can be useful is in certain monetization strategies used by non-practicing entities. While these are mainly associated with the high-tech industry, they can also appear in the life sciences. *See, e.g.*, Robin Feldman & Nicholson W. Price II, *Patent Trolling Why Bio & Pharmaceuticals Are at Risk*, 17 STAN. TECH. L. REV. 773, 22-23 (2014).

[131] Mark Lemley, Doug Lichtman, and Bhaven Sampat, *What to Do about Bad Patents?*, REGULATION, Winter 2005, at 10, 10 (2005) (explaining that examiners spend an average of 18 hours examining each patent, which may not be enough to catch bad patent applications).

[132] Nelson v. Bowler, 626 F.2d 853, 586-87 (CCPA 1980) ("Bowler argues that the…tests are inconclusive showings of pharmacological activity since confirmation by statistically significant means…occurred the critical date [i.e. too late]. But a rigorous correlation is not necessary where the test for pharmacological activity is reasonably indicative of the desired response.").

not possibly work.[133] The PTO acknowledges that these situations are "rare."[134]

However, there are hints in the doctrine that the PTO values reliability of data (which might include replicability) at least to some extent. First, any data used to show that *in vitro* experiments are likely to have *in vivo* utility must be "statistically relevant"[135] – though neither the PTO nor the courts have provided any additional detail on what exactly that means.[136] Second, the PTO instructs examiners to assess whether the provided data is "reasonably predictive of the asserted utility."[137] Examiners should look at factors that are somewhat similar to those thought to promote replicability, for example, "test parameters, choice of animal…relative significance of the data provided" and others.[138] In addition, though the PTO does not instruct examiners to consider the quality of the experiment or its statistical validity,[139] examiners do so at least on occasion. For example, in Application No. 10/628,102, the examiner rejected the application for lack of enablement and wrote that the while the invention was tested on patients, "[i]t is noted there were no control groups shown."[140]

---

[133] MPEP 2107.01(II). Generally this arises in the context of a machine that is physically impossible, such as a perpetual motion machine. *See, e.g.*, Newman v. Quigg, 877 F.2d 1575, 1582 (Fed. Cir. 1989) (upholding the PTO's decision not to grant a patent on a perpetual motion machine).

[134] MPEP 2107.01(II) (advising that in light of "the rare nature of such cases," examiners should be cautious in making a lack of utility rejection on the basis that the utility is not credible."). Further, even if an examiner wanted to examine experimental quality closely, there is no time to do so – examiners spend an average of just 18 hours examining each patent. Robert P. Merges, *As Many as Six Impossible Patents Before Breakfast: Property Rights for Business Concepts and Patent System Reform*, 14 BERKELEY TECH. L.J. 577, 590 (1999).

[135] MPEP 2107.03 (I).

[136] The term "statistically relevant" is used in only four patent cases (as of June 2018), and none elaborate on the term beyond quoting the MPEP. *See* In re '318 Patent Infringement Litigation, 583 F.3d 1317, fn. 10 (Fed. Cir. 2009); CreAgri, Inc. v. Pinnaclife, Inc., 2013 WL 6673676, *21 (N.D.Cal. 2013); Eli Lilly and Co. v. Actavis Elizabeth LLC, 731 F.Supp.2d 348, fn. 15 (D.N.J. 2010), aff'd in part, rev'd in part, 435 F. App'x 917 (Fed. Cir. 2011); Eli Lilly and Co. v. Actavis Elizabeth LLC, 676 F. Supp. 2d 352, 359 (D.N.J. 2009), aff'd in part, rev'd in part, 435 F. App'x 917 (Fed. Cir. 2011).

[137] MPEP 2107.03 (III) *citing* Ex parte Maas, 9 USPQ2d 1746 (BPAI 1987); Ex parte Balzarini 21 USPQ2d 1892 (BPAI 1991).

[138] *Id*.

[139] Instructions for examiners are in section 2164 of the Manual of Patent Examining Procedure. Statistical validity is not mentioned.

[140] Non-Final Rejection, 6 (May 4, 2004) (application filed July 25, 2003). The patent was later granted as U.S. Patent No. 6,987,093 (issued Jan. 17, 2006). Similarly, the examiner rejected Application No. 12/672,963 for lack of enablement, stating that "patients with progressive disease were the only group with a disease assessment used in the comparison. Therefore, it is unpredictable whether determining and comparing the level of the PSPH gene expression can be used to predict that a patient will have any type of response to [the claimed invention.]" In another example, an examiner rejected Application No. 12/324,198 for lack of enablement because the application claimed a particular type of soybean seed and "[s]ince the seed claimed is essential to the…invention, it must be obtainable by a *repeatable* method set forth in this specification or otherwise be readily available to the public." Emphasis added. The patent was granted after a sample of the seed was deposited with the PTO. U.S. Patent No. 8,035,000 (issued Oct. 11, 2011).

While these examples suggest that examiners care about replicability, they are isolated incidents and not representative of PTO policy. Further, even if examiners are looking out for poorly designed experiments and results that might not be replicable, they may not have the expertise to identify these experiments.[141] As with patentee incentives, examiner incentives for enforcing replicability are mixed, suggesting the need for further study to determine how the PTO assesses experimental quality.

3. Effect of Irreplicability on the Patent System

What is the effect of having irreplicable experiments in patents? One might assume that if an experiment does not work, then the patent will be valueless – since it covers non-functional technology – and therefore there is no harm to the public in giving away the exclusive right, since the right does not cover anything useful. However, there are indeed harms to the public from irreplicable experiments in patents, and I set them out here.

a. Waste and Inefficiency

In the context of the scientific literature, much irreplicability arises from poor experimental design and could be avoided by taking proper precautions.[142] These experiments waste resources. It is expensive to purchase labor and materials to conduct experiments, and this is wasted if those experiments do not produce useful results. As mentioned above, economists estimate that a 50% irreplicability rate in pre-clinical research in the life sciences would cost $28 billion in wasted labor and materials.[143] The authors of this study based this finding on the total amount of money spent annually on life sciences research[144] - which includes both experiments reported in papers and experiments reported in patents.[145] And the $28 billion figure does not take into account the cost of filing a patent. Assuming that a patent covering a technology that does not work creates little or no social value, the cost of drafting and filing that patent – between $10,000 and $20,000 per patent[146]

---

[141] Lisa Larrimore Ouellette, *Pierson, Peer Review, and Patent Law*, 69 VAND. L. REV. 1825, 1828 (2016). *See also* Ronald J. Mann, *The Idiosyncrasy of Patent Examiners: Effects of Experience and Attrition*, 92 TEX. L. REV. 2149, 2163 (2014).

[142] *E.g.*, Iain Chalmers and Paul Glasziou, *Avoidable Waste in the Production and Reporting of Research Evidence*, 374 LANCET 86, 86 (2009) ("85% of basic and clinical research is wasted because of poor design, non-publication, and poor reporting."). *See also*, D.G. Altman, *The Scandal of Poor Medical Research*, 308 BRITISH MED. J. 283, 283 (1994) ("Huge sums of money are spent annually on research that is seriously flawed through the use of inappropriate designs, unrepresentative samples, small samples, incorrect methodology, and faulty interpretation.").

[143] Freedman, *supra* note 6, at 1.

[144] *Id*. $114.8 billion in 2015, of which $56.4 billion goes to pre-clinical research.

[145] As well as experiments reported in neither.

[146] Gene Quinn, *The Cost of Obtaining a Patent in the US*, IP WATCHDOG (April 4, 2015), http://www.ipwatchdog.com/2015/04/04/the-cost-of-obtaining-a-patent-in-the-us/id=56485/; Mark Lemley, *Rational Ignorance at the Patent Office*, 95 NW. U. L. REV. 1, 5 (2001).

– is wasted.[147]

There is also waste in our ability to translate pre-clinical research into treatments that benefit humans. The scientific community is deeply concerned that irreplicable pre-clinical experiments in papers are one reason for increasing failures to translate promising pre-clinical findings in animals into effective treatments in humans.[148] Today, the rate of failure to translate drugs from animals to humans is higher than in the 1970s – perhaps because of irreplicability.[149] Better designed pre-clinical studies and more transparent reporting are a path to improving translation from animal studies to human treatment.[150] This may be true in patents as well. Because patents are commonly obtained before trials in humans are possible, most experiments in patents describe pre-clinical findings.[151] The ultimate goal both for the patent system and for the patentee is to translate those pre-clinical findings into drugs that work in humans – but irreplicable experiments may mean this does not happen.

b. Inoperable Patents

A major consequence of irreplicable experiments in patents is that many patents will disclose technologies that do not work. Patent doctrine terms such patents "inoperable."[152] There is a large literature on inoperable patents.[153] This literature generally assumes that inoperable patents are widespread because patents are filed early in the invention lifecycle –

---

[147] Patents on technologies that do not work likely still create private value for the firm that files them. Patents have value as signals of technological accomplishment (whether or not they actually work) and often sheer volume of patents is a source of value for companies. *See, e.g.*, Clarisa Long, *Patent Signals*, 69 U. CHI. L. REV. 625, 626 (2002); Gideon Parchomovsky, R. Polk Wagner, *Patent Portfolios*, 154 U. PENN. L. REV. 1, 3 (2005).

[148] *See, e.g.*, Susan Bridgwood Green, *Can Animal Data Translate to Innovations Necessary for a New Era of Patient-Centred and Indiviualised Healthcare? Bias in Preclinical Animal Research*, 16 BMC MED. ETHICS 53, 53 (2015).

[149] Jack W. Scannell and Jim Bosley, *When Quality Beats Quantity: Decision Theory, Drug Discovery, and the Reproducibility Crisis*, 11 PLOS ONE e0147215, 2 (2016).

[150] F. Daniel Ramirez, et al., *Methodological Rigor in Preclinical Cardiovascular Studies: Targets to Enhance Reproducibility and Promote Research Translation*, CIRCULATION RESEARCH (April 3, 2017), available at http://circres.ahajournals.org/content/early/2017/03/30/CIRCRESAHA.117.310628; Francis Collins, *NIH Plans to Enhance Reproducibility*, 505 NATURE 612, 613 (2014); HB van der Worp, et al, *Can Animal Models of Diseases Reliably Inform Human Studies?*, 7 PLOS MED. E1000245, 2 (2010); P Perel, et. al., *Comparison of Treatment Effects Between Animal Experiments and Clinical Trials: Systematic Review*, 334 BRITISH MED. J. 197, 199 (2007).

[151] Freilich, *supra* note 90, at 19-21.

[152] *E.g.* Brooktree Corp. v. Advanced Micro Devices, Inc., 977 F. 2d 1555, 1571 (Fed. Cir. 1992).

[153] *See, e.g.*, Mark A. Lemley, *Ready for Patenting*, 96 B.U. L. REV. 1171 (2016) (arguing that the patent system privileges untested ideas over inventions that are physically reduced to practice); Michael Risch, 2010 BYU L. REV. 1195, 1198 (2010) (discussing the dimensions of operability); Daniel C. Rislove, *A Case Study of Inoperable Inventions: Why is the USPTO Patenting Pseudoscience?* 2006 WISC. L. REV. 1275 (arguing that patents that rest on "clearly pseudoscientific principles" should not be granted); Sean Seymore, *Making Patents Useful*, 98 MINN. L. REV. 1046, 1092 (2014) (suggesting that the utility doctrine is not necessary to prevent inoperable patents because they can be excluded by the enablement requirement).

while they are still conceptual and before they have been physically tested.[154] While this is undoubtedly one source of inoperable patents, in this Article I emphasize that even inventions that have been physically tested can be inoperable. However, many of the problems with inoperable patents as described in the literature on early-stage patents also apply to the irreplicable patents discussed here.

First, owners of inoperable patents may still seek rents from other inventors. Suppose that I discover a molecule that I believe cures cancer. I test this molecule in mice, and find that I am correct, so I apply for and obtain a patent covering the molecule. Patents of this type typically give the patentee the exclusive right to the molecule for any use at all, so my patent claim is not restricted to using the molecule to treat cancer, but rather covers all possible applications.[155]

It turns out that my experiments in mice are not replicable, and in fact the molecule does nothing whatsoever to treat cancer. However, another inventor, unaware of my findings, discovers that the compound treats HIV. Her results are replicable, and the compound becomes a blockbuster drug. Although I was wrong, she was right, and she did not rely on any information from my patent,[156] she owes me royalties unless she can prove that my patent is invalid – an expensive and uncertain proposition, since granted patents enjoy a presumption of validity.[157] Thus, patents built on irreplicable experiments can drain funds from inventors who did better.

Inoperable patents may also prevent downstream patenting. As in the scenario above, I have patented my newly discovered molecule. After a period of time, my patent expires and I can no longer seek damages from others who use the molecule.[158] Another inventor, again unaware of my findings, discovers that in fact the molecule can cure cancer, only at much higher doses than I used. This inventor tries to pitch his discovery to pharmaceutical companies only to find that nobody is interested because he cannot obtain a patent on his

---

[154] *See, e.g.*, Christopher A. Cotropia, *The Folly of Early Filing in Patent Law*, 61 HASTINGS L.J. 1, 64 (2014); Lemley, *supra* note 153, at 1198; Ouellette, *supra* note 141, at 1830; Ted Sichelman, *Commercializing Patents*, 62 STAN. L. REV. 341, 343 (2010).

[155] Section I.B.1(b), *supra*. *See also* Amy Kapczynski, Chan Park, and Bhaven Sampat, *Polymorphs and Prodrugs and Salts (Oh My!): An Empirical Analysis of "Secondary" Pharmaceutical Patents*, 7 PLOS ONE e49470, 1-3 (2012).

[156] This is because there is no independent invention defense in U.S. patent law. *See*, Mark A. Lemley, *Should Patent Infringement Require Proof of Copying?*, 105 MICH. L. REV. 1525, 1526 (2007); Samson Vermont, *Independent Invention as a Defense to Patent Infringement*, 105 MICH. L. REV. 475, 476 (2006). More drastically than royalties, I could obtain an injunction and start selling her product myself. This may be harder after *eBay v. Mercexchange, LI*, 547, US 388, 393 (2006), which made injunctions more difficult to obtain, particularly for non-practicing entities.

[157] Impax Labs. Inc. v. Aventis Pharm. Inc., 468 F.3d 1366, 1383 (Fed. Cir. 2006).

[158] Patents generally expire 20 years after filing, however, many patents expire earlier for failure to pay maintenance fees to the Patent Office, and this might be particularly likely to occur in a patent covering a technology that does not work. Janet Freilich, *supra* note 90, at 36.

discovery. Patents are only granted on new and nonobvious inventions, and I had previously disclosed that the molecule could cure cancer.[159] Thus, this inventor cannot get a patent on use of the molecule to cure cancer.[160] In this way, irreplicable results can disincentivize later research.

In addition, inoperable patents add volume to an overloaded patent system. They increase the backlog of patents at the PTO, which may reduce quality of examination.[161] Sheer numbers of patents covering a given technology also makes it very hard to search for patents in that area, creating obstacles to conducting freedom-to-operate searches.[162] Similarly, large numbers of patents in a field may create a "thicket" where innovators working in the space must assess, evaluate, and perhaps license large numbers of patents.[163] This creates high transaction costs that might slow or even block innovation.[164] If an innovator wants to develop a new technology but needs licenses to several dozen patents to do so, the project may not be cost effective.

c. Impeding the Goals of the Patent System

Irreplicable experiments do not achieve the foundational goals of the patent system. Patents with irreplicable experiments give their owners an exclusive right, but the public does not get their part of the bargain in return. These patents disclose technologies that do not work, so there is no useful innovation obtained by society. Further, the patents do not communicate useful information, since the experiments are wrong. Just as the replicability crisis has diminished the public trust in science, so too can disclosure of irreplicable information in patents diminish public trust in patents. In theory, patents are supposed to be a public repository of technical information that scientists can access to obtain the details of cutting edge innovations. In practice, scientists already distrust the information provided in patents and think it low quality,[165] although scientists do read patents.[166] If most experiments

---

[159] There are some complexities here, because a previous disclosure will only anticipate a patent if it is enabled, which my invention may not have been (though it also may have been – irreplicability does not mean nonenablement). MPEP 2121. However, a nonenabled disclosure can still render a later patent obvious. MPEP 2121.01; Symbol Techs. Inc. v. Opticon Inc., 935 F.2d 1569, 1578 (Fed. Cir. 1991). And even if the invention is patentable, it may be sufficiently at risk of later invalidation that companies will not invest. Benjamin N. Roin, *Unpatentable Drugs and the Standards of Patentability*, 87 TEX. L. REV. 503, 536 (2008).

[160] He might be able to get a patent on use of the higher dose, but these patents are not as strong. Roin, *supra* note 159 at 548.

[161] Cotropia, *supra* note 154, at 104-105; Roger Ford, *The Patent Spiral*, 164 U. PENN. L. REV. 827, 841-50 (2016).

[162] Janet Freilich, *Patent Clutter*, 103 IOWA L. REV. 101, 113 (2018).

[163] Carl Shapiro, *Navigating the Patent Thicket: Cross Licenses, Patent Pools and Standard Setting*, *in* INNOVATION POLICY AND THE ECONOMY 119, 121 (2001)

[164] *Id*. at 121. *See also*, Michael Heller & Rebecca Eisenberg, *Can Patents Deter Innovation? The Anticommons in Biomedical Research*, 280 SCIENCE 698, 698 (1998).

[165] Lisa Larrimore Ouellette, *Do Patents Disclose Useful Information?*, 25 HARV. J. L. & TECH. 546, 571 (2012).

[166] Lisa Larrimore Ouellette, *Who Reads Patents?*, 35 NATURE BIOTECHNOLOGY 421, 422 (2017).

in patents are wrong, scientists may stop reading patents altogether.

The public also loses the potential to leverage the patent system to incentivize the disclosure of useful information in the future. If company A is granted a patent on the basis of an irreplicable experiment, but later figures out how to make the experiment work, company A can keep those details secret, since they already have a patent. If company B is the one to discover how to make the invention work, company B is also not incentivized to disclose because they may not be able to get their own patent.

d. Experimentation Stops

Some scientists downplay the replicability crisis, saying that irreplicability is just part of the scientific process.[167] In this view, the scientific process naturally addresses irreplicability by encouraging constant testing and development of previous findings. The problem with irreplicability in the context of patents is that once an inventor obtains a patent on their findings, this sort of iterative experimentation stops. Other scientists cannot test or verify the patented findings because doing so would be patent infringement.

There is defense to infringement – the experimental use exception – that may deal with precisely this scenario. The defense was first applied by Justice Story, who explained that it could not be infringement if the defendant had used the invention "for mere purpose of philosophical experiment, or to ascertain the verity and exactness of the specification."[168] For the next two centuries, the experimental use exception would probably have covered a situation where a scientist attempted to replicate an experiment in a patent in order to determine if the experiment worked.[169] However, in 2002 the Federal Circuit decided *Madey v. Duke University*, which significantly narrowed the experimental use exception.[170] Research for business purposes is excluded from the experimental use exception and *Madey* held that research by university faculty is for a business purpose, since universities are in the business of research.[171] *Madey* was widely perceived as destroying the experimental use

[167] *See, e.g.*, A. David Redish et al., *Reproducibility Failures are Essential to Scientific Inquiry*, 115 PROC. NAT'L. ACAD. SCI. 5042, 5046 (2018) ("Many of the current concerns about reproducibility overlook the dynamic, iterative nature of the process of discovery where discordant results are essential to producing more integrated accounts and (eventually) translation. A failure to reproduce is only the first step in scientific inquiry."); Art Markman, *Why Science is Self-Correcting*, PSYCHOLOGY TODAY (Aug. 10, 2010), https://www.psychologytoday.com/us/blog/ulterior-motives/201008/why-science-is-self-correcting.

[168] Sawin v. Guild, 21 F. Cas 554, 555 (C.C.D. Mass. 1813) (Story, J.).

[169] *See, e.g.*, Rebecca S. Eisenberg, *Patents and the Progress of Science: Exclusive Rights and Experimental Use*, 56 U. CHI. L. REV. 1017, 1075 (1989) (arguing that Justice Story's analysis "suggests a function for the experimental use doctrine in the patent system that is analogous to that of replication of scientific experiments…[to] provide a check against fraud or error in research claims by subject research results to potential replication…").

[170] 307 F.3d 1351, 1354 (Fed. Cir. 2002).

[171] *Id*.

exception.[172]

In a post-*Madey* world, it is not clear whether or not attempting to replicate an experiment in a patent would constitute patent infringement.[173] There have been various proposals to institute an exception that would allow replication. These proposals suggest that scholars are skeptical that the post-*Madey* experimental use exception would currently include replication.[174] This uncertainty can chill scientists' willingness to test experiments in patents for replicability. Patent law therefore cuts off one mechanism to improve replicability. Further, this patent doctrine might be making the irreplicability crisis worse not just for patents but also for scientific articles. Because many scientists file patents on the same invention described in scientific articles, patents preclude not only testing of experiments in patents but also testing of experiments in scientific articles.

* * *

Irreplicable experiments harm the objectives and functioning of the patent system. Further, as a theoretical matter, it is plausible that there are many irreplicable experiments in patents. The topic of irreplicable experiments in patents therefore merits more careful study.

II. MEASURING IRREPLICABILITY IN PATENTS: AN EMPIRICAL STUDY

Replicable studies are vital to the transmission of scientific knowledge. Patents, as documents expected to empower the transmission of scientific knowledge, ought to be replicable. While the theoretical case for why patents should be replicable is strong, the doctrine implementing standards of replicability is not. Thus, it is unclear whether experiments disclosed in patents are replicable. It is important for scientists to know whether experiments disclosed in patents are replicable so that scientists can allocate proper weight to information learned from patents. It is also important for policy makers to know whether experiments disclosed in patents are replicable because, if replicability rates are low, it may be worthwhile to make policy changes to improve replicability or adapt to its lack. Studies of replicability in experiments published in scientific journals abound, and the topic has generated enormous debate. There are no empirical studies on replicability in the context of patents, a gap which this Article seeks to fill.

---

[172] *E.g.*, Michelle Cai, Madey v. Duke University: *Shattering the Myth of Universities' Experimental Use Defense*, 19 BERKELEY TECH. L.J. 174, 190 (2004).

[173] Experiments directed specifically towards generating information for FDA approval fall within a statutory experimental use exception. 35 U.S.C. § 271(e)(1). *See also* Amphastar Pharm. Inc. v. Momenta Pharm., Inc., 809 F.3d 610, 615 (Fed. Cir. 2015), *cert. denied sub nom.* Amphastar Pharm. Inc. v. Momenta Pharm., Inc. (U.S. Oct. 3, 2016).

[174] *See, e.g.*, Christina Weschler, *The Informal Experimental Use Exception: University Research After* Madey v. Duke University, 70 N.Y.U. L. REV. 1536, 1547-48 (2004); Tom Saunders, *Rending Space on the Shoulders of Giants:* Madey *and the Future of the Experimental Use Doctrine*, 113 YALE L.J. 261 (2003).

A. Methodology

1. Testing Replicability

Replicability can be tested directly by attempting to re-do an experiment in a lab. However, this is sufficiently expensive that it is rarely done,[175] and has never been done for large numbers of experiments. Instead, much of the literature on the replicability crisis in science has relied not on experimental replication, but on study of the theoretical bases for irreplicability. In particular, many replication studies look at the methodology used to conduct experiments and assess whether the methodology is sufficiently well designed and reported that there is even a chance that the experiment will be replicable.

Instead of testing replication directly, this method tests whether the *methodological predicates* for replicability are present. As I explain in more detail below, I score the methodological quality of experiments in patents as a proxy for whether the experiments are likely to be replicable.

The intuition behind this proxy is that poor methodology often leads to incorrect – and therefore irreplicable – findings. This has been proven repeatedly in the scientific literature. For instance, a study that does not take measures to reduce sources of bias, such as randomization or blinding, is more likely to be irreplicable.[176] Randomization ensures that characteristics are balanced across the treatment and control groups, making it more likely that a difference seen between the groups is the result of the treatment, rather than a confounding variable.[177] Similarly, if a study does no statistical analysis, any difference between the treatment and control groups may be attributable to some chance characteristic about the sample, rather than a true effect that will be replicable in a different group.[178]

Methodology is not a perfect proxy for replicability. First, it relies on the investigator's report of methodology, so it is possible for a study to be done very well but to omit important methodological details in the reporting so that the study appears to be done badly. Second, it is possible for a study to be poorly conducted, and yet also be correct – bad methodology decreases the chance of being correct, but does not eliminate it.

---

[175] For example, a program called the Reproducibility Initiative is seeking to replicate 50 experiments at a cost of $1.3 million. *Trouble at the Lab*, THE ECONOMIST (Oct. 18, 2013), https://www.economist.com/news/briefing/21588057-scientists-think-science-self-correcting-alarming-degree-it-not-trouble.

[176] *See, e.g.*, Story C. Landis *et al.*, *A Call for Transparent Reporting to Optimize the Predictive Value of Preclinical Research*, 490 NATURE 187, 188 (2012).

[177] *See, e.g.*, Hyuna Yang, et al., *Randomization in Laboratory Procedure is Key to Obtaining Reproducible Microarray Results*, 3 PLOS ONE 1, 9 (2008).

[178] *See, e.g.*, Lemuel A. Moye, *Statistical Reasoning in Medicine*, 127 (2006) ("the smaller the p-value, the greater the strength of evidence that the relationship identified in the sample is not due to chance alone.").

However, the use of methodological quality as a proxy for replicability has been well validated in the scientific literature.[179] For example, after C. Glenn Begley and Lee M. Ellis' attempt to replicate experiments directly found many were not replicable, Begley and Ellis reported that, for experiments that could be replicated "authors had paid close attention to [methodology]…and described the complete data set."[180] On the other hand, a host of methodological detail was missing in the experiments that could not be replicated.[181]

In addition, articles finding that many studies have poorly reported methodology are frequently cited as evidence of irreplicability.[182] Further, multiple studies have found that experiments with poor methodology tend to find larger effect sizes than experiments with good methodology.[183] This suggests that experiments with bad methodology are skewed

---

[179] *See, e.g.*, Kenneth R. Hess, *Statistical Design Considerations in Animal Studies Published Recently in Cancer Research*, 71 CANCER RES. 625, 625 (2011) (reviewing 100 articles to determine whether key methodological predicates were present and finding that while "[g]ood statistical design is one hallmark of meritorious research…clearly, the use of essential statistical design features…has room for improvement."); Carol Kilkenny, Nick Parsons, et. al., *Survey of the Quality of Experimental Design, Statistical Analysis and Reporting of Research Using Animals*, 4 PLOS ONE e7824 1, 9 (2009) (reporting a study of 271 publications and finding that although "[a]ccurate and transparent reporting is…vital to allow the reader to assess…the reliability and importance of the scientific findings…[w]e provide evidence that many peer reviewed, animal research publications fail to report important information regarding experimental and statistical methods."); J. Pildal, A. Hrobjartsson, *et al*., *Impact of Allocation Concealment on Conclusions Drawn From Meta-Analyses of Randomized Trials*, 36 INT'L J. EPIDEMIOLOGY 847, 847 (2007) (finding that "inadequate reporting of randomization…is associated with biased estimates of the treatment effect of the order of 20%."). *See also*, O. Steward, P. Popovich, et. al., *Replication and Reproducibility in Spinal Cord Injury Research*, 233 EXP. NEUROL. 597, 597 (2012); H. B. van de Worp & M. R. Macleod, *Preclinical Studies of Human Diseases: Time to Take Methodological Quality Seriously*, 51 J. MOL. CELL. CARDIOL. 449, 449 (2011); M. R. Macleod, *Evidence for the Efficacy of NXY-059 in Experimental Focal Cerebral Ischaemia is Confounded by Study Quality*, 39 STROKE 929, 932 (2008); V. Bebarta, D. Luyten, & K. Heard, *Emergency Medicine Animal Research: Does Use of Randomization and Blinding Affect the Results?* 10 ACAD. EMERG. MED. 684, 686 (2003); H.M. Vesterinen et al., *Improving the Translational Hit of Experimental Treatments in Multiple Sclerosis*, 16 MULTIPLE SCLEROSIS J. 1044, 1050 (2010).

[180] Begley & Ellis, *supra* note 56, at 532.

[181] *Id*.

[182] *See, e.g.*, Story C. Landis, Susan G. Amara, Khusru Asadullah, et al., *A Call for Transparent Reporting to Optimize the Predictive Value of Preclinical Research*, 490 NATURE 187, 188 (2012) ("Several recent articles, commentaries, and editorials highlight that inadequate experimental reporting can result in such studies being un-interpretable and difficult to reproduce."); David Moher, Iveta Simera, et al., *Helping Editors, Peer Reviewers and Authors Improve the Clarity, Completeness and Transparency of Reporting Health Research*, 6 BMC MED. 13, 13 (2008) (Opening the article with the heading title "The reporting of medical research is not clear and transparent: an unacceptable scandal" based on citations to articles finding poor methodological reporting.); David Baker, Katie Lidster, Ana Sottomayor, and Sandra Amor, *Two Years Later: Journals Are Not Yet Enforcing the ARRIVE Guidelines on Reporting Standards for Pre-Clinical Animal Studies*, 12 PLOS BIO. e1001756, 1 (2013) ("Inadequate reporting of key aspects of experimental design may reduce the impact of studies and could act as a barrier to translation by preventing repetition…").

[183] *Dopamine Agonists in Animal Models of Parkinson's Disease: A Systematic Review and Meta-Analysis*, 17 PARKINSONISM AND RELATED DISORDERS 313, 319 (2011) ("we found reported study quality to be limited, and that reported efficacy fell as reported study quality increased."); Emily Sena et al., *How Can We Improve*

towards seeing an effect where no effect, or a smaller effect, actually exists. Moreover, many organizations advocating for improved replicability begin with efforts to improve the reporting of methodology.[184]

As a proxy for replicability, observing the methodological quality of studies likely underreports irreplicability. A study that has well-reported methodology will not necessary be replicable. Thus, the approach used here creates a floor for replicability: whatever the number of patents found to have insufficient methodological quality for replicability, the true number of irreplicable experiments in patents is probably higher.

2. Checklist for Replicability

To assess whether experiments in patents report sufficient methodological detail to be replicable, I use a checklist from the journal *Nature*.[185] Though there are many checklists available, I selected this list because it is general (rather than focusing on a specific type of experiment) and relatively undemanding as compared to other, more detailed checklists. It is therefore an appropriate choice for establishing a floor on replicability.

Although *Nature* did not generate this checklist specifically to test for replicability, the checklist was designed as a bulwark against irreplicable studies.[186] *Nature* explains that:

> This non-exhaustive [check]list summarizes several elements of methodology that are frequently poorly reported. Inconsistent reporting may lead to incorrect interpretation of results and lack of reproducibility. To improve the transparency and the reproducibility of published results, we ask that authors include in their manuscripts relevant details about these elements of their experimental design. During peer review, authors confirm via the Reporting Checklist for Life Sciences Articles that this information is reported.[187]

For each experiment in my sample, I reviewed the experiment to determine if it contains the information in the Table 1, below. In creating this table, I excluded elements of the

---

the Pre-Clinical Development of Drugs for Stroke? 30 TRENDS IN NEUROSCIENCE 433, 433 (2007) ("we show that study-quality and publication bias have substantial effects on published estimates of drug efficacy in animal studies."); Hann M. Vesterinen, et al., *Improving the Translational Hit of Experimental Treatments in Multiple Sclerosis*, 16 MULTIPLE SCLEROSIS 1044, 1045 (2010) ("We have shown that studies reporting measures to avoid bias (random allocation to group and blinded assessment of outcome, both important indicators of internal validity) give substantially lower estimates of efficacy than studies that do not report such measures.").

[184] Landis, Amara, & Asadullah, *supra* note 182, at 188; Moher & Simera, *supra* note 182, at 13; Baker, Lidster, Sottomayor, & Amor, *supra* note 182, at 1.

[185] Nature Publishing Group, *Reporting Life Sciences Research* (April 2015), https://www.nature.com/authors/policies/reporting.pdf.

[186] *Id*.

[187] *Id*.

Nature checklist that were specific to certain types of experiments.[188]

Table 1: *Nature* Checklist

Information on:	Explanation	Example[189]
Sample size	The number of samples used.[190]	"50 female C57/BL mice…were divided into 5 groups of 10 mice per group."[191]
Randomization	Whether samples were randomly assigned to experimental groups.[192]	"When the tumors reached a volume of approximately 200 mm^3 mice were randomized into groups…"[193]
Blinding	Whether investigators were unaware of sample group allocation. [194]	"The study was done blindly, meaning that the treatment and the preparation of the drugs were conducted by separate individuals."[195]
Replication	Whether an experiment was repeated. [196]	"To confirm that indeed immunization with CD86-transfected tumor cells was associated with increased expression of CD200, we repeated the study…"[197]
Statistical tests	Whether statistical tests are used.[198]	"Significance was calculated with Student's t-test or where appropriate one way and two way ANOVA using a standard software package (Origin) with $p<0.05$."[199]

3. Sample

Many patents contain experimental protocols and data, which are called "examples."[200] I compiled a database of all examples in applications filed and patents granted between 2001

[188] Specifically, I excluded elements involving antibodies, cell lines, human clinical trials, and electrophoresis and gel data.

[189] Examples in this column are derived from my data, not from the Nature checklist.

[190] This was coded as present if there was an exact or estimated number of animals. Sample size was also coded as present if the number of animals per group was stated, even if the number of groups was not stated.

[191] U.S. Patent No. 9,629,898, Example 4 (issued April 25, 2017).

[192] This was coded as present if the experiment specified whether or not there was randomization. As a practical matter, all experiments that discussed randomization did so in the context of stating that there was randomization; no experiments stated that there was not randomization.

[193] U.S. Patent No.8,901,136, Example 139 (issued Dec. 2, 2014).

[194] This was coded as present if the experiment specified whether or not there was blinding. As with randomization, all experiments that discussed blinding did so in the context of stating that there was blinding.

[195] U.S. Patent No. 7,396,860, Example 1 (issued July 8, 2008).

[196] This was coded as present if the experiment specified whether or not there were replicates. As a practical matter, all experiments that discussed replication did so in the context of stating that there was replication.

[197] U.S. Patent No. 7,452,536, Example 8 (issued Nov. 18, 2008).

[198] This included any sort of statistical analysis whatsoever, including reporting measures of variance such as standard deviation, standard error, or confidence intervals.

[199] U.S. Patent No. 8,557,788, Example 6 (issued Oct. 15, 2013).

[200] *E.g.*, MPEP 2164.02.

and 2016. I did so by writing an algorithm that identified the examples section of the patent and then broke the section down into individual examples.[201]

Examples in patents come in two varieties: working and prophetic examples. Working examples describe experiments that have actually been conducted whereas prophetic examples describe experiments that are merely hypothetical and have not actually been carried out.[202] Prophetic examples cannot be written in the past tense – doing so is inequitable conduct and can result in the patent being unenforceable.[203] It is therefore conventional to write prophetic examples in the present or future tense and working examples in the past tense.[204]

For this project, I selected only working examples by limiting my sample to examples written in the past tense. While issues of the veracity of prophetic examples are interesting, the term 'replicability' as it is conventionally used requires that the experiment have already been done once, which is applicable only to working examples.

I further limited the sample to pre-clinical animal studies. I made this choice for several reasons. First, animal studies are the penultimate type of experiment conducted by researchers in the life sciences – they precede only human studies.[205] They are also expensive and must be approved by ethics committees. Because of these features, they are not done casually, but are instead the result of careful planning and extensive preparation. Second, the replicability crisis in the scientific literature is thought to be most acute in pre-clinical animal studies.[206] Studying animal experiments in patents makes it easier to compare the results to the scientific literature.

To compile a list of animal studies, I selected working examples that contained at least one of the following words: mouse, mice, rat, rats, hamster, hamsters, guinea pig, guinea pigs, rabbit, rabbits, cat, cats, dog, dogs. These words are derived from a National Research Council and Institute of Medicine study reporting the most commonly used laboratory animals.[207] The resulting list was overinclusive because it included studies on, for example,

---

[201] Full details about the database can be found at Freilich, *supra* note 90, at 26-28. Patent data was obtained from the USPTO's Grant Full Text Database, hosted by Reed Tech. Reed Tech, *USPTO Data Sets; Patent Grant Red Book* (2017), available at http://patents.reedtech.com/pgrbft.php.

[202] Freilich, *supra* note 90, at 8-10. MPEP § 608(p).

[203] Hoffmann-La Roche, Inc. v. Promega Corp., 323 F.3d 1354, 1364 (Fed. Cir. 2003). *See also* Purdue Pharma. L.P. v. Endo Pharmaceuticals Inc., 438 F.3d 1123, 1134 (Fed. Cir. 2006); Novo Nordisk v. Bio-Tech Gen. Corp., 424 F.3d 1354, 1364, 1368 (Fed. Cir. 2005).

[204] Freilich, *supra* note 90, at 28-29.

[205] Food and Drug Administration, *The Drug Development Process* (2018), https://www.fda.gov/forpatients/approvals/drugs/.

[206] *See, e.g.*, Francis S. Collins & Lawrence A. Tabak, *Policy: NIH Plans to Enhance Reproducibility*, 505 NATURE 612, 612 (2014) ("Preclinical research, especially work that uses animal models, seems to be the area that is currently most susceptible to reproducibility issues.").

[207] NATIONAL RESEARCH COUNCIL AND INSTITUTE OF MEDICINE COMMITTEE ON THE USE OF LABORATORY

mouse cells or mouse antibodies. I therefore manually reviewed studies to include only *in vivo* animal studies. This excluded 57% of the original sample. I associated each remaining study with a random number and reviewed studies beginning with the lowest random number and proceeding upwards. Where possible I was blinded.[208]

To ensure that I selected experiments where the elements of the *Nature* checklist would be methodologically appropriate I included only studies that had the format 'does X treatment affect Y outcome.' This excluded 14% of the sample. I additionally excluded continuations and divisionals. I identified continuations and divisionals by searching for patents that had the same priority date and same original assignee.

I reviewed only one experiment per patent. Where the experiment referenced details in other portions of the patent or was a continuation of a previous experiment, I also reviewed those details. Additionally, if the experiment referenced figures, I reviewed those figures, as well as figure legends.[209] If there was a general methodology section outside of a specific experiment, I also reviewed that. Where the patent referenced a study outside the patent, I did not review it.[210]

The example below, from U.S. Patent No. 9,387,199,[211] illustrates how scoring was conducted. I gave this experiment a score of 2 because it disclosed the sample size and included statistical analysis, but did not mention randomization, blinding, or replicates.

ANIMALS IN BIOMEDICAL AND BEHAVIORAL RESEARCH, USE OF LABORATORY ANIMALS IN BIOMEDICAL AND BEHAVIORAL RESEARCH, Table 1 (1988), available at https://www.ncbi.nlm.nih.gov/books/NBK218261/table/ttt00001/?report=objectonly.

[208] I was blind as to whether a particular experiment was part of a treated group (e.g. an Orange Book listed experiment). I was not blinded for the initial assessment of methodological quality in Section B.1, because it was not possible to blind.

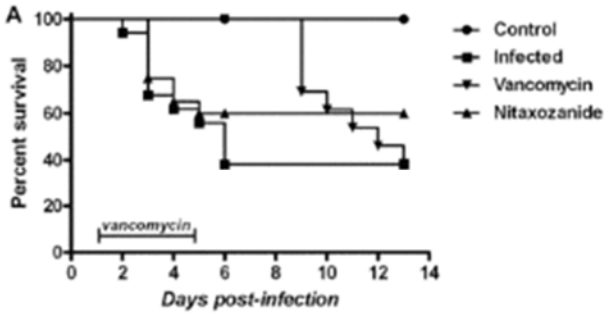
[209] In the 'Brief Description of the Drawings' section of the patent. MPEP 608.01(f).

[210] Patents occasionally cite another study as a source of methodology. I chose not to review those outside studies because while the general design of the experiment was taken from that study, details such as sample size would not necessarily be copied from the study.

[211] U.S. Patent No. 9,387,199 (issued July 12, 2016).

Example 1 Murine Model of *C. difficile* Infection and Treatment

The infection model is a modification of the published protocol of Chen et al. (11). This protocol has been approved by the Center for Comparative Medicine at University of Virginia. C57BL/6 mice, male, 8 weeks old, were used. From 6 to 4 days prior to infection, mice were given an antibiotic cocktail containing vancomycin (0.0045 mg/g), colistin (0.0042 mg/g), gentamicin (0.0035 mg/g), and metronidazole (0.0215 mg/g) in drinking water. One day prior to infection, clindamycin (32 mg/kg of body weight) was injected subcutaneously. The mice were divided into the following groups: control uninfected, control infected, infected and treated with vancomycin (20 mg/kg), and infected and treated with comparator drugs—nitazoxanide, fidaxomicin, and metronidazole (all drugs given at 20 mg/kg/day)…Vancomycin treatment of infected mice was associated with improved mean clinical score versus that for infected controls (1.7±0.3 versus 5.7±0.9; P<0.01), while no difference was seen in nitazoxanide-treated mice (5±1 versus 5.7±0.9) in the first week after infection... Untreated infected mice had an overall survival rate of 38% in this study, and vancomycin prevented 100% of these deaths during the acute infection period (FIG. 1A).



BRIEF DESCRIPTION OF THE DRAWINGSFIG. 1A-B: Effect of vancomycin on *Clostridium difficile*-infected mice during acute infection and posttreatment. C57BL/6 mice were inoculated with VPI 10463 at 10^4 to 10^5 by oral gavage on day 0. The results were data pooled from 94 mice from three 1-week-long and two 2-week-long experiments: 20 uninfected mice, 31 infected mice, 26 mice infected and treated with vancomycin, and 17 mice infected and treated with nitazoxanide. (A) Survival curve. P<0.0001 for uninfected versus infected mice, P=0.0064 for infected mice versus mice infected and treated with vancomycin, and P=NS for infected mice versus mice infected and treated with nitazoxanide by the log rank (Mantel-Cox) test.

Sample size

Statistical Analysis

B. Results

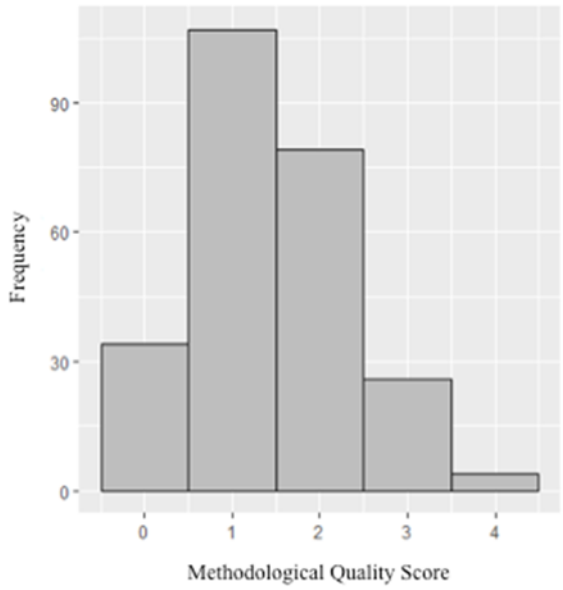
1. Methodological Quality of Experiments in Patents

To obtain an overall measure of methodological quality, I determined how many elements of the *Nature* checklist were disclosed in each experiment, scoring each element of the checklist as a binary (yes/no) variable.[212] I analyzed 250 randomly selected experiments from granted patents. Because there are five elements in the checklist, the maximum possible score is 5, and the minimum possible score is 0. Table 2 shows summary statistics and Figure 1 shows a histogram of scores. The scores ranged from 0 to 4, with no experiments scoring a perfect 5. The median score was 1. 46% of experiments included only one of the elements in the *Nature* checklist.

Table 2: Methodological Scores, Summary Statistics (N=250)

Mean Score	1.4
Median Score	1
Minimum Score	0
Maximum Score	4

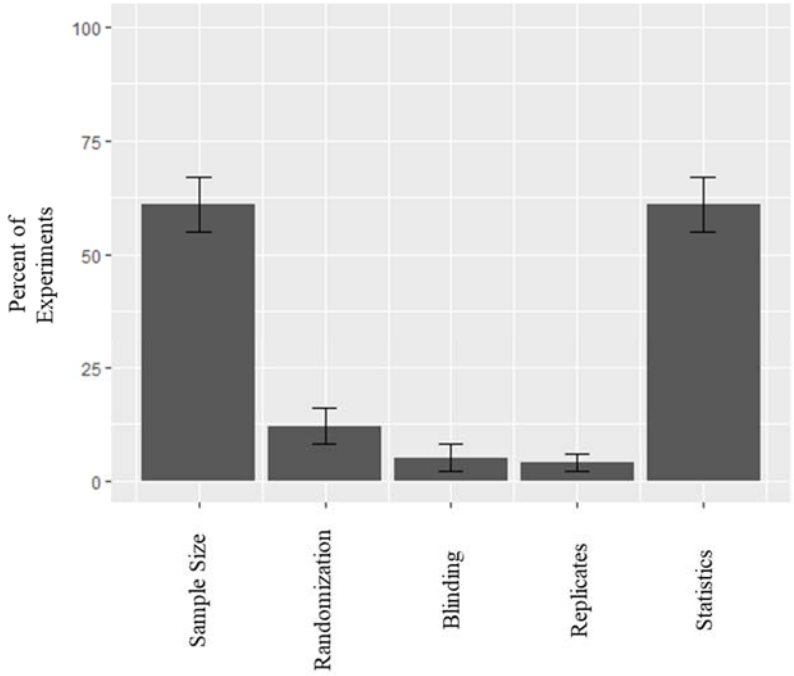
Figure 1: Histogram of Methodological Scores (N=250)



[212] For instance, if sample size was disclosed, the experiment was assigned a '1' for that variable, if no sample size was disclosed, the experiment was assigned a '0'.

Figure 2 breaks the results out by individual element of the checklist. About 60% of experiments disclosed the number of animals used in the experiment and about 60% disclosed some form of statistical analysis. The other checklist elements fared far worse, with 12% of experiments randomizing, 4% blinding, and only 2% disclosing any replicates.

Figure 2: Percent of Experiments in Patents Disclosing Each Element of *Nature* Checklist (N=250)[213]



These numbers do not tell us much in isolation. To understand their significance, Table 3 compares the methodological quality of patents to the methodological quality of scientific articles. For each item on the *Nature* checklist, the item is present less often in patents than in scientific articles. This suggests that replicability in patents is likely also at "crisis" levels. The methodological quality of experiments in scientific articles is associated with unacceptably low rates of replicability – and patents are even worse.

Note that Table 3 presents a range of values for scientific articles, taken from studies of different kinds of biomedical pre-clinical animal experiments. Multiple studies have reviewed methodological quality in scientific articles, and each study uses a somewhat different approach, a different checklist of items, interprets checklist items in

[213] Error bars show 95% confidence intervals.

slightly different ways, studies a different population of journal articles, and is done at a different time. Since no one study done on scientific articles reviewed a population that is directly comparable to patents, I chose to present a range of data from many studies. Additionally, since there are many differences between the types of experiments reviewed in the studies of scientific articles and the study of patents, the numbers should be compared as ballpark estimates, rather than as direct comparisons.

Table 3: Comparing Granted Patents (N=250) and Scientific Articles

	Patents	**Scientific Articles (range)[214]**
Sample Size	62%	70-98%
Randomization	12%	10-22%
Blinding	5%	9-42%
Replicates	4%	--[215]
Statistics	63%	88-100%

2. Validating the Measure

One challenge with the approach used in this article is that it cannot differentiate between failure to conduct an element such as randomization and failure to report the element. It is possible that patents that do not mention randomization do in fact randomize – but the patent attorney drafting the article does not deem it necessary to include the detail. This underreporting would still be a problem because it means that the reader cannot sort high quality studies from low quality studies. However, it would not necessarily imply that the study is irreplicable. While the scientific literature does find some correlation between poor reporting and irreplicable experiments,[216] that correlation may not hold for patents, because the norms and expectations for reporting experiments may be quite different in patents.

[214] David Baker, et al., *Two Years Later: Journals Are Not Yet Enforcing the ARRIVE Guidelines on Reporting Standards for Pre-Clinical Animal Studies*, 12 PLOS BIOLOGY e1001756, 4 (2014); SeungHye Han, *A Checklist is Associated with Increased Quality of Reporting Preclinical Biomedical Research: A Systematic Review*, 12 PLOS ONE e0183591, 7 (2017); Carol Kilkenny, et al., *Survey of the Quality of Experimental Design, Statistical Analysis, and Reporting of Research Using Animals*, 4 PLOS ONE e7824, 4-8 (2009); Kimberley H.J. Ting, et al., *Quality of Reporting of Interventional Animal Studies in Rheumatology: A Systematic Review Using the ARRIVE Guidelines*, 18 INT'L J. RHEUMATIC DISEASES 488, ,493 (2015); Hanna V. Vesterinen, et al., *Systematic Survey of the Design, Statistical Analysis, and Reporting of Studies Published in the 2008 Volume of the Journal of Cerebral Blood Flow & Metabolism*, 31 J. CEREBRAL BLOOD FLOW AND METABOLISM 1064, 1067 (2011).

[215] None of the studies in note 214, *supra*, assessed the number of replicates in the scientific articles studied.

[216] *E.g.*, Jenna Wilson, *Promoting Reproducibility by Emphasizing Reporting: PLOS ONE's approach*, PLOS BLOG (June 14, 2017), http://blogs.plos.org/everyone/2017/06/14/promoting-reproducibility/.

In this section, I seek to show that methodological quality as reported in the patent is correlated with the quality of the experiment itself and not merely attributable to drafting conventions.

a. Association with Lawyers and Clients

If methodological quality as reported in patents is a feature of drafting, rather than of experimental protocol, then particular lawyers should consistently include the same features. By contrast, if methodological quality as reported in patents accurately reflects how the experiment was conducted, then it should vary across lawyers, but the same scientists should consistently include the same features. Essentially, if methodological quality reflects the lawyer's drafting choices, then it should cluster by lawyer, but if methodological quality reflects the scientists' experimental design choices, then it should cluster by client.

Because each lawyer files only a small number of patents, I generated a large sample to test whether methodological quality was associated with lawyers or with clients. To do this, I randomly selected a sample of 7,500 granted patents with animal experiments using the methodology described above and eliminated continuations. 6,529 patents remained. I then associated each remaining patent with the firm that filed the patent and the original assignee using the PatentsView API provided by the PTO and Google Patents.[217]

For each experiment in my sample, I determined if the experiment was randomized. I use randomization because manually determining a methodology score for thousands of experiments is labor intensive, whereas scoring randomization can be semi-automated, making it feasible for large samples.[218] Having classified each experiment as randomized or not randomized, I then used Fisher's Exact Tests to test for an association between firm and randomization and between assignee and randomization. I found a significant association between randomization and *assignee* ($p<0.001$), but not between randomization and *firm* ($p=0.2$). This validates the strategy used to measure replicability in this article because it suggests that the methodology reported in the patent derives primarily from the company conducting the experiment, rather than the lawyer drafting the patent.

---

[217] United States Patent and Trademark Office, *Why Explore Patent Data?*, PATENTSVIEW (2018), http://www.patentsview.org/api/doc.html. Due to limitations of the data, I have data for filing firm, but generally not filing attorney. Thus, I assume that patent drafting style within a firm will be consistent. This assumption is reasonable because attorneys within a firm often work together on patent applications, senior attorneys teach junior attorneys in a firm how to draft patents, and many firms have banks of prior work and templates for attorneys to draw on. I also assume that choice of experimental protocols will be comparable within a given assignee, even though the experiment may have been done by different scientists working for that company.

[218] To determine if an experiment randomized, I selected all patents containing the string ' random', then created a spreadsheet with the text fifty characters before and after the string ' random'. I was able to efficiently review these excerpts to determine if the string ' random' was used in the context of randomizing samples in an animal experiment. I manually reviewed 100 experiments with the string ' random' and my technique categorized the experiment accurately in 96 cases.

b. Patent-Product Link

To further validate the strategy used to measure replicability herein, I ask whether methodological quality is linked to a real-world characteristic: commercialization. Although a patent can fail to result in a commercialized product for many reasons that are entirely unrelated to replicability,[219] if a patent does lead to a commercialized product, it suggests that the technology described in the patent works. This is particularly true in the context of pharmaceutical patents, because commercialized drugs must undergo extensive testing before entering the market.[220] As described below, I find that patents covering commercialized products have better methodological quality scores than matched non-commercialized patents.

In the context of patents covering drug treatments for humans – all patents reviewed for this article – patents that result in commercialized products are listed in the "Orange Book." The Orange Book, officially titled *Approved Drug Products with Therapeutic Equivalence Evaluations*, is maintained by the Food and Drug Administration (FDA).[221] The Orange Book lists patent information for all approved drugs.[222]

I randomly selected 100 animal experiments from Orange Book-listed patents using the same methodology described above.[223] I matched each experiment with a randomly selected experiment from a non-Orange Book-listed patent with a priority date falling in the same year.[224] Experiments from Orange Book-listed patents have considerably better

[219] For instance, a company could run out of funds or a technological advance in a related field could make the patented technology irrelevant. Of particular relevance in the context of pre-clinical animal experiments, an experiment showing that a drug treats a condition in animals could be replicable but still not translate into use in humans.

[220] However, not every drug approved by the FDA works. *See, e.g.*, Sherkow, *supra* note 3, at 846.

[221] Food and Drug Administration, *Orange Book Preface* (Jan. 24, 2018), available at https://www.fda.gov/Drugs/DevelopmentApprovalProcess/ucm079068.htm.

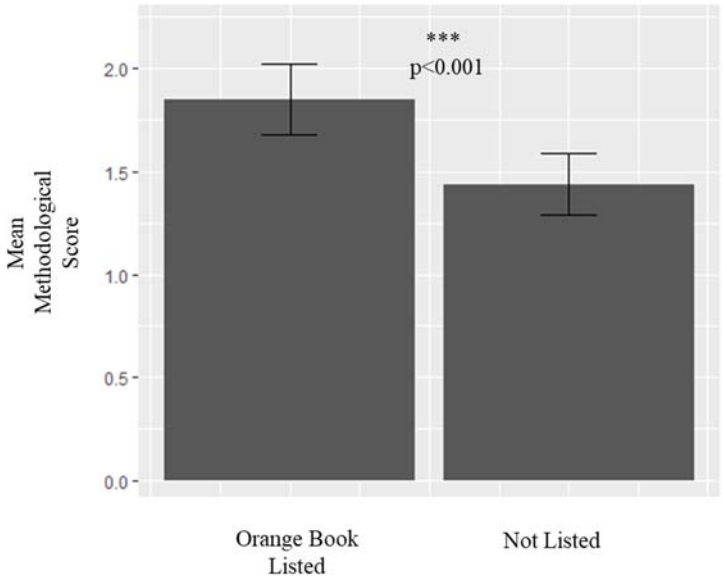
[222] *Id*.

[223] Part II.A, *supra*. Orange Book listed patents were derived from the 2017 and 2018 versions of the Orange Book (on file with the author) as well as archived editions of the Orange Book published between 1985 and 2012. C. Scott Hemphill and Bhaven N. Sampat, *Archival Orange Book Patent Data* (Sept. 29, 2013), available at http://data.nber.org/fda/orange-book/bhaven/documentation_29sep.pdf. Hemphill and Sampat's data file is available at http://data.nber.org/fda/orange-book/bhaven/.

[224] Because commercializing a drug tends to be a lengthy processes, patents listed in the Orange Book are older on average than patents in my general sample. Matching by priority date is important because there is some evidence that methodological quality of scientific publications is improving over time, and this may be true for patents as well. *See, e.g.*, Oscar Florez-Vargas et al., *Bias in the Reporting of Sex and Age in Biomedical Research on Mouse Models*, 5 ELIFE e13615, 4 (2016) (showing trend over time towards more articles reporting the sex and age of mice used in experiments); John Ionnidis et al., *Increasing Value and Reducing Waste in Research Design, Conduct, and Analysis*, 383 LANCET 116, 117 (2014) (showing randomization rates increasing over time). *But see* F. Daniel Ramirez, *Methodological Rigor in Preclinical Cardiovascular Studies: Targets to Enhance Reproducibility and Promote Research Translation*, CIRCULATION RES., 20 (2018), *available at*

methodological scores than experiments from non-Orange Book-listed patents: a mean methodological score of 1.9 as compared 1.4 for matched non-Orange Book-listed patents ($p<0.001$).

Figure 3: Mean Methodological Score for Orange Book and non-Orange Book Listed Patents (N=100)



c. Patent-Paper Pairs

Using methodological quality as a proxy for replicability has been validated in the scientific literature. Therefore, if experiments in patents are written like experiments in scientific articles (at least with respect to methodology) then the proxy should also be effective for patents. Patents and papers are thought to often describe the same experiments.[225] To test the similarity of disclosure of experiments in patents and papers, I matched patents and papers by identity of authors and inventors.[226] I reviewed 100 randomly

http://circres.ahajournals.org/content/circresaha/early/2017/03/30/CIRCRESAHA.117.310628.full.pdf (finding no increase in blinding or randomization rates over time).

[225] Tom Magerman, Bart van Looy, Koenraad Debackere, *Does Involvement in Patenting Jeopardize One's Academic Footprint? An Analysis of Patent-Paper Pairs in Biotechnology*, 44 RES. POL'Y 1702, 1705 (2015); Fiona Murray & Scott Stern, *Do Formal Intellectual Property Rights Hinder the Free Flow of Scientific Knowledge? An Empirical Test of the Anti-Commons Hypothesis*, 63 J. ECON. BEHAVIOR & ORG. 648. 660 (2007).

[226] I obtained inventor names from the PTO's PatentsView API (United States Patent and Trademark Office, *PatentsView* (2018), www.patentsview.org) and then searched PubMed for papers filed by authors with

selected patent-paper pairs. There was no significant difference between the methodological scores for papers and for patents (1.9 vs 2.1, p=0.3).[227] Although patents with paper pairs are not a representative sample of all patents,[228] the similarities in methodological scores between patents and papers are another piece of evidence that the way that experiments are written in the two different media is sufficiently close that the technique that has been validated for use in the scientific literature should also work well in patents.

3. PTO Rejections and Patent Grant

As explained in Section I.B.2, *supra*, whether PTO examiners assess likelihood of replicability is an empirical question. To address this question, I investigated how methodological quality correlated with rejections from the PTO. If the PTO evaluates likelihood of replicability as part of determining patentability, then applications with poor methodological scores should be more likely to be rejected. Specifically, applications with poor methodological scores might be rejected for one or more of the following reasons:

- Lack of utility (35 U.S.C. § 101) because a patent that does not work is not useful.
- Lack of enablement (35 U.S.C. § 112(a)), because an experiment that does not work does not teach others how to make and use the invention.
- Lack of written description (35 U.S.C. § 112(b)), because an experiment that does not work does not prove that the inventor was in possession of the invention.

Such applications might also be less likely to be granted. As described below, there is no correlation between the rejections above and methodological score. There is also no correlation between likelihood of grant and methodological score. This suggests that the PTO does not evaluate patents based on their potential for replicability.

Using the same methodology as described above, I randomly selected 250 animal experiments from patent applications and scored the methodology. The mean aggregate score for patent applications is not significantly different from granted patents (1.38 vs 1.44,

---

the same name as the inventor. I included papers only if they covered roughly similar topics to their patent pair. If multiple papers had authors with the same name as a patent's inventors, I selected the paper filed soonest after the priority date of the patent. For each patent in the pair, I randomly selected an experiment from the patent and manually checked to ensure that it was an animal experiment. I then compared the methodology in each patent-paper pair. For pairs where the randomly selected experiment appeared in both patent and paper, I compared methodology for only that experiment. For pairs that did not have an experiment overlap, I compared methodology across the entire patent and paper.

[227] Paired t-test. The methodological scores are higher here than for other samples studied for this article. The higher scores are likely because methodology was measured on a per-paper or per patent basis, rather than on a per-experiment basis as was done for the remainder of this article. This would occur if, for example, one experiment in a patent randomized but another experiment did not.

[228] For one, they are more likely to be filed by an academic institution.

p=0.5). I obtained data on rejections from the USPTO's Office Action Research Dataset.[229] I obtained data on patent grant from Google Patents.[230] Patent grant was defined as the grant of a U.S. patent that was either directly derived from the application in question or from a continuation or divisional of the application in question. Data on grants was collected in July 2018 and so is current up to that date.

Figure 4 shows the correlation between methodology score and patent grant. Because patent grant often takes many years, the regression includes an offset to control for the number of years since the patent's priority date. There is no significant correlation between methodological quality and likelihood of patent grant.

Figure 4: Correlation Between Methodology Score and Patent Grant (N=250)

	Logistic Regression
Variable	Whether Application Was Granted
Methodology score	-0.01 (p=0.9)
Years since priority date	Yes

Figure 5 shows the correlation between methodology score and likelihood that an application will be rejected for lack of enablement, written description, or utility. Because it takes several years for an application to be processed by the PTO, the first rejection may not occur for several years after the patent is filed, therefore the regression includes an offset to control for years since filing. There is no significant correlation between the methodological quality of an experiment in a patent application and the likelihood that the application will be rejected on the grounds studied.

[229] USPTO, *Office Action Research Dataset* (2017), https://www.uspto.gov/learning-and-resources/electronic-data-products/office-action-research-dataset-patents. For a description of the dataset, see Qiang Lu, Amanda Myers, and Scott Beliveau, *USPTO Patent Prosecution Research Data: Unlocking Office Action Traits*, USPTO Economic Working Paper No. 10 (2017), available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3024621.

[230] Google, *Google Patents* (2018), www.patents.google.com

Figure 5: Correlation Between Methodology Score and PTO Rejections (N=250)
Logistic Regression

Variable	(1) Rejected for lack of enablement	(2) Rejected for lack of written description	(3) Rejected for lack of utility	(4) Rejected for any of lack of enablement, written description, or utility
Methodology score	0.2 (p=0.3)	0.2 (p=0.5)	0.1 (p=0.6)	0.2 (p=0.4)
Years since filing date	Yes	Yes	Yes	Yes

4. Change Over Time

There is evidence that the methodological quality of experiments in scientific papers is improving over time.[231] I tested whether this also held true in patents. It does, but the magnitude of the change is small. Regressing methodology scores on priority year for patent applications shows that the methodology score is improving by 0.03 units per year (p=0.04). While the directionality of the trend is encouraging, progress is slow – at this rate it would take 33 years to improve the mean methodology score in applications by one point.

5. Industry and Academia

The irreplicability debate in the scientific literature pits industry against academia.[232] Academia generates most of the irreplicable articles, while industry must spend millions of dollars verifying results produced by academics.[233] Because patents are filed by both academic institutions and industry, they provide a rare opportunity to compare the experimental design of academic and industry scientists. I manually classified the 250 granted patents in my sample as being filed by either industry or the academy based on the original assignee listed on the patent.[234]

[231] *See* sources cited in note 224, *supra*.

[232] B.R. Jasny, et al., *Fostering Reproducibility in Industry-Academia Research*, 357 SCIENCE 759, 759 (2017) ("many industry researchers distrust quality control in academia and question whether academics value reproducibility as much as rapid publication.").

[233] *Id*. at 760.

[234] Government patents were classified with academic patents. Patents filed by individuals were classified in a separate category.

As shown in Fig. 6, there is no significant difference in methodological quality between academic and industry patents. This suggests that irreplicability is also a problem in industry. These results are interesting because pharmaceutical companies carefully verify the replicability of results before conducting clinical trials, but apparently do not before filing a patent.[235] One explanation for this discrepancy is the relative cost of filing a patent and conducting a clinical trial. Filing a patent costs tens or hundreds of thousands of dollars. Conducting a clinical trial costs tens of millions of dollars (or more).[236] It may be that the cost of patenting is too low to incentivize careful review of data before filing.[237] This suggests that a steep increase in the cost of filing patents might increase the reliability of the data therein.[238] Alternatively, it may be that patents have significant value to the patentee beyond the technical use of the science described in them,[239] or that the pressures to file patents early in the life-cycle of an invention are high enough that companies have no time to ensure replicability.[240]

---

[235] Jasny, *supra* note 232, at 760.

[236] Linda Martin, et al., *How Much Do Clinical Trials Cost?*, 16 NAT. REV. DRUG DISCOVERY 381, 381 (2017).

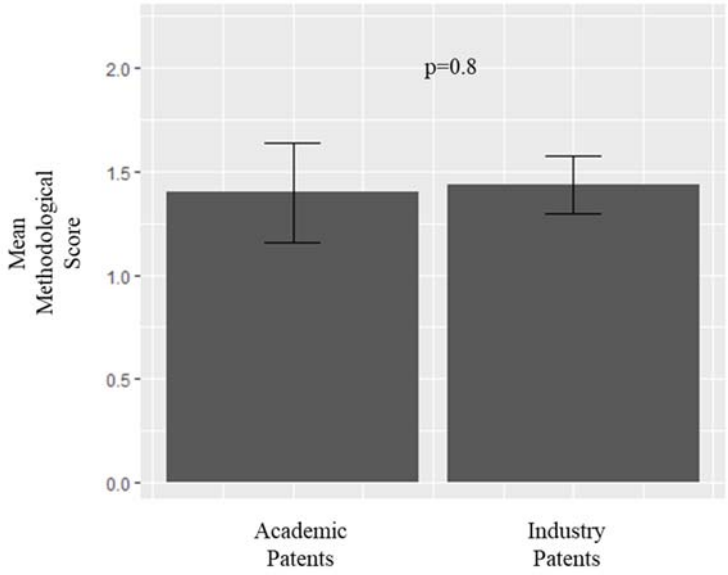
[237] This relates to a widespread debate about the appropriate cost of filing a patent. Many scholars have examined increasing the cost of filing or maintaining a patent as a mechanism to improve patent quality. *See, e.g.*, Brian J. Love, *An Empirical Study of Patent Litigation Timing: Could A Patent Term Reduction Decimate Trolls Without Harming Innovators?* 161 U. PA. L. REV. 1309, 1356 (2013); Jonathan S. Masur, *Costly Screens and Patent Examination*, 2 J. LEGAL ANALYSIS 687, 700 (2010: David S. Olson, *Removing the Troll from the Thicket: The Case for Enhancing Patent Maintenance Fees in Relation to the Size of a Patent Owner's Patent Portfolio*, 68 FLA. L. REV. 519, 520 (2017).

[238] Although, for reasons described in Section III.A, *infra*, I do not recommend this as a policy change.

[239] For instance, as signals or as defensive mechanisms. *See* Section I.B.2, *supra*.

[240] The patent system creates substantial pressure for inventors to file patents as soon as possible. *See, e.g.*, Cotropia, *supra* note **Error! Bookmark not defined.**, at 68; Sichelman, *supra* note **Error! Bookmark not defined.**, at 348-51.

Figure 6: Mean Methodological Score for Industry and Academic Patents (N=250)



C. Mechanism

Why is the methodological quality of experiments in biomedical patents so poor? First, it suggests that patenting is about more than obtaining a patent that works – that there is also value to patentees in obtaining patents that are not functional. This fits with the literature on the value of patents as signals, as defensive mechanisms, and as part of portfolios, where the advantage of the patent lies not in the technology itself, but in the ability of the patentee to claim ownership of a granted patent.[241] If patents provide benefits beyond covering functional technology, then patentees are not incentivized to carefully test technology before filing a patent. As a result, patents will have more irreplicable experiments.

For the same reasons described in Part A, *supra*, with respect to the scientific literature, incentivizing companies to file greater numbers of patents could lead to poor methodology. If the goal is to obtain a finding that looks novel and nonobvious, then there is less incentive to use good methodological techniques such as randomization, blinding, and statistical analysis. Patentees would, of course, prefer to hold patents on working technologies, and I am not suggesting that they are deliberately trying to be wrong. However, implementing better methodology takes time and attention, and will not happen if it is not specifically incentivized.

Second, the experiments in patents are early-stage experiments. The patent system is

[241] Part B.2, *supra*.

strongly oriented towards pushing inventors to file for patent protection as early as possible.[242] Most notably, the patent system recently moved to a "first-to-file" regime, wherein the patent is awarded to the first inventor to file an application with the PTO.[243] When two inventors are both developing similar technologies, the inventor who wins the race to the patent office gets the patent. Since an inventor cannot be sure that no others are working on the same technology, she must file her application quickly.[244]

This means that inventors cannot wait to conduct time-consuming experiments before filing a patent. If the inventor chooses to include an experiment in a patent, that experiment will inevitably be quick and preliminary. Preliminary experiments are, by their nature, less methodologically thorough than subsequent experiments. The purpose of a preliminary experiment is to determine if a technology looks promising – a promise which can be confirmed through more extensive experimentation later. Preliminary experiments might therefore be done with a small number of samples (and perhaps such experiments would not specify the number of samples in order to avoid disclosing a very low number). Further, the inventor would not take the time to conduct replicates. Thus, the patent system's bias towards early-stage experiments could contribute to lower quality experiments disclosed in patents.

III. EFFECTS OF IRREPLICABILITY

Irreplicability rates of experiments in biomedical patents are likely comparable to those in scientific papers – meaning that perhaps up to 90% of these experiments are irreplicable. The irreplicability of experiments, however, is just the tip of the iceberg. Only 45% of biomedical patents have any experimental data at all – the remaining 55% are supported purely by speculative and hypothetical evidence.[245] These speculative patents may be even less likely to be accurate than patents supported by experiments.[246]

Irreplicability creates structural challenges for patent law. Most of the classic theories of patent law rest on the assumption that patents work. If patents are a reward for inventing (reward theory), then we presumably want to reward only inventions that work.[247] If patents are an incentive to create inventions that would not otherwise be developed in the absence of the grant of exclusivity (patent-induced theory), then we should primarily seek to

---

[242] Sichelman, *supra* note **Error! Bookmark not defined.**, at 343.

[243] 35 U.S.C. 102.

[244] David S. Abrams & R. Polk Wagner, *Poisoning the Next Apple? How the America Invents Act Harms Inventors*, 65 STAN. L. REV. 517, 529 (2013) (suggesting that some organizations might, in the absence of a first-to-file system, prefer to wait until the technology is further developed).

[245] Freilich, *supra* note 90, at Table 1 (finding that 523,710 out of 1,160,471 biology and chemistry patents granted between 1976 and 2017 have working examples).

[246] *E.g.*, Cotropia, *supra* note 154, at 123 (suggesting that actual reduction to practice allows the inventor to "gain[] a better handle on whether the invention provides the wanted results.").

[247] A. Samuel Oddi, *Un-Unified Economic Theories of Patents – The Not-Quite-Holy Grail*, 71 NOTRE DAME L. REV. 267, 275 (2014).

incentivize inventions that work.[248] If patents are a prospect through which the patentee can coordinate downstream development (prospect theory), then the patentee must be capable of creating a version that works.[249]

Further, as explored above, the utility, enablement, and written description doctrines all require that the invention works.[250] If not, patents will be granted on useless inventions, given to the wrong inventor, and will not teach others how to make and use the invention. If inventions are irreplicable, then patents on those inventions do not accomplish the basic goals of the patent system: to promote the progress of science through the creation and disclosure of useful, working technology.[251]

Although this Article focuses on irreplicable experiments, these experiments do not necessarily reflect bad science. Certainly, many of the experiments could be better designed, but as explained above, many of these experiments are simply early stage and preliminary. It is the nature of the preliminary experiments to be speculative and often wrong. We would not want to dissuade this – experimenters should be encouraged to try ideas that might not work. To this end, it is good if experiments are done in ways that are cheap and easy – for instance, using a small sample size – even if that reduces the likelihood that the results will be replicable.

The problem is not, therefore, that some experiments in patents are irreplicable. The problem is that the patent system is structured to put unmerited weight on the results of such experiments. We give a powerful legal right – the right to exclude others from making and using the invention[252] – to patentees on the basis of these initial experiments. Though preliminary experimental results are inherently tentative, the patent system uses them as a basis for attaching rights with force and permanence. Experiments are very likely to be wrong; but patent rights are very hard to undo.

There is therefore a fundamental mismatch between how patent theory and doctrine treats patents – as reflecting fully formulated inventions – and what they actually are, which is early stage inventions. In broad terms, there are two potential solutions to this problem. First, we could heighten the evidentiary requirements for patents to a point where most patents would cover inventions that work. Second, we could accept that most patented inventions do not work and adapt the patent system to better reflect that reality. For reasons described below, I advocate for the second option.

---

[248] FREDERIC M. SCHERER, INDUSTRIAL MARKET STRUCTURE AND ECONOMIC PERFORMANCE, ch. 16 (1980).

[249] Edmund W. Kitch, *The Nature and Function of the Patent System*, 20 J.L. & ECON. 265, 269 (1977) ("The patent application need not disclose a device or process of any commercial value, only a version of the invention that will work.").

[250] Section I.B.1, *supra*.

[251] Section I.B.3(c), *supra*.

[252] 35 U.S.C. 271.

A. Patent Law Can't Solve Irreplicability

Improving replicability is appealing in theory, but in practice the cost would be too high to justify. Measures such as randomization are free, but the replicability literature also recommends increased sample size, independent replicates, and testing under different conditions.[253] Requiring such measures would raise the cost of patenting, which might make the system inaccessible to some inventors. It would funnel inventors away from patenting and towards trade secrecy.[254] Although deterring patenting of irreplicable inventions may be no great loss, it is likely that inventors of *replicable* inventions would also be deterred.[255] Further, increasing disclosure requirements for experiments in patents might simply lead inventors to file patents without experiments, which is acceptable to the PTO.[256]

Additionally, the PTO does not have the institutional expertise to require replicability.[257] Few patent examiners have PhDs,[258] and even those that do would not necessarily know how to evaluate whether an experiment was likely to be replicable – particularly since the quantum of evidence necessary to make replicability probable would vary based on the nature of the experiment. Some have proposed peer review of patent applications,[259] and bringing in peer reviewers would increase the level of expertise at the PTO. However, peer reviewers are clearly unable or unwilling to assess replicability in scientific journals, so there is no reason to think that they would function better at the PTO.

Finally, it would be prohibitively expensive for the PTO to evaluate replicability of an invention. Even if the PTO could develop the institutional expertise, sometimes verifying replicability comes down to checking whether an experiment works in the lab – something the PTO does not have the facilities to do.[260] Further, the PTO would have to pay examiners

---

[253] *E.g.*, Prasad Patil, Roger Peng, & Jeffrey Leek, *What Should We Expect When We Replicate? A Statistical View of Replicability in Psychological Science*, 11 PERSPECT. PSYCHOL. SCI. 539, 539 (2016); Benjamin Turner, et al., *Small Sample Sizes Reduce the Replicability of Task-Based fMRI Studies*, 1 COMM. BIO. 1, 2 (2018).

[254] For a discussion of problems surrounding trade secrecy in certain areas of drug development, see W. Nicholson Price II, *Expired Patents, Trade Secrets, and Stymied Competition*, 92 NOTRE DAME L. REV. 1611, 1620 (2017); W. Nicholson Price II, *Making Do in Making Drugs*, 55 B.C. L. REV. 491, 495 (2014).

[255] An inventor deciding whether or not to go to the expense of conducting the experiments necessary to file a patent would not know beforehand whether or not the experiment would be replicable, so replicable experiments might also be lost to trade secrecy.

[256] Freilich, *supra* note 90, at 3.

[257] *See* Jacob S. Sherkow, *And How: Mayo v. Prometheus and the Method of Invention*, 122 YALE L.J. ONLINE 351, 356-57 (2013).

[258] Ouellette, *supra* note 141, at 1828.

[259] *Id*. at 1848. *See also* Beth Noveck, *"Peer to Patent": Collective Intelligence, Open Review, and Patent Reform*, 20 HARV. J. L. & TECH. 123, 143-51 (2006).

[260] This already limits the ability of the PTO to examine applicants' adherence to the enablement requirement. *See, e.g.*, Ex Parte Buchi Reddy Reguri & Sudhakar Sunkari, 2007 WL 2745815 at *7 (Sept. 6, 2007) ("the Office does not have the facilities for examining and comparing the appellants' growth factor with

more to spend more time on each patent in order to thoroughly assess replicability. This cost in increased examination fees may not be worth the benefit in more replicable patents.

In other contexts, scholars have argued that increasing the quality of PTO examination is unlikely to be worth the cost. In an essay addressing the prevalence of bad software patents, Professor Mark Lemley pointed out that since the vast majority of patents are never licensed, enforced, or litigated, many of these PTO errors have little cost.[261] The expense of paying for more examiner hours to reduce the number of erroneously granted patents would likely exceed the cost of those bad patents. Lemley therefore recommends improving mechanisms to deal with these patents *ex post* - in litigation, rather than *ex ante* - in examination.[262]

I take a similar approach here. It would be too expensive for the PTO to ensure replicability *ex ante*, at the examination stage. However, we can create much better mechanisms for dealing with irreplicability *ex post*, after patent grant, in order to mitigate the harms of irreplicable experiments in patents. Below, I propose mechanisms to adapt the patent system to accommodate the realities of irreplicability while staying true to the goal of incentivizing innovation.

B. Adapting Patent Law To An Irreplicable World

In order to adapt patent law to address irreplicability, we must make changes to both theory and policy. Beginning with theory, there are two major ways in which we should revise patent theory to accommodate irreplicability. First, instead of assuming that patents work, we should think of these patents as probabilistic – a roll of the dice.[263] This has been discussed in the context of value to the patentee,[264] but it also applies to value to society: when we grant a patent, there is a significant chance that it will not represent a useful innovation. Nonetheless, the possibility that the patent *will* represent a useful innovation may be big enough that we should keep granting patents.[265]

The irreplicability crisis also necessitates a second shift in our thinking about patents, which relates to what it means for an invention to work. There is a mismatch between patent law's conception of operability and how that concept has been developed in the scientific

---

that disclosed by Walker and by Goldstein. It is therefore entirely proper that appellants should have shouldered their burden of persuasion and made some comparison between the two…"); In re Brandstadter, 484 F.2d 1395 (CCPA 1973) ("Keeping in mind the absence of any facilities in the United States Patent and Trademark Office to test out any device, we are constrained to give full faith and credit to the declarations and the statements made thereby by the declarant…").

[261] Mark A. Lemley, *Rational Ignorance at the Patent Office*, 95 NW. U.L. REV. 1, 1-3 (2001).

[262] *Id*. at 15.

[263] Mark A. Lemley and Carl Shapiro, *Probabilistic Patents*, 19 J. ECON. PERSP. 75, 75 (2005).

[264] *Id*.

[265] Whether the patent system actually accomplishes the goal of increased innovation is an empirical question that is hard to definitively answer.

literature on replicability. Patent law assumes that there is a point in time after which an invention has been "reduced to practice" – meaning that the invention works.[266] In fact, the functionality of the invention is not something that can be ascertained at a set point. Rather, it is a spectrum – as an increasing number of experiments are performed, we can be increasingly sure that an invention works. However, in the context of biomedical inventions – the unpredictable arts – there is no point in time when we are ever absolutely certain that an invention works, no matter how extensive the testing.[267]

These shifts in theory press towards policy change that – instead of assuming that patents work – allows the patent system and third parties to efficiently and inexpensively deal with patents that are not functional. Below, I discuss several potential policy changes to accommodate widespread irreplicability in the patent system. Note that because I have focused here on biomedical patents, I discuss these changes in the context of that industry. However, other industries also struggle with irreplicability,[268] and the policy suggestions are applicable across fields, therefore they may be beneficial beyond the life sciences.

1. Clarify Experimental Use Exception

Since there is a high chance that an experiment in a patent will be irreplicable, we should make it easier for third parties to repeat the experiment in order to test whether or not it works. At present, such an attempt might be patent infringement.[269] The possibility of an infringement lawsuit may deter scientists from trying to verify experiments in patents. There should be a clear experimental use exception – either common law or statutory – for attempted replication. This change would fit comfortably with Justice Story's original vision of the common law experimental use exception, which he believed was necessary "to ascertain the verity and exactness of the specification."[270]

To increase the disclosure value of patents, that exception might be conditioned on the replicator publicly disclosing the results of their verification attempt. Ideally this would be linked the patent – perhaps the PTO could create comment or discussion sections appended electronically to each patent. While public comment sections undoubtedly have their problems, they work to flag replicability problems in the scientific literature. PubPeer, a commonly used commenting system, routinely causes retraction notices to be issued.[271] If

[266] *See, e.g.*, Estee Lauder, Inc. v. L'Oreal S.A., 129 F.3d 588, 593 (Fed. Cir. 1997); King Instrument Corp. v. Otari Corp., 767 F.2d 853, 860 (Fed. Cir. 1985), Wetmore v. Quick, 536, F.2d 937, 942 (CCPA 1976). Note that these cases are in the context of interference proceedings, which do not apply to patents filed after the America Invents Act. However, the concept of reduction to practice is more broadly applicable and remains relevant for many other aspects of current patent law.

[267] *See* Sherkow, *supra* note 22, at 886 (giving several examples of drugs that were approved by the FDA after extensive experimentation, but were still later found not to work.

[268] Ouellette, *supra* note 20, at Supplemental Figure 4.

[269] Section I.3(d), *supra*.

[270] Sawin v. Guild, 21 F. Cas 554, 555 (C.C.D. Mass. 1813) (Story, J.).

[271] Stephen Buranyi, *Anonymous Internet Vigilantes Are Taking Peer Review Into Their Own Hands*,

attempts at replication were linked to patents, it might become clear that certain patents were not replicable. Even if those patents were not formally invalidated, the information could give those who wanted to work in the space covered by the patent's claims some confidence that the patent would not hold up in court.

2. Ease Process of Invalidating Irreplicable Patents

Because granted patents are given a presumption of validity, arguing that a patent is not enabled is a long, expensive, and uncertain process. If most patents contain irreplicable experiments – meaning that they are likely not to be enabled – then the presumption makes little sense. The evidence of irreplicability presented in this Article is a strong argument to remove the presumption of validity, at least when it comes to enablement.[272]

Similarly, it is worth considering faster and cheaper options to invalidate patents that are not enabled or not useful.[273] *Inter partes* review ("IPR") proceedings have significantly brought down the cost of challenging a patent on novelty and nonobviousness grounds.[274] The proceedings could be expanded to utility, enablement, and written description to ease the process of removing irreplicable patents. Further, if patents were easier to invalidate on these grounds, it might incentivize patent applicants to conduct better quality experiments before spending money on a patent application. Expanding IPR proceedings would not be straightforward. The advantages of IPR are their low cost and speed, which rely on the proceeding's limited discovery.[275] The legislative history of the America Invents Act, which created IPRs, demonstrates concern that including enablement in such proceedings would be difficult because of the necessary discovery.[276] Thus, an IPR-like proceeding on utility, enablement, or written description grounds would need to be constrained in order to reduce discovery costs. Despite these challenges, it is important to think about ways to reduce the cost of challenging patents that cover inoperative technology.

MOTHERBOARD (Oct. 3, 2016), https://motherboard.vice.com/en_us/article/pgkxey/anonymous-internet-vigilantes-are-taking-peer-review-into-their-own-hands-pubpeer.

[272] In other contexts, others have suggested rethinking the presumption of validity. *See, e.g.*, John H. Barton, *Reforming the Patent System*, 287 SCI. 1933, 1934 (2000); F. Scott Kieff, *The Case for Preferring Patent-Validity Litigation Over Second-Window Review and Gold-Plated Patents*, 157 U. PA. L. REV. 1937, 1940 (2009); Doug Lichtman & Mark Lemley, *Rethinking Patent Law's Presumption of Validity*, 60 STAN. L. REV. 45, 46 (2007); Seymore, *supra* note **Error! Bookmark not defined.**, at 158.

[273] This is not all patents with irreplicable experiments, since a patent can contain an irreplicable experiment and still be valid (Section I.B.1, *supra*), however, it will include some of the worst offenders.

[274] Ariel D. Multak, *The Big Patent Short: Hedge Fund Challenges to Pharmaceutical Patents and the Need for Financial Regulation*, 23 FORDHAM J. CORP. & FIN. L. 301, 307 (2017)

[275] *E.g.*, Stacy Lewis & Tom Irving, *Very Few Appreciated Just How Bad AIA* Inter Partes *Review (IPRS) Would Be For Patent Owners*, 11 BUFF. INTELL. PROP. L.J. 28, 32 (2015). For another discussion of using PTO proceedings in creative ways, see Dmitry Karshtedt, *Contracting for a Return to the USPTO: Inter Partes Reexamination as the Exclusive Outlet for Licensee Challenges to Patent Validity*, 51 IDEA 309, 309 (2011).

[276] 157 CONG. REC. S1360, S1375 (2011).

3. Disclose Ex Post Data

Irreplicable experiments do not necessarily mean bad science, they just mean that the process of proving that something works is long and difficult. Since patents are inevitably filed before we have clear evidence that an invention works, it does not make sense to halt an inventor's disclosure duties at the time of patent filing. Incorporating ex post data is foreign to the US patent system, but has been suggested in limited circumstances[277] and is sometimes done in other countries.[278] Such a system might require the patentee to update the patent with any data that bear directly on information provided in the patents. This could be done at the same time that maintenance fee payments are made, with the patentee swearing that all proper updates are made. This could create significant additional work for patentees. However, if we are serious about ensuring that patents disclosure a working invention, there are only two possibilities: delay patenting until the invention is sure to work, or include ex post data in the patent. The latter is likely more palatable to patentees.

4. Strategic Ex Ante Improvements

As explained above, I favor strategies that will address irreplicability *ex post*, after patent grant, rather than at the PTO. Although a system where PTO examiners seek to enforce a replicability requirement would be unwieldy, there are certain ways in which the PTO is well positioned to improve replicability. In particular, we should take advantage of structural differences between the PTO and the scientific community. The scientific community's best experts have been struggling with this problem for over a decade – with limited success – but the PTO has two advantages that the scientific community does not: (1) it is centralized and (2) examiners are paid. By contrast, scientific journals operate through a decentralized, norms-based system that relies heavily on volunteer peer reviewers – and consequently journals have found it difficult to enforce guidelines intended to enforce replicability.

The PTO could capitalize on these differences to require more disclosure for experiments. For example, journals believe that increased disclosure would improve replicability, and thousands of journals have tried to implement disclosure checklists to ensure that articles included key methodological details – but they have not yet been successful.[279] Perhaps there is a role for patents. The PTO could adopt a checklist such as the one used for this study and require applicants to disclose the information on the checklist. The PTO would not have to generate its own list but could instead borrow a list already

[277] Jeanne C. Fromer, *Dynamic Patent Disclosure*, 69 VAND. L. REV. 1715, 1722 (2016); W. Nicholson Price II & Arti K. Rai, *Manufacturing Barriers to Biologics Competition and Innovation*, 101 IOWA L. REV. 1023, 1070 (2015).

[278] For example, Australia allows inventors to add working examples to the specification as long as these examples do not encompass matter that was "not in substance disclosed" in the specification as filed. Shann Kerner, Andrej Barbic, & Kyle Robertson, *Examples Requirement for Patentability of Inventions in US and Foreign Jurisdictions*, 3 BLOOMBERG LAW REPORTS, 8, 14 (2009).

[279] *See, e.g*, Baker, et al., *supra* note 83, at 3; Stodden, et al., *supra* note 83, at 1.

created by the scientific community.[280] This change would not require additional work for patent applicants because, having done the experiment, they must know the information already – for instance, what sample size was used or whether the experiment was blinded. The change simply requires writing the details down. At present they may not disclose it because they do not want to (perhaps they are trying to overemphasize the importance of the results, or hide the key details needed to conduct the experiment), because they do not keep good records, or because the attorney does not bother to ask or include the details. None of these are good reasons to avoid disclosure.

Better disclosure of methodological details would not directly lead to improved replicability. An experiment that was poorly done does not become replicable just because readers know it was poorly done. However, it would vastly improve the ability of readers to assess the likelihood of replicability and understand the quality of the experiment. Readers could then discern which experiments appeared promising and worth trying, and which should not be bothered with. This would be a significant improvement over the present state where patent readers must simply guess. Further, it might incentivize use of better methodology, since that methodology would be public.

Disclosure checklists are just one area where the PTO might be better positioned than the scientific community to improve replicability than the scientific community. The patent system and the scientific community have different strengths and should work together to address the replicability crisis that affects them both.

CONCLUSION

* * *

[280] This would mitigate problems caused by the PTO's lack of institutional expertise, although the PTO would still need the expertise to select a proper checklist.