Deceptive Science Maps: Organizational Affiliation and Data Manipulation in Alzheimer's Research*

Matteo Tranchero
The Wharton School

Christian Fons-Rosen
UC Merced

Lee Fleming
UC Berkeley

Abstract

Does corporate authorship increase or decrease deceptive conduct in science? Commercial interests are often viewed as a threat to scientific integrity. Yet when firms build on their science in downstream innovation, fraudulent findings can derail costly development, creating private incentives to produce reliable evidence. We test this logic in Alzheimer's research, using inappropriate image duplication as an observable marker of deceptive conduct. Combining AI-based detection with manual validation, we document a rising incidence of data manipulation. Such cases are nearly absent in corporate-authored papers and less frequent in academic-industry collaborations, but only when firms exercise meaningful control over the production of evidence. The reduction in problematic duplications is largest when firms are likely to build on the results and in settings where deceptive conduct is more prevalent. However, this difference disappears when research concerns drugs already commercialized by the authoring firm. These findings recast the role of commercial interests in science: they can undermine integrity when legitimating existing products, but strengthen it when firms need solid evidence to build on.

*We thank Elizabeth Bik, Justine Boudou, Kevin Boudreau, Colleen Cunningham, Tim Simcoe, Heidi Sofia, Olav Sorenson, John P. Walsh, and participants to the Institutions and Innovation Conference, AOM, SRF, ISA, DRUID, TSE Health Economics Conference, as well as at seminars at Boston University, Harvard, Temple, UC Berkeley, University of Utah, and Wharton for their feedback. We are grateful to Imagetwin for granting us access to their software for research purposes. Any opinions expressed herein are those of the authors and do not represent the views of Imagetwin. Emma Scharfmann and Charlie Guthmann provided excellent research assistance. Matteo Tranchero acknowledges the financial support of the Mack Institute for Innovation Management (2025) and of the Sylff Foundation (2024). Corresponding author: mtranc@wharton.upenn.edu.

1 Introduction

Few scientific claims are treated with more suspicion than those produced by companies that stand to gain from them. This suspicion is rooted in a long history of corporate efforts to bend scientific evidence toward commercial ends (Krimsky, 2019; Legg et al., 2021). Tobacco firms funded research that manufactured doubt about the harms of smoking (Brandt, 2012); fossil-fuel companies contested evidence on climate change (Oreskes and Conway, 2010); and pharmaceutical firms have been shown to shape the evidence base around their products (Lundh et al., 2017; Salandra, 2018). These episodes have coalesced into a "conflict-of-interest" view of corporate science: when firms have commercial stakes in the answer, they have incentives to bias evidence and will produce findings that are more likely to be deceptive.

These arguments offer a compelling account of why corporate findings may be distrusted, but they treat corporate science primarily as a means of influencing external audiences such as regulators or customers. Yet corporate laboratories have long conducted basic research to absorb and extend knowledge relevant to their own innovation (Cockburn and Henderson, 1998; Cohen and Levinthal, 1990; Gambardella, 1992; Rosenberg, 1990). Most importantly, scientific findings can serve as inputs into development (Arora et al., 2021), shaping which technologies firms pursue. This realization calls for a more nuanced view of corporate incentives in science. If a company builds on false evidence, unreliable findings can contaminate internal experiments and divert development down costly paths (Bikard, 2018; Freedman et al., 2015). In this situation, commercial interests create demand not only for favorable evidence, but also for evidence reliable enough to build on.

Motivated by this logic, we develop a framework linking corporate authorship to the risk of deceptive conduct in published research. When firms rely on science to guide their own innovation, they benefit from accurate evidence to inform downstream development. For this reason, corporate involvement could reduce deceptive conduct by giving firms greater control over how evidence is produced and checked before publication. This effect should be strongest when false findings are most costly, such as when research is closer to technological application. However, the same logic also implies a boundary condition. Once research concerns an already developed product, publication becomes less a guide for learning and more a tool for legitimation: favorable evidence can strengthen the product's market position, while unfavorable evidence can threaten sales or

regulatory approval. In sum, corporate authorship should reduce deception when firms need science they can build on, but not when they need science that supports what they have already built.

Empirically testing these ideas is difficult because scientific fraud is rarely observed directly. Most evidence comes from cases exposed through retractions, formal investigations, or replications (Azoulay et al., 2015; Fang et al., 2012; Furman et al., 2012). These cases are highly selected and cannot characterize an entire field (Lacetera and Zirulia, 2011), helping explain why estimates of deceptive research range from fractions of a percent based on retractions to double-digit rates in systematic replication studies (Fanelli, 2009; Bik et al., 2016; Errington et al., 2021). Interpretation is further complicated because some irregularities reflect benign error or reasonable false positives at the knowledge frontier rather than intentional deception (Tranchero, 2026). These limitations make it difficult to assess how corporate involvement shapes scientific integrity.

Here we address these challenges using a transparent, field-wide indicator of deceptive data: inappropriate duplication or manipulation of scientific images (Bik et al., 2016). We focus on Alzheimer's research, where figures provide primary experimental evidence and several high-profile image-related scandals have emerged (Piller, 2025). We assemble the corpus of peer-reviewed, English-language Alzheimer's papers published between 1980 and 2020 and screen their PDFs using Imagetwin, a commercial AI tool that detects within-paper alterations and across-paper plagiarism. Because automated detection also flags legitimate repetitions, such as zoomed panels, two coders independently review each case against the original PDFs. We then classify image problems as potential mistakes or clearly problematic cases, yielding an article-level proxy for deceptive conduct that does not depend on the selected set of caught fraud. This approach provides a systematic basis for examining how organizational context shapes scientific integrity.

Our protocol identifies 818 papers with inappropriate image duplications among 46,548 analyzed, more than sixteen times the number of Alzheimer's papers retracted over the same period. Around 60% of them contain problems indicative of serious concerns with the underlying evidence, while the others could possibly be due to sloppiness or honest mistakes in figure assembly. Validation exercises confirm that the measure captures meaningful data problems: flagged papers are substantially more likely to be discussed on post-publication review platforms and to be retracted.

Duplication rates have risen over time and have recently stabilized at roughly 3% of publications. Papers with duplicate images receive more citations from subsequent scientific articles, clinical studies, and corporate patents, indicating that unreliable evidence can propagate into downstream research and development (R&D).

We find that corporate-authored papers are substantially less likely to contain problematic image duplications, with the reduction concentrated in articles where firms exercised meaningful control over evidence production. Among investigators who move between academia and industry, problematic duplications are also less common during corporate-affiliated periods, suggesting that the results are not driven by scientists' sorting patterns. Consistent with our theory, the negative association between corporate authorship and deceptive conduct is stronger when the science has greater commercial potential, but disappears when research concerns a drug already commercialized by the authoring firm. The reduction is also larger in animal research, where deceptive conduct is more prevalent and the value of corporate control is correspondingly greater. Together, these results show that the association between corporate involvement and deceptive conduct depends on both what firms need science for and how much control they exercise over its production.

This study makes three contributions. First, it contributes to research on organizational wrongdoing (Eaglin, 2026; Greve et al., 2010; Pierce and Snyder, 2008). Specifically, we show that commercial stakes can constrain deceptive conduct when organizations internalize the downstream costs of misconduct (Balasubramanian et al., 2022). Second, the paper has implications for theories of organizational boundaries. Existing explanations for corporate engagement in science emphasize absorptive capacity, hiring, and signaling (Arora et al., 2021; Cohen and Levinthal, 1990; Gambardella, 1992; Sauermann and Roach, 2014). Our findings identify a complementary benefit of internal scientific production: organizational control may improve the reliability of evidence on which downstream innovation depends (Bikard, 2018). Third, the paper contributes a measurement approach for studying scientific misconduct. Because wrongdoing is typically observed only when caught (Wang et al., 2021), prior evidence relies on highly selected cases. Combining AI-based image screening with manual validation, we construct a scalable measure of deceptive conduct and link it to organizational behavior. As the cost of AI tools continues to decrease, this approach opens new avenues for research at the intersection of meta-science, innovation, and firm strategy.

2 Conceptual Framework

2.1 Deceptive Conduct in Science

Science provides a crucial input to corporate inventions (Fleming et al., 2019). Research findings guide firms toward promising ideas and away from dead ends, serving as a map for innovation (Fleming and Sorenson, 2004). Yet there is growing concern about the reliability of this map (Cobey et al., 2024; Freedman et al., 2015; Ioannidis, 2005). Part of the problem inheres in the very process of scientific discovery. Research at the knowledge frontier will naturally generate false negatives and false positives even when scientists act in good faith (Tranchero, 2026). More troubling, however, is the growing evidence that some shaky findings reflect deliberate deception. The rising number of retracted papers, many linked to misconduct, supports this concern (Fang et al., 2012; Van Noorden, 2023). Manual audits similarly document a growing incidence of data fabrication and manipulation (Bik et al., 2016).

Deceptive conduct has consequences because scientific results grow cumulatively and published findings shape downstream research and technologies (Krieger et al., 2024). When deceptive findings enter the scientific record, they can redirect effort toward trajectories built on false evidence, allowing the damage to propagate beyond the original study (Freedman et al., 2015; Michalek et al., 2010). Yet the scale of this issue remains unclear. Most evidence comes from retracted papers — that is, cases of "caught fraud" (Azoulay et al., 2015; Wang et al., 2021). These cases likely differ systematically from undetected problems, limiting what can be inferred about field-wide patterns (Lacetera and Zirulia, 2011). As a result, the prevalence of deceptive conduct in science remains uncertain.[1]

This selection problem limits not only estimates of prevalence, but also explanations of where deceptive conduct originates. Existing accounts focus on visible incidents and the individuals implicated in them (Azoulay et al., 2017; Fang et al., 2012). These cases reveal how scientific

[1]Retraction counts would naively suggest that fraudulent science affects at most 0.2% of publications, but surveys of scientists report rates between 2 and 9% (Fanelli, 2009; John et al., 2012), and manual audits find rates around 3.8% (Bik et al., 2016). In a survey, respondents estimated that 17.1% of other researchers' papers were fraudulent, but only 5% admitted to misconduct in their own work (Bailey et al., 2001). Replication efforts in biomedicine report higher failure rates, between 25 and 89% (Begley and Ellis, 2012; Errington et al., 2021; Prinz et al., 2011), but these estimates capture both honest mistakes and more troubling behavior. Without field-wide baseline measures, it remains difficult to assess how widespread deceptive findings are and how much they distort downstream innovation.

fields respond once fraud is discovered (Furman et al., 2012), but provide limited insight into why deceptive conduct emerges in some settings and not others. The focus on detected cases directs attention toward individual wrongdoing, often by high-profile scientists, even though the widespread nature of the problem points to more systematic causes (Baron et al., 2016). Deceptive science is likely to reflect not only isolated misbehavior, but also the organizational and institutional environments in which research is conducted (Pierce and Snyder, 2008; Walsh et al., 2019).

Individual scientists do not conduct research in a vacuum: their choices reflect how their professional context defines success, allocates rewards, and assigns blame when findings prove unreliable. Despite a shared rhetoric of Mertonian norms, scientific work remains embedded in institutional logics that determine which contributions are valued and which standards of evidence are required (Gittelman and Kogut, 2003; Lacetera, 2009). These logics differ sharply between academia and industry, shaping not only the questions scientists pursue but also how they conduct their research (Bikard, 2018; Boudou, 2025). Explaining variation in deceptive conduct thus requires examining how an organization's goals for scientific findings — and its authority over their production — shape the integrity of published evidence.

2.2 Organizational Goals and Scientific Reliability

Corporate involvement in science is often viewed as a threat to scientific integrity (Krimsky, 2019; Legg et al., 2021). This concern follows from the Mertonian norm of *disinterestedness*, which holds that scientific inquiry should remain uncorrupted by the self-interest of those who produce it (Merton, 1973). Firms, however, often have direct commercial stakes in research conclusions. Favorable findings can strengthen products and competitive positions (Azoulay, 2002; Polidoro and Theeke, 2012; Sood et al., 2014), creating incentives to manipulate evidence or disclose it selectively (Lexchin et al., 2003; Lundh et al., 2017; Salandra, 2018). This logic underlies the conventional conflict-of-interest view of corporate science: compared to academic research, commercial interests can compromise the integrity of research results (Brandt, 2012).

This view treats corporate science largely as a tool for legitimating products to external audiences. Yet firms also conduct research to solve technological problems (Rosenberg, 1990). In this role, scientific findings guide costly development decisions (Evans, 2010; Gambardella, 1992; Arora

et al., 2021). Unreliable evidence can contaminate subsequent R&D and encourage firms to pursue paths that fail only after substantial investment (Freedman et al., 2015). Because firms bear these costs directly, they have a strong private incentive to build upon reliable science (Bikard, 2018). This incentive strengthens when findings are translated into products that must withstand regulatory scrutiny and market use (Balasubramanian et al., 2022). Academic science, by contrast, remains insulated from these downstream demands because its findings are judged primarily by scientific peers rather than by their reliability for technological development (Kuhn, 1970).

This discussion suggests that the relevant question is not whether firms have commercial stakes in science, but rather how they intend to use the research. When science guides technological development, firms have greater incentives to get it right, in order to avoid investing in approaches that will ultimately fail. When science informs product development, direct involvement gives firms greater control over evidence production. Because scientific work is difficult to monitor at arm's length (Azoulay, 2004), conducting research internally allows firms to impose the evidentiary standards required for downstream use. In biomedicine, for instance, companies often replicate academic studies before advancing projects (Freedman et al., 2015), reflecting their reluctance to rely on published findings from academia (Bikard, 2018). Research that guides innovation exposes firms directly to the costs of unreliable evidence, so corporate involvement might reduce deceptive conduct relative to academic research. This logic leads to the following hypothesis:

> ***Hypothesis 1:*** *Corporate authorship is associated with less deceptive conduct in science when research informs the firm's subsequent technological development.*

The preceding logic also implies a boundary condition. Corporate control should improve scientific integrity only when accuracy, rather than a particular conclusion, serves the firm's objectives. When research concerns an existing product, publications serve less as a map for R&D and more as a means of legitimating that product (Azoulay, 2002; Polidoro and Theeke, 2012). At this stage, scientific findings can directly affect commercial outcomes, as unfavorable evidence may threaten sales or regulatory approval (Sood et al., 2014; Salandra, 2018). Firms may then use control over evidence production to shape the published record, consistent with evidence of selective disclosure and sponsor-favorable findings (Lexchin et al., 2003; Lundh et al., 2017). This logic underlies canonical accounts of corporate scientific distortion, including the tobacco industry's efforts to

manufacture doubt about the harms of smoking (Brandt, 2012). Under these conditions, corporate control should reduce deceptive conduct less, leading to our second hypothesis:

> ***Hypothesis 2:*** *The negative association between corporate authorship and deceptive conduct disappears when research concerns a product the firm has already developed.*

Taken together, this discussion illustrates the context-dependent effects of commercial interests on scientific integrity. Our theorizing reconciles conflicting findings: corporate involvement can skew evidence in favor of existing products (Lundh et al., 2017), yet strengthen integrity when firms rely on science to guide development (Guzman et al., 2026; Balasubramanian et al., 2022). The net effect should therefore vary across firms and industries as a function of the relative importance of these two uses of corporate science.

2.3 When is the Value of Corporate Control Greater?

Scientific integrity rests upon organized skepticism (Merton, 1973). Research communities institutionalize this principle through a distributed system of knowledge governance that subjects findings to collective scrutiny (Dasgupta and David, 1994). Before publication, peer review assesses whether claims are supported by the evidence presented (Zuckerman and Merton, 1971); after publication, the system of scientific retractions allows the broader community to identify unreliable findings and redirect subsequent research away from them (Azoulay et al., 2015; Furman et al., 2012). These institutions, however, can only be effective in disciplining scientific conduct when external audiences can effectively evaluate the evidence underlying published claims (Horbach and Halffman, 2019).

Depending on the methods used in the research, such an evaluation may require access to the original data and materials used to generate its findings (Furman and Stern, 2011). But even with such access, tacit laboratory practices can shield critical features of the research process from external scrutiny (Collins, 2001; Owen-Smith, 2001). These limits create an information asymmetry between those who produce evidence and those who evaluate it. As external verifiability declines, manipulating evidence becomes more attractive because it is less likely to be detected and sanctioned (Choi and Valente, 2023; Dyck et al., 2026; Lacetera and Zirulia, 2011). Scientific methods thus differ not only in how they produce evidence, but also in how effectively peer review can expose fraudulent claims before publication (Bik et al., 2016).

Against this backdrop, corporate authorship should become more consequential as the gap between evidence production and its external verifiability widens. When important dimensions of an activity are difficult to specify or monitor at arm's length, doing the research internally gives firms greater control over which activity is performed (Aghion et al., 2008). Applied to research, this control allows firms to observe the procedures behind a result (Azoulay, 2004) and prevent weak evidence from shaping development or entering the scientific record. Firms exercise this oversight through internal review systems; for instance, AstraZeneca requires mandatory sign-off across business units to ensure scientific accuracy (AstraZeneca, 2024), and Novartis mandates reviews to safeguard data integrity (Novartis, 2024). Corporate control thus extends scrutiny to the very process that generated the scientific result.

This reasoning implies that external verifiability shapes the marginal value of corporate control across research settings. When research employs methods that outsiders can readily assess, peer review and other forms of scientific scrutiny provide relatively effective constraints against deception (Horbach and Halffman, 2019), limiting both its prevalence and the additional value of firm oversight. When verification is difficult, by contrast, weaknesses in how evidence is produced are harder for outsiders to detect (Azoulay, 2004), creating greater scope for deceptive conduct. Under these conditions, direct authority over evidence production becomes a more valuable governance arrangement (Pisano, 1990). In this sense, internal organizational control and external verification operate as partial substitutes, leading to the following hypothesis:

> ***Hypothesis 3:*** *The negative association between corporate authorship and deceptive conduct is stronger when the evidence is difficult for external audiences to verify.*

3 Setting: Deceptive Data in Alzheimer's Research

Alzheimer's disease is a neurodegenerative disorder characterized by cognitive and memory decline, followed in later stages by severe functional impairment. In the U.S. alone, nearly 7.4 million people suffer from the disease, a number expected to double within 35 years (Alzheimer's Association Report, 2026). Survival after diagnosis averages four to eight years, although some patients live for as long as two decades. This prolonged course magnifies the public health burden because much of that time is spent with severe disability and dependence. The lifetime cost of caring for a person

with dementia is estimated at $400,000, of which approximately 70% is borne by family members and the remainder by the welfare state (Alzheimer's Association Report, 2026).

Despite decades of work and billions of dollars spent on research, progress in treating Alzheimer's disease has been strikingly limited (Herrup, 2023). A long succession of clinical trials has failed to yield therapies with meaningful clinical benefit (Liu et al., 2019; Kim et al., 2022; Piller, 2025). The eight drugs currently approved by the U.S. Food and Drug Administration (FDA) provide only modest symptomatic relief, with effects that are often imperceptible to patients and caregivers (Seegert, 2025). The most recent approvals, such as lecanemab, have sparked intense debate due to frequent adverse reactions, uncertain long-term efficacy, and high financial costs (Reardon, 2023). The sobering reality is that we still lack a treatment that can reliably prevent or reverse the progression of Alzheimer's disease.

The slow pace of therapeutic progress is troubling in light of growing concerns about scientific fraud. Although the prevalence of deceptive results is difficult to quantify, cases of data malpractice have been documented with disturbing regularity (Begley, 2019; Piller, 2025). A recent example concerns a highly influential 2006 *Nature* study led by Sylvain Lesné, which influenced two decades of subsequent Alzheimer's research (Piller, 2025). The article was retracted in 2024 after concerns were raised that several figures showed signs of manipulation and duplication (Lesné et al., 2024). Investigative reports estimate that tens of thousands of biomedical publications have been shaped by results presenting doctored images like this one (Piller, 2025). The consequences of such issues are both immediate and far-reaching. In the short term, unreliable findings can misdirect research funding and waste resources. Over time, they can undermine public health outcomes and erode trust in the scientific enterprise.

Questions about the soundness of scientific evidence arise across many fields, but they are particularly consequential in Alzheimer's research. Because the causes of the disease remain contested, deceptive data can entrench dominant hypotheses and narrow subsequent inquiry (Hoelzemann et al., 2025). The amyloid-beta hypothesis, for example, has come to dominate the field, accounting for nearly a quarter of clinical trials (Liu et al., 2019) and potentially diverting resources from more fruitful explanations (Begley, 2019). Alzheimer's research thus provides a setting to study how deceptive evidence can distort the direction of an entire field. The coexistence of academic, col-

laborative, and fully corporate research further allows us to examine how organizational incentives shape research integrity.

4 Data and Measurement

4.1 Data Coverage

We assembled the corpus of peer-reviewed, English-language publications on Alzheimer's disease spanning 1980–2020. We stopped our data collection in 2020 to avoid truncation bias when measuring the downstream impact of research. Using NIH's PubTator3 database, which assigns MeSH terms to all PubMed-indexed papers, we retained those tagged with the identifier "D000544 – Alzheimer's Disease". We then merged these records with bibliometric information from iCite (`https://icite.od.nih.gov/`) using PubMed identifier (details in Appendix A). We restricted the sample to English-language articles published in peer-reviewed journals that featured at least 50 Alzheimer's papers during the period. This restriction makes our sample conservative, as it focuses on the field's most relevant outlets. The resulting data includes 50,606 publications.

We then assembled the full-text PDFs for the publications in our sample. We retrieved files from open-access repositories, such as PubMed Central, and online archives, including bioRxiv and medRxiv. The process required thousands of manual downloads, constrained by the coverage and access restrictions of our institutional libraries. In total, we obtained 48,506 PDFs, corresponding to a 96% coverage.[2] The missing papers are similar in publication year to those we recovered (2008.7 versus 2008.3 on average), but appear in journals with lower impact factors (3.93 versus 4.53), which may explain why copies were harder to retrieve through our library subscriptions. Our data spans 262 journals. Of the available PDFs, 17,457 include between one and three figures, 13,522 between four and six, and 7,150 feature seven or more. These files were used exclusively to scan and analyze scientific images, in compliance with their copyright licenses. To our knowledge, this constitutes the largest bibliographic database of full-text articles in Alzheimer's.

Finally, we carried out seven semi-structured interviews with scientists, inventors, and corporate executives. The interviews were not used to test our hypotheses; rather, they served two main

[2]The regression sample is slightly smaller, comprising 46,548 publications. Because the empirical specifications include fixed effects for publication year, journal, and translational stage, singleton observations absorbed by these fixed effects are excluded.

purposes. First, they provided an important reality check on our findings, helping us interpret the patterns we observe and assess their plausibility from the perspective of domain experts. The interviews offered valuable insight into how researchers and executives think about the integrity of scientific discoveries and how it relates to their institutional origins. Second, the interviews helped us identify empirical tests grounded in the field's institutional features.

4.2 Additional Data Sources

We supplemented these data with several sources. Retraction information was drawn from Retraction Watch as of April 2024 (`https://retractionwatch.com/`), which provides records of papers retracted, including the stated reason for each case. We also obtained data from PubPeer (`https://pubpeer.com/`), the leading forum for post-publication peer review. The dataset, graciously shared with us by the blog administrators, lists the DOIs of all articles discussed online as of April 2024. Principal investigator (PI) affiliations and other indicators of research impact are taken from OpenAlex (`https://openalex.org/`). The count of citations with a negative tone received by an article is from the Scite data (Nicholson et al., 2021). Finally, these data were merged with Reliance on Science (Marx and Fuegi, 2022) to obtain citations from patents.

We downloaded from iCite each article's Approximate Potential to Translate (APT), a machine learning measure of the likelihood that its knowledge will be used in later clinical studies (Hutchins et al., 2019). Articles with higher APT scores also receive more clinical and patent citations, consistent with the measure capturing a more applied research orientation (Figure A.1). As a further validation check, we observe that journals emphasizing methods or molecular mechanisms tend to publish articles with lower APT scores, while translational outlets such as the *Journal of the American Medical Association* consistently show much higher values (Panel A of Figure A.2). In the empirical analyses, we include APT as a control to account for the fact that firms tend to produce research closer to application (Panel B of Figure A.2).

We also construct two measures to capture how firms use the science they produce. To test Hypothesis 1, we use the commercial potential measure developed by Masclans et al. (2025), which predicts whether a scientific article contributes to commercially valuable innovation. We classify papers above the median as having high commercial potential, with results robust to

alternative cutoffs. To test Hypothesis 2, we use drug data from Cortellis through July 2020. We identify the chemicals studied in each publication using PubTator3 and link them to their development histories in Cortellis. Combining these records with author affiliations, we determine whether a paper examines a drug that the authoring firm had already commercialized at the time of publication. This measure distinguishes research that may still inform development from research on products already on the market.

Finally, to test Hypothesis 3, we exploit variation in the experimental methods used across publications. Our interviews identified animal studies, or *in vivo* research, as an area where external verification is particularly difficult and firms consequently value greater control over evidence production. As one industry manager explained, companies conducting such studies may want "*full control over study design, execution, and documentation to ensure quality, compliance, and traceability.*" This view is consistent with evidence of limited replicability in animal research (Tsilidis et al., 2013; Errington et al., 2021). Accordingly, we use iCite's MeSH-based classification of methods to identify publications employing animal models, while allowing studies that combine multiple approaches to belong to more than one methodological category (see Appendix A.2).

4.3 AI-Flagged Image Duplications

We focus on a well-defined manifestation of irregular data: the duplication or manipulation of images that report experimental results (Bik et al., 2016; Piller, 2025). Images provide a transparent lens on research integrity because they remain fixed in the published record. In Alzheimer's research, they are central to reporting experimental findings. Our approach offers three empirical advantages. First, it yields a relatively objective indicator that does not rely on subjective judgment or experimental replication, making it scalable to an entire scientific field. Second, because images are pervasive in this domain, we can estimate baseline rates of problematic evidence without relying solely on the selected subset of retracted articles. Third, the heterogeneity of image anomalies allows us to distinguish cases that may be attributable to mistakes from those more consistent with deceptive behavior, as described in the next section.

We partnered with Imagetwin, a company that has developed a leading AI tool to detect irregularities

in scientific images.[3] Imagetwin's software scans all figures in a paper and flags repetitions or alterations that may indicate unreliable evidence. It distinguishes between duplications within the same paper and across different papers, as illustrated in Figure 1. Between March and June 2024, we processed 48,506 PDFs through Imagetwin's API. Because the software flags all instances of visual similarity, including legitimate reuse such as zoomed-in regions, human validation is required. After removing obvious false positives, such as journal logos, we manually reviewed 9,655 papers. Each flagged case was independently assessed by two research assistants, and only instances confirmed by all were retained. Finally, we re-inspected the original PDFs to verify context and to determine whether any duplications were justified by the authors.

We identify 818 Alzheimer's papers containing inappropriate images — 4.4 and 16.7 times the number of papers discussed on PubPeer or retracted, respectively (Figure C.1). Among these, 668 papers contain one duplicated image, 121 contain two, 22 contain three, and 7 contain four or more. We validate this measure in three ways. First, papers flagged by Imagetwin are more likely to be discussed on PubPeer or retracted (Table C.1). Second, using Retraction Watch's classifications of retraction causes, we regress our duplication indicator on indicators for (1) any retraction, (2) retraction due to image problems, and (3) retraction due to image duplication, obtaining monotonically increasing coefficients (Table C.2). Reassuringly, papers retracted specifically for image duplication exhibit the strongest association with our AI-based measure. Finally, duplicated figures are referenced substantially more often in the article text than other figures, suggesting that they tend to report the papers' most important results (Figure C.2 and Table C.3).

Overall, the evidence suggests that our procedure successfully identifies genuine instances of problematic image use while manual checks ensured a low rate of false positives. By integrating automated detection with systematic manual validation, we combine the scale advantages of AI with human examination.[4] This hybrid approach represents an advance over earlier image audit studies, which relied solely on visual inspection (Bik et al., 2016).

[3]Imagetwin's tool has been extensively covered in the popular press; see, for instance, https://www.wsj.com/tech/scientific-papers-imagetwin-proofig-image-scanning. More details about the company are available here: `https://imagetwin.ai/`.

[4]Imagetwin has since introduced an automated filtering feature that removes many irrelevant matches (https://imagetwin.ai/posts/confidence). This update was released after our data collection, but suggests that the precision of AI-based approaches will continue to improve over time.

4.4 Classification of Likely Intent

Imagetwin detects a broad spectrum of image similarities. Some instances involve reusing identical images within the same paper and may reflect honest errors. Others involve transformations such as cropping, color changes, or selective splicing that are harder to reconcile with benign explanations. Although both raise concerns about data quality, the latter provides a more credible indication of deceptive behavior.

Building on prior work that distinguishes among different forms of data irregularities (Bik et al., 2016; Azoulay et al., 2015), we developed a protocol to classify the nature of each duplication. We label cases in which an image is reused to illustrate the same experimental condition as potentially *mistaken*, reflecting errors that may occur during figure assembly. In contrast, when a duplicated panel or portion of it is used to represent different experimental conditions, originates from unrelated papers, or shows alterations, we classify it as *problematic*. Figure 2 provides examples, and the full protocol is reported in Appendix B. Using this approach, we identify 355 papers with plausible mistakes and 495 with at least one problematic duplication suggestive of deliberate manipulation, including some papers with both. The relative proportions of these categories are similar to those reported in other audits, where over half of the irregularities detected reflect intentional deception (Bik et al., 2016; Fang et al., 2012).

We validated our classification protocol in two ways. First, we compared our categories with independent indicators of research integrity, namely, whether a paper was discussed on PubPeer or retracted. Regressions controlling for publication year, journal, and APT show that articles classified as problematic are significantly more likely to be scrutinized on PubPeer and to have been retracted, whereas those labeled as plausible mistakes show no significant association with retractions (Figure C.3). Second, we examined whether reliability concerns recur within author histories. A prior duplication is associated with a 217% higher incidence of deception in subsequent papers, rising to 280% when the earlier duplication was intentionally deceptive (Table C.4). Taken together, these patterns indicate that our coding scheme captures meaningful differences and distinguishes isolated mistakes from recurring (and likely fraudulent) issues.

4.5 Descriptive Statistics

Table 1 reports summary statistics for the 46,548 Alzheimer's publications in the estimation sample. Approximately 1.8% of articles contain at least one manually validated image duplication. About 0.8% contain a duplication classified as potentially mistaken, while 1.1% contain a problematic duplication. Fully corporate-authored papers account for 1% of the sample, while academia-industry collaborations account for about 5.4% of publications, split almost evenly between collaborations led by academic and corporate PIs.[5] The sample also varies in the type and commercial relevance of research: 23% of papers involve animal research, 51% have high commercial potential, and 0.4% study an already commercialized drug. About 3.8% of papers are self-cited in corporate patents. The mean APT score is 0.49, with values ranging from 0.05 to 0.95, indicating substantial heterogeneity in expected translational relevance.

5 Results

5.1 The Growth of Image Duplications in Alzheimer's Research

We begin by examining the overall incidence of image duplication and manipulation in Alzheimer's research. Figure 3 shows a steady increase in publications containing at least one manually validated duplication, from near zero in the 1980s to roughly 3% in the final years of our sample.[6] This estimate remains conservative as a measure of research integrity because image duplication captures only one manifestation of questionable research practices (Fanelli, 2009; Gopalakrishna et al., 2022). Although plausible mistakes have increased in recent years, most detected cases exhibit characteristics consistent with deceptive intent. This pattern parallels the rise in retractions (Van Noorden, 2023) and concerns documented by manual image audits (Bik et al., 2016). Importantly, the average number of figures per publication has increased only gradually, ruling out greater image use as a mechanical explanation.

Articles presenting unreliable data become consequential when their findings diffuse through the

[5] These categories are not exhaustive; papers affiliated with hospitals and other organizations form the omitted category in our regressions. In Section 6, we exploit this institutional variation to further examine how organizational incentives shape research integrity.

[6] A potential concern is left-censoring in Imagetwin's archive, which could make older papers less likely to be flagged for across-paper image reuse. Repeating the analysis using only internal duplications, which are unaffected by archive coverage, yields the same upward trend (Figure C.4).

scientific and technological system. To assess these downstream consequences, we estimate the following article-level OLS specification:

$$Y_i = \alpha + \beta \textit{Duplication}_i + \gamma_{j(i)} + \lambda_{t(i)} + \phi_{\mathrm{APT}(i)} + \varepsilon_i, \tag{1}$$

where Y_i is the outcome of interest for paper i, and $\textit{Duplication}_i$ equals one if the paper contains at least one confirmed image duplication. The terms $\gamma_{j(i)}$, $\lambda_{t(i)}$, and $\phi_{\mathrm{APT}(i)}$ denote journal, year, and APT fixed effects, respectively. Standard errors are clustered by PI. Figure C.5 summarizes the results. Papers containing duplicated images receive more citations from subsequent publications, clinical studies, and corporate patents. These results are unchanged after excluding citations that Scite classifies as negative in tone, suggesting that the higher citation counts do not simply reflect downstream criticism. This pattern echoes evidence that retracted articles accumulate "excess" citations before removal from the literature (Furman et al., 2012), confirming that deceptive research can contaminate subsequent development.

Our measure captures only one manifestation of a broader problem. Freedman et al. (2015) estimate that roughly US$28 billion per year is spent on non-reproducible preclinical research in the United States alone. Although there is no direct estimate of the downstream innovation costs induced, we provide suggestive correlational evidence from our data. We construct an indicator equal to one if an article cites at least one paper containing a duplicated image and estimate OLS regressions with the same controls as in equation 1. Appendix Figure C.6 shows that articles building on such work are more likely to receive negative citations, be discussed on PubPeer, and be retracted. Although descriptive, these patterns suggest that unreliable evidence can create a shaky foundation for follow-on research and propagate problems into later stages of R&D, increasing the risk of delays and failed projects.

5.2 Firms' Authorship and Scientific Integrity

We next test whether deceptive conduct in published science varies with corporate authorship as predicted by our hypotheses. We estimate the following article-level specification:

$$\textit{Duplication}_i = \alpha + \beta \textit{Corporate Authorship}_i + \gamma_{j(i)} + \lambda_{t(i)} + \phi_{\mathrm{APT}(i)} + \varepsilon_i, \tag{2}$$

where *Duplication*$_i$ equals one if paper i contains at least one manually validated image duplication. We estimate separate models for potentially mistaken and problematic duplications, with the latter more directly capturing deceptive conduct rather than errors in figure assembly. *Corporate Authorship*$_i$ indicates whether the paper is authored by researchers affiliated with firms. Like in equation 1, $\gamma_{j(i)}$, $\lambda_{t(i)}$, and $\phi_{\mathrm{APT}(i)}$ denote journal, year, and APT fixed effects.

We begin by exploring the association between organizational affiliation and deceptive conduct. Table 2 reports the results. Academic publications are more likely to contain duplications overall, but this difference is driven largely by apparent errors. Corporate publications exhibit the opposite pattern: they are 87% less likely to contain any duplication, with problematic duplications nearly absent. Corporate authorship is thus associated with substantially less deceptive conduct, a result that is stable over time (Figure C.7). At face value, this finding runs counter to the conventional conflict-of-interest view, which associates corporate involvement with greater risks of scientific distortion. We next examine whether these differences arise under the conditions predicted by Hypotheses 1 and 2.

Our framework predicts that corporate involvement should reduce deceptive conduct more as science moves closer to technological application, where unreliable findings become costly to firms' innovation efforts. Figure 4 tests this prediction using the commercial potential measure developed by Masclans et al. (2025). Corporate-authored papers exhibit fewer duplications at both low and high levels of commercial potential. However, the reduction in deceptive cases is significantly larger among articles more likely to contribute to commercially valuable applications. This result supports Hypothesis 1 and suggests that firms place greater value on scientific integrity when they expect to rely on the research in downstream innovation.

Hypothesis 2 identifies an important boundary condition to this result. Once research no longer guides future development but instead concerns an already commercialized product, firms' incentives shift toward generating evidence that supports an existing asset. To test this prediction, we identify papers studying drugs already commercialized by the authoring firm using PubTator3 and Cortellis. Figure 5 reports the results. The reduction in problematic duplications observed in corporate-authored research largely disappears once the paper studies a drug already on the market. These findings are consistent with a shift in the role of science: once a product is on the market,

scientific publications serve mostly as a tool for legitimizing it to external audiences (Azoulay, 2002; Polidoro and Theeke, 2012).

A final implication of our framework is that firms should rely more heavily on their own science when it serves as an input into development than when it concerns existing products. Table 3 examines this expectation using patent self-citations to the firm's own publications. Corporate papers are nearly twice as likely to be cited in the firm's subsequent patents as similar articles, consistent with evidence that firms conduct in-house science to support their inventions (Arora et al., 2021). This reliance is significantly stronger for research with high commercial potential, while the coefficient for papers concerning already commercialized products is negative but imprecisely estimated. Thus, the conditions under which corporate research exhibits fewer deceptive duplications are also those in which firms appear to make greater use of their own science, linking the integrity advantage of corporate research to its role in innovation.

5.3 Exploration of the Mechanisms

Our theorizing attributes the lower incidence of deceptive conduct in corporate science to firms' control over evidence production, which allows them to impose the evidentiary standards required for downstream applications. This mechanism is difficult to observe in bibliometric data. Academic–industry collaborations, however, offer a useful comparison because both sectors participate in the same project, and the PI's affiliation indicates who leads it. If control is the relevant mechanism, reductions in deceptive conduct should emerge primarily under corporate leadership. Consistent with this prediction, Table 4 shows that papers led by a corporate PI exhibit a 54% lower incidence of problematic duplications, whereas those led by an academic PI do not. Although descriptive, this comparison separates corporate participation from corporate control and suggests that the reduction in deceptive conduct is concentrated where firms exercise meaningful authority over how the research is conducted.

Additional tests in the Appendix sharpen this interpretation. First, neither corporate funding nor the presence of a corporate coauthor is associated with fewer problematic duplications; the reduction appears only in the presence of deeper corporate involvement (Table C.5). Second, the estimates remain remarkably stable as we progressively add demanding controls, making observable

differences in the types of research firms conduct an unlikely explanation for our results (Figure C.8). Finally, we find similar reductions in deceptive data among small biotechnology firms and large pharmaceutical companies (Table C.6). This last result suggests that the pattern is not confined to corporations with greater resources or more elaborate compliance systems.

A remaining concern is that these patterns might reflect the selection of scientists into industry rather than the organizational environment in which they work (Agarwal and Ohyama, 2013; Sauermann and Roach, 2014). We examine this possibility by comparing publications by the same PI before and after moving between academia and industry. Figure 6 shows that PIs exhibit fewer problematic duplications while affiliated with firms, a pattern that persists after controlling for publication year, journal, translational potential, and PI fixed effects (Table C.7). Although such moves are endogenous and relatively infrequent, which limits the precision of our estimates, the within-PI comparison makes stable differences among scientists an incomplete explanation. Instead, the results are more consistent with organizational context shaping how research is conducted.

5.4 Does Corporate Control Substitute for External Verification?

Our final hypothesis predicts that corporate control should matter more when external evaluators, such as peer reviewers, have a limited ability to verify how evidence was produced. We test this prediction by exploiting differences among experimental methods, with a focus on animal research. Our interviews consistently characterized *in vivo* studies as more prone to deception. As one industry scientist explained, "*Animal models are inherently more complex and variable than in vitro systems. This complexity makes exact replication harder and may increase the temptation to reuse or select images that visually support the desired outcome.*" In our data, 3.3% of animal-based publications exhibit problematic images, compared with 0.4% of non-animal publications. Animal-based publications are also more likely to be discussed on PubPeer or retracted (Figure C.9), in line with broader evidence of limited replicability in animal research (Errington et al., 2021).

We next examine whether the association between corporate authorship and scientific integrity is stronger in precisely these studies. Table 5 reports OLS regressions interacting corporate authorship with the use of animal models. The reduction in problematic duplications is substantially larger

in animal research than in other publications, consistent with Hypothesis 3. By contrast, the corresponding interaction for plausible mistakes is small and statistically insignificant, suggesting that firms are not simply making fewer errors in technically complex work. Instead, the result is consistent with external verification and corporate authority acting as partial substitutes: when outsiders have less visibility into how evidence is produced, who controls the research process becomes more consequential for scientific integrity.

Our interviews also pointed to human clinical research as the converse case, with an academic researcher noting that "*If I'm running a basic science lab... it's just me doing my thing, reading my data, and it's there for me to put into an image. Whereas if I'm doing a clinical trial, I've got lots of people inputting data into a database that was locked and scrutinized.*" As suggested by this account, the negative association between corporate authorship and image duplication disappears when we restrict the sample to publications reporting clinical study results (Figure C.10). Taken together, our results suggest that firms gain most from conducting basic research in-house where external verification offers limited assurance that published findings are reliable enough to support downstream innovation.

6 Discussion: Scientific Integrity as an Organizational Outcome

Most visible cases of scientific misconduct invite a simple inference that unreliable science reflects unreliable scientists. Yet organizational research has long cautioned against treating wrongdoing as a property of individuals alone (Frake and Harmon, 2024; Greve et al., 2010; Pierce and Snyder, 2008). Like workers in companies (Pierce et al., 2015), scientists operate within systems that define valued outcomes and allocate rewards tied to scientific production (Baron et al., 2016; Walsh et al., 2019). Our findings reinforce this organizational view: the same PIs exhibit fewer problematic duplications while affiliated with firms, and academic-industry collaborations exhibit a similar pattern only when a corporate PI leads the research. These results point to organizational environments as sources of variation in research integrity (Baron et al., 2016; Walsh et al., 2019).

From this perspective, commercial interests are not inherently corrosive to scientific integrity. The conflict-of-interest view emphasizes firms' incentives to produce favorable findings (Krimsky,

2019), but commercial stakes can also create a private demand for reliable evidence (Freedman et al., 2015; Bikard, 2018). Indeed, recent studies find no greater manipulation in industry-sponsored clinical trials and, in some settings, more evidence of p-hacking among academic or nonprofit sponsors (Balasubramanian et al., 2022; Guzman et al., 2026). Our findings help reconcile these patterns by showing that the effect of corporate involvement depends on the role science plays for the firm. Firms are neither inherently better nor worse producers of science; their incentives for integrity depend on what the research is meant to accomplish.

Academic science, by contrast, operates under a markedly different incentive structure. In corporate research, publications play a relatively small role in career advancement (Liu and Stuart, 2014; Sauermann and Roach, 2014), whereas academic careers reward priority (Merton, 1973). Our interviews repeatedly linked this difference to research reliability. One industry scientist described publication as the "currency" of academic careers, where "*the incentive is to do more, publish more, be cited more.*" In the Appendix, we compare academics whose careers differ in how heavily they depend on publication. Papers led by PIs in medical schools, whose careers often involve clinical responsibilities, exhibit fewer mistaken duplications than those led by PIs in nonmedical units within the same university (Table C.8). At the same time, duplications are not systematically associated with funding availability, supervisory burden, or seniority (Table C.9). These descriptive patterns suggest that research reliability is shaped by the career incentives embedded in scientists' organizational environments.

Nonetheless, these differences should not be read as implying a simplistic ranking of corporate over academic science. A central strength of academic science is openness: publication places findings in the public domain, allowing others to scrutinize and build on them (Merton, 1973; Dasgupta and David, 1994). Corporate researchers face different incentives because disclosure can also reveal valuable knowledge to competitors (Arora et al., 2021). Firms thus publish selectively, and prior work documents both selective reporting and strategic withholding in pharmaceutical research (Kao, 2025; Salandra, 2018). The lower incidence of deception in corporate-authored papers may, therefore, partly reflect greater selectivity in what firms choose to make public. In this sense, scientific integrity is an organizational outcome shaped both by how research is conducted and by what results organizations choose to disclose.

7 Conclusion

This paper investigates the organizational origins of deceptive scientific evidence. Combining AI-based screening with manual validation, we estimate the incidence of data problems in Alzheimer's research. We document a growing number of studies presenting problematic duplications and show that these papers substantially shape downstream applications. We then examine whether deceptive conduct varies across organizational contexts. Corporate-authored papers contain less deceptive evidence when firms have more to gain from reliable science, particularly when research has greater commercial potential. This difference is larger when external audiences have limited ability to verify how the evidence was produced, consistent with corporate control substituting for external verification. Yet the corporate advantage disappears once research concerns an already commercialized product. Commercial incentives, in short, can strengthen scientific integrity when firms need evidence they can build on, but not when science serves other strategic purposes.

Taken together, our findings point to a limit in the conventional division of innovative labor between universities and firms. A long tradition in the economics of innovation showed that universities are well-suited to open-ended discovery, while firms specialize in development and commercialization (Aghion et al., 2008; Lacetera, 2009). Yet, separating the "R" from the "D" works best when published research provides a reliable map for development. When academic incentives reward publication over the careful validation of findings, firms face uncertainty about the reliability of science (Bikard, 2018; Freedman et al., 2015). Our results suggest that this problem becomes acute when science is commercially relevant or hard to verify externally. In these settings, firms rely more on their own scientific publications in patents, as corporate authorship is associated with lower rates of problematic data.

Several limitations of our study point to opportunities for future research. First, our analyses do not causally identify the mechanisms linking organizational incentives to research integrity. The within-researcher evidence moves us closer to separating organizational context from stable differences among scientists, but it does not isolate why the same investigator produces more reliable evidence in one organizational context than in another. Future work could examine the specific practices, incentives, and monitoring systems through which organizations shape evidence production. Second, our measure captures only one observable manifestation of deception. Image

duplication and manipulation provide a revealing but necessarily partial window into research integrity, and other forms of selective reporting or data distortion may follow different organizational patterns.

Finally, our analysis focuses on Alzheimer's disease, a setting characterized by unsettled science and substantial commercial interest. These features may shape both the incentives for deceptive conduct and the value firms place on reliable evidence, potentially strengthening the organizational differences we document. Extending our analysis to scientific domains that vary in their underlying uncertainty, commercial relevance, and external verifiability would help establish the boundary conditions of our findings. The declining cost of AI-assisted screening makes such comparisons increasingly feasible and opens the door to a broader comparative study of scientific integrity, linking variation in deceptive conduct to the contexts in which knowledge is produced.

References

Agarwal, R. and A. Ohyama (2013): "Industry or academia, basic or applied? Career choices and earnings trajectories of scientists," *Management Science*, 59, 950–970.

Aghion, P., M. Dewatripont, and J. C. Stein (2008): "Academic freedom, private-sector focus, and the process of innovation," *The RAND Journal of Economics*, 39, 617–635.

Alzheimer's Association Report (2026): "2026 Alzheimer's disease facts and figures," *Alzheimer's & Dementia*, 22, e71345.

Arora, A., S. Belenzon, and L. Sheer (2021): "Knowledge spillovers and corporate investment in scientific research," *American Economic Review*, 111, 871–898.

AstraZeneca (2024): "Publications - Global Standard," `https://www.astrazeneca.com/sustainability/resources.html`, accessed: 2025-09-25.

Azoulay, P. (2002): "Do pharmaceutical sales respond to scientific evidence?" *Journal of Economics & Management Strategy*, 11, 551–594.

——— (2004): "Capturing knowledge within and across firm boundaries: evidence from clinical development," *American Economic Review*, 94, 1591–1612.

Azoulay, P., A. Bonatti, and J. L. Krieger (2017): "The career effects of scandal: Evidence from scientific retractions," *Research Policy*, 46, 1552–1569.

Azoulay, P., J. L. Furman, J. L. Krieger, and F. Murray (2015): "Retractions," *Review of Economics and Statistics*, 97, 1118–1136.

Bailey, C. N., J. R. Hasselback, and J. N. Karcher (2001): "Research misconduct in accounting literature: A survey of the most prolific researchers' actions and beliefs," *Abacus*, 37, 26–54.

Balasubramanian, P., L. Pierce, and T. Cummings (2022): "Research Validity Across Organizational Forms: Evidence from Phase 2 Oncology Clinical Trials," *Aalborg University, Johns Hopkins University, and Washington University in St. Louis*.

Baron, J., M. King, and O. Sorenson (2016): "S/he blinded me with science: The sociology of scientific misconduct," in *Organizational wrongdoing: Key perspectives and new directions*, Cambridge University Press, 176–202.

Begley, C. G. and L. M. Ellis (2012): "Raise standards for preclinical cancer research," *Nature*, 483, 531–533.

Begley, S. (2019): "The maddening saga of how an Alzheimer's 'cabal' thwarted progress toward a cure for decades," *STAT News*.

Bik, E. M., A. Casadevall, and F. C. Fang (2016): "The prevalence of inappropriate image duplication in biomedical research publications," *MBio*, 7, 10–1128.

Bikard, M. (2018): "Made in academia: The effect of institutional origin on inventors' attention to science," *Organization Science*, 29, 818–836.

Boudou, J. (2025): "Not All That Glitters is Gold: Firm Hiring in the Market for Knowledge Workers," *Harvard Business School*.

Brandt, A. M. (2012): "Inventing conflicts of interest: a history of tobacco industry tactics," *American Journal of Public Health*, 102, 63–71.

Choi, T. J. and M. Valente (2023): "The crisis in local newspapers and organizational wrongdoing: The role of community social connectedness," *Organization Science*, 34, 1777–1799.

Cobey, K. D., S. Ebrahimzadeh, M. J. Page, R. T. Thibault, P.-Y. Nguyen, F. Abu-Dalfa, and D. Moher (2024): "Biomedical researchers' perspectives on the reproducibility of research," *PLoS Biology*, 22, e3002870.

Cockburn, I. M. and R. M. Henderson (1998): "Absorptive capacity, coauthoring behavior, and the organization of research in drug discovery," *The Journal of Industrial Economics*, 46, 157–182.

Cohen, W. M. and D. A. Levinthal (1990): "Absorptive capacity: A new perspective on learning and innovation," *Administrative Science Quarterly*, 35, 128–152.

Collins, H. M. (2001): "Tacit knowledge, trust and the Q of sapphire," *Social Studies of Science*, 31, 71–85.

Dasgupta, P. and P. A. David (1994): "Toward a new economics of science," *Research Policy*, 23, 487–521.

Dyck, A., F. Fang, C. Hebert, and T. Xu (2026): "Venture Fraud," *NBER wp 34868*.

Eaglin, F. C. (2026): "The need for speed: The impact of capital constraints on strategic misconduct," *Management Science*, 72, 1680–1698.

Errington, T. M., M. Mathur, C. K. Soderberg, A. Denis, N. Perfito, E. Iorns, and B. A. Nosek (2021): "Investigating the replicability of preclinical cancer biology," *eLife*, 10, e71601.

Evans, J. A. (2010): "Industry induces academic science to know less about more," *American Journal of Sociology*, 116, 389–452.

Fanelli, D. (2009): "How many scientists fabricate and falsify research? A systematic review and meta-analysis of survey data," *PloS One*, 4, e5738.

Fang, F. C., R. G. Steen, and A. Casadevall (2012): "Misconduct accounts for the majority of retracted scientific publications," *Proceedings of the National Academy of Sciences*, 109, 17028–17033.

Fleming, L., H. Greene, G. Li, M. Marx, and D. Yao (2019): "Government-funded research increasingly fuels innovation," *Science*, 364, 1139–1141.

Fleming, L. and O. Sorenson (2004): "Science as a map in technological search," *Strategic Management Journal*, 25, 909–928.

Frake, J. and D. Harmon (2024): "Intergenerational transmission of organizational misconduct: Evidence from the Chicago Police Department," *Management Science*, 70, 3856–3878.

Freedman, L. P., I. M. Cockburn, and T. S. Simcoe (2015): "The economics of reproducibility in preclinical research," *PLoS Biology*, 13, e1002165.

Furman, J. L., K. Jensen, and F. Murray (2012): "Governing knowledge in the scientific community: Exploring the role of retractions in biomedicine," *Research Policy*, 41, 276–290.

Furman, J. L. and S. Stern (2011): "Climbing atop the shoulders of giants: The impact of institutions on cumulative research," *American Economic Review*, 101, 1933–1963.

Gambardella, A. (1992): "Competitive advantages from in-house scientific research: The US pharmaceutical industry in the 1980s," *Research Policy*, 21, 391–407.

Gittelman, M. and B. Kogut (2003): "Does good science lead to valuable knowledge? Biotechnology firms and the evolutionary logic of citation patterns," *Management Science*, 49, 366–382.

Gopalakrishna, G., G. Ter Riet, G. Vink, I. Stoop, et al. (2022): "Prevalence of questionable research practices, research misconduct and their potential explanatory factors: A survey among academic researchers in The Netherlands," *PloS One*, 17, e0263023.

Greve, H. R., D. Palmer, and J.-E. Pozner (2010): "Organizations gone wild: The causes, processes, and consequences of organizational misconduct," *Academy of Management Annals*, 4, 53–107.

Guzman, J., Z. Jelveh, and B. Kogut (2026): "The SPIRIT of Science: Institutional Responses to the P-Hacking Problem in Academic Clinical Trials," *Available at SSRN 6226781*.

Herrup, K. (2023): *How not to study a disease: The story of Alzheimer's*, Mit Press.

Hoelzemann, J., G. Manso, A. Nagaraj, and M. Tranchero (2025): "The streetlight effect in data-driven exploration," *NBER wp 32401*.

Horbach, S. P. and W. Halffman (2019): "The ability of different peer review procedures to flag problematic publications," *Scientometrics*, 118, 339–373.

Hutchins, B. I., M. T. Davis, R. A. Meseroll, and G. M. Santangelo (2019): "Predicting translational progress in biomedical research," *PLoS Biology*, 17, e3000416.

Ioannidis, J. P. (2005): "Why most published research findings are false," *PLoS Medicine*, 2, e124.

John, L. K., G. Loewenstein, and D. Prelec (2012): "Measuring the prevalence of questionable research practices with incentives for truth telling," *Psychological Science*, 23, 524–532.

Kao, J. (2025): "Information disclosure and competitive dynamics: Evidence from the pharmaceutical industry," *Management Science*, 71, 5948–5970.

Kim, C. K., Y. R. Lee, L. Ong, M. Gold, A. Kalali, and J. Sarkar (2022): "Alzheimer's disease: key insights from two decades of clinical trial failures," *Journal of Alzheimer's Disease*, 87, 83–100.

Krieger, J. L., M. Schnitzer, and M. Watzinger (2024): "Standing on the shoulders of science," *Strategic Management Journal*, 45, 1670–1695.

Krimsky, S. (2019): *Conflicts of interest in science: How corporate-funded academic research can threaten public health*, Simon and Schuster.

Kuhn, T. (1970): *The nature of scientific revolutions*, Chicago University Press.

Lacetera, N. (2009): "Different missions and commitment power in R&D organizations: Theory and evidence on industry-university alliances," *Organization Science*, 20, 565–582.

Lacetera, N. and L. Zirulia (2011): "The economics of scientific misconduct," *The Journal of Law, Economics, & Organization*, 27, 568–603.

Legg, T., J. Hatchard, and A. B. Gilmore (2021): "The science for profit model—how and why corporations influence science and the use of science in policy and practice," *Plos one*, 16, e0253272.

Lesné, S., M. T. Koh, L. Kotilinek, R. Kayed, C. G. Glabe, A. Yang, M. Gallagher, and K. H. Ashe (2024): "Retraction Note: A specific amyloid-β protein assembly in the brain impairs memory." *Nature*, 631.

Lexchin, J., L. A. Bero, B. Djulbegovic, and O. Clark (2003): "Pharmaceutical industry sponsorship and research outcome and quality: Systematic review," *bmj*, 326, 1167–1170.

Liu, C. C. and T. Stuart (2014): "Positions and rewards: The allocation of resources within a science-based entrepreneurial firm," *Research Policy*, 43, 1134–1143.

Liu, P.-P., Y. Xie, X.-Y. Meng, and J.-S. Kang (2019): "History and progress of hypotheses and clinical trials for Alzheimer's disease," *Signal Transduction and Targeted Therapy*, 4, 29.

Lundh, A., J. Lexchin, B. Mintzes, J. B. Schroll, and L. Bero (2017): "Industry sponsorship and research outcome," *Cochrane database of systematic reviews*.

Marx, M. and A. Fuegi (2022): "Reliance on science by inventors: Hybrid extraction of in-text patent-to-article citations," *Journal of Economics & Management Strategy*, 31, 369–392.

Masclans, R., S. Hasan, and W. M. Cohen (2025): "Measuring the commercial potential of science," *Strategic Management Journal*, 46, 2199–2236.

Merton, R. K. (1973): *The sociology of science: Theoretical and empirical investigations*, University of Chicago Press.

Michalek, A. M., A. D. Hutson, C. P. Wicher, and D. L. Trump (2010): "The costs and underappreciated consequences of research misconduct: a case study," *PLoS Medicine*, 7, e1000318.

Nicholson, J. M., M. Mordaunt, P. Lopez, A. Uppala, D. Rosati, N. P. Rodrigues, P. Grabitz, and S. C. Rife (2021): "Scite: A smart citation index that displays the context of citations and classifies their intent using deep learning," *Quantitative Science Studies*, 2, 882–898.

Novartis (2024): "Position on Scientific Publications," `https://www.novartis.com/investors/reporting-and-transparency-hub`, accessed: 2025-09-25.

Oreskes, N. and E. M. Conway (2010): "Defeating the merchants of doubt," *Nature*, 465, 686–687.

Owen-Smith, J. (2001): "Managing laboratory work through skepticism: Processes of evaluation and control," *American Sociological Review*, 66, 427–452.

Pierce, L., D. C. Snow, and A. McAfee (2015): "Cleaning house: The impact of information technology monitoring on employee theft and productivity," *Management Science*, 61, 2299–2319.

Pierce, L. and J. Snyder (2008): "Ethical spillovers in firms: Evidence from vehicle emissions testing," *Management Science*, 54, 1891–1903.

Piller, C. (2025): *Doctored: Fraud, Arrogance, and Tragedy in the Quest to Cure Alzheimer's*, One Signal Publishers/Atria Books.

Pisano, G. P. (1990): "The R&D boundaries of the firm: an empirical analysis," *Administrative science quarterly*, 153–176.

Polidoro, F. J. and M. Theeke (2012): "Getting competition down to a science: The effects of technological competition on firms' scientific publications," *Organization Science*, 23, 1135–1153.

Prinz, F., T. Schlange, and K. Asadullah (2011): "Believe it or not: how much can we rely on published data on potential drug targets?" *Nature Reviews Drug Discovery*, 10, 712–712.

Reardon, S. (2023): "FDA approves Alzheimer's drug amid safety concerns," *Nature*, 613, 277–278.

Rosenberg, N. (1990): "Why do firms do basic research (with their own money)?" *Research Policy*, 19, 165–174.

Salandra, R. (2018): "Knowledge dissemination in clinical trials: exploring influences of institutional support and type of innovation on selective reporting," *Research Policy*, 47, 1215–1228.

Sauermann, H. and M. Roach (2014): "Not all scientists pay to be scientists: PhDs' preferences for publishing in industrial employment," *Research Policy*, 43, 32–47.

Seegert, L. (2025): "Controversial New Alzheimer's Drugs Offer Hope-But at a High Cost," *Nature*, 645, S10–S12.

Sood, A., E. Kappe, and S. Stremersch (2014): "The commercial contribution of clinical studies for pharmaceutical drugs," *International Journal of Research in Marketing*, 31, 65–77.

Tranchero, M. (2026): "Finding diamonds in the rough: Data-driven opportunities and pharmaceutical innovation," *University of Pennsylvania*.

Tsilidis, K. K., O. A. Panagiotou, E. S. Sena, E. Aretouli, et al. (2013): "Evaluation of excess significance bias in animal studies of neurological diseases," *PLoS Biology*, 11, e1001609.

Van Noorden, R. (2023): "More than 10,000 research papers were retracted in 2023—a new record," *Nature*, 624, 479–481.

Walsh, J. P., Y.-N. Lee, and L. Tang (2019): "Pathogenic organization in science: Division of labor and retractions," *Research Policy*, 48, 444–461.

Wang, Y., T. Stuart, and J. Li (2021): "Fraud and innovation," *Administrative Science Quarterly*, 66, 267–297.

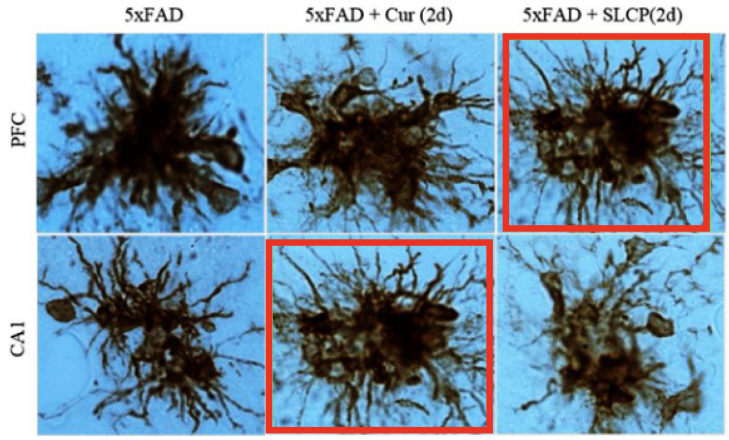
Weber, G. M. (2013): "Identifying translational science within the triangle of biomedicine," *Journal of Translational Medicine*, 11, 126.

Zuckerman, H. and R. K. Merton (1971): "Patterns of evaluation in science: Institutionalisation, structure and functions of the referee system," *Minerva*, 9, 66–100.

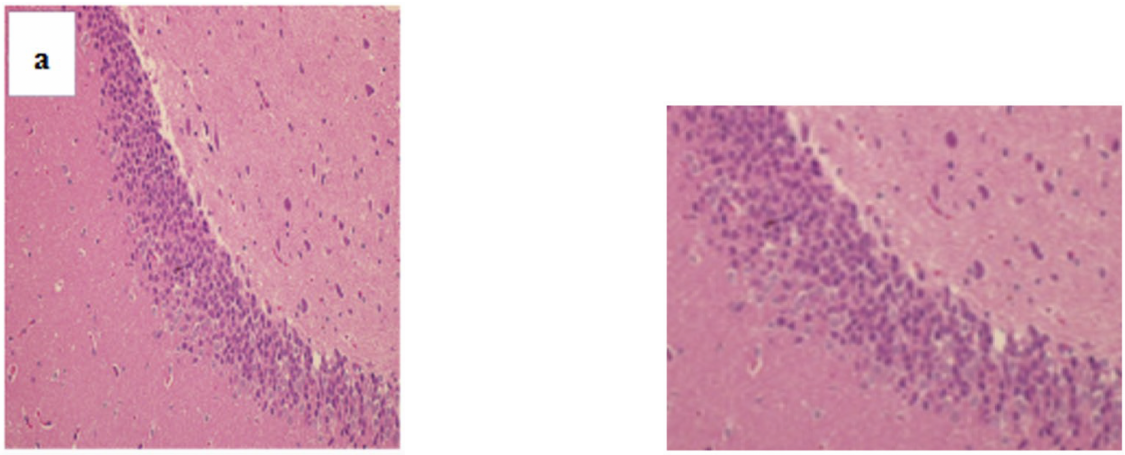
8 Figures and Tables

Figure 1: Examples of unreliable data as captured by Imagetwin's AI software.

Panel A. Example of image duplication within a publication (internal duplication).



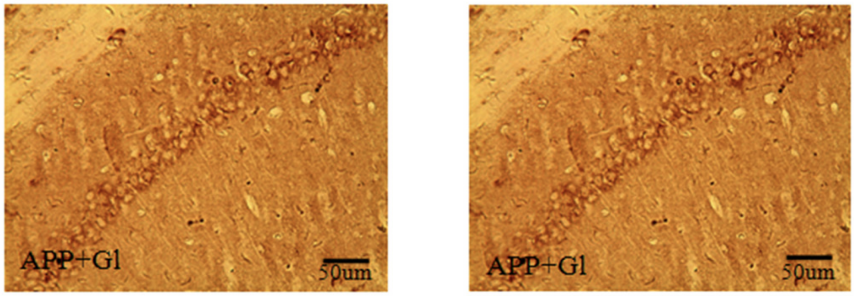
Panel B. Example of image duplication across publications (external duplication).



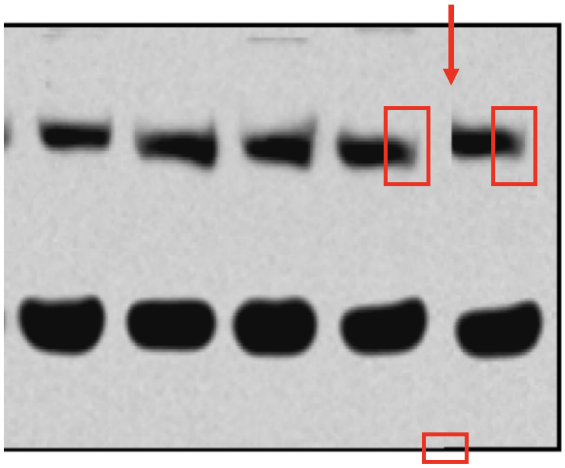
Note: The figure shows two examples of duplicated images captured by our software. Panel A shows Figure 6 from a paper published in 2018. This case presents two identical panels purportedly showing results from different experimental treatments. The article was eventually retracted with the agreement of all the authors: https://bmcneurosci.com/10.1186/s12868-023-00785-5. Panel B shows Figure 6(a) from an article published in 2018 (left) and Figure 1(b) from an article published in 2020 (right). This case presents the same image being cropped and reproduced in another article that does not even cite it. The latter article was eventually retracted upon request of the authors: https://www.degruyter.com/document/doi/10.1515/tnsci-2022-0262/html. All images licensed for reproduction under the CC BY 4.0 license.

Figure 2: Examples of unreliable data manually classified according to our protocol.

Panel A. Example of identical duplication classified as "mistaken".

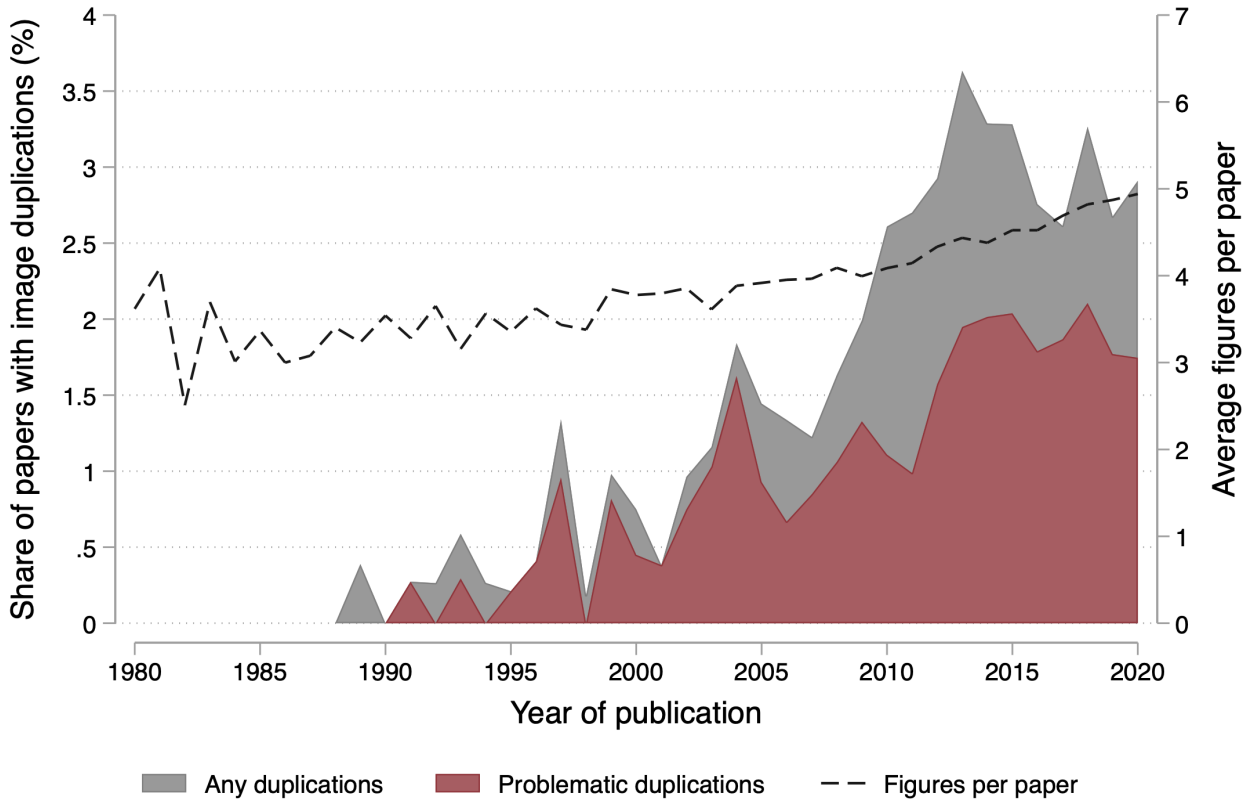


Panel B. Example of cropped duplication classified as "problematic".



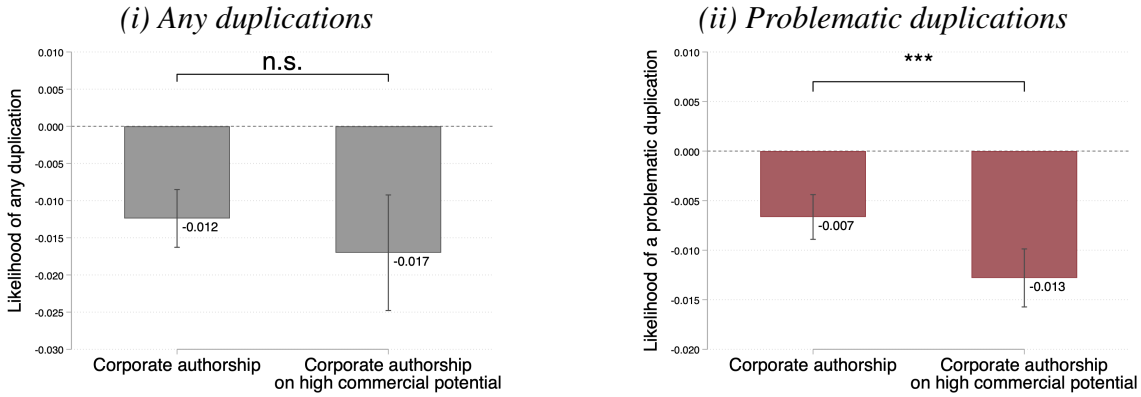
Note: The figure shows examples of image malpractice captured by our software. Panel A shows a portion of Figures 4 and 5 from a paper published in 2016. Since the image shows no signs of alteration and represents the same experimental condition, our protocol gives the authors the benefit of the doubt and classifies it as a potential mistake. However, this illustrates how conservative our approach is, since this type of duplication was enough to result in the retraction of the paper: https://bmccomplementmedtherapies.biomedcentral.com/articles/10.1186/s12906-021-03331-0. Panel B shows a portion of Figure 3 from a paper published in 2014. This figure shows evidence of inappropriate splicing in the rightmost lane of the photo (marked with a red arrow for illustration). This article has received substantial criticism on the PubPeer blog: https://pubpeer.com/publications/62CD938FC80C350418F2FC5CEF3481.

Figure 3: The growing incidence of inappropriate image duplications in Alzheimer's research (1980–2020)



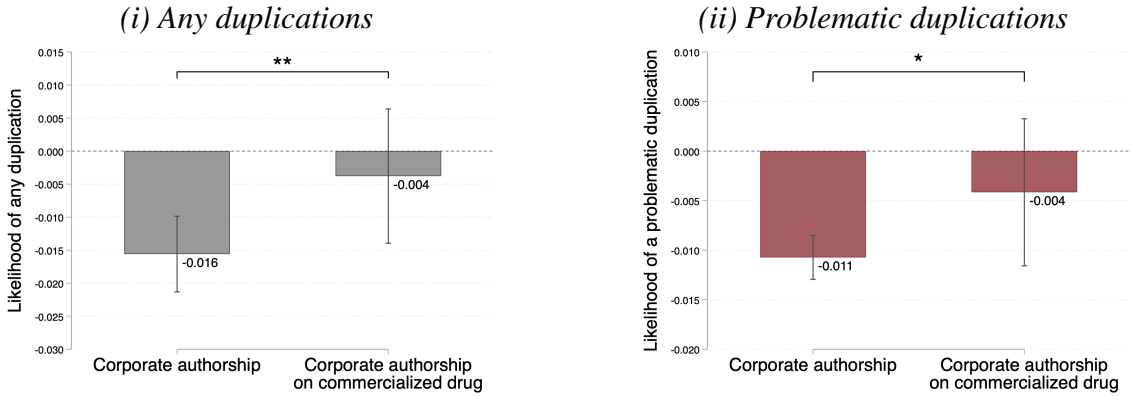
Note: The figure shows trends over time in the incidence of inappropriate image duplications in Alzheimer's research from 1980 to 2020. The left axis reports the annual share of papers containing at least one inappropriate image duplication and the share containing at least one problematic duplication with features suggestive of intentional deception. The right axis reports the average number of figures per publication. The sample includes Alzheimer's publications containing at least one figure.

Figure 4: Corporate-authored papers exhibit fewer deceptive duplications, especially when research is closer to technological application.



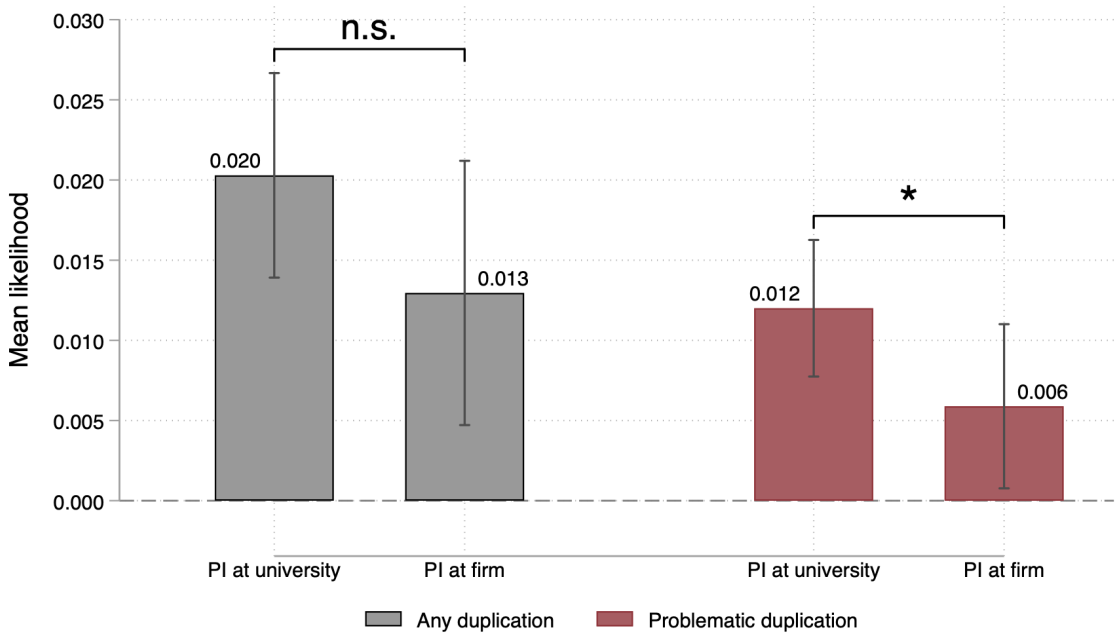
Note: This figure examines whether the association between corporate authorship and image duplications differs by the commercial potential of the research. Each panel reports estimates from a publication-level OLS regression of the form $Y_i = \alpha + \beta_1 Corporate_i + \beta_2 HighCommercialPotential_i + \beta_3 Corporate_i \times HighCommercialPotential_i + \mathbf{Z}_i'\boldsymbol{\delta} + \varepsilon_i$, where Y_i is an indicator equal to one if paper i contains at least one image duplication of the type indicated in the panel title. $Corporate_i$: 0/1 = 1 for papers authored only by researchers affiliated with firms; $HighCommercialPotential_i$: 0/1 = 1 for papers above the median of our commercial potential score. Bars report the implied marginal effect of corporate authorship for papers with low commercial potential, β_1, and for papers with high commercial potential, $\beta_1 + \beta_3$. $\mathbf{Z}_i$ includes fixed effects for publication year, journal, and the Approximate Potential to Translate (APT). Vertical lines represent 95% confidence intervals. Standard errors are clustered at the PI level. The corresponding regression estimates are reported in Table C.10.

Figure 5: The disciplining effect of corporate authorship on deceptive duplications weakens when papers study already commercialized drugs.



Note: This figure examines whether the association between corporate authorship and image duplications differs when the research studies an already commercialized drug. Each panel reports estimates from a publication-level OLS regression of the form $Y_i = \alpha + \beta_1 Corporate_i + \beta_2 CommercializedDrug_i + \beta_3 Corporate_i \times CommercializedDrug_i + \mathbf{Z}_i'\boldsymbol{\delta} + \varepsilon_i$, where Y_i is an indicator equal to one if paper i contains at least one image duplication of the type indicated in the panel title. $Corporate_i$: 0/1 = 1 for papers authored only by researchers affiliated with firms; $CommercializedDrug_i$: 0/1 = 1 if the paper studies a drug already commercialized by the affiliated firm. Bars report the implied marginal effect of corporate authorship for papers not studying a commercialized drug, β_1, and for papers studying a commercialized drug, $\beta_1 + \beta_3$. $\mathbf{Z}_i$ includes fixed effects for publication year, journal, and the Approximate Potential to Translate (APT). Vertical lines represent 95% confidence intervals. Standard errors are clustered at the PI level. The corresponding regression estimates are reported in Table C.11.

Figure 6: Principal investigators who publish from both universities and firms exhibit fewer deceptive duplications when affiliated with firms.



Note: This figure reports raw mean duplication rates for publications by PIs observed with both university and firm affiliations at different moments in their careers. Bars report the mean likelihood that a publication contains any duplicated image or a problematic duplicated image. Vertical lines represent 95% confidence intervals with standard errors clustered at the PI level.

Table 1: Descriptive statistics at the publication level.

	Cross-sectional descriptives					
	Mean	Median	SD	Min	Max	N
Any duplication (0/1)	0.0176	0	0.1314	0	1	46,548
Mistaken duplication (0/1)	0.0076	0	0.0870	0	1	46,548
Problematic duplication (0/1)	0.0106	0	0.1026	0	1	46,548
Academic authorship (0/1)	0.2760	0	0.4470	0	1	46,548
Corporate authorship (0/1)	0.0096	0	0.0975	0	1	46,548
Co-authorship with academic PI (0/1)	0.0289	0	0.1675	0	1	46,548
Co-authorship with corporate PI (0/1)	0.0251	0	0.1565	0	1	46,548
High commercial potential (0/1)	0.5063	1	0.5000	0	1	46,548
Paper on commercialized drug (0/1)	0.0039	0	0.0626	0	1	46,548
Animal research (0/1)	0.2304	0	0.4211	0	1	46,548
Scientific citations	70.6696	32	178.0525	0	12,634	46,548
Clinical citations	1.8086	0	11.6810	0	1,683	46,548
Patent citations	2.1274	0	13.4017	0	669	46,548
Paper self-cited in patent (0/1)	0.0376	0	0.1903	0	1	46,548
APT	0.4926	0.500	0.2940	0.05	0.95	46,548
Year	2008.5	2011	9.2529	1980	2020	46,548

Note: This table lists summary statistics for 46,548 Alzheimer's publications in the estimation sample. *Any duplication*: 0/1 = 1 if the publication presents at least one duplicated image. *Mistaken duplication*: 0/1 = 1 if the publication presents at least one duplicated image classified through manual review as potentially stemming from mistakes. *Problematic duplication*: 0/1 = 1 if the publication presents at least one duplicated image containing features suggestive of intentional deception. *Academic authorship*: 0/1 = 1 for publications authored only by researchers affiliated with universities. *Corporate authorship*: 0/1 = 1 for publications authored only by researchers affiliated with companies. *Co-authorship with academic PI*: 0/1 = 1 for publications co-authored by researchers in academia and industry where the principal investigator is a university researcher. *Co-authorship with corporate PI*: 0/1 = 1 for publications co-authored by researchers in academia and industry where the principal investigator is a company researcher. *High commercial potential*: 0/1 = 1 for publications with an above-median value of the commercial potential score by Masclans et al. (2025). *Paper on commercialized drug*: 0/1 = 1 for publications studying a drug already commercialized by the authoring firm. *Animal research*: 0/1 = 1 for publications classified as animal research by iCite. *Scientific citations*: number of citations received by the publication from other scientific publications. *Clinical citations*: number of citations received by the publication from clinical publications. *Patent citations*: number of citations received by the publication from patents. *Paper self-cited in patent*: 0/1 = 1 if the publication is cited by at least one patent owned by the firm employing at least one of its authors. *APT*: Approximate Potential to Translate score from iCite. *Year*: publication year.

Table 2: Alzheimer's publications authored by corporate researchers present a smaller incidence of deceptive duplications than those authored by academics.

Dependent Variable	Paper with any duplication (0/1)		Paper with mistaken duplication (0/1)		Paper with problematic duplication (0/1)	
	(1)	(2)	(3)	(4)	(5)	(6)
Academic authorship (0/1)	0.00506***		0.00412***		0.00161	
	(0.00185)		(0.00118)		(0.00138)	
Corporate authorship (0/1)		-0.0153***		-0.00502**		-0.0108***
		(0.00282)		(0.00248)		(0.00112)
Year FE	YES	YES	YES	YES	YES	YES
Journal FE	YES	YES	YES	YES	YES	YES
APT FE	YES	YES	YES	YES	YES	YES
Observations	46,548	46,548	46,548	46,548	46,548	46,548
Mean of the DV	0.0176	0.0176	0.00763	0.00763	0.0106	0.0106
Percentage change	**+29%**	**-87%**	**+54%**	**-66%**	+15%	**-101%**

Note: *, **, *** denote significance at the 10%, 5%, and 1% levels, respectively. Cross-sectional OLS regressions at the publication level. Standard errors clustered at the PI level. All models include fixed effects for journal, year, and Approximate Potential to Translate (APT). *Paper with any duplication*: 0/1 = 1 if the publication presents at least one duplicated image. *Paper with mistaken duplication*: 0/1 = 1 if the publication presents at least one duplicated image classified through manual review as potentially reflecting mistakes. *Paper with problematic duplication*: 0/1 = 1 if the publication presents at least one duplicated image containing features suggestive of intentional deception. *Academic authorship*: 0/1 = 1 for publications authored only by researchers affiliated with universities. *Corporate authorship*: 0/1 = 1 for publications authored only by researchers affiliated with companies. Percentage changes are computed relative to the sample mean ($\hat{\beta}/\bar{Y} \times 100$) and rounded to the nearest integer; significant effects are shown in bold.

Table 3: Firms' publications are more likely to be self-cited in patents when they support ongoing innovation rather than commercialized products.

Dependent Variable	Paper self-cited in corporate patent (0/1)		
	(1)	(2)	(3)
Corporate authorship (0/1)	0.0450***	-0.00348	0.0458***
	(0.0140)	(0.0140)	(0.0144)
Corporate authorship (0/1)		0.0702***	
× High commercial potential (0/1)		(0.0234)	
Corporate authorship (0/1)			-0.0190
× Paper on commercialized drug (0/1)			(0.0546)
Year FE	YES	YES	YES
Journal FE	YES	YES	YES
APT FE	YES	YES	YES
Observations	46,548	46,548	46,548
Mean of the DV	0.0376	0.0376	0.0376
Percentage change		**+187%**	-51%

Note: *, **, *** denote significance at the 10%, 5%, and 1% levels, respectively. Cross-sectional OLS regressions at the publication level. Standard errors are clustered at the PI level. All models include fixed effects for journal, year, and Approximate Potential to Translate (APT). *Paper self-cited in corporate patent*: 0/1 = 1 if the publication is cited by at least one patent classified as a self-citation. *Corporate authorship*: 0/1 = 1 for publications authored only by researchers affiliated with companies. *High commercial potential*: 0/1 = 1 for publications in scientific areas with high commercial potential. *Paper on commercialized drug*: 0/1 = 1 for publications studying a drug already commercialized by the authoring firm. Percentage changes for the interaction terms are computed relative to the sample mean ($\hat{\beta}_{interaction}/\bar{Y} \times 100$) and rounded to the nearest integer; significant effects are shown in bold.

Table 4: Academic-industry collaborations are associated with a lower incidence of problematic duplications, but only when led by corporate PIs.

Dependent Variable	Paper with any duplication (0/1)		Paper with mistaken duplication (0/1)		Paper with problematic duplication (0/1)	
	(1)	(2)	(3)	(4)	(5)	(6)
Co-authorship with academic PI (0/1)	-0.000982		-0.00242		0.00142	
	(0.00355)		(0.00214)		(0.00298)	
Co-authorship with corporate PI (0/1)		-0.00470		0.000319		-0.00572***
		(0.00327)		(0.00260)		(0.00201)
Year FE	YES	YES	YES	YES	YES	YES
Journal FE	YES	YES	YES	YES	YES	YES
APT FE	YES	YES	YES	YES	YES	YES
Observations	46,548	46,548	46,548	46,548	46,548	46,548
Mean of the DV	0.0176	0.0176	0.00763	0.00763	0.0106	0.0106
Estimate as % of DV mean	-6%	-27%	-32%	+4%	+13%	**-54%**

Note: *, **, *** denote significance at the 10%, 5%, and 1% levels, respectively. Cross-sectional OLS regressions at the publication level using the full estimation sample. Standard errors are clustered at the PI level. All models include fixed effects for journal, year, and Approximate Potential to Translate (APT). Odd-numbered columns examine academic-industry collaborations led by an academic PI, while even-numbered columns examine academic-industry collaborations led by a corporate PI. *Paper with any duplication*: 0/1 = 1 if the publication presents at least one duplicated image. *Paper with mistaken duplication*: 0/1 = 1 if the publication presents at least one duplicated image classified through manual review as potentially reflecting mistakes. *Paper with problematic duplication*: 0/1 = 1 if the publication presents at least one duplicated image containing features suggestive of intentional deception. *Co-authorship with academic PI*: 0/1 = 1 for publications co-authored by researchers in academia and industry where the PI is a university researcher. *Co-authorship with corporate PI*: 0/1 = 1 for publications co-authored by researchers in academia and industry where the PI is a company researcher. Percentage changes for the interaction terms are computed relative to the sample mean ($\hat{\beta}_{interaction}/\bar{Y} \times 100$) and rounded to the nearest integer; significant effects are shown in bold.

Table 5: The lower incidence of deceptive duplications among firms' publications is particularly pronounced in animal research.

Dependent Variable	Paper with any duplication (0/1)		Paper with mistaken duplication (0/1)		Paper with problematic duplication (0/1)	
	(1)	(2)	(3)	(4)	(5)	(6)
Corporate authorship (0/1)	-0.0153***	-0.0105***	-0.00502**	-0.00403***	-0.0108***	-0.00679***
	(0.00282)	(0.00135)	(0.00248)	(0.000681)	(0.00112)	(0.00103)
Corporate authorship (0/1)		-0.0103*		-0.00219		-0.00872***
× Animal research (0/1)		(0.00567)		(0.00520)		(0.00169)
Year FE	YES	YES	YES	YES	YES	YES
Journal FE	YES	YES	YES	YES	YES	YES
APT FE	YES	YES	YES	YES	YES	YES
Observations	46,548	46,548	46,548	46,548	46,548	46,548
Mean of the DV	0.0176	0.0176	0.0076	0.0076	0.0106	0.0106
Percentage change		**-59%**		-29%		**-82%**

Note: *, **, *** denote significance at the 10%, 5%, and 1% levels, respectively. Cross-sectional OLS regressions at the publication level. Standard errors clustered at the PI level. All models include fixed effects for journal, year, and Approximate Potential to Translate (APT). *Paper with any duplication*: 0/1 = 1 if the publication presents at least one duplicated image. *Paper with mistaken duplication*: 0/1 = 1 if the publication presents at least one duplicated image classified through manual review as potentially reflecting mistakes. *Paper with problematic duplication*: 0/1 = 1 if the publication presents at least one duplicated image containing features suggestive of intentional deception. *Corporate authorship*: 0/1 = 1 for publications authored only by researchers affiliated with companies. *Animal research*: 0/1 = 1 for publications classified as animal research using iCite data. Percentage changes for the interaction terms are computed relative to the mean of the corresponding dependent variable ($\hat{\beta}_{interaction}/\bar{Y} \times 100$) and rounded to the nearest integer; significant interaction effects are shown in bold.

Deceptive Science Maps: Organizational Affiliation and Data Manipulation in Alzheimer's Research

Appendix

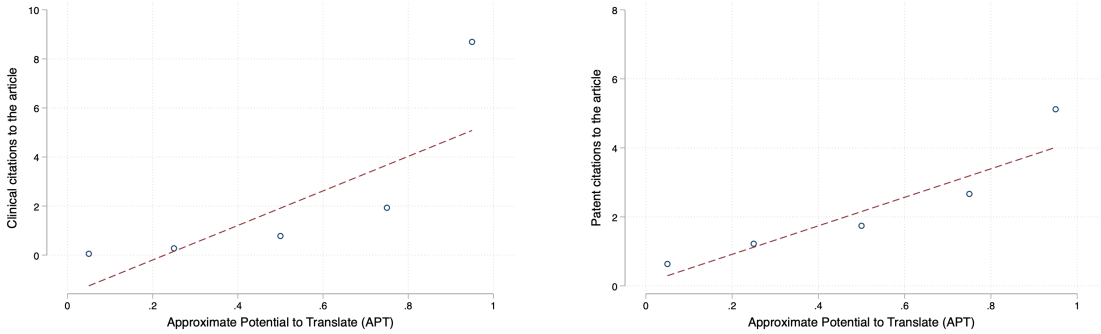
A Additional Data Details

A.1 Approximate Potential to Translate (APT) of an Article

We use the Approximate Potential to Translate (APT) score developed by the NIH as an article-level indicator of expected translational relevance (Hutchins et al., 2019). APT is constructed with a supervised machine learning model that predicts whether a publication will eventually be cited by a clinical trial or guideline. The algorithm produces a predicted probability of receiving a clinical citation for each article. In the iCite data, this probability is assigned to one of five bins, ranging from 0.05 to 0.95. Higher values indicate a higher expected probability that knowledge from the article will later be used in clinical research. In our empirical analysis, APT serves as a control for the ex ante translational orientation of a paper, capturing its position on the basic-to-clinical continuum of biomedical research. Figure A.1 documents that APT is closely related to realized downstream impact. Panel (a) shows that articles that attract more clinical citations have substantially higher ex ante APT scores. Panel (b) shows an analogous pattern for patent citations: articles that are cited more intensively by patents have higher APT values on average.

Figure A.2 examines how APT varies across journals and author affiliations. Panel (a) presents mean APT scores for five illustrative journals. The basic methods outlet *Methods in Molecular Biology* has the lowest mean APT, whereas the clinical journal *JAMA* has the highest, with the multidisciplinary journals *PNAS* and *Nature* in between. *Medical Hypotheses*, which explicitly positions itself as a venue for speculative and theoretical contributions that may lack extensive experimental support, has an average APT of 0.22 in our data, well below the sample mean. This pattern is consistent with the journal's mission, since its articles tend to articulate novel concepts that are relatively far from immediate clinical application. Panel (b) of Figure A.2 reports mean APT scores by primary institutional affiliation of the authors. Articles written by university-based researchers have slightly lower APT values than articles authored by researchers at hospitals or firms. The higher scores for hospital and company affiliations are consistent with the idea that research conducted in healthcare and corporate environments is, on average, closer to clinical use and therefore more likely to be cited in later clinical work.

Figure A.1: Approximate Potential to Translate (APT) correlates with clinical and patent citations received by an article.

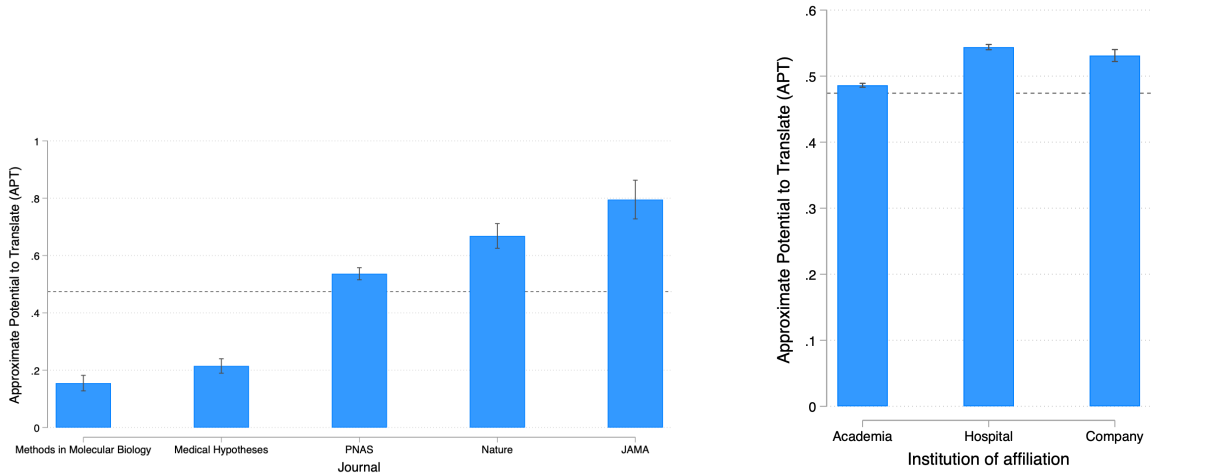


(a) Articles with higher APT receive more clinical citations

(b) Articles with higher APT receive more patent citations

Note: The figure reports raw correlations between the APT of an article and proxies of downstream impact. Panel (a) shows a higher APT for articles that received more clinical citations. Panel (b) shows a higher APT for articles that received more patent citations. See text for more details.

Figure A.2: Articles published in more applied journals or authored by corporate researchers have a higher Approximate Potential to Translate (APT) score.



(a) *Average APT for selected journals*

(b) *Average APT by authors' affiliation*

Note: The figure reports mean APT scores by article type. Panel (a) shows a higher APT for articles published in applied and translational medical journals than for those published in basic science outlets. Panel (b) shows that authors at healthcare institutions or firms publish articles with a higher average APT than academic authors. See text for more details.

A.2 Methods Used in an Article

We use the Human, Animal, and Molecular or Cellular Biology scores from NIH's iCite database to construct article-level measures of the experimental methods employed in each publication

(Weber, 2013; Hutchins et al., 2019). These scores are derived from Medical Subject Headings (MeSH) assigned by the National Library of Medicine and follow the framework developed by Weber (2013), which groups MeSH descriptors into broad domains corresponding to human studies, animal model research, and molecular or cellular work. The Office of Portfolio Analysis refines this approach by computing fractional contributions of an article's MeSH terms to each domain. As a result, papers dominated by human-related descriptors receive high Human scores, papers describing animal model experiments receive high Animal scores, and papers centered on cell- or molecule-level biology receive high Molecular or Cellular scores. For each PubMed-indexed article in our Alzheimer's dataset, we retrieve these three scores from iCite and treat them as standardized, content-based indicators of the study's primary experimental setting.

To convert these continuous measures into discrete methodological categories, we generate inclusive indicator variables at the PMID level. An article is coded as human, animal (in vivo), or cellular (in vitro) if the corresponding iCite score is at least 0.33, allowing papers to be assigned to multiple categories when they employ more than one experimental approach, such as combining animal models with cellular assays. This yields a flexible, MeSH-based classification of experimental methods that can be merged with publication-level data and used to characterize methodological heterogeneity across Alzheimer's research. Our results are robust to alternative codings, including a mutually exclusive rule that assigns each article to the domain with the highest iCite score.

B Detailed Protocol for Classifying Duplication Types

Our conservative approach labels an image as duplicated or manipulated only when two independent coders confirm the Imagetwin-flagged instance. Yet even among inappropriate image duplications, substantial variation exists in both behavior and likely origin. Prior studies of western blot images suggest that roughly one-third of duplications are simple repetitions, consistent with sloppiness (Bik et al., 2016). Others exhibit features such as unnatural background patterns or visibly altered panels that are harder to reconcile with benign explanations and raise more serious concerns about the authors' intent. As Greve et al. (2010) and Walsh et al. (2019) emphasize, the boundary between "reputable" and "disreputable" errors is often socially constructed rather than inherent in the error. Even honest mistakes can be labeled deviant when they fall outside community expectations, because social-control agents such as journal editors and post-publication reviewers play a key role in defining which irregularities merit sanction (Walsh et al., 2019). What matters empirically is whether the duplication or manipulation violates the norms of acceptable practice about how data should be collected or presented. Building on this insight, and drawing on earlier work by Azoulay et al. (2015) and Bik et al. (2016), we developed a manual protocol to classify the type and severity of image duplication, distinguishing plausible assembly errors from patterns that more clearly indicate a deceptive intent.

Protocol for internal inappropriate duplications

1 Establish the type of duplication:

- 1.1 Simple duplication for the same condition: if the same image (complete or in part) is used to represent the same condition but in another experiment, without apparent manipulations.
- 1.2 Simple duplication for different conditions: if the same image (complete or in part) is used to represent a different condition in the same or another experiment, without apparent manipulations.
- 1.3 Image splicing: if the image presents clear evidence of vertical splicing that is not documented or noted in the footnote or text.[7]
- 1.4 Duplication with repositioning and/or alteration: if the same image (complete or in part) is repositioned, rotated, mirrored, shifted, stretched, overlapped, cropped, modified, or if colors are altered,[8] labels changed, or when the image presents unnatural repetition

[7] We followed the methodology of Bik et al. (2016) and do not consider as inappropriate instances of potential beautification, such as cases where part of the background of a western blot shows signs of patching to remove a smudge or stain. Instead, we focused on vertical splicing across lanes of the western blot that are commonly considered inaccurate data representation (e.g., see https://www.frontiersin.org/news/2021/western-blot-image-cropping).

[8] Note that when the duplication was due to different stains of the same picture done on purpose, we did not consider it as an inappropriate duplication at all. We only flagged duplicate images where the color appeared altered, suggesting the image was misrepresenting the result of a different experiment.

patterns in the background or other explicit manipulation in how the image is assembled.

2 Assign the type of likely behavior evinced by the duplication:

2.1 Mistaken duplication: if the duplication is a simple duplication for the same condition.

2.2 Problematic duplication: if the duplication is a simple duplication for different conditions or the image presents splicing, repositioning and/or alteration; category for ambiguous cases that are not simple duplications for the same condition.

Protocol for external inappropriate duplications

1 Establish if the focal paper is the duplicator or the victim of the duplication (assuming that the first published paper is the victim; if two papers are published the same year, refer to the date of submission when reported in the paper).

2 Establish if the duplicated image is properly attributed; if not, establish if the duplicator paper at least cites the paper from which it copies.

3 Establish the type of duplication:

3.1 Poor acknowledgment: if the image is clearly taken from another paper for illustrative purposes, but it fails to explicitly acknowledge the source of the image in the footnote or text.

3.2 External duplication without alteration: if the image (complete or in part) is taken from another paper and used to represent a new experiment.

3.3 External duplication with repositioning and/or alteration: if the image (complete or in part) is taken from another paper and repositioned, rotated, mirrored, shifted, stretched, overlapped, cropped, modified, or if colors are altered, labels changed, or when the image presents unnatural repetition patterns in the background or other explicit manipulation in how the image is assembled.

4 Assign the type of likely behavior evinced by the duplication:

4.1 mistaken duplication: if the duplication is poorly acknowledged or if the original paper is cited.

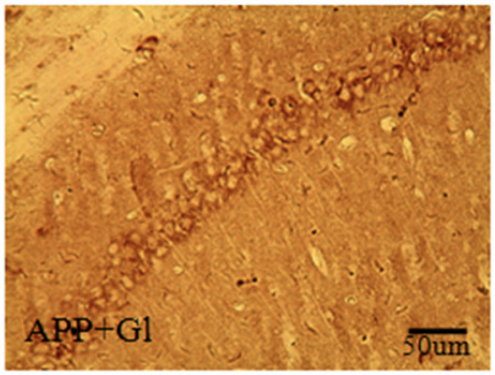
4.2 problematic duplication: if the external duplication is without alterations and does not cite the original paper, or if the external duplication is used for different conditions, with repositioning and/or alteration.

We present a few examples from our data and their corresponding classifications according to our protocol. Examples of duplications are mostly drawn from papers that have been retracted or flagged for issues on the PubPeer forum.

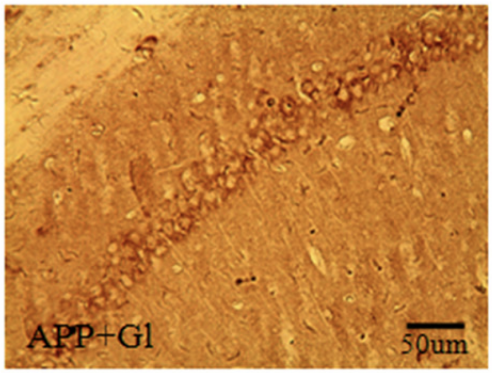
Simple duplication for the same condition Figure B.1 shows two panels from Figure 4 and Figure 5 of a retracted paper.[9] Both figures are composed of six panels each and present instances of repeated identical images. This instance of duplication found with Imagetwin software confirms the conclusions of the journal's editors. The paper has been retracted and the editorial note reported several issues, including that *"Figure 4: the bottom left panel appears to be identical with the bottom left panel of Fig. 5."* Since the image does not present signs of alterations and is used to represent the same experimental condition, our protocol gives the authors the benefit of the doubt and classifies this case of potential sloppiness. However, this illustrates how conservative our approach is, since this type of duplication was enough to impair the credibility of the entire manuscript and result in the retraction of the paper.

Figure B.1: Example of internal duplication manually classified as potentially "mistaken" according to our protocol.

(a) *Panel 4 in Fig. 4 from Shi et al. (2016)*



(b) *Panel 4 in Fig. 5 from Shi et al. (2016)*



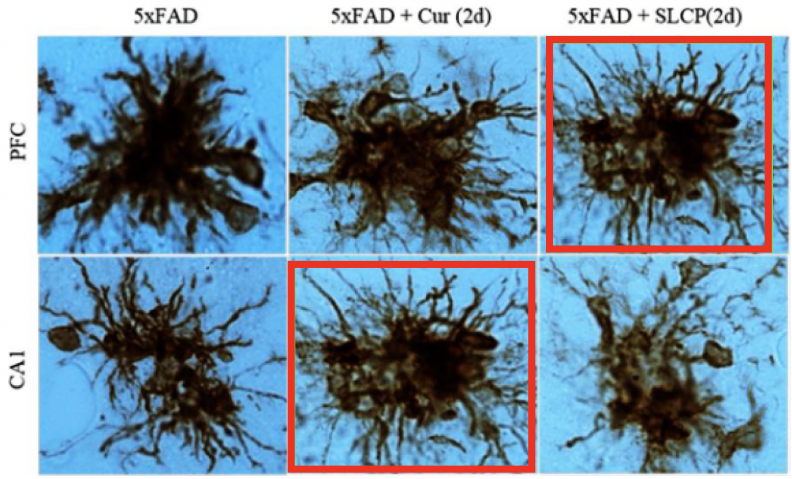
Simple duplication for different conditions Figure B.2 shows a portion of Figure 6 from another retracted paper.[10] Two panels of the figure appear identical despite showing data about different experimental treatments (highlighted by Imagetwin in red in Figure B.2). In particular, the top-left panel supposedly reports 5xFAD mice treated with SLCP for two days, while the middle-bottom panel reports 5xFAD mice treated with Cur for two days. However, the two figures are suspiciously similar. Due to this duplication and a few similar ones, the article was eventually retracted with the agreement of all the authors. Since the duplicated figure represents different experimental conditions, our protocol classifies this as problematic.

[9]The retraction note is available here: https://bmccomplementmedtherapies.biomedcentral.com/articles/10.1186/s12906-021-03331-0.

[10]The retraction note is available here: https://bmcneurosci.biomedcentral.com/articles/10.1186/s12868-023-00785-5

Figure B.2: Example of internal duplication manually classified as "problematic" according to our protocol.

(a) *Portion of Fig. 6 from Maiti et al. (2018); duplicated panels are highlighted in red.*



Internal duplication with alteration Figure B.3 shows a portion of Figure 7 from a retracted article.[11] The green square marked by Imagetwin in the figure highlights a worrying pattern: several areas of the figure show identical repetitive features, which are highly likely to be the result of manipulations. In this specific case, the article was retracted, and the accompanying notice is quite telling: *"After publication, concerns were raised regarding repetitive elements in the images presented in Figs. 6, 7, and 8. The authors have stated that the images were edited to remove artifacts. Due to the number and extent of image modifications, the Editor-in-Chief no longer has confidence in the presented data."* Owing to the presence of alteration of the background of the figure, we classified this instance of duplication as problematic.

Image splicing Figure B.4 shows panel A of Figure 3 from a paper that has not been retracted but has received substantial scrutiny on the PubPeer blog.[12] The figure presents a western blot graph, a common laboratory technique used to detect a specific protein in a blood or tissue sample. However, this figure shows potential evidence of inappropriate splicing in the rightmost lane of the photo (which we marked with a red arrow for illustrative purposes). While cropping in western blots is sometimes warranted, best practices require full disclosure and explanation of the rationale behind it. In this case, neither the text nor the figure footnote provide any justification for the image manipulation. Furthermore, anonymous users on the PubPeer blog have noticed that the cropped lane is an exact duplication of the neighboring one, further compounding doubts about the integrity of this data. According to our protocol, we classified this splicing event as problematic.

[11]The retraction note is available here: https://link.springer.com/article/10.1007/s11011-023-01282-8

[12]The PubPeer blog discussion with detailed evidence on the figure is publicly available here: https://pubpeer.com/publications/62CD938FC80C350418F2FC5CEF3481

Figure B.3: Example of internal duplication manually classified as "problematic" according to our protocol.

(a) *Panel 1a of Fig. 7 from Rezk et al. (2022); duplicated areas are highlighted in green*

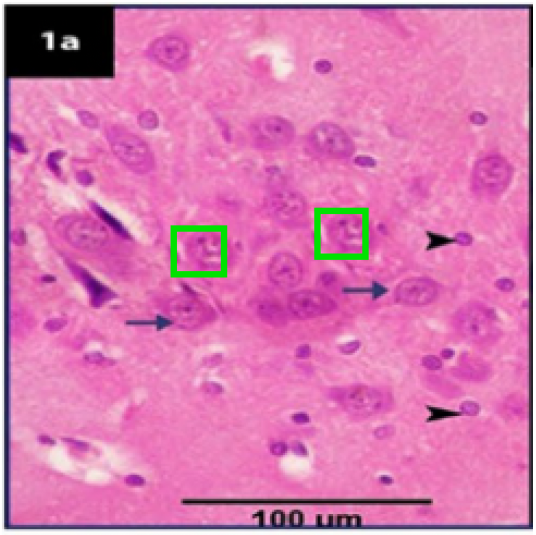
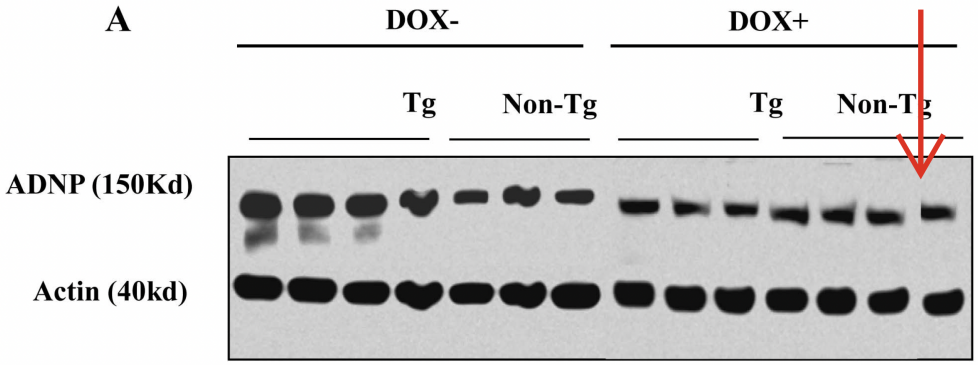


Figure B.4: Example of splicing manually classified as "problematic" according to our protocol. The red arrow is added for illustrative purposes and not from the original figure.

(a) *Panel A of Fig. 3 from Schirer et al. (2014); location of the splice is marked by a red arrow*



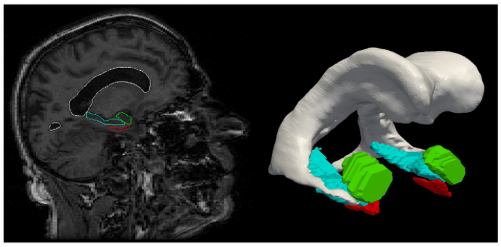
Poor acknowledgment of external duplication Panel (a) of Figure B.5 shows Figure 1 from an article published in 2013,[13] while panel (b) of the same figure shows Figure 2 from an article published in 2015.[14]. The figures are clearly identical, despite the text or footnotes of the later article failing to acknowledge so. However, several reasons lead us to believe that this duplication is likely a case of poor acknowledgment rather than a more serious attempt to deceive. First, the image has a clear illustrative purpose and intent and, as such, does not purport to represent the results of a novel study. Second, the two articles share authors, and the second paper cites the first

[13]The article is available here: https://www.sciencedirect.com/science/article/pii/S2213158213001186
[14]The article is available here: https://onlinelibrary.wiley.com/doi/ftr/10.1002/hbm.22810

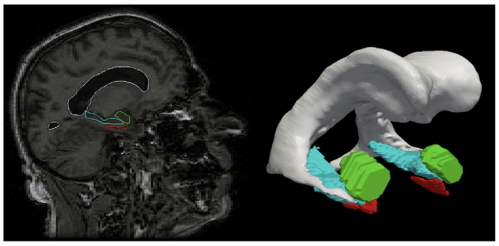
extensively. Both articles present distinct sets of results from the same larger collaborative study, BIOCARD. Therefore, our protocol leads us to classify this instance of external duplication as just due to sloppiness in acknowledgments.

Figure B.5: Example of external duplication manually classified as potentially "mistaken" according to our protocol.

(a) *Fig. 1 from Miller et al. (2013)*



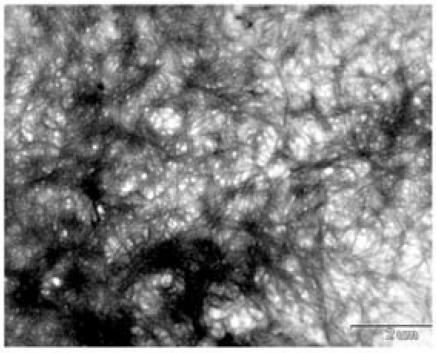
(b) *Fig. 2 from Soldan et al. (2015)*



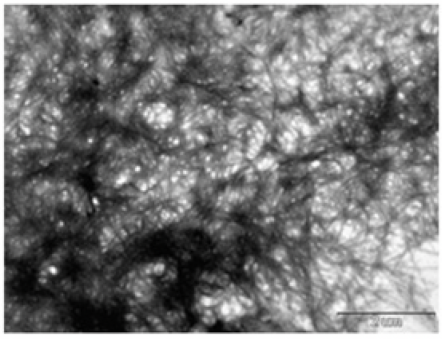
External duplication without alteration Panel (a) of Figure B.6 shows Figure 2(C) from an article published in 2010[15], while panel (b) of the same figure shows Figure 1(D) from a subsequent publication.[16] Also in this case, the subsequent article shares the same authorship as the duplicated one, as well as citing it. However, the article does not acknowledge the use of the same image. This case is more ambiguous because the same image is used to represent the control condition of experiments in different papers. In the first paper, the figure represents the A_{1-42} alone control for a treatment of $A_{1-42} + NQTrp$. In the second paper, the figure represents the A_{1-42} alone control for a treatment of $A_{1-42} + Cl - NQTrp$. Therefore, our protocol leads us to classify this case of duplication as problematic.

Figure B.6: Example of external duplication manually classified as "problematic" according to our protocol.

(a) *Panel C in Fig. 2 from Scherzer-Attali et al. (2010)*



(b) *Panel D in Fig. 1 from Scherzer-Attali et al. (2012)*

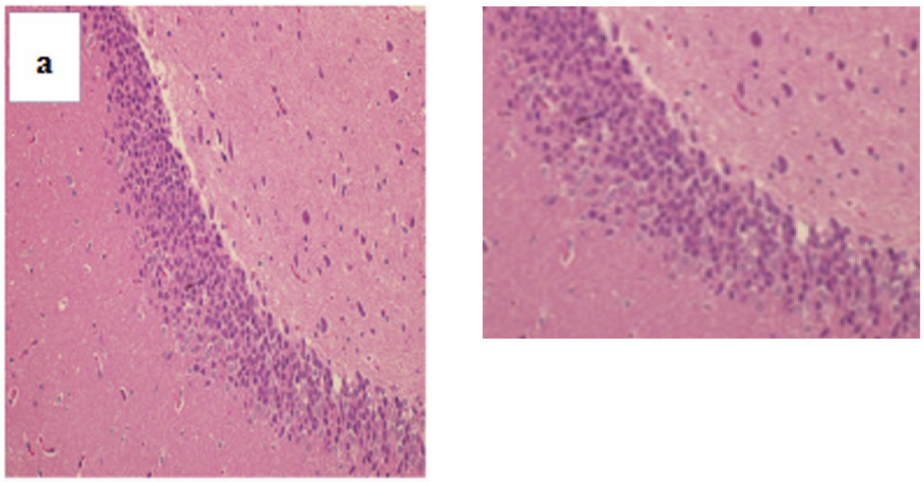


[15]The article is available here: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0011101
[16]The article is available here: https://www.sciencedirect.com/science/article/abs/pii/S0969996112000769

External duplication with alteration Panel (a) of Figure B.7 shows Figure 6(a) from an article published in 2018,[17] while panel (b) of the same figure shows Figure 1(b) from a paper published in 2020.[18] This case presents the same image being cropped and reproduced in another manuscript, without any attribution or citation to the original paper. The papers do not share any authors. Such external duplications would appear less likely to result from honest mistakes in assembling the manuscript. Indeed, the paper in question was eventually retracted because *"The results are flawed and the final conclusions could not be confirmed."*

Figure B.7: Example of external duplication manually classified as "problematic" according to our protocol.

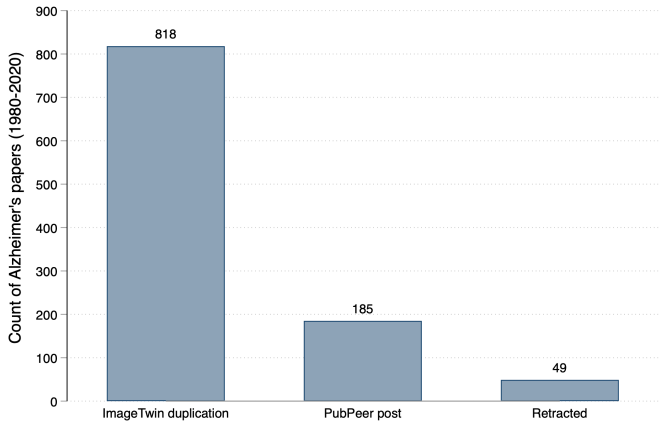
(a) *Panel a in Fig. 6 of Chen et al. (2018)* (b) *Panel b in Fig. 1 of Cheng et al. (2020)*



[17]The article is available here: https://www.nature.com/articles/s41598-018-19732-0
[18]The retraction note is available here: https://www.degruyter.com/document/doi/10.1515/tnsci-2022-0262/html

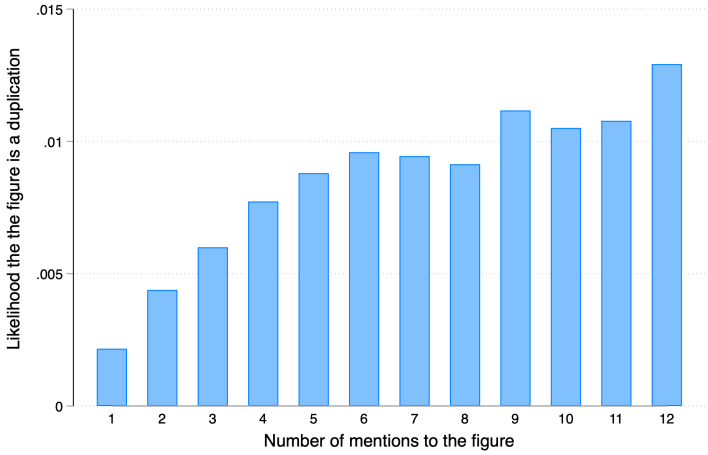
C Additional Results

Figure C.1: Imagetwin detects publications with inappropriate image duplications that have not yet been retracted or discussed on PubPeer.



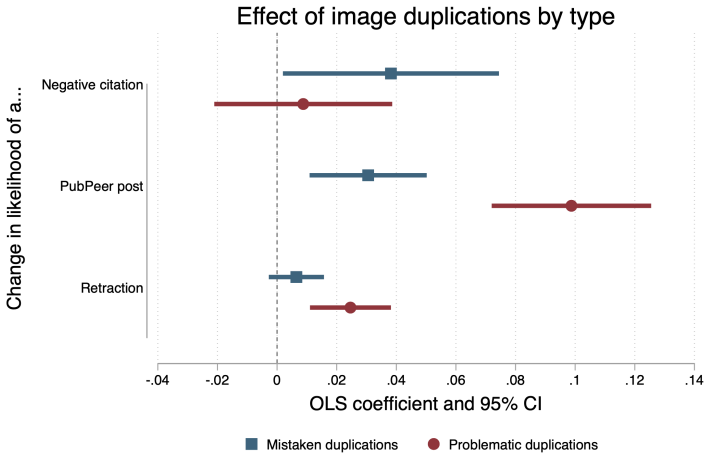
Note: The figure compares the number of Imagetwin-flagged papers identified in this study with the number of papers flagged on the PubPeer blog and the number of retractions reported by Retraction Watch, focusing on publications from 1980 to 2020. See text for more details.

Figure C.2: Duplicated images are more likely to be the central findings of a paper, as proxied by the number of mentions in the articles' text.



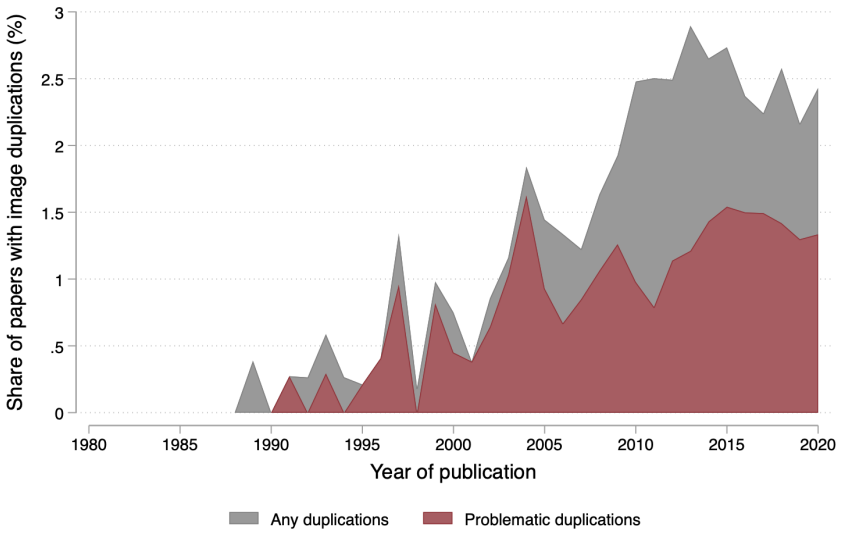
Note: The figure shows the relationship between the number of times an image is mentioned in the article's text and the likelihood that it contains inappropriate duplications. Across all images in the dataset, the average number of mentions in the articles' text is 3.14. See text for more details.

Figure C.3: Publications with duplicated images detected by Imagetwin are more likely to have been flagged on PubPeer or retracted, especially when we code them as problematic.



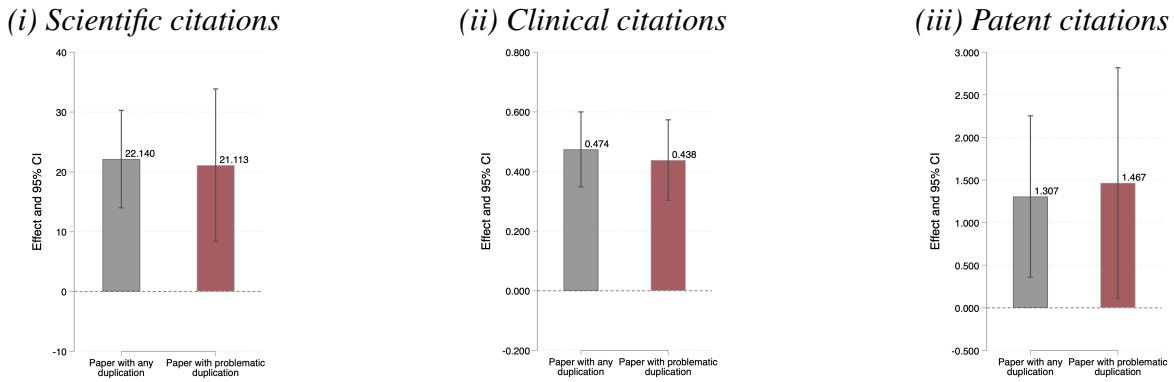
Note: Each bar is estimated from a separate OLS regression at the paper-level: $Y_i = \alpha + \beta\, Duplication_i + \mathbf{Z}_i'\boldsymbol{\delta} + \varepsilon_i$, where Y_i is the outcome of interest for paper i and $Duplication_i$ equals one if the paper contains at least one image issue. $\mathbf{Z}_i$ includes fixed effects for journal, year, and the Approximate Potential to Translate (APT). Bars are 95% confidence intervals. Standard errors clustered at the PI level. *Negative citation*: 0/1 = 1 if the article has received at least one negative citation in the Scite data. *Negative citation*: 0/1 = 1 for articles with at least one mention of negative tone from the Scite data; *PubPeer post*: 0/1 = 1 for articles with at least one discussion thread on the PubPeer post-review blog; *Retraction*: 0/1 = 1 for articles that have been retracted in the Retraction Watch data (as of April 2024). See text for more details.

Figure C.4: Incidence of internal (within-paper) duplications in Alzheimer's research (1980–2020)



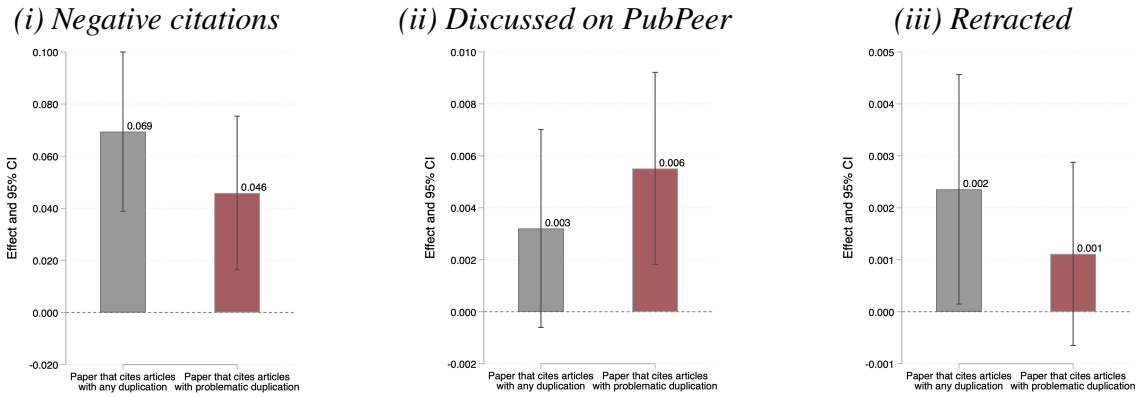
Note: The figure shows trends over time in the incidence of inappropriate image duplications in Alzheimer's research from 1980 to 2020, including only internal duplications. See text for details.

Figure C.5: Articles containing image duplications have a larger downstream impact.



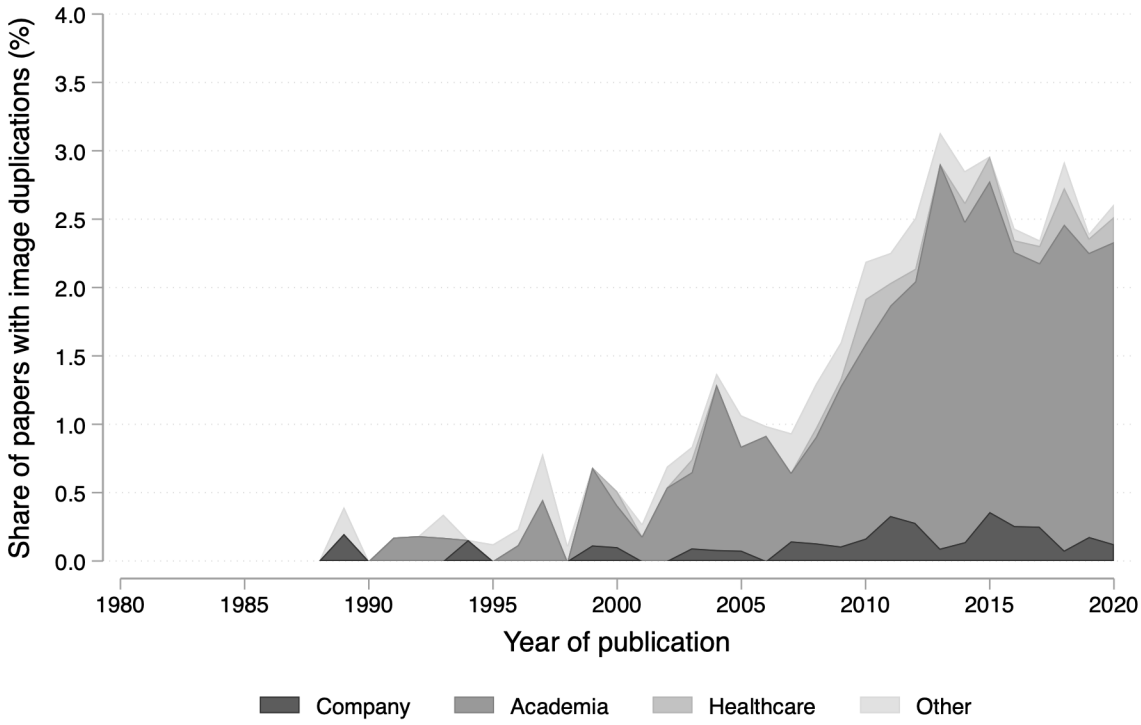
Note: This figure provides visual illustrations of the downstream diffusion of inappropriate image duplications in Alzheimer's research. Each bar is estimated from a separate OLS regression at the paper level: $Y_i = \alpha + \beta\, Duplication_i + \mathbf{Z}_i'\boldsymbol{\delta} + \varepsilon_i$, where Y_i is the outcome of interest for paper i and $Duplication_i$ equals one if the paper contains at least one confirmed image duplication. $\mathbf{Z}_i$ includes fixed effects for journal, year, and the Approximate Potential to Translate (APT). Bars represent 95% confidence intervals. Standard errors clustered at the PI level. The charts report regressions of citations received from (i) scientific publications, (ii) clinical trials, and (iii) USPTO granted patents, on whether the focal paper presents image duplications.

Figure C.6: Papers building on articles with duplicated images appear less scientifically solid.



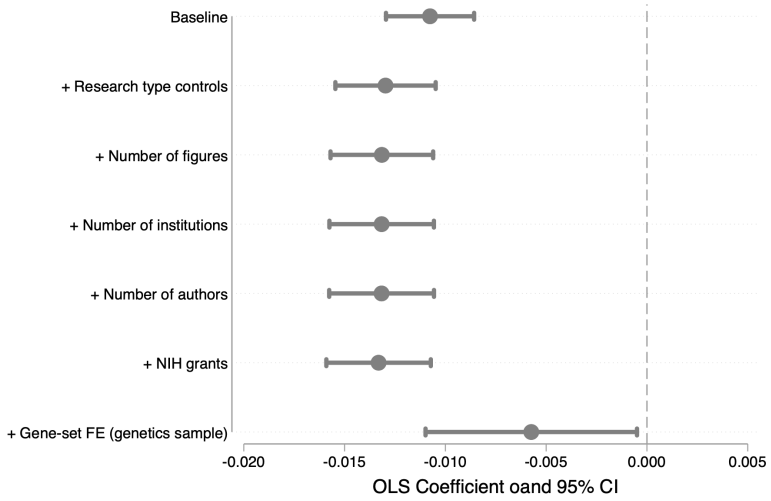
Note: This figure provides visual illustrations of the downstream consequences of unreliable Alzheimer's research. Each bar is estimated from a separate OLS regression at the paper level: $Y_i = \alpha + \beta\, Cites\ an\ Article\ with\ Duplication_i + \mathbf{Z}_i'\boldsymbol{\delta} + \varepsilon_i$, where Y_i is the outcome of interest for paper i and $Cites\ an\ Article\ with\ Duplication_i$ equals one if the paper cites previous work with at least one confirmed image duplication. $\mathbf{Z}_i$ includes fixed effects for journal, year, and the Approximate Potential to Translate (APT). Bars represent 95% confidence intervals. Standard errors clustered at the PI level. The charts report the count of citations with a negative tone (i), the likelihood of being criticized on the PubPeer blog (ii) or being retracted (iii), on whether the focal paper cites articles presenting image duplication. See text for details.

Figure C.7: The organizational origin of papers with image duplications over time (1980–2020)

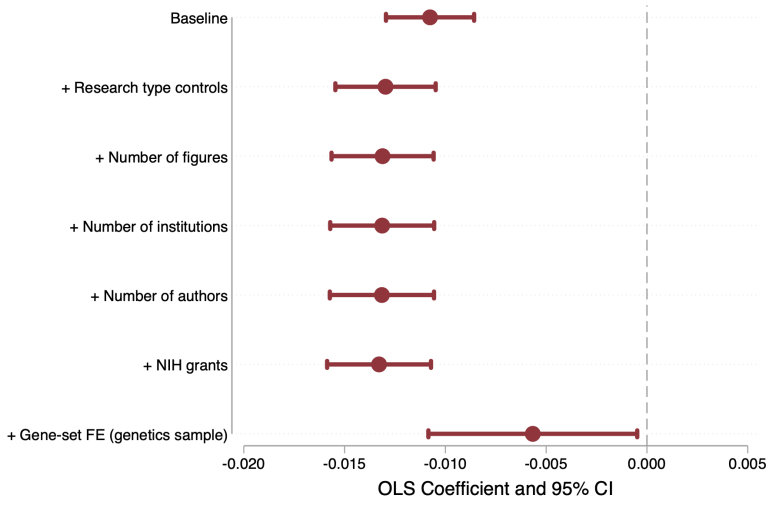


Note: The figure decomposes the annual share of Alzheimer's research papers containing image duplications according to the organizational affiliations of their authors. Papers with any company involvement are classified as Company; among the remaining papers, those with any university involvement are classified as Academia, and so on. The stacked areas sum to the total annual share of papers containing any image duplications. The sample includes all publications on Alzheimer's disease that contain at least one image.

Figure C.8: The relationship between corporate authorship and image duplications is stable across alternative specifications with increasingly tighter controls.



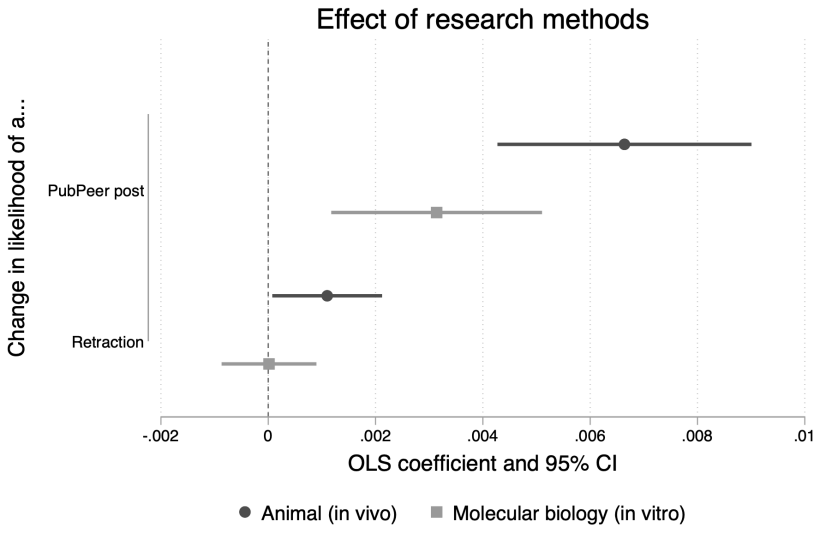
(a) *Any duplications*



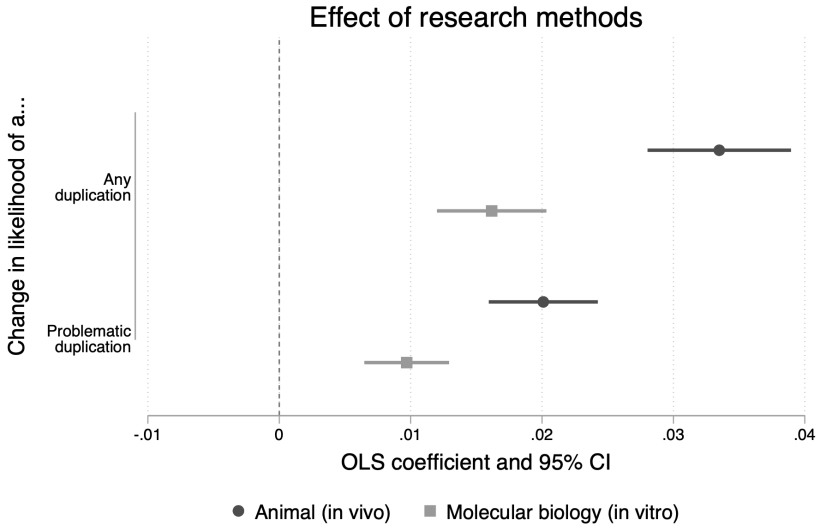
(b) *Problematic duplications*

Note: The figure examines the stability of the relationship between corporate authorship and image problems reported in Table 2 as we progressively introduce additional controls. Each point reports the OLS with 95% confidence intervals. All specifications include journal, publication year, and Approximate Potential to Translate (APT) fixed effects. We then sequentially add controls for research type (Human, Animal, and Molecular/Cellular), the number of figures in the paper, the number of institutions, the number of authors, and the number of NIH grants associated with the article. The final specification additionally includes fixed effects for the set of genetic targets studied in the article obtained from PubTator3. Because gene information is available only for a subset of publications, this final specification is estimated on a smaller sample. Standard errors are clustered at the PI level. See text for more details.

Figure C.9: Articles using animal methods (in vivo studies) are more likely to present deceptive data.



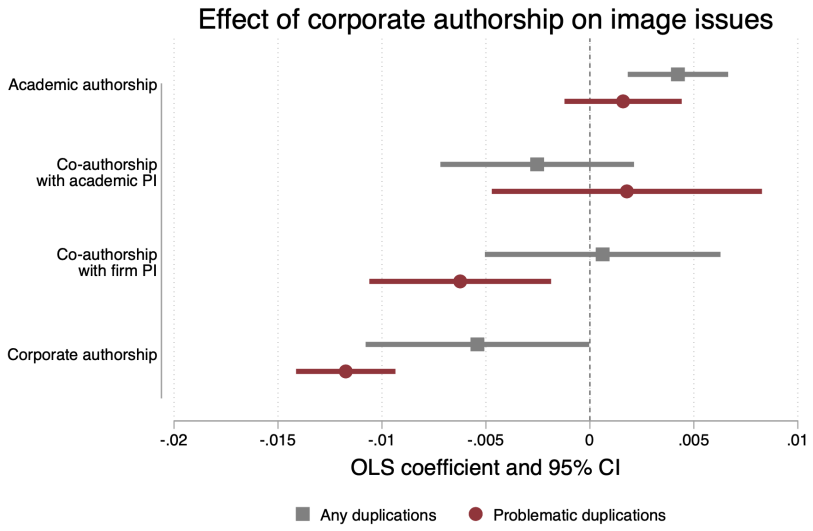
(a) *PubPeer posts and retractions by methods*



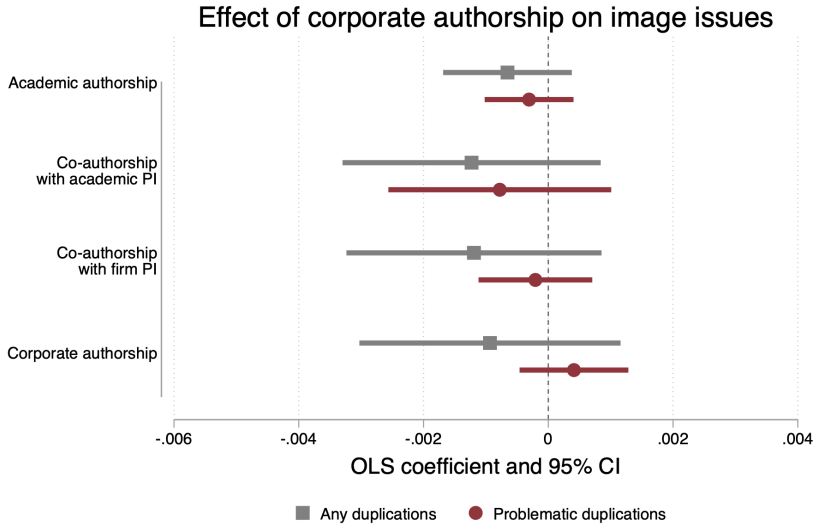
(b) *Image duplications by methods*

Note: The figure reports the association between articles' methods and proxies of research integrity. Each bar is estimated from a separate OLS regression at the paper-level: $Y_i = \alpha + \beta\, \textit{Method Used}_i + \mathbf{Z}_i'\boldsymbol{\delta} + \varepsilon_i$, where Y_i is the outcome of interest for paper i and $\textit{Method Used}_i$ equals one if the paper uses a given experimental method (in vitro or in vivo). $\mathbf{Z}_i$ includes fixed effects for journal, year, and the Approximate Potential to Translate (APT). Bars represent 95% confidence intervals. Standard errors clustered at the PI level. Panel (a) shows that articles using animal models are more likely to be flagged on the PubPeer blog or to be retracted. Panel (b) shows that articles using animal models are more likely to present inappropriate image duplications. See text for more details.

Figure C.10: Placebo: the negative association between corporate authorship and image duplication disappears in clinical studies.



(a) *All articles excluding clinical studies*



(b) *Only clinical studies*

Note: The figure displays the relationship between authorship type and the likelihood of image duplications, reported separately for articles that report clinical study results and all other publications. Each bar is estimated from a separate OLS regression at the paper-level: $Duplication_i = \alpha + \beta\, Affiliation_i + \mathbf{Z}_i'\boldsymbol{\delta} + \varepsilon_i$, where $Duplication_i$ equals one if the article i presents a duplicated image and $Affiliation_i$ captures the institutional affiliation of the authors. $\mathbf{Z}_i$ includes fixed effects for journal, year, and the Approximate Potential to Translate (APT). Bars represent 95% confidence intervals. Standard errors clustered at the PI level. Panel (a) shows the results for all articles in our sample, excluding those reporting the results of a clinical study. Panel (b) shows the results only for articles reporting the results of a clinical study. See text for more details.

Table C.1: Validation of Imagetwin data using external proxies of scientific unreliability.

Dependent Variable	Citations with negative tone (0/1)	PubPeer post (0/1)	Paper retracted (0/1)
	(1)	(2)	(3)
Paper with any duplication (0/1)	0.0268**	0.0735***	0.0156***
	(0.0127)	(0.00998)	(0.00464)
Year FE	YES	YES	YES
Journal FE	YES	YES	YES
APT FE	YES	YES	YES
Observations	46,548	46,548	46,548

Note: *, **, *** denote significance at the 10%, 5%, and 1% levels, respectively. Cross-sectional OLS regressions at the publication level. Standard errors are clustered at the PI level. All models include fixed effects for journal, year, and Approximate Potential to Translate (APT). *Paper with contradicting citation*: 0/1 = 1 if the publication receives at least one contradicting citation in the Scite data. *PubPeer post*: 0/1 = 1 if the publication has at least one discussion thread on PubPeer. *Paper retracted*: 0/1 = 1 if the publication has been retracted. *Paper with any duplication*: 0/1 = 1 if the publication presents at least one duplicated image detected by Imagetwin. See text for details.

Table C.2: Validation of Imagetwin data using the reasons for paper retractions in Retraction Watch.

Dependent Variable	Paper with any duplication (0/1)		
	(1)	(2)	(3)
Paper retracted (0/1)	0.284***		
	(0.0739)		
Paper retracted for issue in images (0/1)		0.620***	
		(0.120)	
Paper retracted for duplication in images (0/1)			0.895***
			(0.0561)
Year FE	YES	YES	YES
Journal FE	YES	YES	YES
APT FE	YES	YES	YES
Observations	46,548	46,548	46,548
Mean of the DV	0.0176	0.0176	0.0176

Note: *, **, *** denote significance at the 10%, 5%, and 1% levels, respectively. Cross-sectional OLS regressions at the publication level. Standard errors are clustered at the PI level. All models include fixed effects for journal, year, and Approximate Potential to Translate (APT). *Paper with any duplication*: 0/1 = 1 if the publication presents at least one duplicated image detected by Imagetwin. *Paper retracted*: 0/1 = 1 if the publication has been retracted in the Retraction Watch data. *Paper retracted for issue in images*: 0/1 = 1 if the publication has been retracted due to image-related issues in the Retraction Watch data. *Paper retracted for duplication in images*: 0/1 = 1 if the publication has been retracted due to image duplications in the Retraction Watch data. See text for details.

Table C.3: Images that are referenced more often in the article's text are more likely to present duplicated images.

Dependent Variable	Number of text mentions		
	(1)	(2)	(3)
Figure with any duplication (0/1)	1.508***	0.701***	0.499***
	(0.132)	(0.104)	(0.0987)
Article FE	NO	YES	YES
Order of figure FE	NO	NO	YES
Observations	164,770	158,932	158,932

Note: *, **, *** denote significance at the 10%, 5%, and 1% levels, respectively. Cross-sectional OLS regressions at the article-image level. Standard errors are clustered at the article level. Column 2 includes article fixed effects, and Column 3 additionally includes fixed effects for the order of appearance of the figure in the article's text (accounting for systematic differences in text mentions across figure position). *Number of text mentions*: count of mentions to a figure in the full text of the article. *Figure with any duplication*: 0/1 = 1 for figures presenting at least one duplicated image detected by Imagetwin. See text for details.

Table C.4: Papers led by PIs with previous image duplications are more likely to present duplications again.

Dependent Variable	Paper with any duplication (0/1)		Paper with problematic duplication (0/1)	
	(1)	(2)	(3)	(4)
PI's any previous duplication (0/1)	0.0452***		0.0230***	
	(0.00812)		(0.00623)	
PI's previous problematic duplication (0/1)		0.0506***		0.0297***
		(0.0108)		(0.00812)
Year FE	YES	YES	YES	YES
Journal FE	YES	YES	YES	YES
APT FE	YES	YES	YES	YES
Observations	27,996	27,996	27,996	27,996

Note: *, **, *** denote significance at the 10%, 5%, and 1% levels, respectively. Publication-level OLS regressions with standard errors clustered at the PI level. All models include fixed effects for publication year, journal, and Approximate Potential to Translate (APT). *Paper with any duplication*: 0/1 = 1 if the publication contains at least one duplicated image detected by Imagetwin and validated through manual review. *Paper with problematic duplication*: 0/1 = 1 if the publication contains at least one duplicated image classified through manual review as problematic. *PI's previous duplication*: 0/1 = 1 if the PI has previously led at least one publication containing an image duplication. *PI's previous problematic duplication*: 0/1 = 1 if the PI has previously led at least one publication containing a problematic image duplication. See text for details.

Table C.5: Problematic duplications are less frequent in fully corporate-authored publications, but not systematically associated with corporate funding or generic corporate co-authorship.

Dependent Variable	Paper with any duplication (0/1)			Paper with mistaken duplication (0/1)			Paper with problematic duplication (0/1)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Corporate funding acknowledged (0/1)	-0.00480*			-0.00297			-0.00189		
	(0.00277)			(0.00201)			(0.00213)		
Any corporate author (0/1)		-0.00272			-0.00180			-0.00145	
		(0.00203)			(0.00135)			(0.00161)	
Corporate authorship (0/1)			-0.0153***			-0.00502**			-0.0108***
			(0.00282)			(0.00248)			(0.00112)
Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Journal FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
APT FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Observations	46,548	46,548	46,548	46,548	46,548	46,548	46,548	46,548	46,548
Mean of the DV	0.0176	0.0176	0.0176	0.00763	0.00763	0.00763	0.0106	0.0106	0.0106
Percentage change	**-27%**	-15%	**-87%**	-39%	-24%	**-66%**	-18%	-14%	**-101%**

Note: *, **, *** denote significance at the 10%, 5%, and 1% levels, respectively. Cross-sectional OLS regressions at the publication level. Standard errors are clustered at the PI level. All models include fixed effects for journal, year, and Approximate Potential to Translate (APT). *Paper with any duplication*: 0/1 = 1 if the publication presents at least one duplicated image. *Paper with mistaken duplication*: 0/1 = 1 if the publication presents at least one duplicated image classified through manual review as potentially reflecting mistakes. *Paper with problematic duplication*: 0/1 = 1 if the publication presents at least one duplicated image containing features suggestive of intentional deception. *Corporate funding acknowledged*: 0/1 = 1 if the publication acknowledges funding from corporations. *Any corporate author*: 0/1 = 1 if at least one author of the publication is affiliated with a company. *Corporate authorship*: 0/1 = 1 for publications authored only by researchers affiliated with companies. Percentage changes are computed relative to the sample mean ($\hat{\beta}/\bar{Y} \times 100$) and rounded to the nearest integer; significant effects are shown in bold.

Table C.6: The lower incidence of problematic image duplications among firms' publications does not vary systematically with firm size.

Dependent Variable	Paper with any duplication (0/1)			Paper with mistaken duplication (0/1)			Paper with problematic duplication (0/1)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Corporate authorship (0/1)	-0.0153***	-0.0137***	-0.0161***	-0.00502**	-0.00538***	-0.00625***	-0.0108***	-0.00852***	-0.0102***
	(0.00282)	(0.00231)	(0.00217)	(0.00248)	(0.00123)	(0.00118)	(0.00112)	(0.00159)	(0.00153)
Corporate authorship (0/1)		0.000299			0.000756			-0.000652	
× Large firm (0/1)		(0.00468)			(0.00368)			(0.00262)	
Corporate authorship (0/1)			0.00515			0.00526			0.000168
× Top 50 pharma firm (0/1)			(0.00550)			(0.00425)			(0.00326)
Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Journal FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
APT FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Observations	46,548	46,548	46,548	46,548	46,548	46,548	46,548	46,548	46,548
Mean of the DV	0.0176	0.0176	0.0176	0.0076	0.0076	0.0076	0.0106	0.0106	0.0106

Note: *, **, *** denote significance at the 10%, 5%, and 1% levels, respectively. Cross-sectional OLS regressions at the publication level. Standard errors are clustered at the PI level. All models include fixed effects for journal, year, and Approximate Potential to Translate (APT). *Paper with any duplication*: 0/1 = 1 if the publication presents at least one duplicated image. *Paper with mistaken duplication*: 0/1 = 1 the publication presents at least one duplicated image classified through manual review as potentially reflecting mistakes. *Paper with problematic duplication*: 0/1 = 1 if the publication presents at least one duplicated image containing features suggestive of intentional deception. *Corporate authorship*: 0/1 = 1 for publications authored only by researchers affiliated with companies. *Large firm (0/1)*: 0/1 = 1 if the firm has more than 500 employees as of 2024. *Top 50 pharmaceutical firm (0/1)*: 0/1 = 1 if the firm is among the 50 largest pharmaceutical firms by revenue as of 2024.

Table C.7: For principal investigators observed at both universities and firms, firm-affiliated publications present fewer problematic image duplications.

Dependent Variable	Paper with any duplication (0/1) (1)	Paper with mistaken duplication (0/1) (2)	Paper with problematic duplication (0/1) (3)
PI's affiliation is a company (0/1)	-0.00724 (0.00577)	-0.00150 (0.00474)	-0.00616* (0.00321)
Year FE	YES	YES	YES
Journal FE	YES	YES	YES
APT FE	YES	YES	YES
Principal Investigator FE	YES	YES	YES
Observations	17,034	17,034	17,034
Mean of the DV	0.0216	0.00933	0.0132
Percentage change	-34%	-16%	**-47%**

Note: *, **, *** denote significance at the 10%, 5%, and 1% levels, respectively. Cross-sectional OLS regressions at the publication level. The sample is restricted to publications for which the PI's last affiliation is either a university or a company. Standard errors are clustered at the PI level. All models include fixed effects for journal, year, Approximate Potential to Translate (APT), and PI. The PI fixed effects absorb time-invariant differences across PIs, so the coefficient compares publications by the same PI when the PI's last affiliation is a company rather than a university. *Paper with any duplication*: 0/1 = 1 if the publication presents at least one duplicated image. *Paper with mistaken duplication*: 0/1 = 1 if the publication presents at least one duplicated image classified through manual review as potentially reflecting mistakes. *Paper with problematic duplication*: 0/1 = 1 if the publication presents at least one duplicated image containing features suggestive of intentional deception. *PI's affiliation is a company*: 0/1 = 1 if the PI's last listed affiliation on the publication is a company; the omitted category is publications where the PI's last listed affiliation is a university. Percentage changes are computed relative to the sample mean ($\hat{\beta}/\bar{Y} \times 100$) and rounded to the nearest integer; significant effects are shown in bold.

Table C.8: Within universities, publications led by PIs affiliated with medical or healthcare units exhibit fewer mistaken duplications.

Dependent Variable	Paper with any duplication (0/1) (1)	Paper with mistaken duplication (0/1) (2)	Paper with problematic duplication (0/1) (3)
PI affiliated with medical or healthcare unit (0/1)	0.00167	-0.00595**	0.00297
	(0.00440)	(0.00243)	(0.00262)
Year FE	YES	YES	YES
Journal FE	YES	YES	YES
APT FE	YES	YES	YES
University FE	YES	YES	YES
Observations	39,633	39,633	39,633

Note: *, **, *** denote significance at the 10%, 5%, and 1% levels, respectively. Publication-level OLS regressions with standard errors clustered at the PI level. All models include fixed effects for publication year, journal, Approximate Potential to Translate (APT), and parent university. University fixed effects absorb time-invariant differences across institutions, so the coefficient compares publications led by PIs on the medical or healthcare side of a university with publications led by PIs on the nonmedical academic side of the same university. *Paper with any duplication*: 0/1 = 1 if the publication contains at least one duplicated image. *Paper with mistaken duplication*: 0/1 = 1 if the publication contains at least one duplicated image classified through manual review as potentially reflecting a mistake. *Paper with problematic duplication*: 0/1 = 1 if the publication contains at least one duplicated image with features suggestive of intentional deception. *PI affiliated with medical or healthcare unit*: 0/1 = 1 if the PI's last listed affiliation belongs to the medical or healthcare side of the parent university.

Table C.9: Post hoc correlates of image duplications among academic-only publications.

Dependent Variable	Paper with any duplication (0/1)				
	(1)	(2)	(3)	(4)	(5)
Linked R01 grants	0.000411				
	(0.00104)				
PI papers supervised in year		0.00217			
		(0.00209)			
Number of authors			0.000206		
			(0.000256)		
Young PI				-0.00187	
				(0.00461)	
PI experience					-0.0000468
					(0.000286)
Year FE	YES	YES	YES	YES	YES
Journal FE	YES	YES	YES	YES	YES
APT FE	YES	YES	YES	YES	YES
Observations	12,844	12,844	12,844	12,844	12,844

Note: *, **, *** denote significance at the 10%, 5%, and 1% levels, respectively. Publication-level OLS regressions with standard errors clustered at the PI level. The sample is restricted to publications authored exclusively by university-affiliated researchers. Each column reports a separate regression. All models include fixed effects for publication year, journal, and Approximate Potential to Translate (APT). *Paper with any duplication*: 0/1 = 1 if the publication contains at least one manually validated image duplication. *Linked R01 grants*: number of distinct NIH R01 grants linked to the PI in the publication year. *PI papers supervised in year*: number of publications for which the PI is listed as last author in the publication year. *Number of authors*: number of authors listed on the publication. *Young PI*: 0/1 = 1 if the PI is within six years of their first observed last-author Alzheimer's disease publication. *PI experience*: number of years since the PI's first observed last-author Alzheimer's disease publication.

Table C.10: The negative association between corporate authorship and problematic image duplication is stronger for research with high commercial potential.

Dependent Variable	Paper with any duplication (0/1)		Paper with mistaken duplication (0/1)		Paper with problematic duplication (0/1)	
	(1)	(2)	(3)	(4)	(5)	(6)
Corporate authorship (0/1)	-0.0153***	-0.0124***	-0.00502**	-0.00605***	-0.0108***	-0.00665***
	(0.00282)	(0.00198)	(0.00248)	(0.00129)	(0.00112)	(0.00115)
Corporate authorship (0/1)		-0.00462		0.00115		-0.00615***
× High commercial potential (0/1)		(0.00438)		(0.00375)		(0.00182)
Year FE	YES	YES	YES	YES	YES	YES
Journal FE	YES	YES	YES	YES	YES	YES
APT FE	YES	YES	YES	YES	YES	YES
Observations	46,548	46,548	46,548	46,548	46,548	46,548
Mean of the DV	0.0176	0.0176	0.0076	0.0076	0.0106	0.0106
Percentage change		-26%		+15%		**-58%**

Note: *, **, *** denote significance at the 10%, 5%, and 1% levels, respectively. Cross-sectional OLS regressions at the publication level. Standard errors clustered at the PI level. All models include fixed effects for journal, year, and Approximate Potential to Translate (APT). *Paper with any duplication*: 0/1 = 1 if the publication presents at least one duplicated image. *Paper with mistaken duplication*: 0/1 = 1 if the publication presents at least one duplicated image classified through manual review as potentially reflecting mistakes. *Paper with problematic duplication*: 0/1 = 1 if the publication presents at least one duplicated image containing features suggestive of intentional deception. *Corporate authorship*: 0/1 = 1 for publications authored only by researchers affiliated with companies. *High commercial potential*: 0/1 = 1 for publications with an above-median value of the commercial potential score by Masclans et al. (2025). Percentage changes for the interaction terms are computed relative to the mean of the corresponding dependent variable ($\hat{\beta}_{interaction}/\bar{Y} \times 100$) and rounded to the nearest integer; significant interaction effects are shown in bold.

Table C.11: The negative association between corporate authorship and problematic image duplication weakens when papers study drugs already commercialized by the authoring firm.

Dependent Variable	Paper with any duplication (0/1)		Paper with mistaken duplication (0/1)		Paper with problematic duplication (0/1)	
	(1)	(2)	(3)	(4)	(5)	(6)
Corporate authorship (0/1)	-0.0153***	-0.0156***	-0.00502**	-0.00523**	-0.0108***	-0.0107***
	(0.00282)	(0.00292)	(0.00248)	(0.00258)	(0.00112)	(0.00113)
Corporate authorship (0/1)		0.0118**		0.00684**		0.00658*
× Paper on commercialized drug (0/1)		(0.00589)		(0.00322)		(0.00386)
Year FE	YES	YES	YES	YES	YES	YES
Journal FE	YES	YES	YES	YES	YES	YES
APT FE	YES	YES	YES	YES	YES	YES
Observations	46,548	46,548	46,548	46,548	46,548	46,548
Mean of the DV	0.0176	0.0176	0.0076	0.0076	0.0106	0.0106
Percentage change		**+67%**		**+90%**		**+62%**

Note: *, **, *** denote significance at the 10%, 5%, and 1% levels, respectively. Cross-sectional OLS regressions at the publication level. Standard errors clustered at the PI level. All models include fixed effects for journal, year, and Approximate Potential to Translate (APT). *Paper with any duplication*: 0/1 = 1 if the publication presents at least one duplicated image. *Paper with mistaken duplication*: 0/1 = 1 if the publication presents at least one duplicated image classified through manual review as potentially reflecting mistakes. *Paper with problematic duplication*: 0/1 = 1 if the publication presents at least one duplicated image containing features suggestive of intentional deception. *Corporate authorship*: 0/1 = 1 for publications authored only by researchers affiliated with companies. *Paper on commercialized drug*: 0/1 = 1 if the paper studies a drug already commercialized by the affiliated firm. Percentage changes for the interaction terms are computed relative to the mean of the corresponding dependent variable ($\hat{\beta}_{interaction}/\bar{Y} \times 100$) and rounded to the nearest integer; significant interaction effects are shown in bold.