From Free Rider to Innovator:

The Rise of China's Drug Development

Panle Jia Barwick Hongyuan Xia Tianli Xia*

January 2026

Abstract

This paper examines China's transition from pharmaceutical "free rider" to global innovator over the last decade. In 2010, China accounted for less than 8% of global clinical trials; by 2020, it had surpassed the US in annual registered clinical trial volume. To study this transformation, we compile a comprehensive, synchronized database spanning the pharmaceutical drug development supply chain, covering scientific publications, clinical trials, drug development milestones for China, the U.S., and Europe, alongside drug sales and government policies over the same period. We provide strong evidence that China's rise was primarily driven by the National Reimbursement Drug List (NRDL) reform, which dramatically expanded the effective market size for innovative drugs. We document a sharp rise in both the quantity (86% increase) and novelty of drug trials post reform, with growth concentrated in reform-exposed disease categories, first- or best-in-class drugs, and among domestic firms. A decomposition exercise reveals that the NRDL reform accounts for 43% of the growth in oncology trial activity, nearly doubling the combined contribution of upstream knowledge accumulation and talent flows (24%), while other government policies play a minor role. Finally, dynamic gains from induced innovation exceed the reform's static gains in consumer access to innovative drugs by threefold, underscoring the importance of accounting for the reform's long-run effects on innovation incentives in addition to near-term improvements in drug affordability.

Keywords: Clinical Trials, Pharmaceutical innovation, NRDL, China

JEL Classification: O38, I18, O31, L65

---

*Barwick: Department of Economics, UW-Madison, NBER, and CEPR, pbarwick@wisc.edu. H. Xia: Department of Economics, Cornell University, hx276@cornell.edu. T. Xia: Simon Business School, University of Rochester, tianli.xia@rochester.edu. We are grateful to SinoHealth for generously providing the data. We thank Shizhe Yu and Congrong Pan for providing stellar research support and Leila Agha, Pierre Azoulay, Travis Baseler, Jack Collison, Rebekah Dix, Florian Ederer (discussant), Yufeng Huang, Charles Hodgson, Hanming Fang, Russell Morton, Ellie Prager, Chris Snyder, Fiona Scott Morton, and seminar participants for helpful comments. Barwick acknowledges the support from UW-Madison's VCGRR with funding from the Wisconsin Alumni Research Foundation.

1 Introduction

Since World War II, the global landscape of the pharmaceutical industry has been defined by a clear hierarchy. The US serves as the undisputed leader in upstream R&D, generating the vast majority of novel drug concepts and first-in-class therapies, with Europe and Japan serving as a strong second pillar. The rest of the world, by contrast, plays a peripheral role, functioning primarily as consumers of existing medicines or producers of generic copies rather than creators of novel technologies.

This hierarchy aligns with the conventional wisdom among health economists: smaller or low-income nations optimally rely on the diffusion of innovation from wealthy nations rather than investing in indigenous creation (Olson and Zeckhauser, 1966; Kyle, 2007; Dubois et al., 2015; Lakdawalla, 2018; Garthwaite, 2025). The logic is straightforward. Pharmaceutical R&D entails high uncertainty and enormous fixed costs, with only 10% of projects ultimately receiving drug approval and a median Phase III trial costing $21.4 million (Martin et al., 2017; Aryal et al., 2024). For most countries where market size is limited, budgets are constrained, and the technological frontier is distant, catching up to the global innovation frontier seems out of reach.

This paper challenges this conventional wisdom and asks two questions. First, can a country move from a downstream consumer to an upstream innovator in drug development? Second, what forces or policy designs contribute to the emergence of a local innovation ecosystem?

To answer the first research question, we draw on the Citeline TrialTrove database, a comprehensive census of global drug development projects that allows us to track the life cycle of clinical R&D with granular detail across therapeutic areas. We document a striking structural shift in the global supply of pharmaceutical innovation: over the past decade, China has grown from a marginal player, accounting for less than 8% of global clinical trials in 2010, to surpassing the U.S. and EU in total annual registered clinical trials by 2020.

To formally evaluate China's catch-up in innovation, in addition to showing the raw data, we employ an event-study framework that benchmarks China's trajectory against the U.S., the global frontier in biomedical research. We use Europe as an alternative counterfactual to ensure robustness. All analyses are conducted at the disease-country-year level, and include disease-by-country fixed effects to absorb time-invariant heterogeneity in country- and disease-specific demand and comparative advantage, along with year (and disease-year) fixed effects to account for global temporal shocks.

The event study confirms a sharp structural break in China's clinical trial activity beginning in 2016, preceded by a flat pre-trend. By 2024, China's annual trial volume had expanded by 172% relative to the US.

China's catch-up has three remarkable features worth highlighting. First, this expansion is not driven by low-value "me-too" testing. By 2024, China's high-novelty trials had increased by 49-123% relative to the U.S. across three novelty measures we construct: (1) Dranove et al. (2022)'s novelty score, (2) trials on new molecules (non-generics), and (3) trials using patented drugs rather than placebos or off-patent drugs as a control arm. Second, the surge in domestic innovation extends well beyond trial initiation. Although many trials have not yet reached completion, we document a two-fold increase in out-licensing transactions to multinational corporations (MNCs), a 36% rise in first-in-world regulatory approvals for China-developed drugs, and a 27% increase in firm entry. Finally, this growth reflects an unprecedented rise in indigenous innovative activities rather than reallocation of the global activities from the US/EU to China: domestic Chinese firms account for 88% of the post-2015 increase in trials, while multinational pipelines remain relatively flat. These results are robust to alternative specifications and sample choices. Together, these patterns mark a qualitative transformation of China's role in global pharmaceutical R&D: from a peripheral adopter of existing therapies to a major source of frontier innovation.

To answer the second research question and to causally examine the underlying drivers of this transformation, we integrate the trial registries with complementary data sources at the country-disease-year level, including (a) global scholarly publications and author affiliations from OpenAlex; (b) disease prevalence rates from the Global Burden of Disease database; (c) national annual drug sales from SinoHealth; (d) central policy documents issued by China's State Council; and (e) trial application backlogs from China's National Medical Products Administration. To our knowledge, this assembled dataset provides the most comprehensive view of the pharmaceutical innovation supply chain in the existing literature.

We argue that China's transformation was not an incidental byproduct of aggregate economic growth, but the consequence of a specific policy shock: the 2016 National Reimbursement Drug List (NRDL) reform. By combining centralized price negotiations with a large-scale expansion in insurance coverage, the reform fundamentally altered the effective market size facing innovative pharmaceutical firms. To examine this mechanism, our empirical strategy exploits variation in the reform's treatment

intensity across disease categories. We first show that the NRDL's quantity expansion dominates its price contraction, generating a revenue increase of more than 400% for covered drugs. Therapeutic areas with greater exposure to the NRDL experienced disproportionately larger increases in clinical trial activity. Crucially, a placebo test confirms that disease categories never exposed to NRDL expansion show *no systematic increase* in trial activity after 2015. Domestic Chinese firms also respond far more strongly to the reform than their multinational counterparts, consistent with lower institutional frictions in navigating local regulations. Finally, a back-of-the-envelope welfare calculation suggests that by catalyzing the development of a new generation of therapies, the NRDL's long-run dynamic gains in social surplus are substantial: conservatively estimated at three times the static gains from improved patient access to innovative drugs alone.

Beyond the NRDL, we conduct a comprehensive analysis of alternative drivers operating through demand- and supply-side channels. On the demand side, we examine regulatory harmonization through China's accession to the ICH (International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use), which aligned China's regulatory framework with international standards and may have lowered barriers to global markets for drugs developed in China. On the supply side, we evaluate upstream knowledge accumulation and international talent flows, specifically the return migration of scientists from the U.S. to China, which provides essential inputs for pharmaceutical innovation. We also consider sector-specific regulatory reforms and industrial policies, most notably the 2015 clinical trial application reform, which reduced approval backlogs, and Made in China 2025, which increased public investment in biomedicine. A decomposition exercise reveals that the NRDL reform accounts for 43% of the observed surge in oncology trials, while upstream knowledge creation and talent inflows contribute 24%. Other policies and regulatory changes play a comparatively modest role, confirming that the NRDL was the dominant catalyst of China's pharmaceutical innovation boom.

China's experience offers a valuable case study. It demonstrates that through deliberate institutional design, specifically leveraging state monopsony power to expand market access while preserving innovation incentives, late-mover economies can move beyond the passive consumption of frontier medicines to become active contributors to the global stock of biomedical knowledge.

Related Literature Our analysis contributes to several strands of literature. First, it relates to the literature on market size and innovation. A large body of work based on U.S. data has established that expected revenue is a key determinant of pharmaceutical R&D, providing empirical support for Schmookler (1966)'s hypothesis. Acemoglu et al. (2006) demonstrates that exogenous market expansions driven by demographic shifts stimulate drug innovation, with subsequent contributions by Yin (2008); Blume-Kohout and Sood (2013); Dubois et al. (2015); Ji and Rogers (2024).

We provide the first evidence from a developing country / emerging economy context. Several features make NRDL a distinctive setting for studying how demand-side policy shapes innovation. First, unlike the broad insurance expansions commonly studied in the literature, such as Medicare Part D, NRDL inclusion is narrowly targeted at innovative drugs and explicitly favors those with high clinical effectiveness. Second, the reform was implemented in an economy that had historically been a consumer rather than a producer of frontier pharmaceutical innovation, offering a rare opportunity to study whether demand-pull incentives can catalyze indigenous R&D in a market with limited pre-existing capacity. Our results indicate that by expanding effective market size while containing drug prices via government monopsony power, the NRDL offers rare evidence that a well-designed insurance policy can help reconcile the tension between affordability and innovation incentives.

Consistent with studies on firms' strategic responses to market-based incentives (e.g., Finkelstein (2004), Budish et al. (2015), Hermosilla (2021), Agha et al. (2022), Cunningham et al. (2025), and Bignon (2026), Liu et al. (2026)), we find that firms reorient clinical trial efforts toward therapeutic areas with higher expected profitability. However, while the existing literature tends to find that demand-side shocks primarily pull incremental innovation (Dranove et al., 2022; Geng and Shi, 2024), the trials induced by the NRDL reform are characterized by high scientific novelty. This likely reflects the NRDL's inclusion criteria, which explicitly reward novelty/originality and prioritize unmet clinical needs.

Second, our paper contributes to the literature on industrial policy and innovation. Much of the existing research focuses on direct supply-side interventions such as government subsidies and tax credits (Branstetter and Sakakibara, 2002; Bronzini and Piselli, 2016; Howell, 2017; Dechezleprêtre et al., 2023; Melnik and Smyth, 2024). In the pharmaceutical sector specifically, Ward and Dranove (1995), Azoulay et al. (2019), and Gross and Sampat (2025) show that public funding (such as U.S. Committee on Medical Research contracts and NIH grants) can increase private-sector patenting and

R&D investment, though Wallsten (2000) cautions that such grants may crowd out firm-financed research. A notable exception is Finkelstein (2004), which finds that government vaccine recommendations (a demand-side intervention) have dynamic effects that stimulate R&D. We extend this literature in two directions: by providing evidence from an emerging economy context, and by comprehensively analyzing both demand- and supply-side drivers within a unified framework rather than examining either channel in isolation.

Third, our study contributes to a growing literature documenting China's rapid technological ascent, while speaking directly to the debates about the quality of this progress.[1] Chen et al. (2021) cautions that reported R&D growth may partly reflect expense relabeling in response to tax incentives; Wei et al. (2023) finds negative average returns and a proliferation of low-quality patents from China's flagship innovation policy (InnoCom); and Li and Branstetter (2024) argues that policy-induced increase in R&D-sales ratio does not translate to patenting or productivity gains. On the other hand, Lerner et al. (2026) documents a dramatic rise in high-potential entrepreneurship and venture capital in China that contributes to innovation. We contribute to this debate by providing the most comprehensive evidence to date on China's biopharmaceutical innovation. Leveraging granular novelty indices and decomposing the sources of growth, we show that the surge in clinical development reflects a genuine reorientation toward the global quality frontier rather than a mere scaling-up of low-value activity.

The remainder of the paper is organized as follows. Section 2 details the institutional background of the Chinese healthcare system and describes the data. Section 3 analyzes both the quantity and quality of clinical trials in China relative to the U.S., as well as downstream business development and drug licensing activities. Section 4 investigates the underlying mechanisms. Section 5 quantifies the long-run vs. short-run welfare implications. Section 6 concludes.

[1] For recent debates in the public media, see NYT (Aug 17, 2025), Washington Post (Sep 22, 2025), and Bloomberg (June 10, 2025).

2 Institutional Background and Data

2.1 Drug Development Process

The regulatory process governing drug development and market authorization in China and the U.S. is similar. In both jurisdictions, a drug sponsor must demonstrate that a product is safe and effective through a series of randomized clinical trials, typically benchmarking the investigational drug against a placebo or the prevailing standard of care (National Institutes of Health, 2017).

The drug development lifecycle begins with the identification of a lead compound. Sponsors generally file patent applications before the commencement of human testing to protect the resulting intellectual property. Before clinical trials begin, sponsors must submit an Investigational New Drug (IND) application, which provides comprehensive pre-clinical evidence on the compound's toxicity and pharmacological profile.

Clinical evaluation proceeds through three main stages before sponsors may apply for market approval. Phase I trials assess basic safety, dosage ranges, and potential side effects, typically enrolling a small group of healthy volunteers. Phase II and Phase III evaluate efficacy, continue to monitor safety, and benchmark the drug against existing therapies. Phase III studies, often involving hundreds of subjects and conducted in a blinded and randomized setting, serve as the "pivotal" trials required for regulatory approval. After approval, sponsors (and other public organizations) may conduct Phase IV studies, which are usually observational and focus on longer-term safety and real-world effectiveness. Clinical trials represent the most capital-intensive part of the R&D pipeline. On average, Phase I trials cost approximately $3.4 million, and Phase III trials cost roughly $21.4 million (Martin et al., 2017). In addition, trials involve high risks of failure (Aryal et al., 2024). In our data, a new drug discovery is associated with an average of 123 clinical trials, while only about 4 percent of Investigational New Drug (IND) projects ultimately result in a successful drug.

2.2 The Chinese Healthcare System

The Chinese healthcare system (CHS) provides nearly universal coverage to over 95% of the population, making it the world's largest public health insurance program (Yu, 2015). Established in 1999, it has since evolved into two main schemes: the urban employee basic medical insurance (UEBMI), which covers formal-sector employees and retirees, and the urban and rural resident basic medical

insurance (URRBMI), which covers the remaining population, including children and the elderly.

As the dominant purchaser in the domestic market, the CHS can exert significant monopsony power. In 2022, the system spent over ¥520 billion ($74 billion) on pharmaceuticals (National Health Commission of the PRC, 2023). In contrast, the private commercial insurance market remains underdeveloped, accounting for less than 7% of total health expenditures. Consequently, inclusion in the National Reimbursement Drug List (NRDL) is the primary gateway to the Chinese pharmaceutical market. Historically, the NRDL prioritized low-cost generics and off-patent molecules due to budget considerations, effectively excluding innovative, high-priced therapies and forcing patients to bear the full cost out-of-pocket when purchasing in the private market.[2]

The NRDL Reform In response to growing public demand for broader access to innovative drugs, the National Healthcare Security Administration (NHSA) launched a landmark reform in 2016, integrating innovative therapies into the national insurance catalog through centralized price negotiations. This ongoing reform represents a fundamental shift from a rigid, budget-capped formulary to a dynamic, value-based reimbursement model.

The negotiation process follows an annual cycle. Each spring, NHSA defines an "eligible list" of drugs, targeting innovative drugs (obtaining market approval within the last five years) that address significant unmet clinical needs.[3] Firms with eligible drugs and NHSA exchange information on costs and medical effectiveness during the pre-negotiation stage, which can last for months. On the day of negotiation, government representatives negotiate with delegates from each drug company simultaneously and independently. Successful negotiations result in inclusion in the NRDL, typically involving average price reductions of 50–60% (Barwick et al., 2025). If negotiations fail, the drug remains on the private market with significantly lower volume. The reform has a huge impact on the drug market and has successfully included 699 drugs in NRDL by 2024, and the accrued spending on innovative drugs has reached ¥460 billion ($67 billion) between 2016 and 2024 (Table 1), dwarfing the ¥66 billion ($ 10 billion) innovative drug market in 2015.[4]

The National Reimbursement Drug List (NRDL) reform possesses three features central to our analyses. First, the uneven and staggered reimbursement coverage across drug classes provides rich

[2]For the 2017 NRDL list, see: NRDL (2017 version).

[3]See: the 2024 Inclusion Criteria.

[4]Source: 2025 National Forum on Medical Insurance & Innovative Drug Development.

temporal and cross-sectional variation. According to the government document, the eligibility criteria for NRDL inclusion strongly favor innovative drugs: more than 80% of negotiated drugs qualified as "new molecule entities approved within the past five years."[5] The final inclusion of drugs often favors popular and effective ones that are first or best in class (Barwick et al., 2025). This priority towards innovative categories with unmet clinical needs can reshape the directions of firm R&D.

Second, the scale of the commitment is massive. The reform represents one of the largest public expansions of insurance coverage for innovative drugs globally, allocating over ¥100 billion to negotiated drugs in 2024 alone.[6] Anecdotal evidence and government town hall meetings indicate that the Chinese government is aware of NRDL's potential (advantageous) effect on innovation.[7] Industry leaders also show enthusiasm towards the reform in news reports. Chengdu Kanghong Pharmaceutical Group, the developer of conbercept (Lumitin®), the first Chinese innovator biologic to receive WHO international nonproprietary name recognition, commented in 2017 after conbercept's inclusion in the NRDL: "The inclusion of innovative drugs in the national health insurance system is of great significance – whether in improving people's well-being, promoting industrial innovation, or supporting pharmaceutical companies' own development."[8] On the contrary, before the reform, Beigene, a leading pharmaceutical firm in China, expressed concerns that the absence of insurance coverage was detrimental to the innovative drug development: "No matter how good a new drug is, it has no opportunity to be included in NRDL. This has been detrimental to public welfare, industrial development, and incentives for innovation."[9]

Third, the policy represents a sophisticated exercise of government monopsony power. Rather than relying on a laissez-faire approach or direct R&D subsidies, the Chinese government utilizes "price-for-volume" negotiations. This institutional design seeks to resolve a classic market failure:

[5]Source: Public Announcement of the Eligible List of Drugs Submitted for Formal Review in the 2020 National Reimbursement Drug List (NRDL). Although other pathways to NRDL inclusion exist, such as "orphan drug eligibility" and "children's drug eligibility," few drugs enter through these criteria.

[6]Total sales of innovative drugs in China are ¥212 billion in 2024, based on authors' calculation using SinoHealth data.

[7]Wang Guodong, Deputy Director of the Medical Insurance Center at NHSA, stated at the "Press Conference on NHSA Measures to Support the High-Quality Development of Innovative Drugs" in July 2025: "In 2024 alone, fund spending on negotiated drugs during the agreement period already exceeded ¥100 billion, demonstrating that the healthcare system is providing real, tangible financial support for innovation." Source: https://www.nhsa.gov.cn/art/2025/7/1/art_14_17065.html.

[8]Source: https://www.mrjjxw.com/articles/2017-08-16/1139072.html.

[9]Source: http://finance.ce.cn/rolling/201603/30/t20160330_9940081.shtml.

the trade-off between static efficiency (patient access through lower prices) and dynamic efficiency (innovation incentives through firm profits). By leveraging its buyer power to guarantee volume in exchange for price concessions, the state aims to find a "sweet spot" that expands access while preserving, or even enhancing, the net present value (NPV) of innovative drug projects.

From an economics perspective, the reform introduces a price-quantity trade-off. On one hand, firms face deep, centralized price cuts that compress profit margins. On the other hand, successful inclusion triggers a massive expansion in effective market size through nearly universal insurance coverage and low coinsurance rates. Whether this reform encourages or deters innovation depends on whether the quantity effect outweighs the price compression. In Section 4.1, we exploit the staggered entry of therapeutic categories into the negotiation rounds to identify the impact of this policy shock on the direction and quality of R&D.

2.3 Data

Our empirical analyses integrate four primary data sources with several auxiliary datasets to track the lifecycle of pharmaceutical innovation and to measure a broad set of demand- and supply-side factors that could drive innovation activity. To our knowledge, this is the first empirical paper that conducts a comprehensive evaluation of both demand- and supply-side policies and regulatory changes in the context of pharmaceutical innovation.

Citeline TrialTrove Database We utilize the Citeline TrialTrove database (2010–2024) to construct a comprehensive census of clinical trials initiated globally. TrialTrove aggregates data from official national registries (e.g., ClinicalTrials.gov in the U.S. and the ChinCTR in China) alongside public disclosures. For each trial, we observe the tested drug, sponsor (private firm vs. public institution), therapeutic area, disease/indication, trial start date, completion status, trial phase, objective, methodology, duration, and trial population. Table 2 reports the number of trials across nine therapeutic areas and phases in China and the United States.

Citeline Pharmaprojects Database To capture the commercial realization of R&D, we link trials to the Citeline Pharmaprojects database, which tracks key development milestones including patenting, out-licensing, and regulatory approvals (Cunningham et al., 2021). These events serve

as proxies for the "quality" of trials, as licensing deals, particularly those acquired by multinational firms, reflect market recognition of the underlying innovation.

SinoHealth Drug Sales Data We obtain granular pharmaceutical sales data from SinoHealth (2016–2024), a leading health-data provider in China (Xia, 2024). Unlike other databases, SinoHealth provides superior coverage of both urban and rural hospital channels, through which nearly all innovative drugs in China are dispensed. The data are at the province-quarter-SKU level and report quantities dispensed and monetary sales at the SKU (stock keeping unit) level.

OpenAlex Publications To measure upstream knowledge creation, we utilize OpenAlex, an open, large-scale scholarly database that systematically tracks global scientific publications and the related metadata. It provides comprehensive information about authors, citations, institutions, journals, and concepts (i.e., keywords). We focus on biomedical publications with valid PubMed IDs, where the first and last authors are affiliated with either the Chinese or U.S. institutions.

Auxiliary Dataset: IND Application and Approval Dates The 2015 clinical trial application reform in China streamlined regulatory procedures by clearing the backlog of pending IND (investigational new drug) applications and is therefore likely to have boosted clinical trial activity. To measure the backlog days between application and approval, we collect the application and approval dates for all IND applications from the Center for Drug Evaluation of China's NMPA. We further assign each IND to a disease so that we can calculate backlog days for each disease category in our data. The data construction process is documented in Appendix D.2.

Auxiliary Dataset: Disease Prevalence Data We additionally collect disease prevalence data from the Global Burden of Disease (GBD) database.[10] This dataset provides standardized disease-level prevalence measures across countries and years.

Auxiliary Dataset: Policy Documents We have scraped all biopharma-related policies in China from the State Council Policy Document Library. We identify 830 policy documents targeting the

[10] Available from https://vizhub.healthdata.org/gbd-results/.

biopharmaceutical sector. We use GPT to classify these documents into four categories: (a) policies that promote advanced manufacturing; (b) innovation policies that reduce reliance on foreign technologies; (c) capital market reform; and (d) fiscal subsidies. See Appendix D.3 for details.

Data Contributions A significant contribution of our data work is linking these disparate sources systematically. Appendix E.1 details the procedures undertaken to link drug names in the SinoHealth sales data with the disease categories in the Citeline databases, including GPT-based semantic mapping and manual verification. Appendix E.2 explains the graph embedding approach we use to link keywords in OpenAlex scientific publications with disease categories in Citeline. By representing both keywords and disease categories as points in a high-dimensional vector space, we systematically identify the OpenAlex keywords most closely related to each Citeline disease category.

Another significant contribution of our data work is the development of three distinct metrics to quantify clinical trial novelty, addressing a central challenge in the economics of innovation: distinguishing between frontier R&D and incremental ("me-too") development. As explained in detail below, these quality metrics identify novel trials that target previously unexploited biological pathways, trials that are based on New Molecular Entities (NMEs), and trials that use innovative drugs as control arms rather than placebos or generics. These measures allow us to more precisely track China's movement toward the global technological frontier. Appendix C provides details about how these novelty measures are constructed.

3 Empirical Results

3.1 Descriptive Evidence of the Time Trend of Trials and Drugs

We begin our empirical analyses by documenting the aggregate trends in clinical trial initiations in China, the U.S., and Europe. Beyond simple counts, we investigate whether this growth represents a move toward the global technological frontier or merely a quantitative expansion.

R&D Volume and Novelty Figure 1 illustrates the time trends in the number of clinical trials launched in China, the U.S., and Europe from 2010 to 2024. Panel (a) shows that while trial activity in the U.S. remained relatively stable and Europe began to decline after 2021, China underwent a

dramatic structural shift. Starting from a baseline of fewer than 1,000 annual trials in 2010, China's volume surged after 2015, ultimately surpassing the U.S. and Europe in 2020. By 2024, China was initiating more than 5,000 clinical trials annually.

A volume-based comparison may be confounded by routine or low-value R&D. To address this, Panels (b) to (d) utilize three complementary measures to evaluate trial novelty. First, following Dranove et al. (2022), we construct a novelty index that measures how frequently a drug's mechanism of action (i.e., its molecular target) has been explored in prior research. Trials scoring above the global median of this index are classified as "high-novelty" projects.[11]

Second, we utilize a fine-tuned GPT-4o-mini model to perform a functional classification of trial objectives based on textual registry data. This classification distinguishes between (i) new molecular entities (NMEs), (ii) novel drug combinations, (iii) new indications for existing molecules, and (iv) bioequivalence tests for generic drugs.[12]

Third, for Phase III trials, novelty is related to the type of control used in the trial design. Phase III trials are typically conducted with both a treatment group and a control group, where the control may consist of an innovative drug, a generic drug, or a placebo. Trials that use an innovative drug as the control are positioned against the existing clinical frontier and therefore reflect a higher degree of novelty relative to trials compared against generics or a placebo. This multi-measurement approach ensures that our analyses are benchmarked against the global clinical development frontier.

The results in Panel (b) of Figure 1 demonstrate that China's expansion is not limited to incremental projects. The volume of high-novelty trials rose from around 500 in 2015 to over 2,400 by 2021. Furthermore, Panel (c) confirms that the surge is primarily driven by trials investigating novel mechanisms, new combinations, or new indications, while generic bioequivalence testing accounts for a diminishing share of the total growth. Panel (d) shows a substantial increase in Phase III trials that use an innovative drug as the control, indicating that a growing number of late-stage trials in

[11] To be more specific, Dranove et al. (2022) compute novelty of a trial i with mechanism $m(i)$ at stage $j(i)$ as:

$$Novelty_i = \frac{1}{1 + \text{Number of prior trials at stage } j(i) \text{ using mechanism } m(i)}.$$

This measure distinguishes drugs with mechanisms rarely explored before from those relying on well-established pathways. We calculate the novelty median within each year, pooling observations across all countries to benchmark novelty against the global clinical development frontier rather than country-specific baselines.

[12] We manually labeled a training set of trials and used the model to classify the remaining trials. Fine-tuning reduces hallucination risk and enables consistent large-scale categorization of trial novelty.

China are explicitly benchmarked against the existing clinical development frontier. Collectively, these patterns indicate that China's rapid catch-up encompasses both the quantity and the quality of innovative clinical development.

Post Trial Outcomes Whether this increased investment in trial activity translates into progress at later stages of the drug development lifecycle is an open question. Given that the majority of observed trials were initiated within the last decade and the average pharmaceutical development cycle spans 8–12 years, many long-run product-market outcomes remain unobserved. To partially bridge this gap, we examine intermediate milestones along the commercialization pathway using multiple data sources. In particular, we analyze license-out transactions and new drug approvals (NDAs) that are recorded in the Citeline Pharmaprojects database, which capture market-based valuation and final regulatory validation of clinical success, respectively. We further examine new firm entry into clinical development, measured as the number of firms conducting their first clinical trials.

Figure 2 compares these downstream milestones for China and the US. Panel (a) shows a gradual rise in license-out events for Chinese-originated drugs starting several years after 2016, a pattern consistent with the expected lag between early-stage trials and subsequent commercialization activity. Panel (b) reveals a more pronounced acceleration in global first approvals of drugs originated by Chinese firms, which rose from fewer than 5 before 2016 to over 40 by 2024.[13] Panel (c) displays a sharp increase in the number of new drugs approved by the Chinese administration, rising from fewer than 20 before 2016 to around 100 by 2024, exceeding the corresponding number of approvals by the U.S. administration. Panel (d) shows a rapid increase in the number of firms conducting clinical trials for the first time in China after 2016, with annual counts rising from fewer than 100 before 2016 to over 300 by 2024.

These findings are corroborated by an industry report indicating that China has overtaken the EU4+UK in total new drug approvals over the past five years, trailing only the U.S. (IQVIA Institute for Human Data Science, 2024).[14] If the trend continues, China may surpass the U.S. in new drug access in the near future. These trends indicate that China's catch-up is not confined to trial

[13]For each drug, we only count the first approval anywhere in the world.

[14]Over the past five years, the U.S. approved 267 new drugs, China approved 192 drugs, and EU4+UK approved 187 drugs, and many drugs approved China are first-in-world and first-in-class.

initiations. The rise in licensing and approvals indicates that the increased R&D effort is beginning to reshape both the scientific pipeline and the commercialization landscape of drug development.

3.2 Structural Break Test

The descriptive evidence suggests a notable inflection point in China's clinical trial activity relative to the US (and Europe) during the mid-2010s. To identify the timing of this shift, we employ a structural break test following the framework established by Quandt (1960) and Andrews (1993). Specifically, we implement a supremum Wald (Sup-W) test, which identifies an unknown break date by maximizing the Wald statistic across a series of Chow tests for all possible partitions of the sample period.

The test results yield clear, statistical evidence of a regime shift in China's trial output around 2016-2017.[15] Since a break identified in 2016 reflects a divergence in the data-generating process beginning that year, we designate 2015, the final year of the pre-break regime, as the reference period ($t = -1$). This data-driven identification procedure provides the empirical basis for the treatment period used in the event-study and difference-in-differences specifications presented in the following sections.

3.3 DID and Event Study

The descriptive evidence suggests a sharp expansion in China's clinical trial activity relative to the U.S. (and Europe) post 2015. To formally quantify the magnitude of this shift, we exploit the granular nature of our micro-data by estimating regressions at the disease–country–year level. This high-frequency, disaggregated approach is motivated by three considerations. First, drug development is frequently catalyzed by global scientific breakthroughs that are disease-specific, such as the emergence of checkpoint inhibitors in oncology. Our framework accounts for these time-varying, disease-specific global shocks that might otherwise confound the estimates. Second, a disaggregated analysis addresses potential compositional differences. If China's growth were merely driven by "low-hanging fruit," such as me-too drugs in the crowded antibiotics sector, a country-level aggregate would not distinguish between increases in trial volume from changes in the R&D mix. Finally, the therapeutic area serves

[15]The p-value of the chi-square test is 0.046 for 2016 and 0.0003 for 2017. P-values for earlier candidate break years are significantly higher.

as the economic unit of decision-making, as pharmaceutical firms typically allocate R&D budgets and manage pipelines along specific disease classifications.

We use the U.S. as our primary comparison group for three reasons. First, it serves as a proxy for the time-varying global innovation frontier. Since scientific knowledge is a global public good, the supply of new drug targets is simultaneously available to all countries; using the U.S. as a control allows us to isolate China-specific changes in innovation activity from global technological shifts. Second, the U.S. represents a stable, mature pharmaceutical market whose clinical trial trends parallel those in Europe, supporting its suitability as a benchmark. Third, since China has historically been a follower, we would expect a stable or widening gap between the two nations. A narrowing gap, or a structural break, suggests a transition toward indigenous innovation. In robustness checks, we also use Europe and the U.S. + Europe combined as alternative comparison groups.

We estimate the following difference-in-differences specification at the disease-country-year level, using 2015 as the baseline year:

$$Y_{d,c,t} = \beta_c \mathbf{1}(\text{China})_c \times \text{Post2015}_t + \phi_{d,c} + \xi_t + \epsilon_{d,c,t}, \tag{1}$$

where the dependent variable is the log number of trials (or drug projects) for disease d in country c and year t. Consistent with the results of our structural break test, Post2015_t is an indicator equal to one for all years from 2016 onward. We include disease-by-country fixed effects ($\phi_{d,c}$) to control for time-invariant factors, such as country-specific comparative advantages in certain therapeutic areas. Year fixed effects (ξ_t) absorb global temporal trends in drug development. These fixed effects ensure that identification comes from relative changes within disease categories after 2015 rather than shifts in disease composition or global trends. Under the assumption of parallel trends, which we test formally in the following section, β_c identifies the differential growth in China's clinical output.

R&D Volume and Novelty Table 3 reports the baseline difference-in-differences estimates (we examine robustness in Section 3.4 below). Across all specifications, the coefficients are positive and statistically significant at the 1% level, indicating a substantial acceleration in China's clinical trial output relative to the U.S. The estimate in Column (1) implies that the total number of trials in China increased by approximately 86.1% ($e^{0.62} - 1$) post 2015.

Beyond volume, we find evidence of a corresponding surge in the quality and novelty of these trials. We examine three novelty measures. In Columns (2) and (3), high-novelty trials, defined as those with above-median novelty scores according to Dranove et al. (2022) and those investigating non-generic compounds, increased by 58% and 61%, respectively. In Column (4), we restrict the sample to Phase III trials and define novel trials as those using a *patented* drug, rather than a placebo or off-patent drug, as the control arm.[16] These trials increased by 32%, a sizeable change given the stringent threshold this measure imposes. Together, these results indicate that China's post-2015 growth reflects a broad-based strengthening of clinical development capabilities across therapeutic areas, rather than a concentrated shift toward low-value, imitative projects.

Figure 3 plots the event-study coefficients underlying the DID results in Table 3. It provides a visual test of the parallel-trends assumption and quantifies the dynamic effects following the structural break. The estimated coefficients before 2016 are small in magnitude and statistically indistinguishable from zero, suggesting that China and the U.S. followed comparable growth trajectories in clinical trial activity during the pre-treatment period, as shown in Panel (a).

Beginning in 2017, we observe a clear divergence. By 2024, the relative increase reaches approximately 1.0 log point, representing a 172% expansion in China's trial volume relative to the US. Panels (b) and (c) reveal analogous patterns for high-novelty and non-generic trials: both series exhibit flat pre-trends followed by significant post-break growth of 0.8 log points (a 123% increase). Panel (d) shows a significant positive increase in Phase III trials that use an innovative drug as the control arm, though the magnitude at 0.4-0.6 log points (a 49%-82% increase) is somewhat smaller. This partly reflects the smaller share of Phase III trials in overall clinical activity and the stringent bar imposed by the control-arm criterion. Together, these patterns suggest that the expansion in China's clinical activity is not a transitory shock but a sustained shift toward higher-intensity, more innovative R&D.

Post-trial Outcomes Figure A1 examines the transmission of China's surge in clinical trial activity to downstream outcomes in drug development and industry entry, including license-out transactions, new drug approvals, and the entry of new firms. Across Panels (a) through (d), the estimated coefficients exhibit a clear post-2015 upward trajectory. By 2024, out-licensing transactions to multinational corporations (MNCs) had doubled, China's NMPA drug approvals had tripled, first-in-world

[16] Out of 332 disease categories, only 11 do not contain any non-generic drugs.

regulatory approvals for China-developed drugs (in any country) had risen by 36%, and firm entry had increased by 27%, all relative to 2015 and benchmarked against the control group (the U.S.). In addition, license-out activity and approvals of Chinese-originated drugs exhibit a marked upturn beginning around 2020–2021. This timing is consistent with the multi-stage nature of pharmaceutical development, whereby downstream milestones such as licensing deals and regulatory approvals respond with a lag to earlier increases in clinical trial activity. Relative to the trial-level results, the estimated coefficients are more dispersed, reflecting the sparsity of licensing, approval, and entry events at the disease–year level. Nevertheless, the broadly rising post-2015 patterns across all four panels indicate that China's expansion in clinical activity is translating into tangible downstream outcomes, spanning commercialization, regulatory success, and new firm participation in the pharmaceutical innovation ecosystem. We examine the long-run welfare benefits of induced drug innovation in Section 5.

3.4 Robustness and Heterogeneity

In this subsection, we conduct robustness checks using alternative specifications and sample choices and investigate heterogeneity across trial stages and sponsor type (private or public entities). These results inform the mechanisms analysis in the following section.

Domestically-led R&D A natural question is whether the rapid acceleration in clinical activity is driven by domestic Chinese firms or by multinational (MNC) incumbents operating within China. Distinguishing between these two groups is essential because growth led by Chinese firms reflects an expansion of China's indigenous innovative capacity, whereas a surge led by international firms might simply represent a geographic reallocation of global R&D activities without local capability building.

Figure A2 compares the trial initiations by nationality of sponsoring firms. Panel (a) focuses on trials conducted in China and shows that the post-2015 expansion is driven almost exclusively by domestic firms. Domestic Chinese firms account for 88% of the post-2015 increase in trials conducted by private organizations. In contrast, non-Chinese firms account for only a small fraction of total trials in China, and their activities remained relatively stable over the sample period. Panel (b) examines the presence of Chinese sponsors in trials conducted in the U.S. While their initial volume was limited relative to non-Chinese firms, we document a steady increase in Chinese-sponsored trials

in the U.S. after 2015, peaking around 2021. This trajectory suggests a nascent "internationalization" of Chinese pharmaceutical firms, though the aggregate magnitude of their overseas activity remains modest.[17]

Figure A3 conducts an event study and decomposes trial activity in China by firm origin.[18] The event-study estimates for both groups exhibit stable, near-zero pre-trends, further validating our empirical design. However, the post-2015 trajectories diverge sharply: while the relative growth for domestic Chinese firms climbs to approximately 1.2 log points by 2024 (a 232% increase), the estimates for international (non-Chinese) firms remain relatively flat. This asymmetry, along with a flat pre-trend, suggests that the post-2015 acceleration is primarily an endogenous phenomenon driven by domestic firms (indigenous capacity building), rather than a result of global firms reallocating their R&D pipelines to China.[19]

Alternative Specifications We further evaluate the sensitivity of our baseline estimates to alternative specifications and sample definitions. Table A2 assesses robustness to different sets of fixed effects and functional form choices. Column (1) replicates the baseline specification from Column (1) of Table 3. In Column (2), we augment the baseline model with disease-by-year fixed effects, which control for time-varying, disease-specific shocks at the global level. Columns (3) to (5) address observations with zero trials and functional form choices. Column (3) follows Chen and Roth (2024) and defines the dependent variable as log(number of trials) when at least one trial is observed in a country-disease-year cell and as -1 otherwise.[20] To ensure that our results are not driven by the log transformation of the dependent variable, Columns (4) and (5) report estimates using the raw trial counts as the outcome, employing both OLS and Poisson Pseudo–Maximum Likelihood (PPML) estimators. Across all specifications, the post-2015 upstick remains robust to fixed-effect choices and functional form assumptions.

[17] The relatively modest increase in U.S.-based trials initiated by Chinese firms does not imply a lack of innovativeness. Conducting clinical trials in the U.S. is costly, and Chinese firms face additional institutional and regulatory frictions when operating overseas. As a result, Chinese firms commonly enter the U.S. market through out-licensing arrangements or collaborations with multinational pharmaceutical companies rather than directly sponsoring trials.

[18] The equation is: $Y_{d,c,t} = \beta_{c_1} \mathbf{1}(\text{ChineseFirmInChina})_{d,c} \times Post2015_t + \beta_{c_2} \mathbf{1}(\text{NonChineseFirmInChina})_{d,c} \times Post2015_t + \phi_{d,c} + \xi_t + \epsilon_{d,c,t}$.

[19] The divergence between domestic and foreign firms may be attributed to lower institutional frictions or a greater ability to navigate the local regulatory and negotiation landscape (Javorcik and Wei, 2009).

[20] This approach assumes that the effect of moving from zero to one trial is equivalent to 100 log points. Results are robust to alternative cutoffs of 0 and -0.1 and slightly larger with a cutoff of -3, as expected.

In Table A3, we address potential concerns regarding the DID design. Globalization of R&D can lead to multicenter trials. To avoid any spillovers across the treated and control groups, Column (2) excludes trials conducted in both the U.S. and China, thereby ensuring that the estimated effect is not driven by the increasing number of global trials that include a Chinese site. Furthermore, Columns (3) and (4) verify that our results are not sensitive to the choice of the U.S. as the control group by using Europe (and U.S. + Europe) as alternative control groups. The consistency of the coefficients across these samples suggests that our findings reflect a distinct shift in the treatment group (China) rather than idiosyncratic trends in the control group (U.S.).

Heterogeneity Across Trial Stages and Sponsor Types Figure A4 plots event-study estimates for clinical trials by development phase. We observe an immediate and statistically significant increase in Phase I trials following 2015, whereas the rises in Phase II and Phase III trials emerge only after a lag of approximately two to four years. This timing pattern is consistent with new drug development timelines, in which early-stage discoveries require several years to progress to later phases. The delayed response in advanced-phase trials suggests that the post-2015 expansion reflects the entry of genuinely novel projects rather than the sheer resumption of previously suspended or terminated pipelines, which are likely to be more marginal.

Table A4 splits trials by stages and sponsor types. Trial activity increased across all stages, with Phase I trials registering the highest increase ($e^{0.667} - 1 = 95\%$ annual growth in the post period), followed by Phase II ($e^{0.332} - 1 = 39\%$) and Phase III ($e^{0.165} - 1 = 18\%$). This ordering is consistent with the high risk of failure and attrition inherent in the trial process.

Across sponsor types, the patterns reveal significant variation in magnitudes that align with the comparative advantages of different institutions. Private sponsors drove the largest increases in Phase I and Phase III trials, stages most critical for establishing safety, efficacy, and securing regulatory approval. In contrast, public sponsors (hospitals, universities, etc.) exhibited their strongest growth in Phase IV trials. Phase IV trials focus on post-approval observational studies and real-world evidence generation, which are less commercially driven. These patterns reflect a structural division of labor: private firms prioritize high-stakes, market-entry R&D in early and pivotal phases, while public institutions focus more heavily on post-market surveillance and broader clinical evidence generation following a drug's commercial launch.

Collectively, these results provide some of the first empirical evidence of a developing economy transitioning into a global frontier in pharmaceutical innovation.[21] In the next section, we turn to the institutional and policy drivers behind this growth, assessing the extent to which this transition was catalyzed by deliberate regulatory reforms versus broader economic and demographic shifts in a growing society.

4 Mechanisms

In this section, we investigate the institutional and economic drivers behind China's rapid ascent in pharmaceutical R&D. We consider both demand-side and supply-side factors identified in the literature (Finkelstein, 2004; Acemoglu et al., 2006; Kremer and Snyder, 2018; Bloom et al., 2019). On the demand side, we examine the NRDL reform that began in 2016, coinciding with the year of the structural trend break identified in Section 3.2, which dramatically expanded the effective market size for innovative drugs. We also evaluate China's accession to the ICH (International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use) in June 2017, which aligned China's regulatory framework with international standards and potentially broadened access to foreign markets. On the supply side, we evaluate the role of knowledge creation and human capital, the 2015 clinical trial application (CTA) reform that cleared a substantial backlog of pending applications and accelerated review timelines (Jia et al., 2024), and "Made in China 2025" (MIC2025), a national strategic plan aimed at upgrading China's manufacturing sector, including biomedicine. We also assess complementary state-led initiatives that may affect both the demand for and supply of pharmaceutical innovation. We conclude with a decomposition exercise to quantify the relative importance of each driver.

4.1 The NRDL Reform

NRDL's Effect on Market Size To evaluate the effect of NRDL on market size, we first estimate the direct effect of NRDL inclusion on drug sales using the doubly robust estimator from Callaway and Sant'Anna (2021). Figure 4 decomposes the impact on prices, quantities, and total revenue by comparing the included drugs (treated group) with the eligible but not included drugs (control group)

[21] Branstetter and Kwon (2018) shows that favorable exchange rates contributed to South Korea's rise in manufacturing innovation in the 80s and 90s.

before and after the inclusion. Following successful negotiation, retail prices drop by an average of 50% (and 66% for oncology drugs). However, this price compression is more than offset by a massive market expansion: quantities sold increase by 350% on average, with cancer drugs seeing a nearly ten-fold (900%) increase in volume.

The net effect on producer revenue (and profit) is overwhelmingly positive. Panels (e) and (f) show a 100% average increase in revenue post-negotiation, with oncology drugs experiencing a 500% surge. Given that marginal production costs for drugs are typically low relative to their price (Barwick et al. (2025) estimate marginal costs at approximately 18% of the pre-negotiated price), these revenue gains translate into a significant increase in producer surplus. Appendix Figure A5 documents a sharp expansion of the innovative drug market between 2016 and 2024. The number of NRDL-covered drugs increased from zero in 2015 to 699 in 2024, public insurance spending surpassed ¥100 billion in 2024, and hospital market sales quadrupled over the period. These results confirm that the NRDL reform fundamentally altered the profitability of the Chinese pharmaceutical market, providing the financial impetus for the clinical trial surge documented in the next section.

Effect on Clinical Trials To formally test whether the NRDL-induced market expansion catalyzed the surge in clinical activity, we construct disease-specific measures of policy exposure. By merging the Citeline TrialTrove dataset with granular drug-level sales data from SinoHealth (2017–2024), we exploit variation in the intensity of reimbursement coverage across 332 distinct disease categories. This allows us to quantify the "market size" effect on innovation, a central theme in the economics of R&D.

We define four complementary measures of disease-level NRDL exposure to proxy for the size of effective market expansion: (a) the log number of covered drugs, (b) the log revenue of covered drugs, (c) the share of covered drugs within the disease category, and (d) the revenue share of covered drugs. We estimate the relationship between reimbursement exposure and clinical activities using the following specification:

$$log(\text{Trials})_{d,c,t} = \beta \text{NRDL Exposure}_{d,c,t} + \phi_{d,c} + \xi_t + \epsilon_{d,c,t} \tag{2}$$

where $\text{Trials}_{d,c,t}$ is the number of trials in disease category d in country c at time t and coefficient β

identifies the semi-elasticity of trial activity with respect to NRDL-driven market expansion.

Table 4 reports these results with two key findings. First, disease categories with greater NRDL exposure experienced significantly larger increases in clinical trial activity across all four measures. Both the extensive margin (coverage or not) and the intensive margin (the scale of coverage) boost R&D activity. Second, the magnitudes are economically significant. The coefficient in Column (3) implies that a 100% increase in NRDL-covered drug revenue is associated with a 32.3% increase in clinical trials ($e^{0.28}-1$). This estimated elasticity is comparable to the findings of Dubois et al. (2015), who examine the responsiveness of drug development to market size using cross-disease variations in market sizes across OECD countries.[22]

Placebo Test To mitigate concerns that our results reflect broad national trends rather than the NRDL specifically, Figure 5 presents a placebo test using the same sample and specification as Figure 3(a), except that it restricts to disease categories that did not experience NRDL expansion during our sample period. We exclude oncology and infectious diseases to ensure a clean control group because the former is subject to high expectations of future coverage, and the latter was confounded by the COVID-19 pandemic. In addition to disease-by-country and year fixed effects, we also control for disease-year level revenue from volume-based procurement (VBP), which is China's public procurement of non-innovative drugs and is relevant for these disease categories (Cao et al., 2024). As expected, these "never-treated" categories show no systematic increase in trial activity after 2015. The sharp contrast with the patterns presented in Figure 3(a) – minimal expansion in non-reimbursed therapeutic areas alongside a pronounced surge in reimbursed ones – strongly suggests that the NRDL reform played a central role in reorienting China's pharmaceutical R&D toward innovative development.

Heterogeneity across Disease Category There is considerable heterogeneity in the scale of NRDL expansion across disease categories, which allows for a more granular examination of how shifts in market size influence the direction of R&D. Figure 6 visualizes this relationship by plotting the change in clinical trial activity against the expansion of NRDL coverage (the number of drugs included in NRDL) for each disease category between 2016 and 2024. Panel (a) displays the raw

[22] Dubois et al. (2015) estimate an elasticity of 25%, and the previous literature mostly finds elasticity around 50%.

data, while Panel (b) plots the estimated increase in trials attributable to NRDL exposure, using the elasticity of 0.164 from Table 4: $\Delta\text{Trials}_d = \sum_{t=2016}^{2023} \text{Trials}_{d,t} \times 0.164 \times \frac{\text{NRDL}_{d,t+1} - \text{NRDL}_{d,t}}{\text{NRDL}_{d,t}}$.[23]

The disease categories with the most pronounced expansion in NRDL exposure include non–small cell lung cancer, liver cancer, and renal cancer, diseases that also witnessed some of the sharpest increases in trial volume (Panel (a)). As shown in Panel (b), the upward-sloping fitted line reveals a strong positive association between reimbursement expansion and R&D response. The fact that the regression-based estimates closely track the raw data suggests that our baseline results are not driven by functional form assumptions or outliers, but rather reflect a robust cross-sectional relationship where innovation follows the expansion of the effective market size.

These patterns underscore the role of government reimbursement in overcoming under-investment in R&D. When demand for innovative drugs is relatively inelastic among high-income consumers but elastic among the broader population, private market pricing may lead to a sub-optimal volume of innovation. In this context, institutional intervention through "price-for-volume" negotiations can effectively expand the market, shifting the demand curve outward and restoring incentives for innovation (Kremer and Snyder, 2018).

Importantly, our results suggest that the NRDL reform influences not only the scale but also the direction of pharmaceutical innovation. By disproportionately incentivizing R&D in therapeutic areas with high unmet needs and large patient populations and rewarding clinical effectiveness in the inclusion criteria, the reform has steered clinical activity toward novel drug development. This reorientation represents a critical second-order benefit: while the immediate impact of the NRDL is to increase static efficiency through expanded patient access, its long-run impact is to enhance dynamic efficiency by catalyzing the development of the next generation of therapies. We quantify these long-run benefits in Section 4.7.

4.2 ICH Harmonization

China joined the ICH (International Council for Harmonization of Technical Requirements for Pharmaceuticals for Human Use) in June 2017, aligning its drug regulatory framework with international technical standards. By harmonizing requirements across jurisdictions, this alignment lowers barri-

[23] To avoid undefined percentage changes when $\text{NRDL}_{d,t} = 0$, the calculation is restricted to observations with $\text{NRDL}_{d,t} > 0$.

ers to cross-border drug development and approval, potentially expanding the effective international market size for drugs developed in China.

To test whether ICH accession encouraged China-based trials targeting foreign (U.S.) markets, we interact U.S. disease prevalence with a post-ICH indicator (country-specific disease prevalence is added as a control).[24] If firms increase clinical trial activity in response to an expanding international market, we should expect a disproportionate rise in trials for diseases with higher U.S. prevalence. However, as shown in Column (2) of Table 5, the estimated effect is economically small and statistically insignificant: a one-percentage-point increase in U.S. prevalence is associated with only 1% more Chinese trials after ICH harmonization. Meanwhile, the NRDL coefficient increases slightly from 0.51 in Column (1) to 0.54 in Column (2), suggesting that easier access to the U.S. market does not attenuate the domestic demand-side channel. This result is perhaps unsurprising because the FDA strongly prefers direct evidence from trials conducted in the U.S. or in populations relevant to the U.S. context, particularly for pivotal Phase III studies.[25]

4.3 Knowledge Creation and Talent Flow

A competing hypothesis is that China's clinical surge was driven by a supply-side expansion in innovation inputs – specifically, the accumulation of scientific knowledge and the return of high-skilled talent from abroad. Existing literature shows that the stock of public science provides the "fertile soil" for private-sector R&D (Azoulay et al., 2019; Bloom et al., 2013). Furthermore, high-skilled scientist mobility has been shown to enhance invention, with immigration, diaspora networks, and return migration facilitating the cross-border diffusion of ideas (Kerr, 2008). In the Chinese context, global integration has been linked to the country's surge in patenting (Hu and Jefferson, 2009).

To test this mechanism, we leverage the OpenAlex database to construct disease–country-year measures of knowledge production and talent flows. We proxy for knowledge accumulation using total PubMed-indexed publications and talent inflows by identifying "returnees": scientists with at

[24] We focus on U.S. disease prevalence, as it is the largest pharmaceutical market in the world.

[25] Although the FDA permits the use of foreign clinical data and multi-regional clinical trials (MRCTs), pivotal trials supporting New Drug Applications must establish safety and efficacy in populations relevant to the U.S. context (e.g., 21 CFR 314.106, see https://www.ecfr.gov/current/title-21/chapter-I/subchapter-D/part-314/subpart-D/section-314.106). In practice, this requirement makes it difficult to rely primarily on Chinese-based trials for U.S. approval.

least three publications affiliated with U.S. institutions who subsequently published under a Chinese affiliation.

Figure 7 presents the temporal trends for China and the US. While China's scientific output has grown substantially, the timing does not align with the inflection point observed in clinical trials. Panel (a) shows a steady, secular increase in publication volume since 2010, with no detectable structural break around 2016. Similarly, Panel (b) shows that the inflow of US-trained talent has followed a smooth, long-term upward trajectory. These patterns suggest that while China's scientific capacity has been deepening, these supply-side factors lack the discontinuity needed to explain the post-2015 acceleration in trial activity.

We formally evaluate the contribution of these factors in Table 5 by estimating the effect of knowledge accumulation and talent inflow alongside the NRDL exposure. Column (3) shows that knowledge volume is positively correlated with trial activity, with every additional 1,000 papers contributing to 18% more trials in a disease category. Column (4) finds a significant positive relationship between returnee talent and trial growth, consistent with the role of international expertise in fostering domestic innovation.

Crucially, the coefficient on NRDL exposure remains large, positive, and statistically significant across all specifications. The elasticity of trial activity with respect to reimbursement revenue is remarkably stable even after controlling for these supply-side indicators. This yields an important conclusion: while knowledge and talent provided the necessary infrastructure and innovative capacity, the market-size expansion triggered by the NRDL reform acted as the primary catalyst. Absent the demand-side shock to expected returns, the accumulated supply of knowledge and talent would likely have translated into a more gradual increase in clinical activity rather than the dramatic shift documented in this paper.

4.4 Trial Application Reform

We next consider several government policies that may have complemented the rise in drug innovation. The first is the 2015 clinical trial application (CTA) reform, which streamlined regulatory procedures by clearing the backlog of pending IND (investigational new drug) applications required before initiating clinical trials (Jia et al., 2024).

We collect application and approval dates for all IND submissions from the Center for Drug

Evaluation of China's NMPA to calculate backlog duration. Two findings emerge. First, while the reform provided an immediate boost by alleviating administrative bottlenecks, its effects appear transitory: the backlog that had accumulated before 2014 was resolved rapidly, and approval times flattened after 2019 (Figure 8). The reform alone cannot account for the persistent upward trajectory in trial activity observed after 2015, nor can it explain the heterogeneous increases across therapeutic categories, firm nationality, and trial novelty. Second, Column (5) of Table 5, which controls for both NRDL and application backlog, shows that a 100-day reduction in backlog is associated with a 4.0% increase in clinical trials within a disease category, implying an 8% increase for the observed 200-day reduction from 2016 to 2024. Taken together, these results indicate that the CTA reform, by reducing approval days, likely complemented the rise in trial activity but is unlikely to have been its primary driver.

4.5 Made in China 2025

Made in China 2025 (MIC2025) is a national strategic plan released by the State Council in May 2015, aimed at shifting China's growth model from large-scale, low-cost manufacturing toward technology-intensive production with higher value added. Biomedicine and pharmaceuticals are designated as one of its strategic industries.[26] As part of the policy implementation, MIC2025 designates a set of pilot cities and city clusters as key sites for experimentation and policy rollout, providing them with targeted support including fiscal subsidies, preferential financing, and coordinated regulatory arrangements.[27]

To assess whether the surge in R&D is driven by targeted support under MIC2025, Panel (a) of Figure 9 excludes all trials sponsored by firms headquartered in MIC2025 pilot cities. Both the trajectory and the magnitude of the increase are comparable to the baseline event study using all trials (Figure 3(a)), indicating that the rise in clinical activity is not confined to cities receiving direct MIC2025 support. Consistent with this, the NRDL coefficient in Column (7) of Table 5,

[26]The official State Council announcement: https://www.ndrc.gov.cn/xxgk/zcfb/qt/201505/t20150520_967388.html.

[27]These cities include: Ningbo, Quanzhou, Shenyang, Changchun, Wuhan, Wuzhong, Qingdao, Chengdu, Ganzhou, Guangzhou, Hefei, Huzhou, and city clusters: Sunanshi (Nanjing, Wuxi, Changzhou, Suzhou, Zhenjiang), Pearl River West Bank (Zhuhai, Foshan, Shunde, Zhongshan, Jiangmen, Yangjiang, Zhaoqing), Chang-Zhu-Tan-Heng (Changsha, Zhuzhou, Xiangtan, Hengyang), Zheng-Luo-Xin (Zhengzhou, Luoyang, Xinxiang). Source: http://industry.people.com.cn/n1/2017/1027/c413883-29613030.html.

which excludes trials from firms headquartered in MIC2025 pilot cities, is comparable to the baseline estimate in Column (1) of Table 5, suggesting that both pilot and non-pilot cities exhibit significantly greater and comparable growth in R&D activity in disease categories with high NRDL exposure.

4.6 Other Policy Reforms

Finally, we examine whether other government policies (which consist of either demand or supply factors) contribute to the rise in clinical trial activity. We scrape all policy documents from the State Council Policy Document Library and use GPT-based classification to identify those relevant to biopharma innovation, grouped them into four categories: (a) advanced manufacturing policies, which aim to strengthen domestic production capacity; (b) innovation independence policies, which seek to reduce reliance on foreign technology; (c) capital market reform policies, which improve firms' access to financing; and (d) tax or subsidy policies, which provide direct fiscal incentives for R&D.

Panel (b) of Figure 9 plots the number of policies in each category from 2010 to 2024. Innovation independence policies rose sharply after 2018, consistent with heightened geopolitical tensions and the Chinese government's emphasis on technological self-sufficiency. Manufacturing, capital market reform, and tax and subsidy policies show more moderate upward trends. Column (6) of Table 5 quantifies the relationship between these policies and clinical trial activity. Although the cumulative count of biopharma-related policies is positively and significantly correlated with trial growth, the coefficient on NRDL exposure remains large and stable after controlling for these measures. This suggests that while broader state-led initiatives may have contributed at the margin, they do not account for the sharp post-2015 increase in clinical trials.

In sum, across all the mechanisms examined in this section, the NRDL reform consistently emerges as the dominant driver of China's R&D expansion, while other demand-side and supply-side factors (ICH accession, knowledge accumulation, talent inflows, regulatory streamlining, and industrial policies) played complementary but secondary roles.

4.7 Decomposition of Different Mechanisms

The preceding analyses identify several empirically important drivers of China's clinical expansion. Because supply-side and demand-side changes occurred simultaneously, we adopt a unified framework to quantify their relative importance. We proceed in two steps. First, we estimate semi-elasticities of

clinical trial activity with respect to each mechanism using the most saturated specification, reported in Column (8) of Table 5, which regresses trial activity on NRDL exposure, knowledge accumulation, talent flows, and other policy variables, controlling for disease-by-country and year fixed effects, as well as disease-country specific prevalence rates.

Second, we combine these coefficient estimates with observed changes for each mechanism from 2016 to 2024 to translate the semi-elasticities into implied changes in trial counts.[28] For each of the nine therapeutic areas reported in Table 6, we aggregate the contribution of mechanism M across the constituent disease categories from the full set of 332:

$$\Delta Y_d(M) = \beta_M \times \sum_{t=2016}^{2023} \sum_{j \in d} (M_{j,t+1} - M_{j,t}) \times \text{TrialNumber}_{j,t}, \tag{3}$$

where β_M denotes the coefficient on each mechanism as estimated in Column (8) of Table 5. This decomposition quantifies the contributions of demand-pull and supply-push mechanisms, providing a ranking of the drivers behind China's surge in pharmaceutical innovation.

Table 6 presents the results. To illustrate the economic magnitudes, we focus on Oncology, the largest therapeutic area in our sample. Out of the 1,004 additional oncology trials initiated in China in 2024 relative to the 2016 baseline, NRDL accounts for approximately 43%. Knowledge accumulation and talent flows together explain 24%, while other policies account for a negligible share (less than 1%); the remaining variation is attributed to unobserved factors. Across all therapeutic areas, the NRDL consistently emerges as the most influential driver, explaining at least 20% of the aggregate increase in clinical trial activity. These results reinforce a central finding of this paper: while secular growth in scientific knowledge and international talent flows provided a critical foundation for China's innovative capacity, the demand-side shock from the NRDL reform was the principal catalyst of the post-2015 acceleration, with other policies playing a comparatively modest role.

5 Welfare Implications

NRDL integrates innovative therapies into the national insurance catalog through centralized price negotiations, improving patient access to innovative drugs and generating large social surplus gains

[28] We exclude changes between 2015 and 2016, as this period predates the NRDL reform.

(Barwick et al., 2025). In this section, we provide a back-of-the-envelope calculation to benchmark the NRDL's long-run welfare gains from induced drug development against its short-run gains from expanded market access.

To fix ideas, consider a simple model in which each patented drug is produced by a monopolist setting price p_j^m. Patient demand is $q_j(p_j^m, \gamma_j)$, where γ_j denotes the coinsurance rate. If the drug is excluded from the national reimbursement list, $\gamma_j = 1$, and consumers bear the full cost out of pocket. The NRDL reform lowers the effective price faced by patients through two channels: the introduction of insurance coverage (with an average coinsurance rate of $\gamma_j^{\text{NRDL}} \approx 0.35$) and negotiated price reductions ($p_j^{\text{NRDL}} < p_j^m$). Together, these reduce out-of-pocket prices from p_j^m to $\gamma_j^{\text{NRDL}} p_j^{\text{NRDL}}$ and raise demand to $q_j(p_j^{\text{NRDL}}, \gamma_j^{\text{NRDL}})$. The resulting short-run welfare gain reflects increased utilization of existing innovative drugs, generating static surplus gains by mitigating monopoly markups.

The long-run welfare gain operates through innovation incentives. Before the reform, a pharmaceutical company invests in drug development if expected profit (the clinical success rate η_j multiplied by the long-run monopolist profit conditional on success) exceeds R&D cost c_j:

$$\eta_j \Pi_j(p_j^m; \gamma) > c_j. \tag{4}$$

Post-reform, the NRDL expands demand, raising expected profits to $\Pi_j(p_j^{\text{NRDL}}, \gamma_j^{\text{NRDL}})$. This increase in expected profitability spurs drug development and contributes to the observed rise in clinical trial activity. The long-run consumer surplus gains are computed as:

$$\sum_{k \in \text{Future drug projects}} \eta_k \mathbb{1}[\eta_k \Pi_k(p_k^{\text{NRDL}}; \gamma_k^{\text{NRDL}}) > c_k] CS_k(p_k^{\text{NRDL}}, \gamma_k^{\text{NRDL}})$$

where CS_k denotes consumer surplus and η_k denotes the clinical trial success rate.[29] Short- and long-term changes in firm profit and government expenditure are calculated analogously.

We borrow the estimates of consumer surplus, firm profit, and government expenditure for all 57 cancer drugs included in the NRDL during 2016–2022 from Barwick et al. (2025) to recover the static gains. Focusing on cancer drugs alone, the static consumer surplus gain is approximately ¥35.6

[29] In contrast, short-run consumer surplus gains are $\sum_{j \in \text{Existing drugs}} CS_j(p_j^{\text{NRDL}}, \gamma^{\text{NRDL}}) - CS_j(p_j^m, \gamma)$. Ceteris paribus, the dynamic gain per new drug exceeds the static gain per drug because the static gains reflect expanded access to an existing drug, whereas the dynamic gains capture the full consumer surplus of a previously nonexistent drug.

billion ($5.1 billion) per year. Government spending is of a similar magnitude, while firms earn ¥22.9 billion in profits, yielding an annual social surplus gain of ¥22.52 billion.

To quantify the dynamic gains from new drugs following successful clinical trials, we apply the decomposition in Section 4.7, which attributes 43% of the increase in oncology trials to the NRDL. We begin by calculating the cumulative discounted increase in oncology trials relative to the pre-2016 level. As reported in Appendix Table A1, the observed cumulative increase from 2016 to 2024 totals 7,084 trials.[30] Using a conservative assumption that the annual number of oncology trials remains at the 2024 level going forward, which is 1,004 above the pre-2016 baseline, and applying a 90% discount rate to future trials, the cumulative discounted increase from 2016 onward is $7{,}084 + 1{,}004 \times \frac{1}{1-0.9} = 17{,}124$.[31] Attributing 43% of this increase to the NRDL reform yields a cumulative discounted increase of 7,363 oncology trials. Given an average success rate of $\bar{\eta} = 1/123$ (123 trials per successful new drug in the oncology sector during our sample period), this translates to approximately 60 new oncology drugs.

The social value of these future drugs is uncertain ex ante. We proxy for it using the realized welfare gains from recent new drug launches in comparable oncology categories, providing an empirical benchmark for the expected innovation surplus induced by the NRDL. We treat each new drug as a monopolist and abstract from substitution with existing therapies. This assumption is reasonable in our setting because the innovative cancer drugs we evaluate are highly differentiated in their molecular targets and clinical applications, often addressing distinct genetic mutations, signaling pathways, or tumor subtypes. As a result, these drugs expand the effective treatment frontier rather than merely duplicating existing therapies (Barwick et al., 2025). Under the assumption that new drugs are comparable to those currently on the market, we use sales, demand, and firm profit estimates for innovative cancer drugs in 2022 to approximate the dynamic gains.[32] As a robustness check, allowing new drugs to compete with existing drugs and incorporating the substitution patterns reported in Barwick et al. (2025) deliver similar welfare estimates.

[30] This is the cumulative sum of the year-over-2016 differences in trial counts (i.e., the difference between 2017 and 2016, between 2018 and 2016, and so on).

[31] This estimate represents a lower bound for several reasons. First, our 2024 trial count excludes ongoing trials that have not yet concluded. Second, we hold future trial counts flat despite a significant upward trend, a conservative assumption. Third, we apply a high discount rate, which further reduces the present discounted value of the number of future trials.

[32] Average annual sales in 2022 are ¥0.78 billion per drug, with an interquartile range of ¥[0.15, 1.04] billion.

Table 7 presents the results. The dynamic consumer surplus gain is estimated at ¥67.59 billion per year (interquartile range: ¥[11.97, 99.64] billion), and the dynamic social surplus gain is ¥65.83 billion per year (interquartile range: ¥[12.07, 107.67] billion). Even under conservative assumptions, the dynamic gains are roughly three times the static efficiency gains, underscoring that the NRDL reform's most consequential welfare impact may operate through its effect on innovation incentives. A common concern with demand-side policies is that they primarily pull incremental innovation (Dranove et al., 2022; Geng and Shi, 2024). The NRDL appears to be a notable exception: the trials it induces are characterized by high scientific novelty, likely because the NRDL's inclusion criteria explicitly reward originality and prioritize unmet clinical needs. By tying market access to therapeutic advances, the NRDL does not merely enlarge the pie for existing drug categories; it reshapes the direction of innovation itself.

6 Conclusion

This paper documents a historic shift in the global landscape of pharmaceutical innovation. Using granular data on clinical trials and an event study framework, we provide the first empirical evidence of China's transition from a follower in generic manufacturing to a contributor at the global frontier of clinical R&D. By 2024, China's clinical trial activity has surged by over 170% relative to 2015 and benchmarked against the control group (the U.S.), driven primarily by domestic firms and increasingly concentrated in high-novelty, non-generic compounds.

Our analyses demonstrate that this transition was not a passive consequence of aggregate economic growth. We identify the 2016 National Reimbursement Drug List (NRDL) reform as the primary catalyst: by leveraging government monopsony power to guarantee market volume in exchange for price concessions, the reform not only expanded patient access but also fundamentally altered the expected returns to innovation. A growth-accounting decomposition reveals that this demand-pull effect explains a substantially larger share of the R&D surge than supply-side factors such as knowledge accumulation or talent inflows. Our welfare exercise further indicates that the long-run gains from induced innovation are at least three times larger than the short-run gains from expanded access.

These findings offer a new perspective on the policy studied here. While the pharmaceutical

literature often frames the trade-off between static efficiency (patient access) and dynamic efficiency (innovation) as close to a zero-sum game, the Chinese experience suggests that a sufficiently large and well-designed market expansion can mitigate this tension, as the same reform that drove down out-of-pocket costs for hundreds of millions of patients simultaneously pulled frontier R&D into the domestic market. More broadly, China's trajectory offers a proof of concept that economies lagging behind in innovation need not wait to inherit yesterday's medicines – it can, through deliberate policy design, become a source of tomorrow's.

References

Acemoglu, Daron, David Cutler, Amy Finkelstein, and Joshua Linn, "Did medicare induce pharmaceutical innovation?," *American Economic Review*, 2006, *96* (2), 103–107.

Agha, Leila, Soomi Kim, and Danielle Li, "Insurance Design and Pharmaceutical Innovation," *American Economic Review: Insights*, June 2022, *4* (2), 191–208.

Andrews, Donald WK, "Tests for parameter instability and structural change with unknown change point," *Econometrica: Journal of the Econometric Society*, 1993, pp. 821–856.

Aryal, Gaurab, Federico Ciliberto, Leland E. Farmer, and Ekaterina Khmelnitskaya, "Valuing Pharmaceutical Drug Innovations," April 2024. arXiv:2212.07384 [econ, q-fin].

Azoulay, Pierre, Joshua S Graff Zivin, Danielle Li, and Bhaven N Sampat, "Public R&D investments and private-sector patenting: evidence from NIH funding rules," *The Review of economic studies*, 2019, *86* (1), 117–152.

Barwick, Panle, Ashley Swanson, and Tianli Xia, "A Double Dose of Reform: Insurance and Centralized Negotiation in Drug Markets," March 2025. Working paper.

Bignon, Lea, "Pharmaceuticals and Digital Health: How Data-driven Insights May Reshape the Insulin Market," 2026.

Bloom, Nicholas, John Van Reenen, and Heidi Williams, "A Toolkit of Policies to Promote Innovation," *Journal of Economic Perspectives*, August 2019, *33* (3), 163–84.

_ , Mark Schankerman, and John Van Reenen, "Identifying technology spillovers and product market rivalry," *Econometrica*, 2013, *81* (4), 1347–1393.

Blume-Kohout, Margaret E. and Neeraj Sood, "Market size and innovation: Effects of Medicare Part D on pharmaceutical research and development," *Journal of Public Economics*, January 2013, *97*, 327–336.

Branstetter, Lee G and Mariko Sakakibara, "When do research consortia work well and why? Evidence from Japanese panel data," *American Economic Review*, 2002, *92* (1), 143–159.

Branstetter, Lee G. and Namho Kwon, "South Korea's transition from imitator to innovator: The role of external demand shocks," *Journal of the Japanese and International Economies*, 2018.

Bronzini, Raffaello and Paolo Piselli, "The impact of R&D subsidies on firm innovation," *Research Policy*, 2016, *45* (2), 442–457.

Budish, Eric, Benjamin N Roin, and Heidi Williams, "Do firms underinvest in long-term research? Evidence from cancer clinical trials," *American Economic Review*, 2015, *105* (7), 2044–2085.

Callaway, Brantly and Pedro H.C. Sant'Anna, "Difference-in-Differences with multiple time periods," *Journal of Econometrics*, 2021, *225* (2), 200–230. Themed Issue: Treatment Effect 1.

Cao, Shengmao, Lisa Xuejie Yi, and Chuan Yu, "Competitive Bidding in Drug Procurement: Evidence from China," *American Economic Journal: Economic Policy*, August 2024, *16* (3), 481–513.

Chen, Jiafeng and Jonathan Roth, "Logs with zeros? Some problems and solutions," *The Quarterly Journal of Economics*, 2024, *139* (2), 891–936.

Chen, Zhao, Zhikuo Liu, Juan Carlos Suárez Serrato, and Daniel Yi Xu, "Notching R&D Investment with Corporate Income Tax Cuts in China," *American Economic Review*, July 2021, *111* (7), 2065–2100.

Cunningham, Colleen, Florian Ederer, and Song Ma, "Killer Acquisitions," *Journal of Political Economy*, March 2021, *129* (3), 649–702. Publisher: The University of Chicago Press.

Cunningham, Collen, Florian Ederer, Charles Hodgson, and Zhichun Wang, "Disclosure and the Pace of Drug Development," *Working Paper*, 2025.

Dechezleprêtre, Antoine, Elias Einiö, Ralf Martin, Kieu-Trang Nguyen, and John Van Reenen, "Do tax incentives increase firm innovation? An RD design for R&D, patents, and spillovers," *American Economic Journal: Economic Policy*, 2023, *15* (4), 486–521.

Dranove, David, Craig Garthwaite, and Manuel Hermosilla, "Does consumer demand pull scientifically novel drug innovation?," *The RAND Journal of Economics*, 2022, *53* (3), 590–638.

Dubois, Pierre, Olivier de Mouzon, Fiona Scott-Morton, and Paul Seabright, "Market size and pharmaceutical innovation," *The RAND Journal of Economics*, 2015, *46* (4), 844–871. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/1756-2171.12113.

Finkelstein, Amy, "Static and Dynamic Effects of Health Policy: Evidence from the Vaccine Industry," *Quarterly Journal of Economics*, 2004, *119* (2), 527–564.

Garthwaite, Craig, "Economic Markets and Pharmaceutical Innovation," *Journal of Economic Perspectives*, May 2025, *39* (2), 3–26.

Geng, Hao and Ce Matthew Shi, "Health policy, price regulation, and innovation: Evidence from China's vaccine industry," *Journal of Development Economics*, March 2024, *167*, 103229.

Gross, Daniel P and Bhaven N Sampat, "The Therapeutic Consequences of the War: World War II and the 20th-Century Expansion of Biomedicine," Technical Report, National Bureau of Economic Research 2025.

Grover, Aditya and Jure Leskovec, "node2vec: Scalable feature learning for networks," in "Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining" 2016, pp. 855–864.

Hermosilla, Manuel, "Rushed Innovation: Evidence from Drug Licensing," *Management Science*, January 2021, *67* (1), 257–278. Publisher: INFORMS.

Howell, Sabrina T, "Financing innovation: Evidence from R&D grants," *American economic review*, 2017, *107* (4), 1136–1164.

Hu, Albert Guangzhou and Gary H Jefferson, "A great wall of patents: What is behind China's recent patent explosion?," *Journal of Development Economics*, 2009, *90* (1), 57–68.

IQVIA Institute for Human Data Science, "Global Trends in R&D 2024: Activity, Productivity, and Enablers," 2024. Accessed 2025.

Javorcik, Beata Smarzynska and Shang-Jin Wei, "Corruption and cross-border investment in emerging markets: Firm-level evidence," *Journal of International Economics*, 2009, *78* (2), 432–443.

Ji, Yunan and Parker Rogers, "The Long-Run Impacts of Regulated Price Cuts: Evidence from Medicare," October 2024.

Jia, Ruixue, Xiao Ma, Jianan Yang, and Yiran Zhang, "Improving Regulation for Innovation: Evidence from China's Pharmaceutical Industry," *Working Paper*, 2024.

Kerr, William R, "Ethnic scientific communities and international technology diffusion," *The Review of Economics and Statistics*, 2008, *90* (3), 518–537.

Kremer, Michael and Christopher M. Snyder, "Worst-Case Bounds on R&D and Pricing Distortions: Theory with an Application Assuming Consumer Values Follow the World Income Distribution," October 2018.

Kyle, Margaret, "Pharmaceutical Price Controls and Entry Strategy," *Review of Economics and Statistics*, 2007.

Lakdawalla, Darius N., "Economics of the Pharmaceutical Industry," *Journal of Economic Literature*, June 2018, *56* (2), 397–449.

Lerner, Josh, Junxi Liu, Jacob Moscona, and David Y. Yang, "Appropriate Entrepreneurship? The Rise of China and the Developing World," January 2026.

Li, Guangwei and Lee G. Branstetter, "Does "Made in China 2025" work for China? Evidence from Chinese listed firms," *Research Policy*, July 2024, *53* (6), 105009.

Liu, Chenyuan, Yi Lu, Ronghe Sun, and Wanyu Yang, "Expected Insurance Coverage and Pharmaceutical Innovation: Evidence from China's National Drug Price Negotiation Policy," 2026.

Martin, Linda, Melissa Hutchens, Conrad Hawkins, and Alaina Radnov, "How much do clinical trials cost?," *Nature Reviews Drug Discovery*, 2017, *16* (6), 381–382.

Melnik, Walter and Andrew Smyth, "R&D tax credits and innovation," *Journal of Public Economics*, 2024, *236*, 105157.

National Health Commission of the PRC, "2022 Health Statistics Yearbook," 2023. Accessed: 2024-07-30.

Olson, M. and R. Zeckhauser, "An Economic Theory of Alliances," *Review of Economics and Statistics*, 1966.

Priem, Jason, Heather Piwowar, and Richard Orr, "OpenAlex: A fully-open index of scholarly works, authors, venues, institutions, and concepts," *arXiv preprint arXiv:2205.01833*, 2022.

Quandt, Richard E, "Tests of the hypothesis that a linear regression system obeys two separate regimes," *Journal of the American statistical Association*, 1960, *55* (290), 324–330.

Schmookler, Jacob, *Invention and Economic Growth*, Harvard University Press, 1966. Google-Books-ID: eQozA-AAAMAAJ.

Wallsten, Scott J., "The Effects of Government-Industry R&D Programs on Private R&D: The Case of the Small Business Innovation Research Program," *RAND Journal of Economics*, 2000, *31* (1), 82–100.

Ward, Michael R. and David Dranove, "The Vertical Chain of Research and Development in the Pharmaceutical Industry," *Economic Inquiry*, 1995, *33* (1), 70–87.

Wei, Shang-Jin, Jianhuan Xu, Ge Yin, and Xiaobo Zhang, "Mild Government Failure," NBER Working Papers 31178, National Bureau of Economic Research, Inc 2023.

Xia, Tianli, "Welfare Effects of Resale Price Maintenance: Evidence from the Chinese Pharmaceutical Industry," April 2024. Working paper.

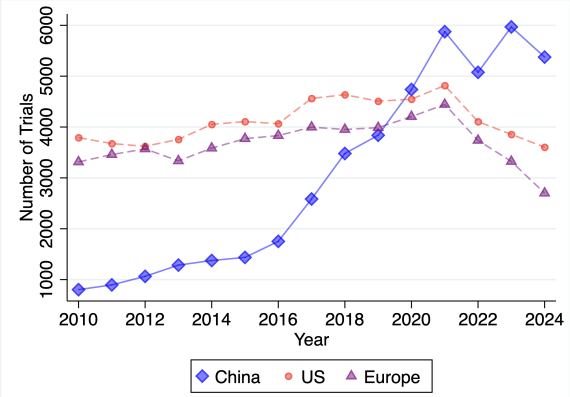
Yin, Wesley, "Market incentives and pharmaceutical innovation," *Journal of Health Economics*, July 2008, *27* (4), 1060–1077.

Yu, Hao, "Universal health insurance coverage for 1.3 billion people: What accounts for China's success?," *Health Policy*, 2015, *119* (9), 1145–1152.

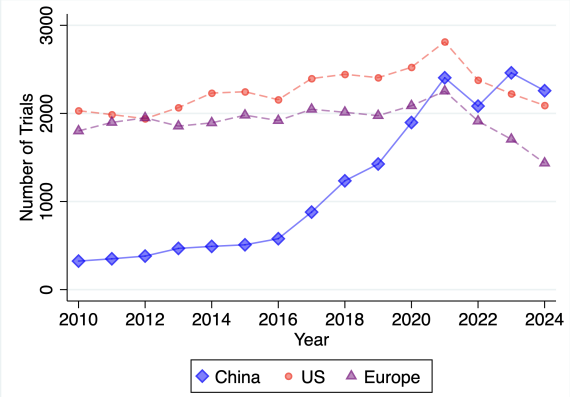
Figures

Figure 1: Time Trend - Trial Quantity and Novelty

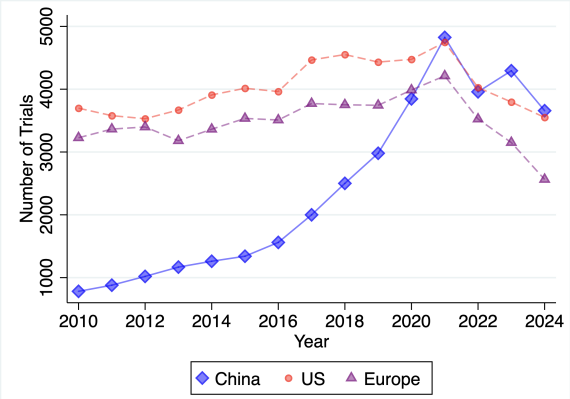
(a) All Trials
China vs US vs Europe



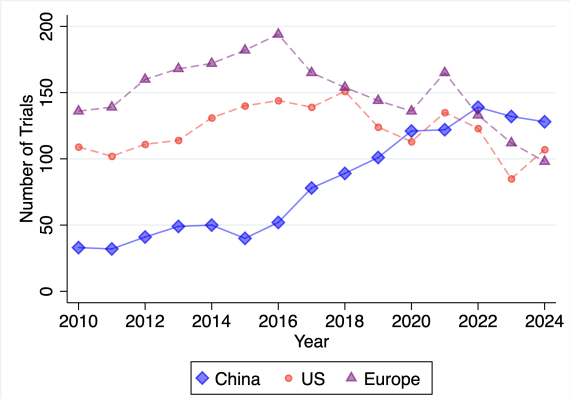
(b) Novelty Score > Median
China vs US vs Europe



(c) Non-Generic Trials
China vs US vs Europe



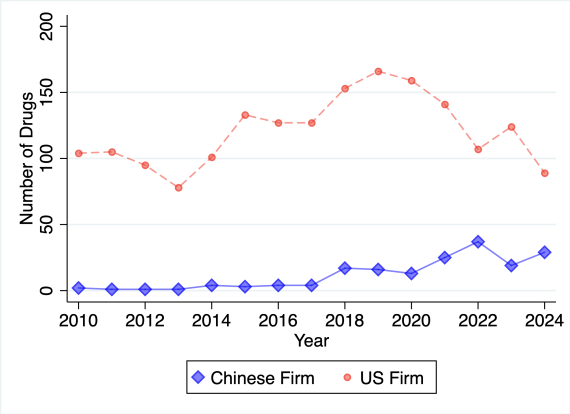
(d) Phase III Trials w. Innovative Control Arm
China vs US vs Europe



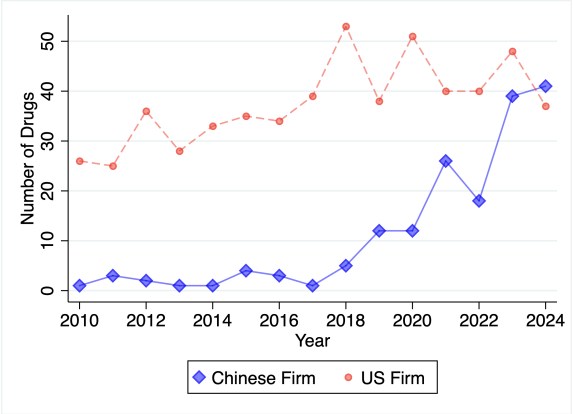
Notes. These figures plot the annual number of clinical trials launched in each year in China, the United States, and Europe. The vertical axis measures trial counts, and the horizontal axis indicates calendar years (2010–2024). Panel (a) compares all trials conducted in the three regions. Panel (b) restricts the sample to trials with above-median novelty scores, calculated following Dranove et al. (2022). Panel (c) restricts to trials investigating non-generic drugs, and Panel (d) limits the sample to Phase III trials that use a patented drug as the control arm. Data source: Citeline TrialTrove database.

Figure 2: Time Trend - Number of Drugs and New Firms

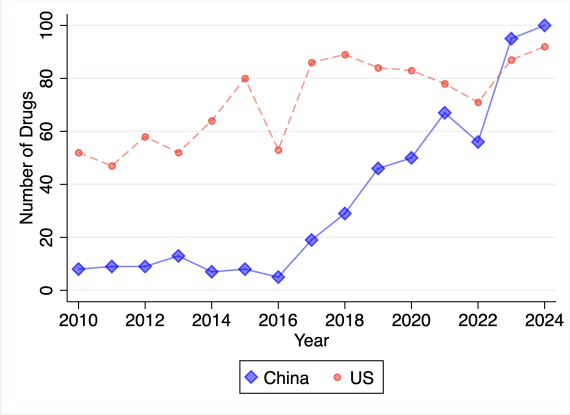
(a) Number of Licensed-Out Drugs
Chinese Firms vs US Firms



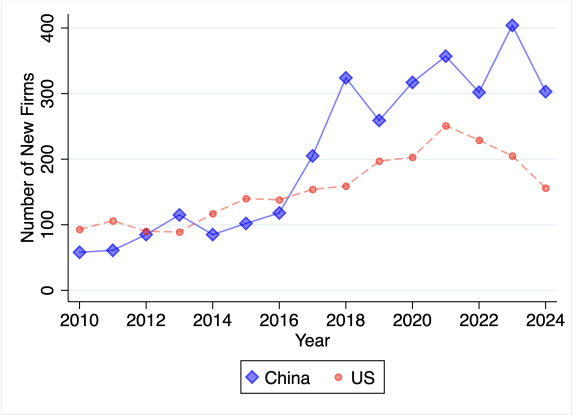
(b) First-in-world Regulatory Approval of Drugs
by Firms Headquartered in China vs US



(c) Number of Drug Approvals
by Chinese vs US Administration



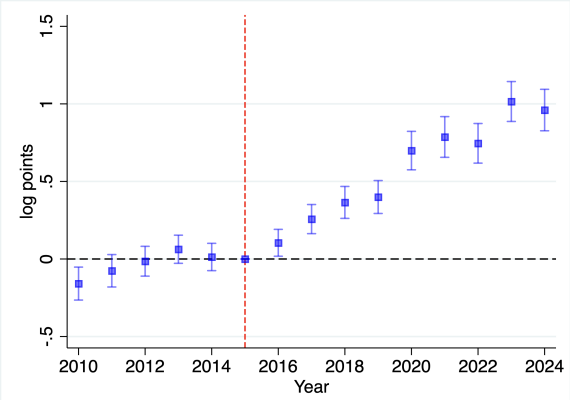
(d) Number of New Firms
China vs US



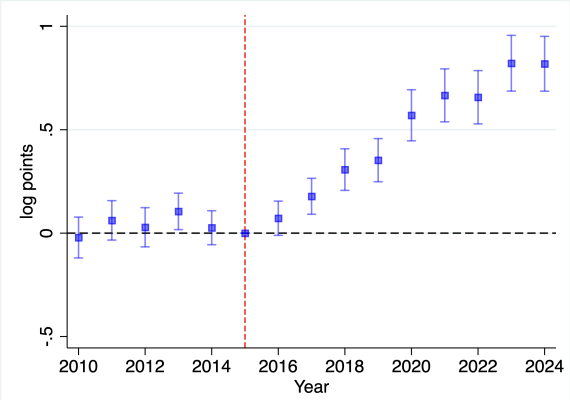
Notes. Panel (a) compares drugs licensed out to multinational corporations by firms headquartered in China and the U.S. Panel (b) compares first-in-world regulatory approvals for drugs developed by Chinese- vs. U.S.-headquartered firms. Panel (c) compares newly approved drugs by the Chinese and US regulatory authorities. Panel (d) compares the number of firms conducting clinical trials for the first time in each country. Data source: Citeline Pharmaprojects database.

Figure 3: Event Study - Trial Quantity and Novelty

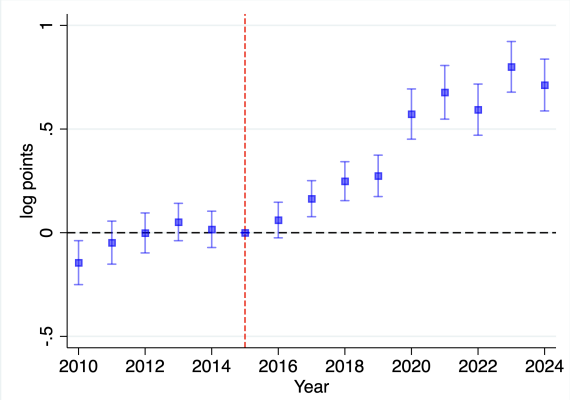
(a) All Trials
China vs US



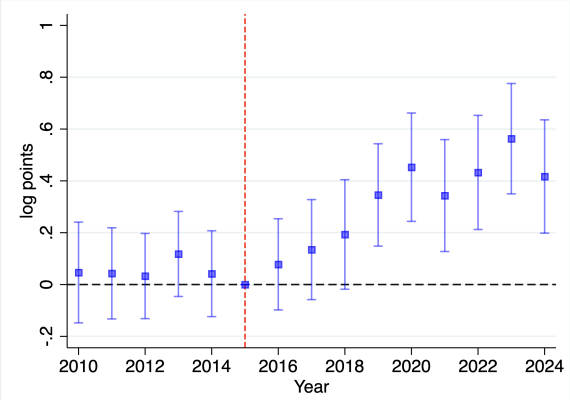
(b) Novelty Score > Median
China vs US



(c) Non-Generic Trials
China vs US



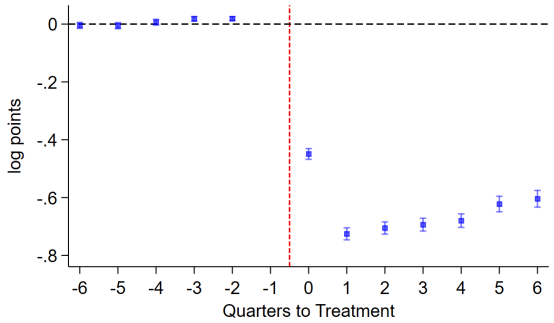
(d) Phase III Trials w. an Innovative Control Arm
China vs US



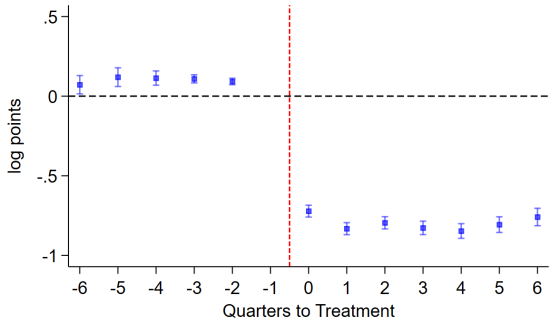
Notes. This figure presents an event-study analysis of clinical trial activity at the disease–country level, with 2015 as the base year and the US as the control group. The vertical axis plots the estimated coefficients (in 100 log points) on the interactions between the China indicator and event-year dummies from the dynamic difference-in-differences specification, along with 95% confidence intervals. The horizontal axis indicates calendar year (2010–2024). Panel (a) includes all trials conducted in China and the U.S., Panel (b) restricts to trials with above-median novelty scores, and Panel (c) restricts to trials investigating non-generic compounds. These three panels use $\log(\text{trials}+1)$ as the dependent variable, where trials is the number of trials launched in that year. Panel (d) focuses on Phase III trials that use a patented drug as the control arm. Because the large number of zero trials in this subsample attenuates estimates (relative to the raw data shown in Figure 1(d)), we follow Chen and Roth (2024) and define the dependent variable as log(trials) when at least one trial is observed in a country-disease-year cell and as -1 otherwise. Results robust to alternative cutoffs. All specifications include disease-by-country and year fixed effects. Robust standard errors are clustered at the disease × country level. Data source: Citeline TrialTrove Database.

Figure 4: Effects of the NRDL Reform on Price and Quantity

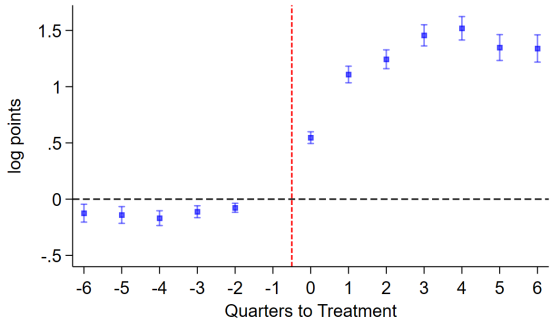
(a) Log Retail Price: All Drugs



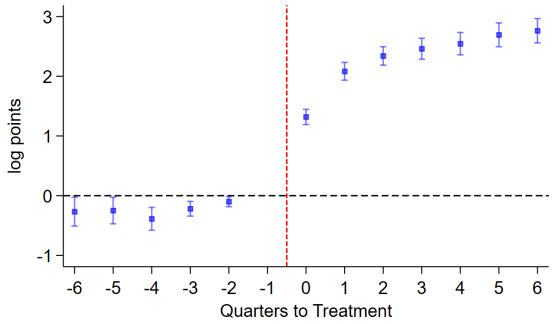
(b) Log Retail Price: Cancer Drugs



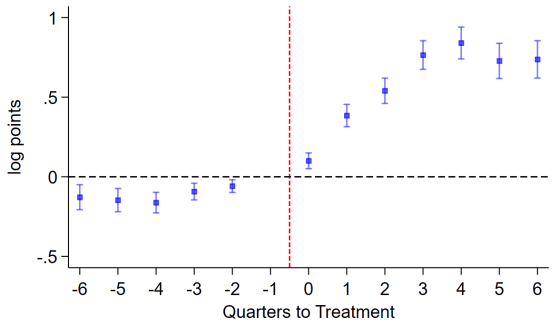
(c) Log Quantity: All Drugs



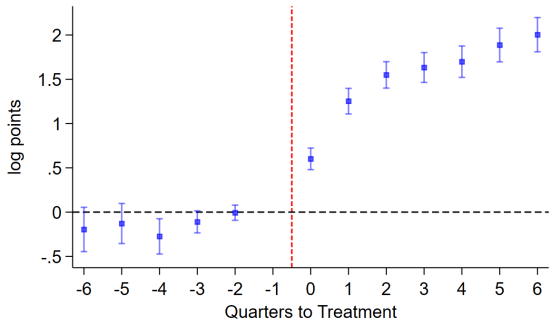
(d) Log Quantity: Cancer Drugs



(e) Log Revenue: All Drugs

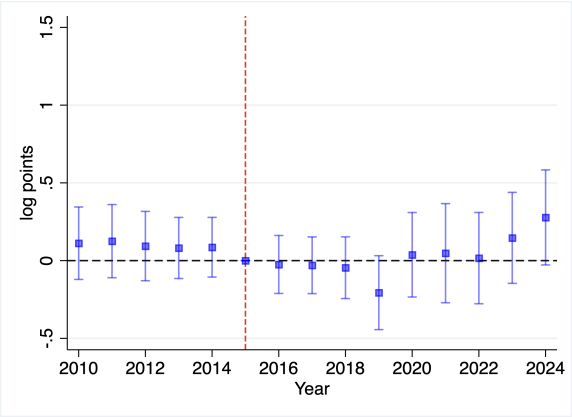


(f) Log Revenue: Cancer Drugs



Notes. This figure reports the Callaway and Sant'Anna (2021) estimated effects of the NRDL Reform on log(retail prices) in Panels (a) and (b), log(quantities) in Panels (c) and (d), and log(revenue) in Panels (e) and (f) using data from 2017 to 2023. The treated group is the innovative drugs included in NRDL, and the control group is the eligible yet not included drugs. The horizontal axis denotes quarters relative to the NRDL inclusion. The vertical axis reports estimated dynamic treatment effects in 100 log points and 95% confidence intervals using the CSDID method from Callaway and Sant'Anna (2021). Data source: Sinohealth.

Figure 5: Disease Categories without NRDL Exposure (Placebo Test)

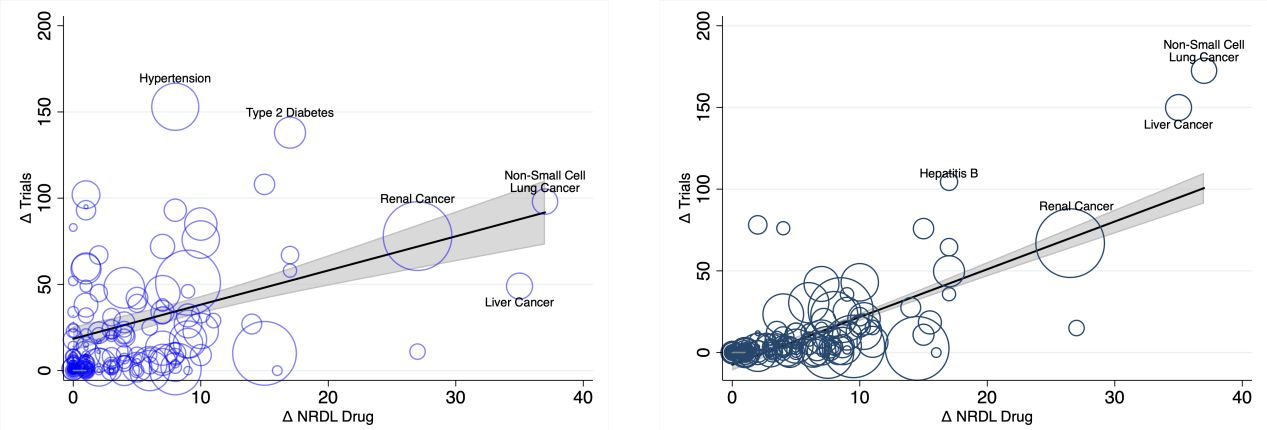


Notes. This figure presents a placebo test using the same specification as Figure 3(a), except that it restricts to disease categories that did not experience NRDL expansion during our sample period. Oncology and infectious disease are excluded due to high expectations of future coverage (oncology) and COVID (infectious disease). The specification includes disease-by-country and year fixed effects and controls for disease-year level revenue from volume-based procurement, China's public procurement of non-innovative drugs. Robust standard errors are clustered at the disease × country level. Data source: Citeline TrialTrove Database.

Figure 6: Mechanism - Effect of NRDL across Disease Categories

(a) Raw Data

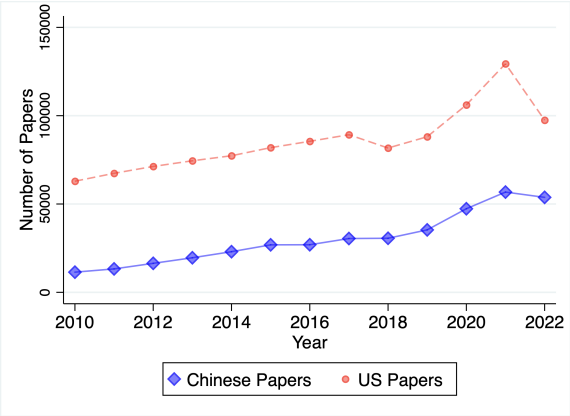
(b) Reg. Results:
Increased Trials vs NRDL Exposure



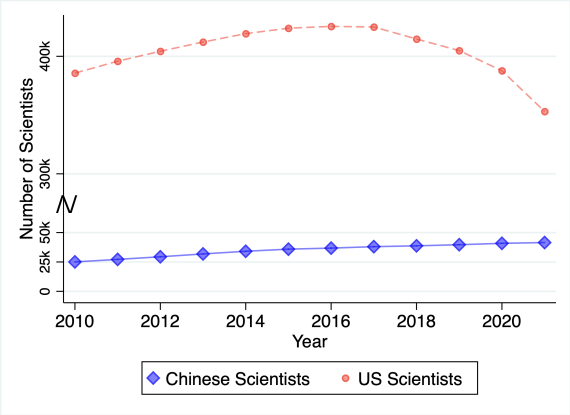
Notes. This figure plots, for each disease category, the change in clinical trial activity against the change in the number of NRDL-included drugs between 2016 and 2024. Panel (a) displays the raw data. Panel (b) presents the estimated increase in trials associated with NRDL exposure (measured as the number of NRDL-included drugs) for each disease category. Circle size is proportional to the total number of drugs in the category. Data source: Citeline TrialTrove Database.

Figure 7: Mechanism - Time Trend of Scientific Publications and Talent Inflows

(a) Number of Publications

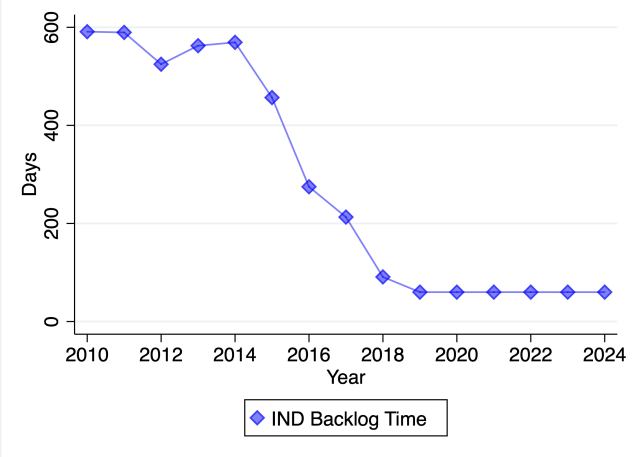


(b) Number of Scientists w. US Experience



Notes. This figure plots trends in knowledge production and international talent flows for China and the U.S. Panel (a) reports the annual number of publications (restricted to papers with PubMed IDs). Panel (b) reports the number of scientists with U.S. experience, defined as those with at least three publications under a U.S. affiliation. For China, this measure captures "returnees": scientists with at least three U.S.-affiliated publications who subsequently publish under a Chinese affiliation. Data source: OpenAlex publication database.

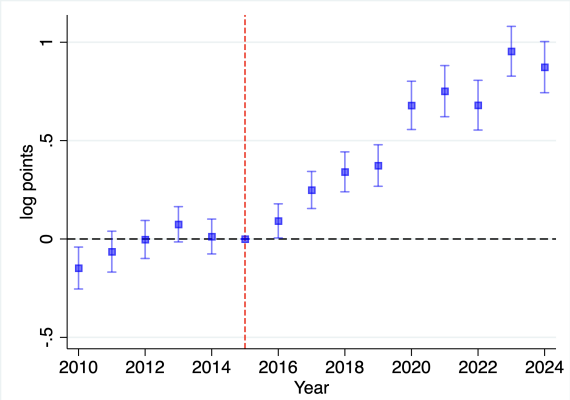
Figure 8: Mechanism - Application Backlog Time in Days



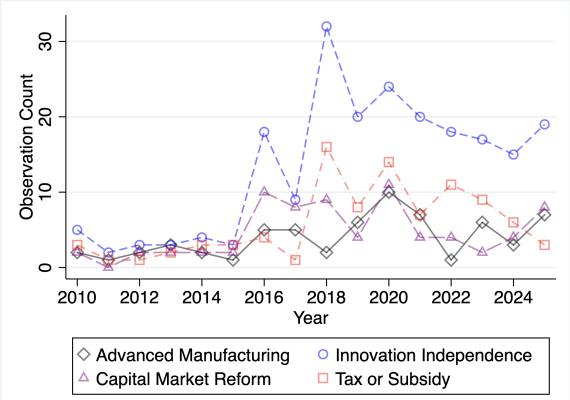
Notes. This figure plots the mean application backlog in days (i.e. approval date - application date) by year. The vertical axis reports the average regulatory backlog, while the horizontal axis indicates calendar year. Data source: National Medical Products Administration (NMPA).

Figure 9: Mechanism - "Made in China 2025" and Other Policies

(a) Exclude Trials in "Made In China 2025" Cities



(b) Number of Different Policies



Notes. Panel (a) presents an event-study analysis of clinical trial activity at the disease–country-year level, with 2015 as the base year. It replicates the baseline specification from Figure 3(a), but restricts the sample to trials that are conducted outside the "Made in China 2025" cities or city clusters, with U.S. trials as the control group. The specification includes disease-by-country and year fixed effects. Robust standard errors are clustered at the disease × country level. Panel (b) plots the annual number of policy documents, classified into four types: (i) advanced manufacturing policies promoting industrial production capacity; (ii) innovation independence policies reducing reliance on Western countries; (iii) capital market reforms alleviating barriers to fundraising; and (iv) tax or subsidy policies incentivizing innovation (not limited to pharma innovation). Data sources: Citeline TrialTrove Database and the State Council Policy Document Library.

Tables

Table 1: Number of New Drugs included in NRDL 2016-2025

Year	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	All
# Drugs Added to NRDL	3	36	17	97	119	94	121	121	91	114	813

Notes. This table reports the total number of chemical drugs and biologics included in NRDL since 2016. Data source: National Health Security Administration (NHSA) website.

Table 2: Descriptive Statistics - Number of Trials by Disease and Country (2010-2024)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	China				US			
	Phase I	Phase II	Phase III	Phase IV	Phase I	Phase II	Phase III	Phase IV
Autoimmune/Inflammation	2061	1154	731	920	1529	2730	1758	1060
Central Nervous System (CNS)	2327	508	441	2029	2091	3004	2030	2746
Cardiovascular	2726	343	534	1572	797	1053	854	1005
Genitourinary	416	102	96	428	136	188	184	288
Infectious Disease	2198	853	618	1306	1735	1853	1145	1256
Metabolic/Endocrinology	2291	462	572	1294	1296	1589	1154	905
Oncology	4293	7711	2528	1463	6579	10241	1892	348
Ophthalmology	85	92	90	142	130	499	311	204
Vaccines (Infectious Disease)	33	13	53	42	36	48	34	12

Notes. This table reports the total number of clinical trials by therapeutic area and trial phase (Phase I–IV), aggregated across 2010–2024 for China and the U.S. Data source: Citeline TrialTrove Database.

Table 3: Difference in Differences Regression - Trial Quantity and Novelty

	(1)	(2)	(3)	(4)
	(Log) Number of All Trials	(Log) Novelty Score > Median Trials	(Log) Non-Generic Trials	(Log) Phase III Trials w. Innovative Control
China × (Year ≥ 2016)	0.621***	0.460***	0.477***	0.282***
	(0.045)	(0.044)	(0.043)	(0.057)
Constant	1.251***	0.923***	1.233***	-0.608***
	(0.012)	(0.012)	(0.012)	(0.015)
Disease × Country FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Adjusted R-squared	0.864	0.829	0.862	0.512
Observations	9030	9030	9030	3990

Notes. This table reports difference-in-differences estimates of Equation (1). The key independent variable is an indicator for post-2015 years interacted with an indicator for China, with the U.S. as the control group. Dependent variable is log(all trials + 1) for Column (1), log(trials with novelty score above median + 1) for Column (2), and log(trials investigating non-generic compounds + 1) for Column (3). Column (4) restricts to Phase III trials that use a patented drug as the control arm. Because the large number of zero-valued observations in this subsample attenuates estimates (relative to the raw data shown in Figure 1(d)), we follow Chen and Roth (2024) and define the dependent variable as log(trials) when at least one trial is observed in a country-disease-year cell and as −1 otherwise. All specifications include disease-by-country fixed effects and year fixed effects. Robust standard errors are clustered at the disease × country level. Data source: Citeline TrialTrove Database.

Table 4: Mechanism Test - NRDL Intensity and Trial Counts

	(1)	(2)	(3)	(4)	(5)
	Mean of Independent Var	(Log) Number of Trials			
Log(Number of NRDL-Included Drugs)	0.428	0.164***			
		(0.038)			
Log(Revenue of NRDL-Included Drugs)	0.107		0.280***		
			(0.084)		
Ratio of NRDL-Included Drugs	0.059			0.454***	
				(0.165)	
Revenue Ratio of NRDL-Included Drugs	0.104				0.512***
					(0.143)
Disease × Country FE		Yes	Yes	Yes	Yes
Year FE		Yes	Yes	Yes	Yes

Notes. This table estimates the effect of NRDL exposure on clinical trial activity using disease-country-year level observations. NRDL exposure is measured in four ways: (1) log number of NRDL-included drugs, (2) log revenue of NRDL-included drugs, (3) share of NRDL-included drugs among all drugs in the disease area, and (4) revenue ratio of NRDL-included drugs relative to total disease-level drug revenues. The dependent variable is log(number of trials + 1). All specifications include disease-by-country and year fixed effects and control for the number of drugs in a given disease-by-country-year. Robust standard errors are clustered at the disease × country level. Data sources: Citeline TrialTrove database and Sinohealth.

Table 5: Summary of Mechanisms

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Mechanisms	NRDL	ICH Harmonization	Papers	Talents	Application Backlog	Other Policies	MadeInChina 2025	All
D.V.	(Log) Number of Trials							
Revenue Ratio of NRDL-Included Drugs	0.512***	0.542***	0.495***	0.500***	0.454***	0.502***	0.453***	0.478***
	(0.143)	(0.145)	(0.140)	(0.143)	(0.141)	(0.143)	(0.141)	(0.140)
US Prevalence (Per Hundred) × Post-ICH		0.010					0.004	0.004
		(0.007)					(0.007)	(0.008)
Number of Papers (Thousands)			0.182**				0.245**	0.243**
			(0.082)				(0.114)	(0.114)
Number of Scientists w. US Experience (Thousands)				0.024***			-0.034	-0.032
				(0.008)			(0.024)	(0.024)
Application Backlog Days (Hundreds)					-0.040***		-0.030***	-0.032***
					(0.007)		(0.007)	(0.007)
Number of Policies						0.005***	0.002**	0.002**
						(0.001)	(0.001)	(0.001)
Constant	1.299***	1.017***	1.271***	1.286***	1.376***	1.251***	1.023***	1.044***
	(0.008)	(0.140)	(0.015)	(0.009)	(0.014)	(0.013)	(0.127)	(0.129)
Disease × Country FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R-squared	0.870	0.871	0.872	0.870	0.872	0.871	0.874	0.874
Observations	9030	9030	9030	9030	9030	9030	9030	9030

Notes. This table examines the contribution of various demand- and supply-side mechanisms to clinical trial activity, using disease-country-year level observations. Column (1) replicates Column (5) of Table 4 that estimates the effect of NRDL exposure, measured as the revenue share of NRDL-included drugs relative to total disease-level drug revenues. Column (2) adds ICH harmonization, measured as the interaction between U.S. disease prevalence and a post-ICH indicator, controlling for country-specific disease prevalence. Column (3) examines the effect of scientific publications, and Column (4) the effect of returnees – scientists with US experience. Column (5) examines the trial application reform, where the key independent variable is the average backlog days (approval date minus application date) by disease category in China. Column (6) examines other policies and uses the annual number of (industrial) policies as the key regressor. Column (7) tests robustness to the Made in China 2025 initiative by excluding all trials conducted in cities or city clusters designated under the program. Column (8) includes all mechanisms jointly. The dependent variable is log(trials + 1). All specifications include disease-by-country and year fixed effects. Robust standard errors are clustered at the disease × country level. Data sources: Citeline TrialTrove, Sinohealth, OpenAlex, and State Council Policy Document Library.

Table 6: Quantification of Different Mechanisms

	Trials in 2024 – Trials in 2016	NRDL	ICH Harmonization	Knowledge & Talent	Application Reform	Other Policies	Unexplained
Oncology	1004	438.93	0.97	240.48	2.22	0.54	320.86
Autoimmune / Inflammation	609	37.26	6.27	10.78	0.88	0.14	553.66
Central Nervous System	572	47.40	2.95	38.62	0.00	0.15	482.89
Metabolic / Endocrinology	496	52.77	12.24	17.51	0.00	0.14	413.34
Cardiovascular	467	29.11	4.56	45.46	2.14	0.19	385.54
Infectious Disease	152	48.59	7.46	35.92	0.09	0.16	59.78
Genitourinary	58	1.54	1.54	3.79	0.08	0.04	51.01
Ophthalmology	22	1.89	0.32	0.28	0.00	0.01	19.50
Vaccines (Infectious Disease)	-14	0.00	0.02	0.02	0.00	0.01	-14.04
Total	3366	657.49	36.33	392.86	5.41	1.37	2272.53

Notes. This table quantifies the contribution of each mechanism to the observed increase in clinical trials in the nine therapeutic areas. The contribution of each mechanism is computed by multiplying the observed change in the mechanism by its corresponding coefficient from Table 5, scaling by the total trial counts within each disease category, and aggregating to the therapeutic-area level. The procedure follows Equation 3.

Table 7: Welfare Decomposition: Static vs. Dynamic Gains in Billions of ¥

	Static gain	Dynamic gain	
		Mean	$[P_{25}, P_{75}]$
Consumer Surplus	35.66	67.59	[11.97, 99.64]
Government Spending	36.04	37.98	[6.98, 51.68]
Producer Surplus	22.90	36.22	[7.08, 59.70]
Social Surplus	22.52	65.83	[12.07, 107.67]

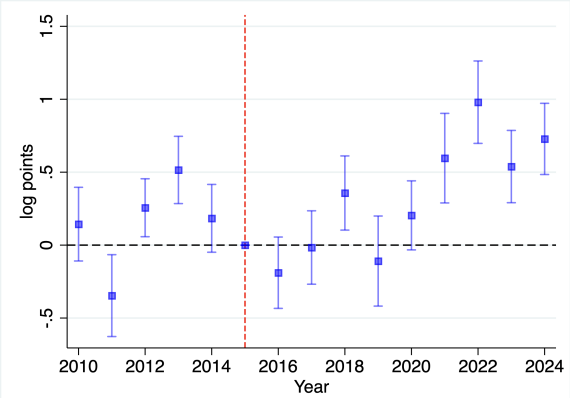
Notes: This table compares the static (short-term) and dynamic (long-run) welfare gains from the NRDL reform. Static gains reflect expanded patient access to existing drugs. Dynamic gains capture long-run welfare effects through induced pharmaceutical innovation and future drug development. We approximate dynamic gains using the mean and interquartile range (25th and 75th percentiles) of the estimated value of existing innovative cancer drugs. All values are reported in billions of ¥.

Online Appendix

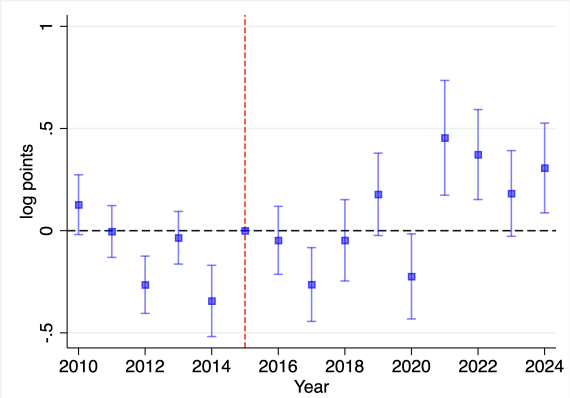
A Online Figures

Figure A1: Event Study - Drug Pipelines and New Firms

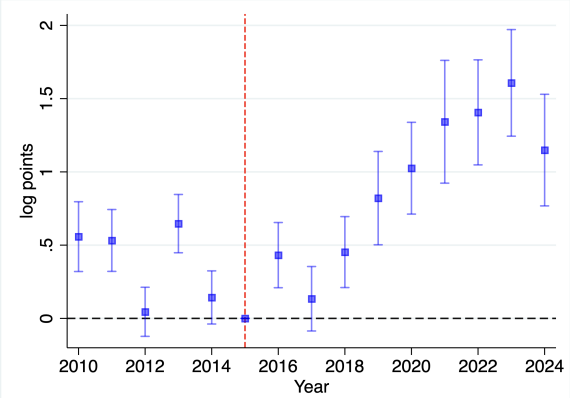
(a) Number of Licensed-Out Drugs
Chinese Firms vs US Firms



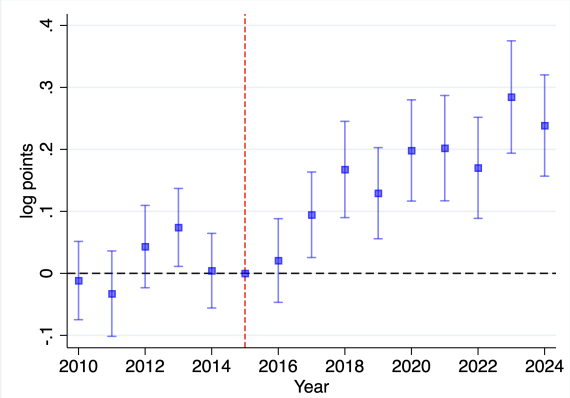
(b) First-in-world Regulatory Approval of Drugs
by Firms Headquartered in China vs US



(c) Number of Drug Approvals
by Chinese vs US Administration



(d) Number of New Firms
China vs US

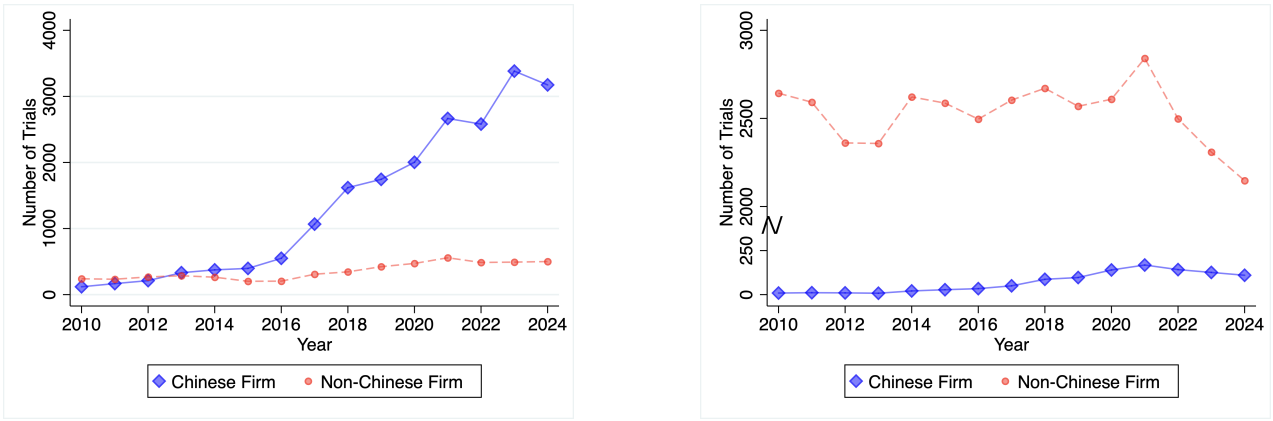


Notes. This figure presents event-study analyses of post-trial outcomes and firm entry at the disease-country-year level, with 2015 as the base year and the U.S. as the control group. Panel (a) examines out-licensing transactions to multinational corporations for drugs developed by Chinese- vs U.S.- headquartered firms. Panel (b) examines first-in-world regulatory approvals (i.e., the first approval in any country) for drugs developed by firms in each country. Panel (c) compares newly approved drugs by the Chinese and US regulatory authorities. Panel (d) compares the number of new entrants (firms conducting clinical trials for the first time) in each country. The dependent variable is $\log(y+1)$, where y is the corresponding count outcome. All specifications include disease-by-country and year fixed effects. Robust standard errors are clustered at the disease $\times$ country level. Data source: Citeline Pharmaprojects Database.

Figure A2: Time Trend - Number of Trials by Chinese Firms and Non-Chinese Firms

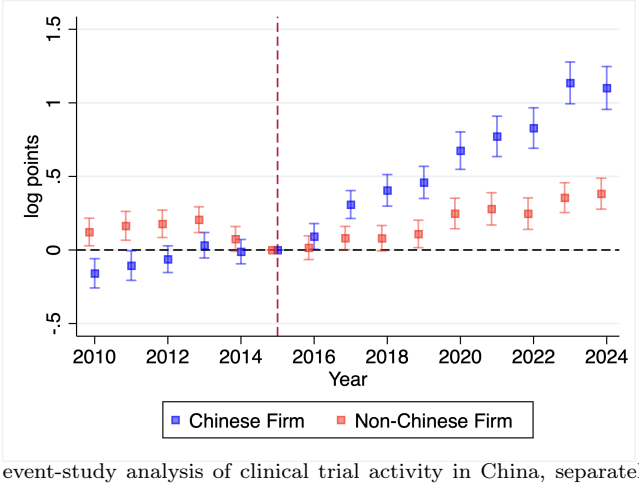
(a) Trials in China
Chinese Firms vs Non-Chinese Firms

(b) Trials in the US
Chinese Firms vs Non-Chinese Firms



Notes. This figure plots the annual number of clinical trials conducted in China and the U.S., distinguishing between Chinese and Non-Chinese firms. Trials sponsored by public entities are excluded. Panel (a) reports trials conducted in China, while Panel (b) reports trials conducted in the U.S. Data source: Citeline TrialTrove database.

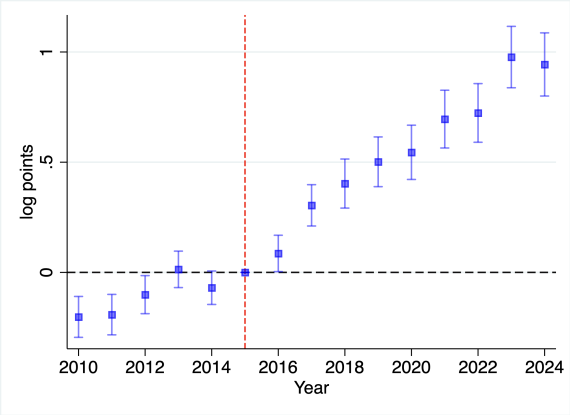
Figure A3: Event Study - Trials by Chinese Firms vs Non-Chinese Firms



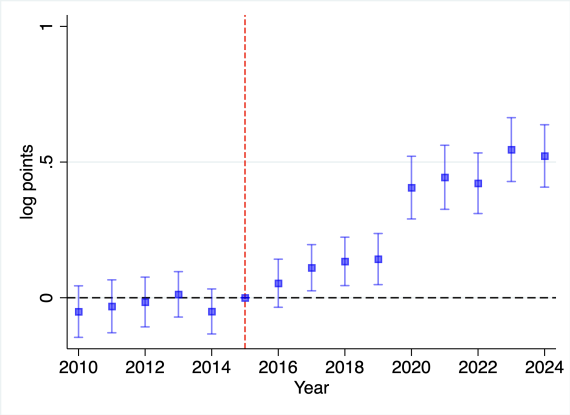
Notes. This figure presents an event-study analysis of clinical trial activity in China, separately for Chinese and non-Chinese firms, with 2015 as the base year and U.S. trials as the control group. The blue line plots the coefficients on the interactions between ChineseFirmInChina and event-year indicators, while the red line plots the coefficients on the interactions between NonChineseFirmInChina and event-year indicators. The vertical bars represent 95% confidence intervals. The dependent variable is $\log(\text{trials}+1)$. The specification includes disease-by-country and year fixed effects. Robust standard errors are clustered at the disease $\times$ country level. Data source: Citeline TrialTrove Database.

Figure A4: Event Study - Trials of Different Phases

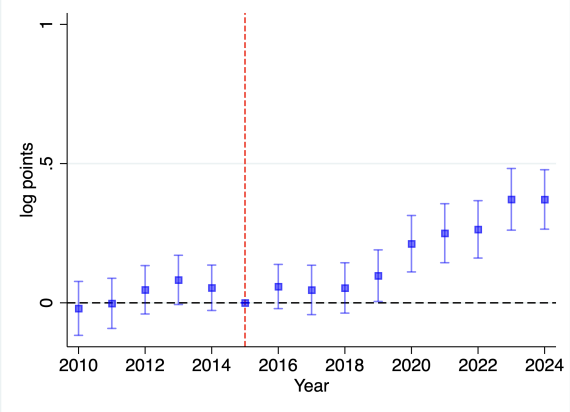
(a) Phase I



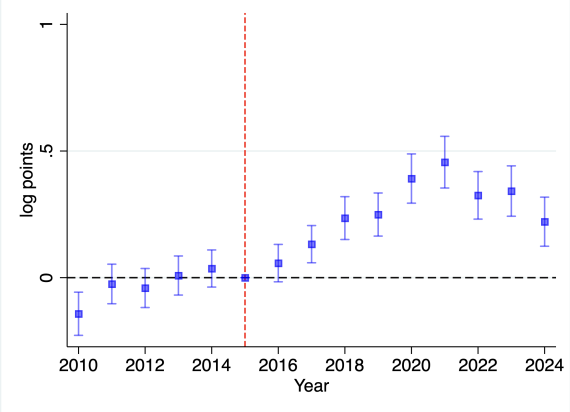
(b) Phase II



(c) Phase III

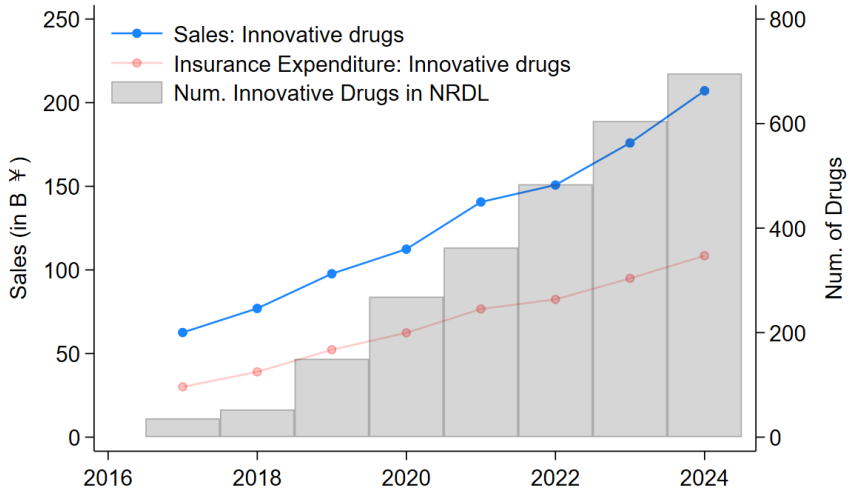


(d) Phase IV



Notes. This figure presents event-study analyses of clinical trial activity at the disease–country-year level, with 2015 as the base year and the U.S. as the control group. The vertical axis plots the estimated coefficients (in 100 log points) on the interaction term from the difference-in-differences specification, along with 95% confidence intervals. Panels (a) to (d) report results separately for Phase I to IV trials. The dependent variable is $\log(y+1)$, where y counts the number of trials launched in that year. All specifications include disease-by-country and year fixed effects. Robust standard errors are clustered at the disease $\times$ country level. Data source: Citeline TrialTrove Database.

Figure A5: Overview of NRDL and the Chinese Pharmaceutical Market



Notes. This figure reports annual sales of innovative (patented drugs) in China, public expenditure on innovative drugs through the NRDL, and the number of innovative drugs covered by the NRDL from 2017 to 2024. Data source: SinoHealth.

B Online Tables

Table A1: Number of Oncology Trials and Difference Relative to 2016

Year	2016	2017	2018	2019	2020	2021	2022	2023	2024	All
Panel A: China										
Number of Trials	759	919	1171	1442	1779	2262	1878	1942	1763	
- Difference from 2016		160	412	683	1020	1503	1119	1183	1004	7084
Panel B: US										
Number of Trials	1267	1464	1386	1343	1309	1448	1266	1199	1163	
- Difference from 2016		197	119	76	42	181	-1	-68	-104	442

Notes. This table reports the annual number of oncology trials conducted in China and the United States from 2016 to 2024. For each country, we also report the difference relative to the 2016 baseline year, calculated as the number of trials in a given year minus the number of trials in 2016. Data source: Citeline TrialTrove Database.

Table A2: Robustness - Different FEs and Functional Forms

	(1)	(2)	(3)	(4)	(5)
	(Log) Number of Trials			Number of Trials	
Model	OLS	OLS	OLS	OLS	PPML
China × (Year ≥ 2016)	0.621***	0.643***	0.757***	10.808***	1.303***
	(0.045)	(0.046)	(0.055)	(1.349)	(0.079)
Constant	1.251***	1.354***	0.719***	7.642***	2.940***
	(0.012)	(0.014)	(0.015)	(0.370)	(0.030)
Percentage Change	86.1%	90.2%	113.2%	138.0%	268.1%
Disease × Country FE	Yes	Yes	Yes	Yes	Yes
Disease × Year FE	No	Yes	No	No	No
Year FE	Yes	No	Yes	Yes	Yes
Observations	9030	9030	9030	9030	9030

Notes. This table reports robustness checks for the baseline difference-in-differences specification from Column (1) of Table 3. The key independent variable is an indicator for post-2015 years interacted with an indicator for China, with the U.S. as the control group. Column (1) replicates the baseline specification with log(number of trials + 1) as the dependent variable. Column (2) adds disease × year fixed effects. Columns (3) to (5) address zero-valued observations and functional form choices. Column (3) follows Chen and Roth (2024) and defines the dependent variable as log(number of trials) when at least one trial is observed in a country-disease-year cell and as −1 otherwise. Columns (4) and (5) use trial counts instead of log(trials+1) as the dependent variable, estimated via OLS and Poisson pseudo-maximum likelihood (PPML), respectively. Robust standard errors are clustered at the disease × country level. Data source: Citeline TrialTrove Database.

Table A3: Robustness - Different Treatment or Control Groups

	(1)	(2)	(3)	(4)
	(Log) Number of Trials			
China × (Year ≥ 2016)	0.621***	0.619***	0.643***	0.630***
	(0.045)	(0.046)	(0.044)	(0.045)
Constant	1.251***	1.224***	1.184***	1.432***
	(0.012)	(0.013)	(0.012)	(0.012)
Disease × Country FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Exclude Trials Conducted in US & China	No	Yes	No	No
Control Group	US	US	EU	EU+US
Adjusted R-squared	0.864	0.862	0.867	0.882
Observations	9030	9030	9030	9030

Notes. This table reports robustness checks using alternative samples and control groups. The key independent variable is an indicator for post-2015 years interacted with an indicator for China. Column (1) replicates the baseline specification from Column (1) of Table 3. Column (2) excludes trials conducted in both the U.S. and China. Column (3) uses trials conducted in Europe as the control group, and Column (4) uses trials conducted in either Europe or the U.S. as the control group. The dependent variable is log(number of trials + 1). All specifications include disease-by-country fixed effects and year fixed effects. Robust standard errors are clustered at the disease × country level. Data source: Citeline TrialTrove Database.

Table A4: Heterogeneity - Different Sponsor Types

	(1)	(2)	(3)	(4)
	(Log) Trials by			
	Any Sponsors	Private Sponsors	Public Sponsors	Hybrid Sponsors
All Trials				
China × (Year ≥ 2016)	0.621***	0.575***	0.393***	0.327***
	(0.045)	(0.048)	(0.035)	(0.031)
Phase-I Trials				
China × (Year ≥ 2016)	0.667***	0.630***	0.153***	0.185***
	(0.050)	(0.050)	(0.023)	(0.024)
Phase-II Trials				
China × (Year ≥ 2016)	0.332***	0.186***	0.245***	0.163***
	(0.038)	(0.033)	(0.030)	(0.022)
Phase-III Trials				
China × (Year ≥ 2016)	0.165***	0.128***	0.073***	0.044***
	(0.031)	(0.029)	(0.016)	(0.011)
Phase-IV Trials				
China × (Year ≥ 2016)	0.295***	0.054***	0.261***	0.117***
	(0.028)	(0.017)	(0.026)	(0.015)
Disease × Country FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Observations	9030	9030	9030	9030

Notes. This table examines heterogeneity across trial phases and sponsor types. The key independent variable is an indicator for post-2015 years interacted with an indicator for China, with the U.S. as the control group. The dependent variable is log(number of trials + 1), where trial counts are computed by sponsor types and trial phase within each disease-country-year cell. Public sponsors are trial sponsors that are public organizations, such as hospitals, universities, or government agencies, while private sponsors are trial sponsors that are firms. Trials with hybrid sponsors are trials sponsored jointly by public organizations and firms. All specifications include disease-by-country fixed effects and year fixed effects. Robust standard errors are clustered at the disease × country level. Data source: Citeline TrialTrove Database.

C Trial Novelty Calculation

We construct measures of trial novelty in several ways. First, we calculate a novelty score based on the extent to which a target has not yet been explored, following Dranove et al. (2022). Second, we use GPT to classify a trial's objective into one of four categories: investigation of a new drug, a new combination of existing drugs, a new indication for an existing drug, or a bioequivalence test for a generic drug. Third, for phase III trials, we further classify the control group as an innovative drug, a generic drug, or a placebo. This appendix provides the technical details underlying the construction of these measures.

C.1 Novelty Score

The first novelty measure follows the approach proposed by Dranove et al. (2022), adapted to the structure of the Citeline database. Whereas Dranove et al. (2022) constructs novelty using "target-based actions" (TBA) from the Cortellis database, our data contain mechanisms of action (MOA), which are conceptually similar. MOAs describe the specific scientific strategy through which a molecule produces a pharmacological effect. Each MOA corresponds to a pair consisting of the biological entity targeted by the molecule (e.g., p38 MAP kinase) and the mechanism used to modulate that entity (e.g., inhibition). In total, there are 4,053 unique MOAs in the data. MOA information is available for 379,993 out of 463,412 trials in the database, covering approximately 82% of all trials. These trials include not only those conducted in China and the United States but also trials conducted in other countries, ensuring that the novelty measure reflects the global frontier of drug development. For the 107,256 trials conducted in China and the United States between 2010 and 2024 (our main analysis sample), the required information is available for 103,959 trials.

We compute translational novelty at the trial level by assessing the extent to which the MOA deployed in a given trial has previously been used. Consider a trial i that deploys a single MOA k. Let $D_{k(i)}$ represent the number of previous deployments of k in stages that are at least as advanced as that of trial i. For example, if i is a Phase II trial, $D_{k(i)}$ would represent the count of other trials associated with k that have been tested in Phases II or III before the initiation of trial i, that is, Phase I deployments of k are not counted. This restriction follows Dranove et al. (2022): evidence generated at more advanced stages (e.g., Phase III) is more informative about translational potential

than evidence generated earlier, and knowledge flows are asymmetric (Phase III → Phase I, but not vice versa).

The translational novelty score for trial i is computed as:

$$\text{Novelty}_i = \frac{1}{1 + D_{k(i)}} \in (0, 1]$$

Thus, the score equals one only if trial i deploys k for the first time ($D_{k(i)} = 0$). If $D_{k(i)} = 1$, the score is 0.5; if $D_{k(i)} = 2$, the score is 0.33; and so on. That is, the score decreases as the number of k's previous trials increases, approaching zero in the limit. Many trials deploy multiple MOAs. In such cases, we compute a trial-level novelty score by taking the average of the MOA-specific novelty values for that trial:

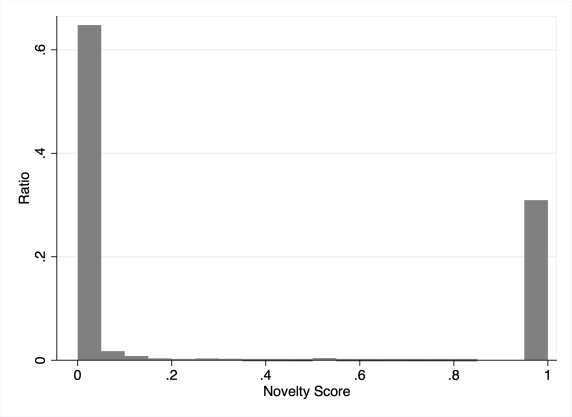
$$\text{Novelty}_i = \frac{1}{K_i} \sum_{k \in \text{MOAs}(i)} \frac{1}{1 + D_{k(i)}},$$

where K_i denotes the number of MOAs associated with trial i. Averaging yields a conservative measure of novelty for multi-target trials, balancing the contribution of highly novel targets and more established ones.

It is important to note that, except for trials with no prior deployments, our inference will not exploit information contained in the specific value of scores and instead relies on the implied ordering.

Figure A6 plots the distribution of the novelty scores, which is highly right-skewed: the mean novelty score is 0.32, while the median is close to zero, reflecting that most trials rely on previously well-explored mechanisms. This distribution closely matches the pattern documented in Dranove et al. (2022), despite relying on MOAs from Citeline rather than TBAs from Cortellis. To classify novel trials, we first compute the median novelty score for each year, pooling trials across all countries. We then designate a trial as "novel" if its novelty score exceeds the median in its corresponding year.

Figure A6: Distribution of Novelty Score (Dranove et al. (2022))



Notes. Data source is Citeline TrialTrove Database.

C.2 GPT-classified Novelty Objective and Control Group

We also construct an innovation classifier using GPT to determine whether a clinical trial investigates (i) a new drug, (ii) a new combination of existing drugs, (iii) a new indication for an existing drug, or (iv) a bioequivalence test for a generic drug. This measure allows us to capture innovative content that is not directly observable from structured fields in Citeline. The classifier takes as input a rich set of trial-level information—including the trial title, phase, primary tested drug, mechanism of action, trial objective, primary endpoint, study design, treatment plan, and investigator notes. An illustrative example of the variables fed into the model is shown as follows.

```
TrialID: 322799

StartDate: 2015-07-06

TrialTitle: Intratumoral Phase Ib Clinical Trial Of IIb, IIIc, And IV M1a,
M1b Malignant Melanoma With Recombinant Human GM-CSF Herpes Simplex Virus
Injection (OrienX-010)

TrialPhase: Phase I (Ib)

PrimaryTestedDrug: OrienX-010

PrimaryTestedDrugMechanism: Immuno-oncology therapy

TrialObjective: To evaluate the safety and tolerability of continuous
administration of OrienX-010 intratumoral injection in patients with stage
IIIb, IIIc, and IV M1a/M1b malignant melanoma who failed standard therapy,
and to evaluate the biodistribution characteristics and biological effects
of continuous administration of OrienX-010 intratumoral injection in the
same patient population.
```

```
PrimaryEndpoint: Adverse events, immunogenicity, neutralizing antibody
response, safety and tolerability

StudyDesign: Single-arm, non-randomized, open-label study

TreatmentPlan: Recombinant human GM-CSF herpes simplex virus injection,
intratumoral administration, specification 2.0-8.0 × 10^7 PFU (1.0 ml) per
injection site, injection every two weeks for 16 consecutive weeks, maximum
dose not exceeding 10 ml (8 × 10^8 PFU), adjusted according to lesion size

TrialNotes: Registry ChinaDrugTrials; registration number CTR20150881;
sponsor Beijing Aoyuan Heli Biotechnology Co., Ltd.; recruitment completed;
first informed consent signed on 2015-07-06; trial end date 2017-05-24
```

However, large language models can hallucinate or infer information that is not explicitly present in the input, which is problematic when the classification must be strictly tied to verifiable trial attributes. To reduce hallucinations and ensure consistency, we develop a supervised fine-tuned model based on GPT-4o-mini. The fine-tuning process proceeds in three steps.

First, we randomly sample 200 clinical trials from the full dataset and manually label each trial's treatment group, control group, and innovation category using the detailed textual descriptions described above. Two healthcare domain experts independently label all trials and reconcile discrepancies, ensuring reliability and high inter-annotator agreement. Second, we fine-tune GPT-4o-mini on the OpenAI platform using these 200 manually verified examples as the training set.[33] Fine-tuning forces the model to anchor its reasoning on the specific structure and terminology of Citeline data rather than relying on external assumptions, which substantially reduces hallucination and improves the reproducibility of trial-level classifications. Third, we deploy the fine-tuned classifier across the full set of clinical trials in our sample, generating for each trial: (i) a predicted innovation category, (ii) the treatment drug, and (iii) the control drug. This produces a scalable measure of whether a trial contains at least some innovative content.

The prompt used for classification is provided below. In total, the model successfully classified all 107,256 trials, of which 35% were labeled as new-drug trials, 23% as new-combination trials, 32% as new-indication trials, and 10% as generic bioequivalence trials.

```
# GPT prompt
message_content = (
    "We provide the trial titles, targeted drug information, and summary below "
```

[33] Additional fine-tuning documentation: https://platform.openai.com/docs/guides/supervised-fine-tuning

```
    "in the format of information name: information, seperated by ';'. "
    "For each trial, follow a step-by-step process: "
    "Step 1: Answer the question: What drugs are used in the treatment versus "
    "control arms of the clinical trial described below? "
    "Extract treatment versus control arms of the clinical trial, If you do "
    "not find treatment drug or control drop, just say 'n/a'. "
    "If multiple drugs are in the treatment group, say 'drugAname+drugBname'. "
    "For durg with both code names and branded names, list both "
    "(branded names in brackets)"
    "Step 2: Answer the question: what type is each trial? Classify the "
    "treatment drugs into one of the four catogories:"
    "'new drug', 'existing drug new indication', "
    "'existing drug new combination', 'generic drug'. "
    "'new drug' is a trial on a drug that is not yet on the market"
    "'existing drug new indication' is a investigatory trial to test whether "
    "an exisiting drug can treat a new disease"
    "'existing drug new combination' is a trial to test whether drugs can be combined"
    "'generic drug' is a trial to test bioequivalence of a drug to an exisiting drug"
    "Put all the information in two steps on one line, in the following format: "
    "'Treatment: [treatment drug]. Control: [control drug]. "
    "Classification: [treatment drug catogeory]'  "
    f"{text}"
)

# Send message_content to the GPT model and retrieve output
completion = call_gpt_api(
    client=client,
    model=MODEL,
    messages=[
        {"role": "system", "content": "You are a healthcare economist"},
        {"role": "user", "content": message_content},
    ],
    temperature=0.5
)

gpt_output = completion.choices[0].message.content.strip()
```

Moreover, Phase III trials are typically designed with both a treatment arm and a control arm. The control arm may involve an innovative drug, a generic drug, or a placebo. Trials that use an innovative drug as the control are benchmarked against the existing clinical frontier and therefore represent a higher degree of novelty than those using generics or placebos. We classify the control arm into these three categories—innovative drug, generic drug, or placebo—based on the treatment drug name, the control drug name, and the trial start date. The prompt used for this classification is provided below. Among Phase III trials in our sample, 17.77% use an innovative drug as the control, 12.04% use a generic drug, and 70.19% use a placebo.

```
# GPT prompt
message_content = (
```

```
    f"You will be given one clinical trial record with (i) start date and
    (ii) control group drug(s).\n\n"
    f"Your task is to determine whether the CONTROL group drug(s) should be
    considered a 'generic drug' or an 'innovative drug' "
    f"AT THE TIME of the start date.\n\n"
    f"Use your background pharmaceutical knowledge to judge whether the
    control drug is typically an "
    f"originator/brand/innovator product or a generic version, and consider
    whether the timing (start date) "
    f"matters for generic entry.\n\n"
    f"Definitions (for the control drug at the start date):\n"
    f"- generic drug: a non-originator generic version of an existing drug.\n"
    f"- innovative drug: originator/brand/innovator product (including
    biologics and patented products).\n\n"
    f"IMPORTANT RULES:\n"
    f"- Base your judgment on the situation at the start date, not today.\n"
    f"- If multiple control drugs exist, classify based on the main comparator.\n"
    f"- Always choose one of: generic drug, innovative drug. Do not output 'unknown'.\n"
    f"- Do NOT output anything except the final answer line.\n\n"
    f"Output format (ONE LINE ONLY):\n"
    f"id: {row['id']}; control_type: <generic drug OR innovative drug>\n\n"
    f"Input:\n"
    f"startdate: {row.get('StartDate', 'n/a')}\n"
    f"control_group: {row.get('control', 'n/a')}\n"
)
completion = call_gpt_api(
    client=client,
    model=MODEL,
    messages=[
        {"role": "system", "content": "You are a healthcare economist"},
        {"role": "user", "content": message_content},
    ],
    temperature=0.5
)
```

D Auxiliary Data Cleaning and Variable Construction

D.1 Papers and Scientists

OpenAlex provides comprehensive, high-quality metadata on scientific publications, including publication year, author names, affiliations, citations, journal identifiers, and concept tags that summarize a paper's scientific content (Priem et al., 2022). To study knowledge production relevant to biopharma innovation, we focus on papers whose first or last authors are affiliated with institutions in China or the United States and restrict the sample to publications with valid PubMed IDs to ensure biomedical relevance. In total, there are 33,809,530 OpenAlex publications with PubMed IDs. Among these, 9,734,217 are affiliated with the U.S., and 2,321,944 are affiliated with China.

After each paper is linked to one or more Citeline disease categories,[34] we construct disease–year level measures of knowledge quantity and international talent flows. These measures aggregate paper-level or author-level information to the disease-country-year unit of analysis used in the main text.

Knowledge quantity (number of papers). Knowledge quantity is defined as the total number of PubMed-indexed papers in a given disease-country-year. If a paper maps to multiple diseases, we assign fractional counts based on the number of disease assignments. The distribution of paper counts varies substantially across diseases, with a median of 31 papers and a mean of 199 papers at the disease-country-year level.

Talent flows (weighted number of scientists with U.S. experience). Talent flows capture the presence of researchers with international training or collaboration. A scientist is classified as having U.S. experience if they have published at least three paper with a U.S. affiliation on or before year t. We then compute, for each disease–year, the weighted sum of such scientists, where each scientist's weight equals the share of their total publications associated with that disease category. For example, if 50% of a scientist's publications map to Disease A and 20% to Disease B, they contribute weights of 0.5 and 0.2, respectively. There are 7,552,293 unique scientists in OpenAlex, and among them, 2,539,289 have prior experience in the United States.

Together, these measures allow us to quantify how scientific knowledge and international talent evolve across diseases and over time, and to examine whether these upstream factors help explain the rise of clinical trial activity documented in the main analysis.

D.2 Application Backlog Time

The clinical trial application reform in 2015, studied in Jia et al. (2024), streamlined regulatory procedures by clearing the backlog of pending IND (investigational new drug) applications, which are required before initiating clinical trials for new drugs under investigation. To calculate application backlog days across diseases and over time, we collect the application and approval dates for almost all IND applications from the Center for Drug Evaluation of China's NMPA. The detailed data collection and cleaning process is as follows.

First, we collect data on all IND applications from the Center for Drug Evaluation (CDE) of China's National Medical Products Administration (NMPA). The NMPA publicly discloses all drug

[34] Details on the linkage can be found in Appendix E.2.

applications on its website,[35] including information such as the *Application ID*, *Drug Name*, *Drug Type*, *Application Type*, *Registration Category*, *Company Name*, and *Submission Date*. In addition, the NMPA provides the approval date for each application on a separate webpage,[36] which we merge to construct the full application and approval timeline.

Second, we restrict the sample to applications whose fourth character is the letter "L," indicating clinical trial applications, and whose "Application Type" is 1 or 2, which correspond to innovative drugs and are therefore relevant for our focus on IND applications. Moreover, in July 2018, the NMPA introduced the implied approval (silent approval) system, under which applications submitted for more than 60 days are automatically approved. This policy change is reflected in the data: after its introduction, most applications no longer received mailed approval documents, indicating that the silent approval mechanism was effectively in operation. However, because some applications were still rejected, we additionally scraped the database of silent approvals and matched entries from this database to the main dataset using the application acceptance number.

Finally, we link these IND drugs to their corresponding therapeutic areas and disease categories. We first use large language models to classify diseases based on drug names, company information, and drug types. We then search the drug names in the Citeline Pharmaprojects database, which provides drug name histories, corresponding diseases, and textual descriptions. We feed this additional information back to a large language model to perform reclassification. This procedure yields 15,079 IND applications with classified diseases and calculated application backlog days.

```
# GPT prompt
def classify_direct(drug_name, drug_type, company):
    prompt = f"""
        You are a senior pharmaceutical and clinical indication expert.
        Please determine the therapeutic area and disease category based on
        the drug name I provide, and return the confidence level of this judgment.
        First choose ONE therapeutic_area.
        Then choose ONE disease_category ONLY from that area.

def classify_with_evidence(drug_name, drug_type, company, candidates):
    evidence = "\n\n".join(f"[{i+1}] {c['summary']}" for i, c in enumerate(candidates))
    prompt = f"""
        You are a senior pharmaceutical and clinical indication expert.
        Please determine the therapeutic area and disease category based on
        the drug name and evidence I provide, and return the confidence level
        of this judgment.
```

[35] https://www.cde.org.cn/main/xxgk/listpage/9f9c74c73e0f8f56a8bfbc646055026d

[36] https://www.nmpa.gov.cn/zwfw/zwfwpjfbzs/pjfbzsyp/index.html

```
        First choose ONE therapeutic_area.
        Then choose ONE disease_category ONLY from that area.
```

D.3 Policy Documents and Classification

We also scraped all biopharma-related policies in China from the State Council Policy Document Library. There are 840 policy documents relevant to biopharma innovation. We use GPT to classify these documents into four categories: (a) advanced manufacturing policies that promote the scale-up and modernization of industry production capacity; (b) innovation independence policies that aim to strengthen domestic R&D capabilities and reduce reliance on foreign technologies; (c) capital market reform policies that lower financing barriers and improve firms' access to external funding; and (d) tax and subsidy policies that provide fiscal incentives to support innovation. The detailed steps are as follows:

To construct this dataset, we first scraped all policy documents available in the State Council Policy Document Library, obtaining 18,488 entries. We then implemented a two-stage classification procedure using GPT-4o. In the first stage, we used a broad screening prompt to filter out all policies potentially related to the four categories of innovation policy described above. This step substantially reduced the number of documents requiring manual or fine-grained review. In the second stage, we applied a more detailed prompt to each remaining document, asking the model to assign (i) a relevance score capturing how closely the policy pertains to biopharma innovation and (ii) a brief explanatory note summarizing the policy's purpose and mechanism. The first stage returned 3,474 policy documents relevant to innovation. These documents were further scored in the second stage. Among them, 840 are related to biopharma innovation, and 273 have a relevance score greater than 0.5. These outputs were used both to validate relevance and to facilitate the final four-way policy categorization described above.

The two prompts used in this procedure are provided below:

```
# GPT prompt
SYSTEM_PROMPT = (
    "You are a careful Chinese policy classification expert."
    "You will receive UP TO 10 Chinese policy documents at once."
    "For EACH document, you must extract:"
    "1) the policy date (政策日期) from the text "
    "(any explicit date of issuance/publishing/promulgation; "
```

```
    "return as the original string; if none, return \"\")."
    "2) whether it clearly relates to these 4 themes (0/1):"
    "   - high_end_mfg: 高端制造业/高端装备/战略性新兴产业/高端生物医药制造 等。"
    "   - innovation_independence: "
    "     自主创新/科技自立自强/提升原始创新能力/创新平台与人才基地建设 等。"
    "   - capital_market_reforms: 支持通过资本市场/多层次资本市场/股权基金/债券/上市/并购/"
    "     融资工具来促进上述产业或创新的政策。"
    "   - tax_subsidy: 对医药/药物研发/临床试验/"
    "     创新平台等提供的减税、税收优惠、财政补贴、进口免税、费用减免、贷款贴息等。"
    "Return STRICT JSON ONLY, as a list of objects, in the SAME ORDER as input."
)
```

```
# GPT prompt
SYSTEM_PROMPT = (
    "You are a senior policy evaluation expert "
    "working inside a pharmaceutical company in China."
    "You will be shown Chinese policy documents."
    "Your task is to judge **specifically from the perspective "
    "of a drug/biopharmaceutical R&D enterprise** whether this policy creates "
    "incentives to invest more in drug R&D."
    "When you evaluate, consider the following channels:"
    "1) The policy directly reduces R&D cost "
    "(税收优惠、财政补贴、进口免税、费用加计扣除、贷款贴息等)."
    "2) The policy increases the expected revenue/profit from new drugs "
    "(医保支付、带量采购配套激励、市场准入加速、知识产权保护等)."
    "3) The policy expands or improves the domestic human capital / innovation"
    "ecosystem for biopharma (人才计划、平台建设、国家重点实验室、创新载体、医工结合)."
    "4) The policy explicitly includes '高端生物医药制造', '生物医药', '医药产业' "
    "within broader high-end manufacturing / strategic emerging industries "
    "and provides concrete support measures."
    "Important:"
    "- Some policies talk about 'high-end manufacturing', "
    "'strategic emerging industries', or 'innovation-driven development'. "
    "You must judge whether biopharma / drug R&D is "
    "**actually covered or very likely covered** by that scope."
    "- If the policy is very general, very macro, or only about administration"
    "without concrete support instruments, the incentive should be low."
    "- Output a score between 0 and 1 (inclusive). "
    "0 = no meaningful incentive for drug R&D. 1 = very strong and direct incentive. "
    "Values in between (e.g. 0.3, 0.5, 0.8) represent your strength of belief."
    "Return STRICT JSON ONLY, in a list of objects, in the same order as input."
)
```

E Linkage across Datasets

E.1 Linking Citeline Data to SinoHealth and GBD Data

Citeline provides detailed disease information for each clinical trial and drug project, covering nine therapeutic areas and 332 granular disease categories. Therapeutic areas include Oncology, Autoimmune or Inflammation, Central Nervous System (CNS), Metabolic Endocrinology, Cardiovascular,

Infectious Disease, Genitourinary, Ophthalmology, and Vaccines. Disease categories, for example, are Non–Small Cell Lung Cancer, Type 2 Diabetes, Rheumatoid Arthritis. In contrast, the SinoHealth sales data report molecule names (and ATC classification) and sales volumes, but does not contain disease classifications. There are 1,377 unique molecules. To merge the two datasets, we perform a multi-step mapping procedure that combines structured ATC information with LLM-assisted classification and expert review.

First, we construct candidate disease information for each drug. For every drug in the SinoHealth dataset, we extract (i) the drug name, including all synonyms, and (ii) its ATC (Anatomical Therapeutic Chemical) classification. We also assemble the complete list of 332 disease categories from Citeline, which will serve as the candidate set to which each SinoHealth drug can be assigned. Second, we use GPT-4o to generate disease assignments based on drug profiles. We provide GPT-4o with the drug name, its ATC code, and the full set of disease categories. The model is instructed to determine the most likely disease indications for each drug and to output at most five disease categories in ranked order of relevance. This step leverages GPT's ability to interpret pharmaceutical naming conventions, ATC structure, and therapeutic context. Third, we implement manual clinical validation by domain experts.

One drug might be linked to several diseases. For instance, a typical drug is mapped to three disease categories, reflecting the common presence of multiple approved or off-label indications. We divide the drug sales equally across the linked diseases.

Together, this pipeline enables us to reliably map sales data to disease categories at a level of granularity comparable to Citeline, allowing us to measure disease-specific market expansion and relate it to clinical trial activity. We apply the same pipeline to link Citeline data with disease prevalence measures from the Global Burden of Disease database.

The prompt used for GPT-4o is as follows:

```
# GPT prompt
message_content = (
    "We provide information about drugs and diseases. "
    "For each drug, we include details in Chinese, such as the drug name "
    "and its ATC classification. "
    "Please primarily rely on the most granular information available, "
    "such as the fields ‘shiyingzheng_fenlei’ and ‘atc4_fenlei’, "
    "when making your judgment. "
    "For diseases, we provide a list of disease names (i.e., category) "
```

```
    "and their category in a format starting with a numeric disease id. "
    "Please determine which disease(s) this drug is likely used to treat, "
    "based on the drug's information and your knowledge. "
    "Select AT MOST 5 relevant diseases from the disease list, "
    "ranked in descending order of relevance. "
    "Please ONLY output from the disease list in its original format. "
    "Do NOT include any other information. "
    "Return in the format starting with: 'Answer:' followed by the diseases "
    "separated by periods. "
    "Example output: "
    "Answer: 1-(therapeutic area: ...); (disease category: ...). "
    "2-(therapeutic area: ...); (disease category: ...)."
    " "
    f"{text}"
)

# Send message_content to the GPT model and retrieve output
completion = call_gpt_api(
    client=client,
    model=MODEL,
    messages=[
        {"role": "system", "content": "You are a healthcare economist"},
        {"role": "user", "content": message_content},
    ],
    temperature=0.5
)

gpt_output = completion.choices[0].message.content.strip()
```

E.2 Linking Citeline Data to OpenAlex Data

Linking OpenAlex publications to Citeline disease categories is challenging for two main reasons. First, the scale of the data makes direct paper-level classification infeasible. Our sample contains 33,809,530 publications, and mapping each paper's abstract directly to diseases using large language models would impose a computational burden. We therefore rely on OpenAlex concepts as an intermediate layer.

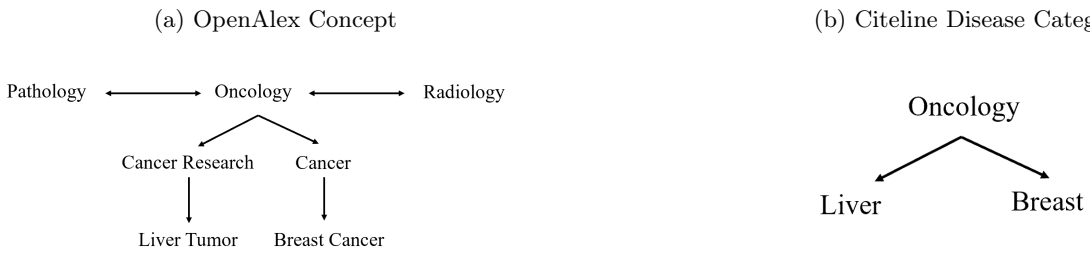
Second, the two databases use entirely different classification systems. OpenAlex concepts are algorithm-generated scientific keywords (e.g., "protein folding," "mental health," "immunotherapy"), whereas Citeline assigns each clinical trial or drug project to a structured hierarchy of nine major therapeutic areas and 332 granular disease categories. Moreover, the relationship between OpenAlex concepts and Citeline disease categories is inherently many-to-many. Each paper is associated with multiple concepts, concepts exist at different levels of abstraction, and a given concept may be relevant to multiple disease categories, while each disease typically corresponds to many distinct concepts. Finally, many concepts are semantically ambiguous or not directly tied to a specific disease,

so surface-level keyword matching fails to capture their relevance.

To bridge these systems systematically and avoid ad hoc manual matching, we construct a unified semantic space, using graph embedding, in which both OpenAlex concepts and Citeline disease categories can be jointly represented and compared.

First, we build graphs separately in OpenAlex and Citeline. In OpenAlex, we extract all concepts associated with biomedical papers. Separately, we build an analogous graph for Citeline diseases, where nodes represent therapeutic areas or disease categories and edges represent their structured parent–child hierarchy (e.g., "oncology → non–small cell lung cancer"). An example can be found in Figure A7.

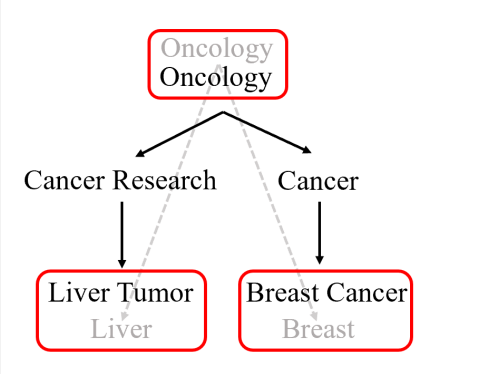
Figure A7: Example Graph - OpenAlex Concepts and Citeline Disease Categories



Notes. OpenAlex provides two types of relationships among concepts. The first is ancestor–descendant relationships (directed edges), which capture hierarchical scientific topics (e.g., "oncology → breast cancer"). The second is related relationships (undirected edges), which link concepts that frequently appear together in the scientific literature, such as oncology and radiology. We construct a graph in which each concept is represented as a node, and each hierarchical or related relationship forms a connection between nodes.

Second, we embed concepts and diseases into a vector space using *node2vec* (Grover and Leskovec, 2016). We convert each graph into a high-dimensional vector representation using the node2vec algorithm, which embeds nodes so that concepts frequently appearing together or closely connected in the graph lie near each other. Disease categories in Citeline are similarly embedded based on their hierarchical structure. This produces two embedding spaces: one for OpenAlex concepts and one for Citeline diseases. Embeddings learned from different graphs of databases (i.e., OpenAlex and Citeline) are not directly comparable because their coordinates are identified only up to an arbitrary rotation. We therefore align the two embedding spaces using a set of anchor pairs to provide common reference points across graphs. This alignment preserves within-graph structure while making cross-graph similarity measures interpretable. One example is shown in Figure A8.

Figure A8: Example - Alignment of Two Embedding Spaces



Note: In this example, *Liver Tumor* and *Breast Cancer* are OpenAlex concepts under the broader concept *Oncology*, while *Liver* and *Breast* are Citeline disease categories under the therapeutic area *Oncology*. We treat (*Liver Tumor*, *Liver*) and (*Breast Cancer*, *Breast*) as anchor pairs. These anchors provide shared reference points that allow us to rotate the two embedding spaces into a common coordinate system. After alignment, nodes that are close to these anchors in their respective graphs—illustrated by the local structure around *Oncology* and its neighbors—become comparable across databases, so that cross-graph similarity between a Citeline disease category and an OpenAlex scientific concept has a consistent interpretation.

Lastly, we link OpenAlex concepts and Citeline diseases using the anchor pairs we constructed. For each Citeline disease category, we search OpenAlex concepts for lexical matches, synonyms, or closely related terms, yielding 15,958 initial linkages across the 332 categories. Each linkage is then evaluated through two layers of validation: manual review by research assistants and a GPT-4o assessment of semantic similarity. After filtering, 12 disease categories had no suitable matches, and 23 had only one; for these 35 cases, research assistants identified additional synonyms or clinically equivalent concepts. Following these adjustments, all disease categories received valid OpenAlex concept matches.

After this process, we retrieve the embeddings of all concepts associated with each publication, align them using the anchor pairs, and compute similarity scores. Specifically, for each publication, we first obtain its concept embeddings and align them via the anchor pairs. We then compute a weighted average of these concept embeddings using the associated confidence scores. The resulting publication-level vector is compared with all disease vectors using cosine similarity, and the five diseases with the highest similarity scores are selected. About 60% of publications have at least one disease whose cosine similarity with the publication-level vector is greater than or equal to 0.5.

```
prompt = f"""
    You are a medical terminology audit expert. Please evaluate the relationship
    between a disease category and a medical concept, using the following criteria:

    - If the concept is a broader or more general area (a hypernym) of the
    category → Output: 0.2, Reason: concept is broader
    - If the concept is a narrower or more specific instance (a hyponym) of the
    category → Output: 0.8, Reason: concept is more specific
    - If the concept is at the same level (e.g., treatment method, diagnostic
    tool, body location) → Output: 0.5, Reason: same-level concept

    Output strictly in the following format:
    Match Score: [0.8 / 0.5 / 0.2]
    Reason: [in Chinese, 10-20 characters]

    therapeutic_area: {area}
    disease_category: {cat}
    concept: {concept}
"""
```