Does AI Accelerate Drug Development:

A Global Perspective

Preliminary draft; please do not cite or circulate.

Panle Jia Barwick Hongyuan Xia Tianli Xia Jing Yi *

October 2026

Abstract

Does artificial intelligence change how new drugs are discovered and developed? We assemble a global dataset linking investigational drug projects in the United States, Europe, and China from 2010 to 2025 to firms' AI human capital, AI-related biomedical research, and biological evidence on molecular targets. The relationship between AI employment and drug development is stable before 2018 and strengthens thereafter. Doubling a firm's AI workforce is associated with 15 percent more new drug projects and a 4 percentage point increase in the probability that a preclinical candidate advances to Phase I, from a baseline of 22 percent. These findings are robust to instrumental-variable estimation. The gains are concentrated among established firms, are most closely associated with AI applications in drug discovery, and are larger where complementary scientific knowledge and firm experience are abundant. AI broadens the range of diseases, modalities, and targets firms pursue, but we find limited evidence of greater drug novelty, faster development, or improved progression in later clinical stages. A counterfactual analysis attributes approximately 1,200 additional Phase I entries to AI between 2018 and 2025, about 60 percent from higher candidate progression rates rather than more candidates.

*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. Yi: College of Economics and Management, China Agricultural University, jing.yi@cau.edu.cn.

1 Introduction

Few forms of innovation match new medicines in economic scale or in contribution to social welfare. Health spending amounts to roughly 10 percent of world GDP, or $9.8 trillion in 2022 (World Health Organization, 2024), and spending on medicines alone approaches $2 trillion a year and is projected to reach $2.4 trillion by 2029 (IQVIA Institute for Human Data Science, 2025). The returns to this spending are substantial. New drugs account for nearly three-quarters of the increase in life expectancy at birth across 30 developing and high-income countries between 2000 and 2009 (Lichtenberg, 2014), and for about a third of the 3.3-year increase in US life expectancy between 1990 and 2015 even after accounting for public health and other medical care (Buxbaum et al., 2020). HIV antiretrovirals, statins, cancer drugs, and, more recently, gene therapies are among the most prominent examples.

Yet drug development is slow, expensive, and risky. Bringing a single drug to market takes 10 to 15 years and costs between $1 billion and $2.6 billion, and only one in ten candidates that enter clinical testing eventually succeed (DiMasi et al., 2016; Wouters et al., 2020). The number of new drugs approved per dollar of research and development spending fell by half approximately every nine years between 1950 and 2010, a pattern known as Eroom's law (Scannell et al., 2012). Can artificial intelligence (AI) alter how medicines are discovered and developed? This matters not only for the pharmaceutical industry but for population health and social welfare more broadly.

AI tools are now available at every stage of drug discovery, including protein structure prediction (AlphaFold), compound generation (REINVENT), activity and toxicity screening (Chemprop), and clinical trial matching (TrialGPT). Nearly all of these tools were developed in or after 2018, the year in which AlphaFold first demonstrated that deep learning could predict protein structures with an accuracy approaching that of experimental methods, and which we therefore treat as the beginning of the modern AI era in drug development. Although these tools have demonstrated gains in specific tasks, whether such gains translate into more or better drugs at the scale of the industry remains an open question, and the

available evidence consists largely of individual cases that point in opposite directions.[1] What is missing is systematic evidence on the population of drug projects.

This paper addresses four questions. First, does AI increase the number of investigational new drug (IND) projects that firms initiate? Second, does AI raise the probability that a preclinical candidate advances to human trials? Third, how does AI change the nature of drug development, including its scope, novelty, and speed, and what determines where it is most effective? Fourth, how large are the resulting aggregate gains, if any?

Answering these questions poses three challenges. The first is measurement of outcomes: drug candidates must be tracked across stages of development, and across countries, because AI adoption and drug development are both global. The second is measurement of AI exposure: because the value of AI depends on both firm capabilities and field-specific knowledge, exposure should be quantified at the level of firms, diseases, and molecular targets rather than for the industry as a whole. The third is identification: firms that invest in AI, and diseases that attract AI research, may differ systematically from others in ways that also affect drug development, so the empirical design must separate the effect of AI from these pre-existing differences.

We address the first two challenges by constructing a harmonized dataset covering drug development in China, the United States, and Europe from 2010 to 2025. Our primary source, Citeline's Pharmaprojects, tracks investigational drug projects, including sponsors, disease indications, molecular targets, and development milestones, supplemented by clinical trial data from TrialTrove. We link these records to three measures of AI-related inputs at different levels. At the firm level, we use job titles and skills from Revelio Labs' professional profiles to construct firm-year counts of AI employees. At the disease level, we identify AI-related research in PubMed using an ontology of AI methods and named systems. At the

[1] Illustrating the optimism, *Fortune* reports Pfizer's claim that AI shortened the development of Paxlovid from four years to four months (Fortune, 2026). More cautiously, medicinal chemist Derek Lowe describes himself as a short-term pessimist and a long-term optimist, emphasizing unresolved challenges in target selection and the prediction of human toxicity (Bio-IT World, 2025). A *Nature* editorial similarly cautions that "AI's potential to accelerate drug discovery needs a reality check" (Nature, 2023).

target level, we use OpenTargets to measure cumulative biological evidence on molecular targets.

We address the third challenge in three ways. First, we exploit the timing of AI diffusion, comparing the relationship between AI employment and drug development before and after 2018. Second, we include rich fixed effects: firm and country-by-year fixed effects for new drug projects, and firm-by-disease, disease-by-year, and country-by-year fixed effects for candidate progression. These specifications absorb persistent firm differences and common disease- and country-specific shocks. Third, we construct a shift-share instrument that combines firms' pre-2010 university connections with subsequent growth in AI graduates at those universities. Because these connections predate the modern AI era, the instrument provides variation in AI talent supply less likely to reflect firms' anticipated drug development opportunities. The IV estimates confirm a positive relationship between AI employment and both outcomes.

Our first set of results uses firm-year observations and examines the number of new drug projects. Firm AI human capital has little relationship with new drug projects before 2018, but a positive and growing relationship thereafter. In the pooled specification, the elasticity of new drug projects with respect to AI employment is 0.21, implying that a 10 percent increase in AI employees is associated with 2.1 percent more projects. The year-specific elasticity reaches approximately 0.28 by 2024, compared with a relatively stable relationship before 2018. A placebo test using generic-drug trials shows no comparable post-2018 increase, suggesting that the relationship reflects innovative drug development rather than a general expansion of firm activity. The gains are larger among established firms.

Our central finding concerns the success of individual drug candidates. In the raw data, AI employment is unrelated to Phase I entry in 2010, whereas by 2025 candidates from firms in the top decile of AI employment reach Phase I at a rate of 63 percent, compared with 20 percent for those in the bottom decile. This pattern persists in drug-candidate-level regressions controlling for sponsor-by-disease, disease-by-year, and country-by-year fixed effects, as well as firm size, prior development experience, and biomedical research. These specifi-

cations compare candidates developed by the same firm for the same disease over time, net of disease- and country-specific shocks. Doubling a firm's AI workforce is associated with a 4.1 percentage point increase in Phase I entry, a 19 percent gain relative to the sample mean of 22 percent. The relationship is statistically indistinguishable from zero in every year before 2018 and strengthens thereafter, reaching approximately 6 percentage points by 2025. Instrumental-variable estimates exploiting firms' pre-2010 university connections and subsequent growth in AI graduates further support a positive relationship between AI employment and both new drug projects and candidate progression.

We next examine how AI contributes to drug development and where its benefits are greatest. Using workers with both AI and domain-specific skills to proxy for AI applications, we find that AI employment related to drug discovery is positively associated with both new drug projects and Phase I entry, whereas AI employment related to clinical development or development processes does not show comparable gains. These relationships are stronger where complementary knowledge is abundant. Both outcomes exhibit positive interactions between firm AI employment and disease-level AI research, prior firm experience in the disease, and biological evidence on molecular targets. For candidate progression, human and molecular evidence is particularly complementary to AI, while clinical evidence is not. AI also broadens firms' drug portfolios: a 10 percent increase in AI employment is associated with approximately 2.2 percent more disease categories and therapeutic modalities, as well as more distinct molecular targets, while reducing concentration across diseases and modalities. These findings suggest that AI primarily improves firms' ability to use existing scientific knowledge to explore more therapeutic opportunities and advance promising candidates into clinical testing.

In contrast to these gains, we find limited evidence that AI improves other dimensions of drug development. AI employment is not significantly associated with molecular target or target-combination novelty. Nor do we find significant effects on development speed, measured by the time from preclinical development to Phase I and the duration of clinical trials.

AI employment is also not significantly associated with progression from Phase I to Phase II or from Phase II to Phase III, although many post-2018 candidates have not yet reached these stages. These findings suggest that AI's contributions to date are concentrated in expanding and improving the preclinical pipeline, rather than generating fundamentally more novel drugs, accelerating development, or overcoming later-stage clinical bottlenecks.

Finally, we quantify the aggregate contribution of AI by combining its effects on the number of candidates that firms initiate and on the probability that each candidate reaches Phase I. For each firm-year from 2018 to 2025, we predict the number of Phase I entries under observed AI employment and under a counterfactual in which AI employment is set to zero, holding all other covariates and fixed effects at their observed values. The difference is approximately 1,200 Phase I entries. Of these, about 480 arise because firms with AI employees initiate more candidates and about 740 because their candidates are more likely to progress, so that about 60 percent of AI's contribution operates through candidate selection rather than through a larger candidate pool. Applying historical rates of progression from Phase I to approval, these entries correspond to approximately 200 additional approved drugs over the following 10 to 15 years.

Related Literature This paper contributes to three strands of literature. The first strand studies AI in pharmaceutical innovation. A long-standing concern in drug development is the decline in R&D productivity (Scannell et al., 2012). Scannell and Bosley (2016) argue that improving candidate selection, rather than simply increasing screening volume, is central to raising productivity. Recent work asks whether AI can improve this process, documenting favorable early clinical outcomes for AI-discovered molecules (Jayatunga et al., 2024) and assessing AI applications across different stages of drug development (Bender and Cortés-Ciriano, 2021; Zhang et al., 2025; Dranove and Garthwaite, 2022). At the firm level, Lou and Wu (2021) show that AI capabilities increased the number of preclinical drug projects before 2018, particularly for compounds with known mechanisms and moderately novel chemistry.

We differ from Lou and Wu (2021) in three main ways. First, our data extend into the post-2018 period, when modern AI tools became widely available, and we find a much stronger relationship between firm AI capabilities and drug development. Second, we study both new drug projects and progression into clinical trials, and find that improved progression is an important part of AI's overall impact. Third, we examine how AI affects firms' scope of drug development, including the set of disease areas they pursue.

The second strand of literature studies AI as an input into scientific discovery. A central idea in this literature is that AI changes how researchers search over and evaluate potential discoveries, while its returns depend on complementary data, knowledge, and organizational capabilities (Cockburn et al., 2018; Agrawal et al., 2018a, 2019a, 2024a; Gans, 2025b,a). A growing empirical literature shows that AI has diffused across scientific fields and increased scientific output (Bianchini et al., 2022; Ludwig and Mullainathan, 2024; Gao and Wang, 2024; Agrawal et al., 2026). Closest to our study, Hill and Stein (2026) show that AlphaFold increased research on proteins without known structures, but find no corresponding increase in patenting. Our results suggest one reason why these findings can coexist: AI need not generate new targets to affect downstream innovation. Firms with more AI human capital expand into more disease areas while using a similar set of molecular targets, suggesting that AI can create value by applying existing biological knowledge more broadly.

The third strand studies AI and productivity. Experiments document substantial task-specific gains in writing, customer support, consulting, and software development (Noy and Zhang, 2023; Brynjolfsson et al., 2025; Dell'Acqua et al., 2026; Cui et al., 2026). Despite rapid adoption and broad exposure (Bick et al., 2026; Eloundou et al., 2024), aggregate effects remain modest (Acemoglu, 2025; Humlum and Vestergaard, 2025). One explanation is that the gains from AI depend on complementary knowledge, organizational capabilities, and other bottlenecks in production (Brynjolfsson et al., 2021; Agrawal et al., 2018b, 2019b, 2024b; Jones, 2026). We provide industry-wide evidence from drug development. We find that AI improves productivity primarily through project selection and progression rather

than faster development, and that these gains are larger when complementary scientific knowledge and firm experience are available.

The rest of the paper proceeds as follows. Section 2 describes the institutional background and data. Section 3 presents the empirical strategy. Section 4 reports the main results, including heterogeneity and robustness checks. Section 5 examines the mechanisms and the scope and limitations of AI in drug development. Section 6 quantifies AI's contribution to Phase I entries. Section 7 concludes.

2 Institutional Background and Data

2.1 AI Usage in Drug Discovery

Figure 1 illustrates how AI can support each stage of drug development. In target discovery, researchers identify biological mechanisms whose modification could treat disease. PandaOmics integrates molecular data to prioritize targets, while structure models such as AlphaFold 2 help identify potential drug-binding sites. In molecule design, researchers develop compounds that act on these targets. Generative models such as REINVENT and GENTRL propose small molecules, RFdiffusion designs proteins, and AlphaFold 3 predicts target–molecule complex structures (Abramson et al., 2024), helping prioritize candidates for experimental testing.

In preclinical testing, laboratory and animal studies assess biological activity, toxicity, and how drugs move through the body. Models such as DeepTox, Chemprop, and ADMET-AI predict these properties to flag liabilities and improve candidate selection. Clinical development then evaluates safety and dosing in Phase I, preliminary efficacy in Phase II, and benefits and risks in larger Phase III trials. AI tools support trial planning, patient recruitment, and outcome prediction, while language models assist with regulatory documentation.

Industry reports highlight AI's potential to improve efficiency, speed, and success in drug discovery. Insilico Medicine reports nominating 30 preclinical candidates through programs

requiring 12–18 months and 60–200 synthesized molecules each (Insilico Medicine, 2026); Pfizer attributes a reduction in Paxlovid discovery time from four years to four months to AI-assisted design (Fortune, 2026); and XtalPi reports that AI and laboratory automation raised synthesis success rates from 10–30 to 80–90 percent (ByDrug, 2025). Skeptics question whether such task-level gains address the central bottlenecks in drug development, particularly target selection and human toxicity (Bio-IT World, 2025; Bender and Cortés-Ciriano, 2021). Novartis's chief executive similarly cautions that faster discovery has yet to alter the industry's five- to ten-year outlook (Shaywitz, 2025). Whether AI's gains in individual tasks translate into broader improvements in drug development therefore remains an empirical question that this paper addresses.

2.2 Data

Our analysis combines five data sources on drug development, firm capabilities, scientific knowledge, and biological evidence. Citeline's Pharmaprojects and TrialTrove provide our core outcome data, which we link to Revelio Labs, PubMed, and OpenTargets. Our data cover drug development in China, the United States, and Europe from 2010 to 2025, while our main regression analysis focuses on US and European firms.[2] Appendix Table C.1 summarizes the data layers, and Appendix C documents their construction. We describe each source in turn.

Pharmaprojects. Pharmaprojects tracks the global development history of individual drug candidates. For each candidate, it records the sponsoring firm, disease indications, molecular targets, therapeutic modalities, and key development milestones from discovery through clinical development. We use these records to identify when firms initiate new drug projects, whether candidates progress to later stages, and the timing of these transitions. We also construct measures of firms' prior drug development experience, the scope of their drug

[2]Revelio's coverage of pharmaceutical employment is relatively more limited in China than in the United States and Europe. Therefore, we restrict our main firm- and drug-level regressions to US and European firms.

portfolios, and the novelty of individual candidates.[3]

TrialTrove. TrialTrove provides complementary information on clinical trials, including the sponsoring firm, trial phase, disease indications, and trial dates. We use these data primarily to supplement the Pharmaprojects records and to construct additional trial-based measures, including generic-drug activity.

Revelio Labs. Revelio Labs provides worker-level professional histories constructed from public professional profiles. The data contain information on workers' employers, employment histories, job titles, and skills, allowing us to characterize firms' employment, age, and AI-related human capital over time. We also use workers' domain-specific skills to examine AI applications across different stages of drug development.

PubMed. PubMed provides a comprehensive record of biomedical research publications. The data include publication text, author information, and affiliations, allowing us to measure the evolution of AI-related biomedical knowledge across diseases, countries, and years.

OpenTargets. OpenTargets integrates multiple sources of biological evidence on molecular targets, including genetic, experimental, clinical, and literature evidence. We aggregate these data to construct cumulative evidence measures at the target-year level and link them to individual drug candidates through their molecular targets. These measures capture the stock of biological knowledge available for each target over time.

2.3 Measuring AI Exposure

We measure AI-related inputs at three complementary levels. Firm-level AI human capital captures firms' capacity to use AI in drug development, while disease-level AI papers capture

[3]Two features of the Citeline data are worth noting. First, the date on which a drug is added to Pharmaprojects ("Drug Added") records when the drug enters the database rather than its actual initiation date; first IND filing dates are available only for some drugs. Second, targets recorded as "Unidentified pharmacological activity" indicate undisclosed or unidentified targets, and their missingness depends on publicly available information.

the AI-related scientific knowledge available in research and development.[4] Finally, we use OpenTargets to measure the biological evidence available for individual drug candidates.

Firm-level AI human capital. Our firm-level measure of AI adoption comes from Revelio Labs. Revelio records more than 30,000 worker skills organized into skill clusters. We classify workers as AI employees if they report at least one skill in the AI cluster, including skills such as machine learning and deep learning, and aggregate the number of AI employees to the firm-year level. We then match Revelio companies to Citeline sponsors, accounting for subsidiaries and name changes (Appendix C.5).

A limitation is that Revelio reports worker skills as of April 2026 rather than when they were acquired. As a result, skills acquired later in a worker's career may be assigned retrospectively to earlier years. This feature of the skill-based measure is also useful for our empirical design: if it simply captures firms that eventually accumulated AI talent, it should predict drug development even before modern AI tools became widely available. We therefore use the pre-2018 period as a placebo. Moreover, as a robustness check, we use an alternative measure based on Revelio's occupation taxonomy, which classifies workers by their time-varying job roles rather than reported skills.

Disease-specific AI research. We use biomedical publications in PubMed to measure AI-related scientific knowledge relevant to each disease. We identify AI-related papers using the PharmAI ontology, a curated vocabulary of AI methods and named systems with safeguards against terms that have non-AI meanings. We map each paper to one of 332 Citeline disease indications, grouped into nine therapeutic areas, and assign papers to countries using author affiliations (Appendices C.2–C.4). Our main measure aggregates AI-related publications at the disease-year level using publications from the preceding three years. In some analyses, we additionally construct measures at the disease-country-year level.

[4]We also construct a patent-based measure of AI activity. Patents capture inventive activity but are less suited to measuring contemporaneous AI adoption because applications are observed with a lag and standard classification codes can both over- and under-identify AI use in biomedicine—for example, by confusing biological "neural networks" with machine learning (Fang et al., 2026). We therefore use a text-based classifier and treat patent-based measures as complementary evidence.

Target evidence. Finally, we use OpenTargets to construct cumulative measures of biological evidence at the target-year level, distinguishing among different types of evidence, including genetic, clinical, and literature evidence. These measures capture the stock and composition of scientific knowledge available for each molecular target over time and are used to study complementarities between AI and existing biological evidence.

2.4 Descriptive Patterns

Figure 2 shows the trends in our two main outcomes. Panel (a) shows that the number of new drug (IND) projects increased from about 1,800 in 2010 to 4,800 in 2021, before declining to about 3,600 in 2025. Most of the increase comes from China, where the number of projects grew from about 60 in 2010 to 1,400 in 2025. Panel (b) shows a similar increase in the number of drug candidates that subsequently enter Phase I. The total number increased from about 150 in 2010 to around 1,000 per year by the early 2020s. This growth was particularly pronounced in China, where the number of candidates entering Phase I rose from 23 in 2010 to more than 400 in 2025.

Figure 3 shows AI human capital in the sample firms. Panel (a) indexes AI employees and total employees to 100 in 2010. Both grow, but AI employment grows faster throughout, and the gap widens sharply after 2020, when total employment flattens while AI employment keeps rising. By 2025, AI employment stands at about three times its 2010 level, compared with less than twice for total employment, so the AI share of the workforce rose from under 3 percent to about 4.5 percent. Adoption is uneven across firms (Appendix Figure B.1). In 2010, about half of the firms had an AI employee share below 1 percent; by 2025, this fraction had fallen to about one-third, and the mass of the distribution had shifted to shares between 2 and 6 percent, with a thin right tail of AI-intensive firms.

Figure 3, panel (b), shows the growth of AI-related biomedical research. The number of AI papers increased from fewer than 100 per year in each region in 2010 to about 1,000 in the United States, 800 in Europe, and more than 2,000 in China in 2025. China experienced

the fastest growth and overtook the United States around 2022. Under the stricter PharmAI classifier, AI-related papers account for 2.55 percent of biomedical publications in China, 1.89 percent in the United States, and 1.60 percent in Europe in 2025 (Appendix C.2).

Table 1 reports the mean, standard deviation, median, minimum, and maximum of key variables at the firm-year level (Panel A) and the drug-candidate level (Panel B) for the US and European firms in our regression sample.

3 Empirical Strategy

We organize the empirical analysis primarily at two levels. At the firm-year level, we examine whether firms with greater AI human capital initiate more drug projects. At the drug-candidate level, we examine whether candidates developed by firms with greater AI human capital are more likely to progress from preclinical development to Phase I.

For some analyses, we also use firm–disease–year and firm–target–year data to study where firms develop new drugs and disease–country–year data to study firm entry into new diseases. We also examine additional outcomes, including the scope of firms' drug portfolios, development speed, and complementarities between AI and scientific knowledge. Regression analysis focuses on US and European firms.

Firm-level analysis. We first examine whether firm AI human capital is associated with the number of new drug projects a firm initiates. Our baseline specification is a Poisson model of the form

$$\mathrm{E}[\mathrm{NewDrugs}_{ft} \mid \mathrm{AI}_{ft}, X_{ft}, \alpha_f, \delta_{ct}] = \exp\left(\beta \log(1 + \mathrm{AI}_{ft}) + X'_{ft}\gamma + \alpha_f + \delta_{ct}\right), \qquad (1)$$

where $\mathrm{NewDrugs}_{ft}$ is the number of new drug projects initiated by firm f in year t, AI_{ft} is the number of AI employees, and α_f and δ_{ct} are firm and country-by-year fixed effects. The controls X_{ft} include log total employment and log prior launched drugs. We estimate the

model using PPML, with standard errors clustered at the firm level.

Drug-level analysis. We next examine whether AI affects the progression of individual drug candidates. Our main outcome is an indicator for whether a preclinical candidate reaches Phase I. We estimate linear probability models of the form

$$\text{PhaseI}_{fdt} = \beta \log(1 + \text{AI}_{ft}) + X'_{fdt}\gamma + \alpha_{fd} + \delta_{dt} + \lambda_{ct} + \varepsilon_{fdt}, \tag{2}$$

where the candidate is developed by firm f for disease d in country c, and t is the year the candidate enters development. The specification includes firm-by-disease, disease-by-year, and country-by-year fixed effects, as well as controls for firm size, the firm's prior drug launches in the disease, and disease-level scientific research. Standard errors are clustered at the firm–disease level.

For both outcomes, we also estimate specifications that interact log AI employment with year indicators to examine how the relationship between AI human capital and drug development evolves over time. We report the implied year-specific coefficients.

Timing and identification. The year-specific coefficients provide a placebo test for our design. Modern AI tools were not widely used in drug development before 2018, so AI exposure should not predict outcomes in the pre-period if it primarily captures the effect of these technologies rather than persistent differences across firms or diseases. The fixed effects further absorb time-invariant heterogeneity and common shocks.

However, these specifications do not rule out post-2018 shocks that jointly affect AI adoption and drug development. We therefore complement this approach with an instrumental-variable strategy, following the intuition of Babina et al. (2024). Firms tend to recruit from universities with which they have established connections. As universities expand their AI training, firms with stronger pre-existing ties to these universities should experience greater growth in AI labor supply, independently of their own drug development opportunities.

We instrument firm AI employment using cumulative AI graduates from universities attended by the firm's early employees, weighted by each university's share among those employees:

$$\text{IV}_{ft} = \sum_{u} s_{fu}^{\text{early}} \times \text{AIGraduates}_{ut}. \tag{3}$$

Here, s_{fu}^{early} is the share of firm f's pre-2010 employees who attended university u, and AIGraduates_{ut} is the cumulative number of AI graduates from that university by year t. The exclusion restriction is more plausible because these university connections predate the modern AI era and are unlikely to have been formed in anticipation of future AI advances.

We use control-function IV-PPML for new drug projects and 2SLS for Phase I entry. In the control-function specification, we include the first-stage residual in the PPML regression and compute standard errors from 100 firm-clustered bootstrap replications.

Other outcomes and specifications. We use related specifications to examine several additional outcomes. At the firm level, we study the number and concentration of disease categories, disease subcategories, therapeutic modalities, and molecular targets in a firm's portfolio. At the drug-candidate level, we examine target, combination, and structure novelty; progression to Phase II and Phase III; and development speed, including the time from preclinical development to Phase I and the duration of clinical trials. We also examine heterogeneity by firm age and country.

To investigate potential mechanisms, we examine complementarities at the firm–disease–year, firm–target–year, and drug-candidate levels. Specifically, we interact firm AI human capital with disease-level AI research, the firm's prior drug launches in the disease, and cumulative target-level biological evidence.

4 Main Results

We examine two outcomes: (i) the number of new drug projects initiated by a firm and (ii) the probability that a drug candidate enters Phase I.

4.1 Raw Correlations

We begin with the firm-level analysis, asking whether firms with greater AI human capital initiate more new drug projects. Figure 4 presents the raw relationship between AI employees and drug development in 2010 and 2025. In 2010, there is little relationship between AI employees and new drug projects, with a slope of 0.006 (standard error 0.139) in panel (a). By 2025, firms in the highest AI-employee bins initiate roughly three times as many new projects as firms with no AI human capital, with a slope of 0.132 (standard error 0.031) in panel (b).

Whether this translates into greater innovative output, however, depends on what happens to those projects after they are initiated. Drug development is highly selective, and most candidates fail before reaching clinical testing. Panels (c) and (d) show the raw relationship between firm AI human capital and candidate progression. In 2010, the probability that a preclinical candidate reaches Phase I is largely unrelated to the firm's AI employment, with a slope of -0.002 (standard error 0.007). By 2025, candidates developed by firms with the most AI employees are more than three times as likely to reach Phase I as those developed by firms with no AI employees, with a slope of 0.067 (standard error 0.009). Both new drug projects and Phase I entry thus exhibit stronger positive correlations with AI employment in 2025.

4.2 Baseline and IV Estimates

Table 2 reports the baseline and instrumental-variable estimates. Column 1 shows an elasticity of 0.207: a 10 percent increase in AI employees is associated with approximately 2.1

percent more new drug projects, conditional on firm and country-by-year fixed effects, total employment, and prior drug-development experience.

Column 3 shows that the relationship with candidate progression persists after accounting for differences across firms, diseases, countries, and years. The coefficient on log firm AI employees is 0.059. Doubling AI employment is therefore associated with a 4.1 percentage point increase in the probability that a candidate reaches Phase I ($0.059 \times \ln 2$), relative to a sample mean of 22 percent. The specification includes sponsor-by-disease, disease-by-year, and country-by-year fixed effects, so the estimate compares candidates within the same firm and disease while absorbing disease-specific and country-specific shocks over time.

Figure 5 examines how these relationships evolve over time. For both outcomes, the year-specific coefficients exhibit no systematic trend before 2018, but increase substantially thereafter. For new drug projects (panel a), the coefficients remain relatively stable before 2018 and rise to approximately 0.28 by 2024. For Phase I entry (panel b), the coefficients are small and statistically insignificant before 2018, but increase steadily thereafter, reaching approximately 0.09 by 2025. The absence of a systematic pre-2018 trend supports our identification strategy, suggesting that the post-2018 increase is not simply driven by persistent differences between AI-intensive firms and other firms.

Instrumental-variable estimates. Despite the rich fixed effects and the absence of differential pre-trends, AI employment may still be endogenous to firms' drug development opportunities. For example, firms pursuing more promising research projects may simultaneously expand their AI workforce, leading to upward bias in the baseline estimates. Conversely, firms may invest in AI in response to declining R&D productivity, potentially attenuating the estimates. The direction of the overall bias is therefore ambiguous.

Columns 2 and 4 report the instrumental-variable estimates based on equation (3). Both estimates are positive and significant, although substantially larger than the baseline estimates. The control-function IV-PPML elasticity of new drug projects is 2.594 (column 2), and the 2SLS coefficient on Phase I entry is 0.254 (column 4). The first-stage F-statistics

are 11.90 and 19.71, respectively. The larger IV estimates are consistent with attenuation in the baseline estimates, although differences in the variation used for identification may also contribute to the gap.

4.3 Heterogeneity and Robustness

We next examine how the relationship between AI human capital and drug development varies across firms and assess the robustness of our main findings.

Heterogeneity. Table 3 examines heterogeneity by firm age and country. Columns 1 and 3 show that the association between AI employment and drug development is weaker among young firms (five years old or younger). For new drug projects, the coefficient on AI employment is 0.246, with an interaction of −0.171 for young firms. For Phase I entry, the corresponding coefficients are 0.059 and −0.027. Both interactions are statistically significant, indicating larger gains among established firms. Columns 2 and 4 show no significant differences between US and European firms, with interaction coefficients of −0.026 for new drug projects and 0.011 for Phase I entry.

Placebo. Figure 6 repeats the event-study specification using generic-drug trials, identified with our GPT-based classifier. Because generic development does not require the discovery of new drug candidates, it provides a placebo for AI-driven drug discovery. We find no systematic relationship between AI employment and generic trials, suggesting that the main results do not simply reflect broader expansion in firms' development activity.

Alternative AI measures. Appendix Table A.1 replaces skill-based AI employment with workers in AI roles, data roles, or holding CS PhDs. AI and data roles are identified using Revelio's occupation taxonomy, which classifies workers based on their job titles and descriptions rather than reported skills.[5] The results remain positive and significant for both new drug projects and Phase I entry. In contrast, IT employment has no significant effect

[5]We prefer the skill-based measure because AI skills may be used across a broader range of occupations, including roles not explicitly classified as AI-related. The role-based measure is more restrictive but varies with workers' employment histories, providing a useful robustness check against the retrospective assignment of skills.

after controlling for AI employment, suggesting that the relationship is driven by AI-specific rather than general IT human capital.

Alternative instruments. Appendix Table A.2 considers three alternative instruments. The first replaces AI graduates with CS PhD graduates. The other two use the share of a firm's pre-2010 employees who attended the top 200 AI departments in CSRankings (based on 2000–2009 publications), interacted with cumulative AI or CS PhD graduates from these departments since 2010. The IV estimates remain broadly consistent with the baseline results. Phase I entry coefficients range from 0.288 to 0.317 and are significant at the 1 percent level, while new drug project elasticities range from 2.332 to 2.705, although the estimate using CS graduates is not statistically significant.

Alternative dependent variable. Appendix Table A.3 uses log(1 + new drug projects) instead of PPML. The coefficient on AI employment remains positive and significant (0.029), and the heterogeneity patterns by firm age and country are similar to those in Table 3.

5 Mechanisms

We have shown that firms with greater AI human capital initiate more drug projects and that their candidates are more likely to advance to Phase I trials. What explains these two findings? These gains could arise through different channels, depending on where AI is applied in the development process, how it interacts with existing scientific knowledge, and whether it enables firms to pursue new therapeutic opportunities. We examine these mechanisms by studying AI applications across development activities, complementarities with scientific knowledge and firm experience, and changes in the scope, novelty, and speed of drug development.

5.1 How AI Matters: Applications in Drug Development

AI can support multiple stages of drug development, from target identification and molecule design to clinical trials and development processes. To identify which applications are most relevant to our findings, we use workers who possess both AI and domain-specific skills as proxies for AI adoption in different activities. Table 4 distinguishes AI workers with skills in drug discovery, clinical development, and development processes, along with other AI workers.

The results point primarily to drug discovery. AI workers with drug-discovery skills are positively associated with both new drug projects and Phase I entry, with coefficients of 0.292 and 0.063, respectively. These relationships remain positive and significant when all AI worker categories are included jointly, with coefficients of 0.350 and 0.064. In contrast, AI workers with clinical development skills are negatively associated with both outcomes, while those with development-process skills show no significant relationship. These patterns suggest that AI's contribution to drug development is most closely associated with its applications in discovery and preclinical research, where firms identify, design, and evaluate drug candidates. The results provide less evidence that AI applications in clinical development or development processes explain these gains.

5.2 Where AI Matters: Complementary Knowledge

AI may be particularly valuable when firms have access to scientific knowledge and experience that can inform the discovery and evaluation of drug candidates. We examine three potential complementarities: AI-related research in the relevant disease, the firm's prior drug-development experience in that disease, and cumulative biological evidence on the candidate's molecular target.

Table 5 examines how complementary knowledge moderates the relationship between AI employment and drug development. Columns 1–2 estimate firm–disease–year PPML regressions of new drug projects, while column 3 uses firm–target–year observations. The interac-

tions between AI employment and disease-level AI research, prior firm–disease launches, and cumulative target evidence are all positive and significant, with coefficients of 0.080, 0.098, and 0.081, respectively. Columns 4–6 examine the same three complementarities for Phase I entry using drug-candidate-level regressions. All three interactions are again positive and significant (0.029, 0.038, and 0.005). When AI employment is interacted with disease-level AI research, its main coefficient falls to 0.021 and becomes statistically insignificant. These results suggest that AI is more strongly associated with new drug development and candidate progression where relevant scientific knowledge and firm experience are more abundant.

We further examine which types of biological evidence are most complementary to AI. Table 6 distinguishes cumulative target evidence from scientific literature, human genetic and molecular studies, and clinical research. AI employment interacts positively with literature evidence (0.007) and human and molecular evidence (0.014), but not significantly with clinical evidence. When all three interactions enter jointly, the interaction with human and molecular evidence remains strong and precisely estimated (0.013), while that with literature evidence becomes smaller (0.004) and clinical evidence remains insignificant. These findings suggest that AI complements existing scientific knowledge and firm experience, helping firms identify and advance promising drug candidates.

5.3 What AI Can and Cannot Do

Our results suggests that AI improves drug discovery by drawing on existing scientific knowledge and firm capabilities. But how far do these gains extend? We examine whether AI changes the scope, novelty, and speed of drug development, and whether its benefits extend beyond preclinical development to later clinical stages.

Scope of drug development. Table 7 examines the breadth and concentration of firms' drug portfolios across disease categories, subcategories, therapeutic modalities, and molecular targets. A 10 percent increase in AI employment is associated with approximately 2.2

percent more disease categories, with similar increases in disease subcategories and therapeutic modalities. AI employment is also associated with lower portfolio concentration across these dimensions, indicating that firms pursue a broader range of therapeutic opportunities. The number of distinct molecular targets increases as well, although target concentration does not change significantly. These findings suggest that AI enables firms to explore a wider range of applications for existing scientific knowledge.

Drug novelty. Expanding the scope of drug development does not necessarily imply more novel discoveries. Table 8 examines three dimensions of novelty: molecular targets, target combinations, and chemical structures. We find no significant relationship between AI employment and either target or combination novelty, nor does either measure significantly moderate the relationship between AI employment and Phase I entry. Structure novelty is weakly positively associated with AI employment (0.070, significant at the 10 percent level), but this result is based on a smaller sample of 1,336 candidates. Structure novelty also does not significantly moderate the relationship between AI employment and Phase I entry. Thus, AI is associated with a broader set of targets and therapeutic applications, but not with systematically more novel drug candidates.

Later clinical stages. An improvement in preclinical candidate progression may not translate into higher success rates in clinical development. Columns 1–2 of Table 9 examine progression to Phase II conditional on Phase I entry and progression to Phase III conditional on Phase II entry. Neither outcome is significantly associated with AI employment, although most post-2018 cohorts have not yet reached these stages. We therefore find no evidence so far that AI improves progression beyond Phase I.

Development speed. AI may also accelerate drug development without changing the probability of success. Columns 3–4 of Table 9 examine the time from preclinical development to Phase I and the duration of clinical trials. Neither outcome is significantly associated with

AI employment, suggesting that AI's benefits have not yet translated into faster development.

These findings help define the current limits of AI in drug development. AI is associated with more drug projects, broader therapeutic exploration, and higher rates of progression into Phase I, but there is limited evidence of gains in fundamental novelty, later-stage clinical progression, or development speed.

6 Decomposition: Additional Phase I Entry from AI

We next quantify the aggregate contribution of AI to drug development. AI can increase the number of candidates reaching Phase I through two margins: firms initiate more drug projects, and a greater share of these projects progress to Phase I. We estimate firm-year Poisson models for new drug projects and Phase I entries on a common sample and predict each outcome under observed AI human capital and under a counterfactual in which every firm has zero AI employees, holding other covariates and fixed effects unchanged. We aggregate these predictions over 2018–2025.

To separate the two margins, let $\widehat{N}^{AI}$ and $\widehat{N}^{0}$ denote the predicted numbers of new drug projects under observed and zero AI employment, and let $\widehat{p}^{AI}$ and $\widehat{p}^{0}$ denote the corresponding implied Phase I conversion rates. The total contribution of AI can be written as

$$\Delta\text{PhaseI}_{AI} = (\widehat{N}^{AI} - \widehat{N}^{0})\widehat{p}^{0} + \widehat{N}^{AI}(\widehat{p}^{AI} - \widehat{p}^{0}). \tag{4}$$

The first term captures additional Phase I entries arising from more drug candidates, holding the conversion rate at its no-AI level. The second captures the contribution of a higher Phase I conversion rate, evaluated at the number of candidates under observed AI adoption.

Figure 7 presents this decomposition. This calculation implies 1,223.2 additional Phase I entries attributable to AI between 2018 and 2025. Of these, 484.0 arise because firms initiate more drug candidates, while 739.1 arise because candidates are more likely to progress to Phase I. Thus, about 60 percent of the estimated increase in Phase I entries operates through

higher candidate progression rather than through the expansion of the candidate pool.

These calculations quantify AI's contribution to the clinical pipeline rather than welfare itself. Translating additional Phase I entries into eventual approvals and health gains requires assumptions about downstream success and the social value of new drugs.

7 Conclusion

This paper examines whether AI improves drug development using global data on investigational drug projects, firm AI human capital, and biomedical research. We find that AI employment is associated with both more new drug projects and a higher probability that preclinical candidates advance to Phase I trials, with these relationships strengthening after 2018. Our counterfactual analysis attributes approximately 1,200 additional Phase I entries to AI between 2018 and 2025, with about 60 percent arising from higher candidate progression rates rather than a larger pool of candidates.

Our findings suggest that AI may improve pharmaceutical R&D productivity primarily through better discovery and candidate selection rather than faster development. These gains depend on complementary scientific knowledge and firm experience. AI is associated with broader exploration across diseases and therapeutic modalities, but we find limited evidence of greater drug novelty, improved progression in later clinical stages, or shorter development times. This distinction matters because drug development is constrained not only by the number of candidates firms can generate, but also by their ability to identify promising candidates and successfully evaluate them in humans. AI appears to have improved the former stages of this process without yet overcoming the bottlenecks in clinical development.

These results have broader implications for the economic returns to AI. First, AI's contribution to innovation may depend less on generating fundamentally new scientific knowledge than on improving the use of existing knowledge. Its returns are larger where relevant research, biological evidence, and domain expertise are already available, suggesting that ac-

cess to complementary knowledge and capabilities may shape which firms and scientific fields benefit most. Second, improvements in intermediate research outcomes need not translate immediately into greater final innovation output. Drug development offers a particularly useful setting for studying this distinction because the path from discovery to market approval is long and subject to substantial scientific and regulatory uncertainty.

Several limitations remain. Our AI measure captures firms' AI human capital rather than the direct use of AI in individual drug projects, and the empirical estimates may still reflect unobserved changes in firms' innovation opportunities. Moreover, the recent diffusion of modern AI tools means that many candidates have not yet completed clinical development. Our counterfactual estimates therefore measure AI's contribution to the clinical pipeline rather than to eventual drug approvals or social welfare. Whether the additional candidates entering clinical trials ultimately produce more effective treatments, and how the resulting benefits compare with the costs of AI adoption, remain important questions for future research.

References

Abramson, Josh, Jonas Adler, Jack Dunger, Richard Evans, Tim Green, Alexander Pritzel, Olaf Ronneberger, Lindsay Willmore, Andrew J Ballard, Joshua Bambrick et al., "Accurate structure prediction of biomolecular interactions with AlphaFold 3," *Nature*, 2024, *630* (8016), 493–500.

Acemoglu, Daron, "The Simple Macroeconomics of AI," *Economic Policy*, 2025, *40* (121), 13–58.

Agrawal, Ajay, John McHale, and Alexander Oettl, "Finding Needles in Haystacks: Artificial Intelligence and Recombinant Growth," Working Paper 24541, National Bureau of Economic Research 2018.

———, ———, **and** ———, "Artificial Intelligence, Scientific Discovery, and Commercial Innovation," 2019. Working Paper.

———, ———, **and** ———, "Artificial Intelligence and Scientific Discovery: A Model of Prioritized Search," *Research Policy*, 2024, *53* (5), 104989.

———, **Joshua Gans, and Avi Goldfarb**, *Prediction Machines: The Simple Economics of Artificial Intelligence*, Harvard Business Review Press, 2018.

———, **Joshua S. Gans, and Avi Goldfarb**, "Artificial Intelligence: The Ambiguous Labor Market Impact of Automating Prediction," *Journal of Economic Perspectives*, 2019, *33* (2), 31–50.

———, ———, **and** ———, "Artificial Intelligence Adoption and System-Wide Change," *Journal of Economics & Management Strategy*, 2024.

Agrawal, Ajay K, John McHale, and Alexander Oettl, "AI in Science," Technical Report, National Bureau of Economic Research 2026.

Babina, Tania, Anastassia Fedyk, Alex He, and James Hodson, "Artificial intelligence, firm growth, and product innovation," *Journal of Financial Economics*, 2024, *151*, 103745.

Bender, Andreas and Isidro Cortés-Ciriano, "Artificial Intelligence in Drug Discovery: What Is Realistic, What Are Illusions? Part 1: Ways to Make an Impact, and Why We Are Not There Yet," *Drug Discovery Today*, 2021, *26* (2), 511–524.

Bianchini, Stefano, Moritz Müller, and Pierre Pelletier, "Artificial Intelligence in Science: An Emerging General Method of Invention," *Research Policy*, 2022, *51* (10), 104604.

Bick, Alexander, Adam Blandin, and David Deming, "The Rapid Adoption of Generative AI," *Management Science*, 2026.

Bio-IT World, "Derek Lowe on AI in Drug Discovery: Between Hype and Hope," Bio-IT World, April 15, 2025, https://www.bio-itworld.com/news/2025/04/15/derek-lowe-on-ai-in-drug-discovery-between-hype-and-hope 2025.

Brynjolfsson, Erik, Daniel Rock, and Chad Syverson, "The productivity J-curve: How intangibles complement general purpose technologies," *American Economic Journal: Macroeconomics*, 2021, *13* (1), 333–372.

———, **Danielle Li, and Lindsey Raymond**, "Generative AI at Work," *The Quarterly Journal of Economics*, 2025, *140* (2), 889–942.

Buxbaum, Jason D., Michael E. Chernew, A. Mark Fendrick, and David M. Cutler, "Contributions of Public Health, Pharmaceuticals, and Other Medical Care to US Life Expectancy Changes, 1990–2015," *Health Affairs*, 2020, *39* (9), 1546–1556.

ByDrug, "Synthetic Chemistry Meets the Intelligent Revolution: Pioneering Scholars Discuss the New Paradigm of AI and Robotics Research (in Chinese)," Interview with Peiyu Zhang, Chief Scientific Officer, XtalPi. ByDrug (Pharmcube), August 18, 2025, https://bydrug.pharmcube.com/news/detail/a41c3305c0e1d912d8faba5ec028e4e7 2025.

Cockburn, Iain M., Rebecca Henderson, and Scott Stern, "The Impact of Artificial Intelligence on Innovation," Working Paper 24449, National Bureau of Economic Research 2018.

Cui, Kevin Z., Mert Demirer, Sonia Jaffe, Leon Musolff, Sida Peng, and Tobias Salz, "The Effects of Generative AI on High-Skilled Work: Evidence from Three Field Experiments with Software Developers," *Management Science*, 2026.

Dell'Acqua, Fabrizio, Edward McFowland III, Ethan Mollick, Hila Lifshitz-Assaf, Katherine C. Kellogg, Saran Rajendran, Lisa Krayer, François Candelon, and Karim R. Lakhani, "Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of Artificial Intelligence on Knowledge Worker Productivity and Quality," *Organization Science*, 2026.

DiMasi, Joseph A., Henry G. Grabowski, and Ronald W. Hansen, "Innovation in the Pharmaceutical Industry: New Estimates of R&D Costs," *Journal of Health Economics*, 2016, *47*, 20–33.

Dranove, David and Craig L Garthwaite, "Artificial intelligence, the evolution of the healthcare value chain, and the future of the physician," Technical Report, National Bureau of Economic Research Cambridge, MA, USA 2022.

Eloundou, Tyna, Sam Manning, Pamela Mishkin, and Daniel Rock, "GPTs Are GPTs: Labor Market Impact Potential of LLMs," *Science*, 2024, *384* (6702), 1306–1308.

Fang, Hanming, Xian Gu, Hanyin Yan, and Wu Zhu, "AI Patents in the United States and China: Measurement, Organization, and Knowledge Flows," Technical Report, National Bureau of Economic Research 2026.

Fortune, "Pfizer CEO Albert Bourla Stared Down the COVID-19 Pandemic. Now He Has His Eyes Set on

Cancer," Fortune, January 27, 2026, https://fortune.com/article/pfizer-ceo-albert-bourla-fortune-500-titans-and-disruptors-of-industry-covid-trump-cancer-ai/ 2026.

Gans, Joshua S., "Growth in AI Knowledge," Working Paper 33907, National Bureau of Economic Research 2025.

———, "A Quest for AI Knowledge," Working Paper 33566, National Bureau of Economic Research 2025.

Gao, Jian and Dashun Wang, "Quantifying the Use and Potential Benefits of Artificial Intelligence in Scientific Research," *Nature Human Behaviour*, 2024, *8*, 2281–2292.

Hill, Ryan R and Carolyn Stein, "How Artificial Intelligence Shapes Science: Evidence from AlphaFold," Technical Report, National Bureau of Economic Research 2026.

Humlum, Anders and Emilie Vestergaard, "Large Language Models, Small Labor Market Effects," Working Paper, National Bureau of Economic Research 2025.

Insilico Medicine, "Insilico Medicine to Showcase AI-Driven Innovation at BIO 2026 International Convention," Press release, May 29, 2026, https://www.eurekalert.org/news-releases/1130202 2026.

IQVIA Institute for Human Data Science, "The Global Use of Medicines 2025: Outlook to 2029," Technical Report, IQVIA Institute 2025.

Jayatunga, Madura K. P., Margaret Ayers, Lotte Bruens, Dhruv Jayanth, and Christoph Meier, "How Successful Are AI-Discovered Drugs in Clinical Trials? A First Analysis and Emerging Lessons," *Drug Discovery Today*, 2024, *29* (6), 104009.

Joachimiak, Marcin P., Mark A. Miller, J. Harry Caufield, Ryan Ly, Nomi L. Harris, Andrew Tritt, Christopher J. Mungall, and Kristofer E. Bouchard, "The Artificial Intelligence Ontology: LLM-Assisted Construction of AI Concept Hierarchies," *Applied Ontology*, 2024, *19* (4), 408–418.

Jones, Charles I., "A.I. and Our Economic Future," Working Paper 34779, National Bureau of Economic Research 2026.

Lichtenberg, Frank R., "Pharmaceutical Innovation and Longevity Growth in 30 Developing and High-Income Countries, 2000–2009," *Health Policy and Technology*, 2014, *3* (1), 36–58.

Lou, Bowen and Lynn Wu, "AI on drugs: Can artificial intelligence accelerate drug development? Evidence from a large-scale examination of bio-pharma firms," *MIS quarterly*, 2021, *45* (3), 1451–1482.

Ludwig, Jens and Sendhil Mullainathan, "Machine Learning as a Tool for Hypothesis Generation," *The Quarterly Journal of Economics*, 2024, *139* (2), 751–827.

Nature, "AI's Potential to Accelerate Drug Discovery Needs a Reality Check," *Nature*, 2023, *622* (7982), 217. Editorial.

Noy, Shakked and Whitney Zhang, "Experimental Evidence on the Productivity Effects of Generative

Artificial Intelligence," *Science*, 2023, *381* (6654), 187–192.

Scannell, Jack W., Alex Blanckley, Helen Boldon, and Brian Warrington, "Diagnosing the Decline in Pharmaceutical R&D Efficiency," *Nature Reviews Drug Discovery*, 2012, *11* (3), 191–200.

——— **and Jim Bosley**, "When Quality Beats Quantity: Decision Theory, Drug Discovery, and the Reproducibility Crisis," *PLoS ONE*, 2016, *11* (2), e0147215.

Shaywitz, David, "Novartis CEO Vas Narasimhan: Drawn to Analytics, Grounded Expectations for AI," Timmerman Report, September 5, 2025, https://timmermanreport.com/2025/09/novartis-ceo-vas-narasimhan-drawn-to-analytics-grounded-expectations-for-ai/ 2025.

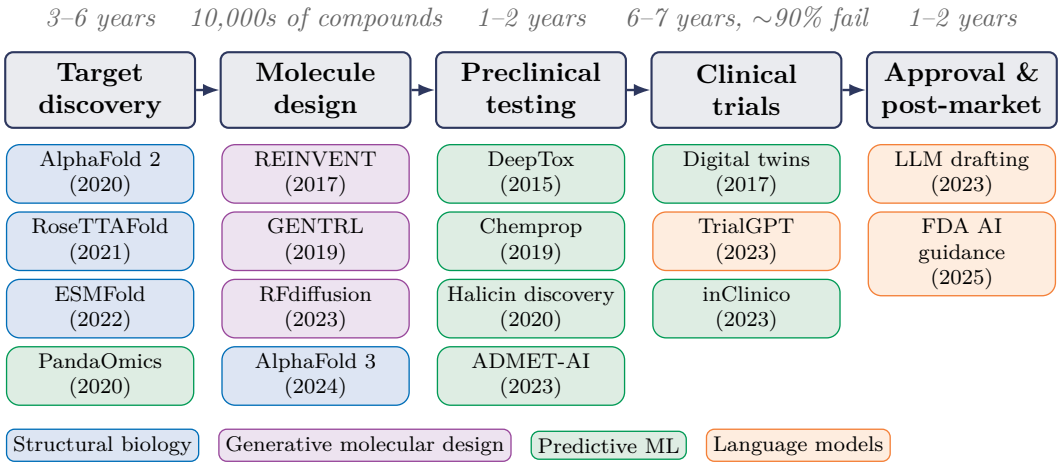
World Health Organization, "Global Spending on Health: Emerging from the Pandemic," Technical Report, World Health Organization, Geneva 2024.

Wouters, Olivier J., Martin McKee, and Jeroen Luyten, "Estimated Research and Development Investment Needed to Bring a New Medicine to Market, 2009–2018," *JAMA*, 2020, *323* (9), 844–853.

Zhang, Kang, Xin Yang, Yifei Wang, Yunfang Yu, Niu Huang, Gen Li, Xiaokun Li, Joseph C Wu, and Shengyong Yang, "Artificial intelligence in drug development," *Nature medicine*, 2025, *31* (1), 45–59.

Figures

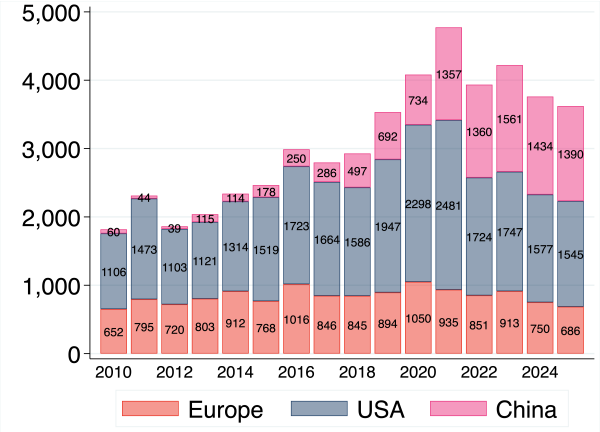
Figure 1: AI Tools at Each Stage of Drug Development



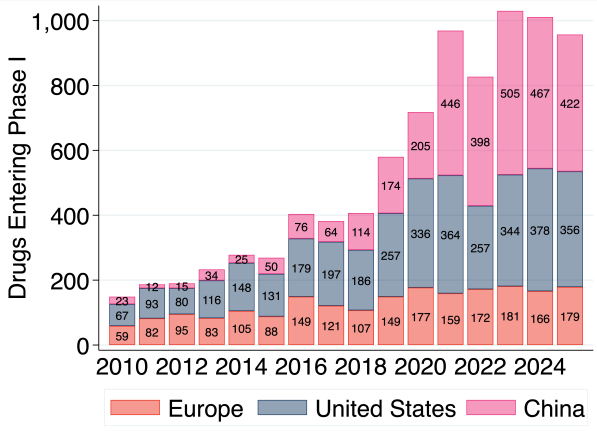
Notes: Each column lists AI tools that have been applied at that stage of development, with the year the tool was introduced. Colors indicate the kind of AI: structural biology models, generative molecular design, predictive machine learning, and language models. The typical duration or difficulty of each stage without AI, shown above the boxes, is approximate.

Figure 2: New Drug Projects and Drug Candidates Entering Phase I

(a) New Drug Projects



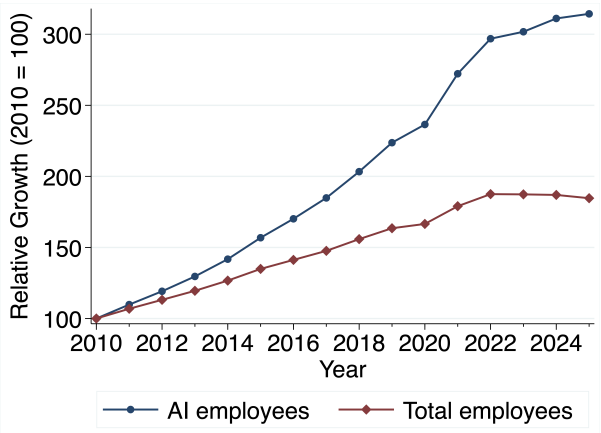
(b) Drug Candidates Entering Phase I



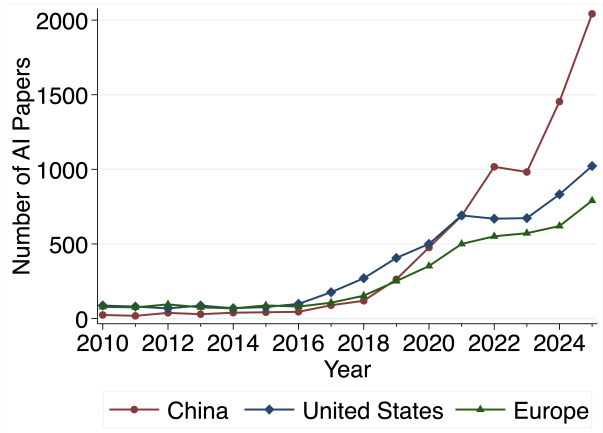
Notes: Panel (a) reports the annual number of new drug projects first recorded in Citeline Pharmaprojects. Panel (b) reports how many candidates in each annual cohort subsequently enter Phase I. Candidates are assigned to the headquarters region of their originator firms. Europe pools European countries, and bar labels indicate counts. Source: Citeline Pharmaprojects, 2010–2025.

Figure 3: AI Employees and AI Research Papers

(a) AI and total employees, 2010 = 100



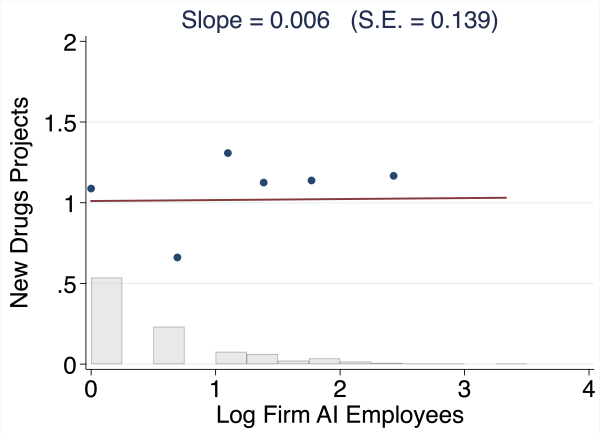
(b) AI-related biomedical papers



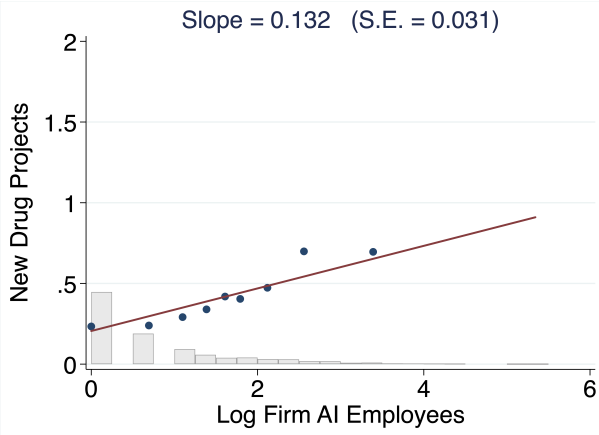
Notes: Panel (a) indexes AI and total employment at pharmaceutical firms headquartered in the United States, Europe, and China to 100 in 2010. AI employees are Revelio workers reporting at least one AI-related skill, such as machine learning or deep learning. Panel (b) shows annual AI-related biomedical publications classified using the PharmAI ontology (tier 2), assigned to countries by author affiliation. Europe pools European countries.

Figure 4: AI Employees and Drug Development: Raw Correlations

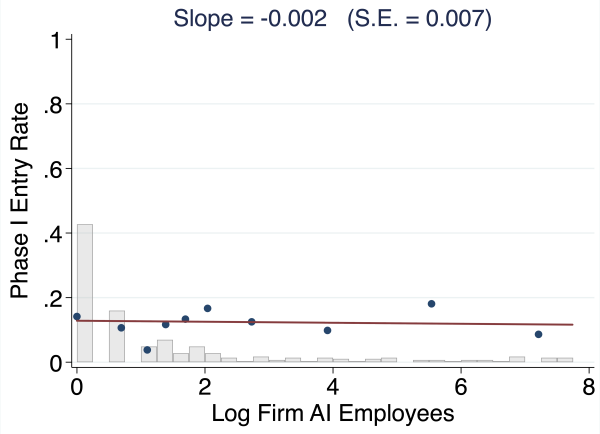
(a) New Drug Projects, 2010



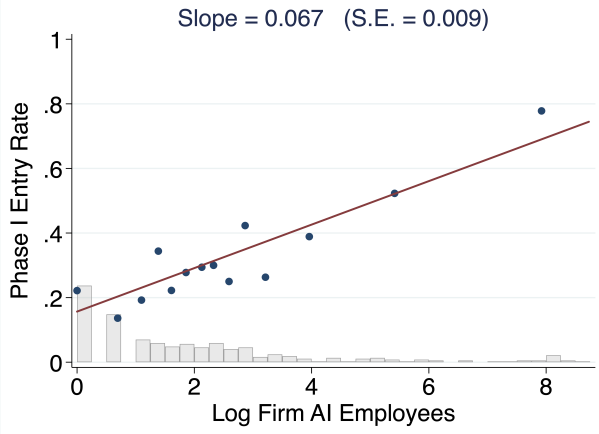
(b) New Drug Projects, 2025



(c) Phase I Entry, 2010



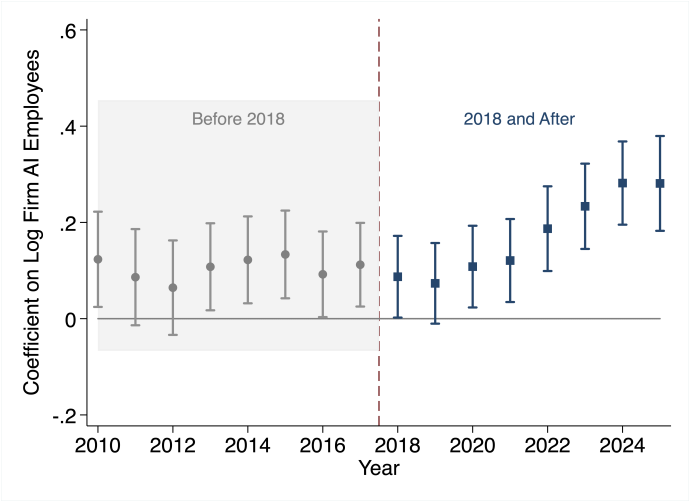
(d) Phase I Entry, 2025



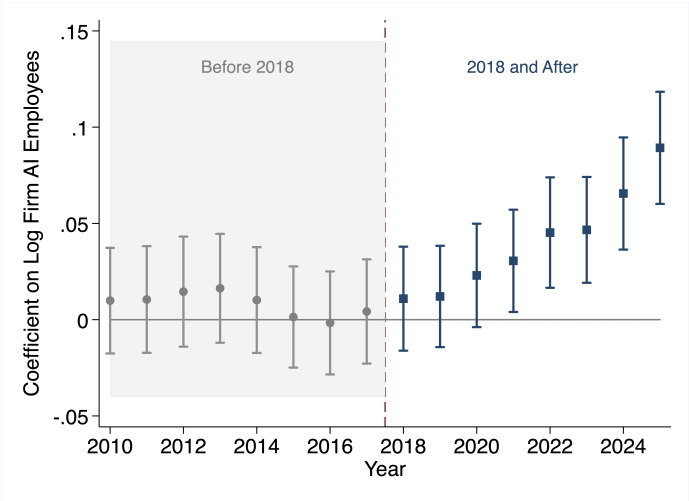
Notes: Panels (a) and (b) plot the number of new drug projects initiated by firms in 2010 and 2025 against log(1 + AI employees); panels (c) and (d) plot the fraction of those projects subsequently entering Phase I. Firms are grouped into 20 quantile bins of log AI employment, and points show within-bin means. Solid lines show OLS fits; reported slopes have heteroskedasticity-robust standard errors. Histograms show the distribution of log AI employment. The sample includes Revelio-matched US and European firms averaging 5–1,000 employees; panels (c) and (d) require at least one new project in the respective year.

Figure 5: AI Employees and Drug Development: Event Study

(a) New Drug Projects

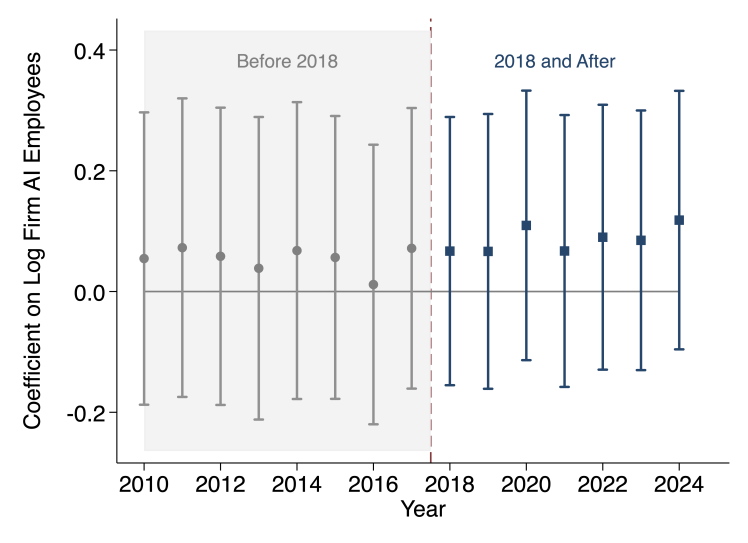


(b) Phase I Entry



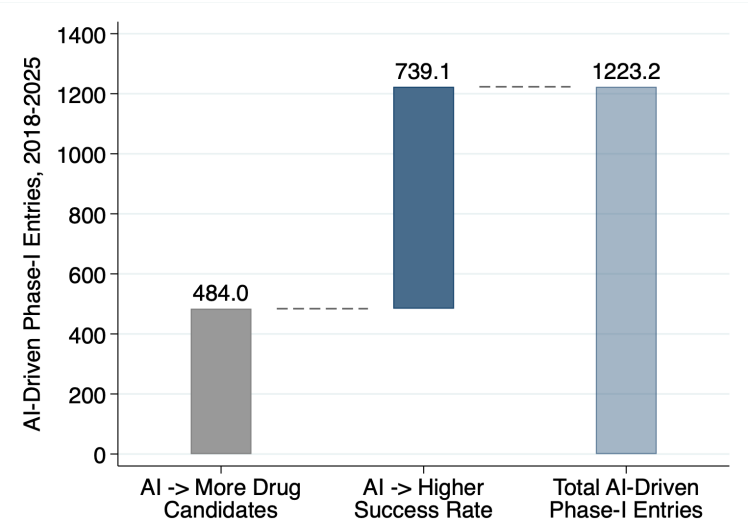
Notes: Panel (a) reports year-specific PPML coefficients of new drug projects on log(1 + firm AI employees), estimated by interacting AI employment with year indicators. Panel (b) reports the corresponding coefficients from drug-candidate-level linear probability models of Phase I entry. Both omit the 2018 interaction; points show the main AI coefficient plus the respective year interaction. Panel (a) includes firm and country-by-year fixed effects and controls for log total employment and log prior launched drugs. Panel (b) includes firm-by-disease, disease-by-year, and country-by-year fixed effects and the controls in Table 2, column 3. The sample covers Revelio-matched US and European firms averaging at least five employees, from their founding years onward. Vertical bars are 95 percent confidence intervals, based on standard errors clustered by firm in panel (a) and by firm–disease in panel (b).

Figure 6: Placebo: AI Employees and Generic Trials



Notes: The figure repeats the year-specific PPML specification in Figure 5, panel (a), using the number of generic-drug trials initiated by a firm as the dependent variable. Generic trials are identified using a GPT-based classifier. The regression controls for log total employment and log prior trials; standard errors are clustered by firm. The sample covers 2010–2024.

Figure 7: Decomposition of AI-Driven Phase I Entries, 2018–2025



Notes: The figure decomposes the predicted AI-driven increase in Phase I entries over 2018–2025 into the contribution of additional drug candidates and that of higher Phase I progression rates. The first component holds the progression rate at its no-AI counterfactual; the second evaluates the change in progression rates at the predicted number of candidates under observed AI adoption. The third bar reports their sum. The counterfactual sets AI employment to zero while holding other covariates and fixed effects at observed values.

Tables

Table 1: Summary Statistics

Variable	N	Mean	SD	Median	Min	Max
Panel A: Firm-Year Level						
New Drug Projects	21,447	0.82	2.31	0	0	60
AI Employees	21,447	51.33	333.25	1	0	6,360
Total Employees	21,447	1,887.03	11,140.31	42	0	164,056
Prior Launched Drugs (Firm)	21,447	1.74	10.59	0	0	205
Panel B: Drug-Candidate Level						
Entering Phase I	20,586	0.22	0.41	0	0	1
AI Employees	20,586	307.93	822.17	2	0	6,360
Total Employees	20,586	11,652.67	28,883.88	79	0	164,056
Prior Launched Drugs (Firm-Disease)	20,586	3.83	9.09	0	0	64
AI Research Papers (Disease, 3 Years)	20,586	757.81	834.13	451	5	3,579
Research Papers (Disease, 3 Years)	20,586	36,261.71	19,816.35	31,561	697	65,542
Entering Phase II	20,586	0.15	0.35	0	0	1
Entering Phase III	20,586	0.07	0.25	0	0	1
Preclinical to Phase I (Days)	5,457	466.70	656.43	176	0	4,660

Notes: Panel A reports firm-year summary statistics, and Panel B reports drug-candidate-level statistics. The sample includes Revelio-matched US and European firms with at least five employees on average.

Table 2: AI Employees, New Drug Projects, and Phase I Entry

	Firm-Year Level		Drug Level	
	(1)	(2)	(3)	(4)
	PPML	IV-PPML	OLS	IV-2SLS
Log AI Employees	0.207***	2.594**	0.059***	0.254**
	(0.043)	(1.094)	(0.013)	(0.125)
Log Total Employees	0.227***	-0.934*	-0.121***	-0.234***
	(0.045)	(0.539)	(0.013)	(0.071)
Log Prior Launched Drugs (Firm)	0.283***	-0.409		
	(0.098)	(0.355)		
Log Prior Launched Drugs (Firm-Disease)			0.088***	-0.021
			(0.030)	(0.070)
Constant	-1.077***		0.851	
	(0.233)		(4.112)	
Sponsor FE	Yes	Yes		
Sponsor-Category FE			Yes	Yes
Category-Year FE			Yes	Yes
Country-Year FE	Yes	Yes	Yes	Yes
First-stage F		11.90		19.71
Observations	21427	21427	18301	18301

Notes: Columns 1–2 report firm-year models of the number of new drug projects, estimated using PPML and control-function IV-PPML, respectively. Columns 3–4 report drug-candidate-level linear probability models of entry into Phase I, estimated using OLS and 2SLS. Log AI Employees is log(1 + AI employees). The instrument uses cumulative AI graduates since 2010 from the universities attended by a firm's early employees, weighted by each university's share among those employees. Column 2 includes the first-stage residual in the PPML regression; standard errors are based on 100 firm-clustered bootstrap replications. The first-stage F-statistic is the squared first-stage t-statistic in column 2 and the Kleibergen–Paap Wald F-statistic in column 4. Firm-year specifications control for log total employment and log prior launched drugs; drug-level specifications additionally account for prior firm–disease launches and disease-level research. Firm-year models include firm and country-by-year fixed effects; drug-level models include firm-by-disease, disease-by-year, and country-by-year fixed effects. The sample includes Revelio-matched US and European firms averaging at least five employees, from their founding years onward. Standard errors are clustered by firm in columns 1, 2, and 4 and by firm–disease in column 3.

Table 3: Heterogeneity by Firm Age and Country

	Number of New Drugs		Entering Phase I	
	(1)	(2)	(3)	(4)
	Young Firms	United States	Young Firms	United States
Log AI Employees	0.246***	0.225***	0.059***	0.052**
	(0.043)	(0.067)	(0.013)	(0.025)
Log AI Employees × Young Firm	-0.171***		-0.027**	
	(0.041)		(0.013)	
Log AI Employees × US Firm		-0.026		0.011
		(0.076)		(0.027)
Log Total Employees	0.239***	0.228***	-0.103***	-0.121***
	(0.043)	(0.045)	(0.013)	(0.013)
Log Prior Launched Drugs (Firm)	0.191*	0.283***		
	(0.099)	(0.098)		
Log Prior Launched Drugs (Firm-Disease)			0.043	0.089***
			(0.030)	(0.030)
Constant	-1.130***	-1.080***	0.453	0.845
	(0.229)	(0.233)	(4.072)	(4.099)
Sponsor FE	Yes	Yes		
Sponsor-Category FE			Yes	Yes
Category-Year FE			Yes	Yes
Country-Year FE	Yes	Yes	Yes	Yes
Observations	21427	21427	18301	18301

Notes: Columns 1–2 extend the firm-year PPML specification in Table 2, column 1; columns 3–4 extend the drug-level OLS specification in column 3. Log AI employees is interacted with an indicator for firms no more than five years old (columns 1 and 3) or for US-headquartered firms (columns 2 and 4), with European firms as the comparison group. Controls, fixed effects, sample restrictions, and clustering follow the corresponding baseline specifications.

Table 4: AI Applications Across Drug Development Stages

	Firm-Year Level		Drug Level	
	(1)	(2)	(3)	(4)
	PPML	PPML	OLS	OLS
Log AI–Discovery Workers	0.292***	0.350***	0.063***	0.064**
	(0.043)	(0.079)	(0.013)	(0.026)
Log AI–Clinical Workers		-0.128**		-0.038**
		(0.058)		(0.018)
Log AI–Process Workers		-0.024		0.002
		(0.078)		(0.026)
Log AI–Other Workers		0.160***		0.078***
		(0.050)		(0.016)
Log Total Employees	0.222***	0.222***	-0.114***	-0.119***
	(0.041)	(0.042)	(0.013)	(0.013)
Log Prior Launched Drugs (Firm)	0.265***	0.203**		
	(0.095)	(0.102)		
Log Prior Launched Drugs (Firm-Disease)			0.088***	0.052*
			(0.030)	(0.030)
Constant	-1.278***	-1.317***	1.064	0.642
	(0.211)	(0.211)	(4.017)	(4.058)
Sponsor FE	Yes	Yes	No	No
Sponsor-Category FE	No	No	Yes	Yes
Category-Year FE	No	No	Yes	Yes
Country-Year FE	Yes	Yes	Yes	Yes
Observations	21,427	21,427	18,301	18,301

Notes: Columns 1–2 report firm-year PPML regressions of new drug projects; columns 3–4 report drug-candidate-level linear probability models of Phase I entry. Because firms' specific AI applications are not directly observable, we proxy for AI use at different stages of drug development using workers who possess both AI skills and domain-specific skills in drug discovery, clinical development, or drug development processes, along with a residual category of other AI workers. Each measure is defined as the log of one plus the corresponding number of workers. Columns 1 and 3 include only AI–discovery workers, while columns 2 and 4 jointly include all four categories. All specifications include the corresponding controls from Table 2. Columns 1–2 include firm and country-by-year fixed effects, with standard errors clustered by firm. Columns 3–4 include firm-by-disease, disease-by-year, and country-by-year fixed effects, with standard errors clustered by firm–disease. Chinese firms are excluded from the sample.

Table 5: AI Employees and Complementary Knowledge

	Number of New Drugs			Entering Phase I		
	(1) AI Papers	(2) Prior Launches	(3) Target Evidence	(4) AI Papers	(5) Prior Launches	(6) Target Evidence
Log AI Employees	0.162***	0.202***	0.097***	0.021	0.048***	0.051***
	(0.036)	(0.037)	(0.034)	(0.013)	(0.013)	(0.018)
Log AI Employees × AI Research Papers (Disease)	0.080***			0.029***		
	(0.011)			(0.003)		
Log AI Employees × Prior Launched Drugs (Firm-Disease)		0.098***			0.038***	
		(0.025)			(0.008)	
Log AI Employees × Target Evidence			0.081***			0.005***
			(0.012)			(0.001)
Constant	1.116	2.116	-3.838***	-0.656	0.028	-4.295
	(10.617)	(11.266)	(0.171)	(4.203)	(4.287)	(6.834)
Sponsor-Category FE	Yes	Yes		Yes	Yes	Yes
Category-Year FE	Yes	Yes		Yes	Yes	Yes
Sponsor-Target FE			Yes			
Target FE			Yes			
Country-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	58837	58837	117997	18301	18301	9140

Notes: Columns 1–2 report firm–disease-category–year PPML regressions of new drug projects; column 3 reports firm–target–year PPML regressions; and columns 4–6 report drug-candidate-level linear probability models of Phase I entry. Log AI employees is interacted with AI-related publications in the disease category over the preceding three years (columns 1 and 4), the firm's prior launched drugs in the disease (columns 2 and 5), or OpenTargets evidence (columns 3 and 6). AI publications and the target evidence measure are standardized, and interaction variables are mean-centered. Column 6 includes only candidates matched to OpenTargets. Columns 1–2 and 4–6 include firm-by-disease, disease-by-year, and country-by-year fixed effects, while column 3 includes firm-by-target, target, and country-by-year fixed effects. All specifications include the corresponding controls from Table 2, as applicable. Standard errors are clustered by firm–disease (columns 1–2 and 4–6) or firm–target (column 3).

Table 6: AI Employees and Target Evidence

	Entering Phase I			
	(1) Literature	(2) Human/Molecular	(3) Clinical	(4) All Evidence
Log AI Employees	0.051***	0.040**	0.058***	0.041**
	(0.019)	(0.019)	(0.019)	(0.019)
Log AI Employees × Literature Evidence	0.007***			0.004*
	(0.002)			(0.002)
Log AI Employees × Human/Molecular Evidence		0.014***		0.013***
		(0.002)		(0.002)
Log AI Employees × Clinical Evidence			0.001	-0.002
			(0.002)	(0.002)
Constant	-5.754	-4.211	-4.876	-3.569
	(7.543)	(7.431)	(7.610)	(7.398)
Sponsor-Category FE	Yes	Yes	Yes	Yes
Category-Year FE	Yes	Yes	Yes	Yes
Country-Year FE	Yes	Yes	Yes	Yes
Observations	9140	9140	9140	9140

Notes: The table reports drug-candidate-level linear probability models of Phase I entry. Log AI employees is interacted with three types of target evidence from Open Targets: scientific literature (column 1), human and molecular evidence (column 2), and existing drug evidence (column 3). Column 4 includes all three interactions simultaneously. Human and molecular evidence combines genetic associations, genetic literature, somatic mutations, RNA expression, and affected pathways. All evidence measures are transformed using log(1 + evidence count) and standardized. The sample is restricted to drug candidates matched to Open Targets. All specifications include the corresponding controls from Table 2, firm-by-disease, disease-by-year, and country-by-year fixed effects. Standard errors are clustered at the firm–disease-category level.

Table 7: AI Employees and the Scope of Drug Development

	Number of Categories / Targets				Concentration (HHI)			
	(1) Disease Category	(2) Disease Subcategory	(3) Modality	(4) Target	(5) Disease Category	(6) Disease Subcategory	(7) Modality	(8) Target
Log AI Employees	0.218***	0.228***	0.215***	0.164***	-0.019**	-0.021*	-0.014**	0.016
	(0.037)	(0.045)	(0.035)	(0.054)	(0.008)	(0.011)	(0.006)	(0.015)
Log Total Employees	0.094**	0.044	0.136***	0.222***	0.006	0.018	-0.006	-0.016
	(0.037)	(0.047)	(0.035)	(0.059)	(0.009)	(0.012)	(0.006)	(0.014)
Log Prior Launched Drugs (Firm)	0.439***	0.405***	0.326***	0.222*	-0.010	-0.011	-0.019	-0.021
	(0.082)	(0.091)	(0.084)	(0.114)	(0.018)	(0.024)	(0.017)	(0.030)
Constant	-0.958***	0.114	-1.403***	-1.211***	0.801***	0.529***	0.947***	0.736***
	(0.169)	(0.231)	(0.163)	(0.327)	(0.038)	(0.051)	(0.027)	(0.067)
Sponsor FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	21427	21427	21427	17047	6490	6490	6490	3949

Notes: Columns 1–4 report firm-year PPML regressions of the number of distinct disease categories, disease subcategories, therapeutic modalities, and molecular targets among newly initiated drug projects. Columns 5–8 report OLS regressions of the corresponding Herfindahl–Hirschman concentration indices, restricted to firm-years with new drug projects; columns 4 and 8 also require target information. All regressions include firm and country-by-year fixed effects, with controls for log total employment and log prior launched drugs. Standard errors are clustered by firm.

Table 8: AI Employees and Novelty

	Target Novelty		Combination Novelty		Structure Novelty	
	(1) Novelty	(2) Entering Phase I	(3) Novelty	(4) Entering Phase I	(5) Novelty	(6) Entering Phase I
Log AI Employees	0.004	0.056***	-0.007	0.055***	0.070*	0.012
	(0.008)	(0.018)	(0.016)	(0.018)	(0.036)	(0.050)
Log AI Employees × Target Novelty		-0.002				
		(0.008)				
Log AI Employees × Combination Novelty				0.005		
				(0.005)		
Log AI Employees × Structure Novelty						-0.011
						(0.020)
Constant	2.462	-2.329	5.177	-2.752	-3.102	-33.406*
	(2.906)	(5.722)	(4.689)	(5.644)	(10.344)	(17.603)
Sponsor-Category FE	Yes	Yes	Yes	Yes	Yes	Yes
Category-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Country-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	9689	9689	9689	9689	1336	1336

Notes: Columns 1, 3, and 5 regress target novelty, combination novelty, and structure novelty, respectively, on log AI employees. Columns 2, 4, and 6 examine Phase I entry and interact log AI employees with the corresponding novelty measure. All specifications are drug-candidate-level linear probability models with the fixed effects, controls, and clustering of Table 2, column 3. The sample is restricted to candidates with the relevant novelty information.

Table 9: AI Employees, Later Clinical Stages, and Development Speed

	Other Outcomes			
	(1) Enter Phase II	(2) Enter Phase III	(3) Preclinical to Phase I	(4) Trial Duration
Log AI Employees	-0.016	-0.014	-0.265	0.031
	(0.030)	(0.041)	(0.173)	(0.031)
Constant	-4.562***	2.281	7.913	-1.202
	(1.356)	(2.175)	(7.224)	(1.576)
Sponsor-Category FE	Yes	Yes	Yes	Yes
Category-Year FE	Yes	Yes	Yes	Yes
Country-Year FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Observations	2755	1646	3900	61298

Notes: Columns 1–3 report drug-candidate-level regressions of progression to Phase II conditional on Phase I entry, progression to Phase III conditional on Phase II entry, and log(1 + days from preclinical development to Phase I), respectively. Column 4 reports log(1 + trial duration in days), weighted so each trial has total weight one. Controls follow Table 2, column 3, except that column 4 uses prior firm–disease trials instead of prior launched drugs. All columns include firm-by-disease-category, disease-category-by-year, and country-by-year fixed effects. Standard errors are clustered by firm and disease category.

A Additional Tables

Table A.1: Robustness: Alternative Measures of AI Human Capital

	Number of New Drugs				Entering Phase I			
	(1) AI Roles	(2) Data Roles	(3) CS PhDs	(4) IT Roles	(5) AI Roles	(6) Data Roles	(7) CS PhDs	(8) IT Roles
Log AI Employees (AI Roles)	0.208***				0.069***			
	(0.047)				(0.011)			
Log Data Employees		0.126***				0.028**		
		(0.041)				(0.011)		
Log CS PhD Employees			0.433***				0.105***	
			(0.077)				(0.022)	
Log IT Employees				0.020				0.017
				(0.043)				(0.012)
Log Total Employees	0.363***	0.296***	0.330***	0.219***	-0.077***	-0.099***	-0.091***	-0.129***
	(0.037)	(0.039)	(0.036)	(0.046)	(0.011)	(0.012)	(0.012)	(0.014)
Log Prior Launched Drugs (Firm)	0.232**	0.326***	0.281***	0.281***				
	(0.102)	(0.098)	(0.099)	(0.098)				
Log Prior Launched Drugs (Firm-Disease)					0.053*	0.109***	0.085***	0.087***
					(0.031)	(0.030)	(0.030)	(0.030)
Constant	-1.702***	-1.458***	-1.681***	-1.073***	0.803	0.795	0.469	1.029
	(0.214)	(0.210)	(0.206)	(0.232)	(3.997)	(4.064)	(3.983)	(4.104)
Sponsor FE	Yes	Yes	Yes	Yes				
Sponsor-Category FE					Yes	Yes	Yes	Yes
Category-Year FE					Yes	Yes	Yes	Yes
Country-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	21427	21427	21427	21427	18301	18301	18301	18301

Notes: Columns 1–4 repeat the firm-year PPML specification in Table 2, column 1, and columns 5–8 repeat the drug-level OLS specification in column 3, using alternative measures of AI-related human capital. These are log(1 + employees in AI occupations), log(1 + employees in data occupations), log(1 + employees with a computer-science PhD), and log(1 + employees in IT occupations). Occupations are based on Revelio job classifications rather than reported skills. Other controls, fixed effects, and clustering follow the corresponding baseline specifications.

Table A.2: Robustness: Alternative Instruments

	Number of New Drugs			Entering Phase I		
	(1) CS Graduates	(2) A-CSR200	(3) C-CSR200	(4) CS Graduates	(5) A-CSR200	(6) C-CSR200
Log AI Employees	2.647	2.705*	2.332**	0.309***	0.317***	0.288***
	(1.618)	(1.399)	(1.185)	(0.102)	(0.117)	(0.107)
Log Total Employees	-0.958	-0.982	-0.800	-0.266***	-0.270***	-0.254***
	(0.770)	(0.697)	(0.590)	(0.060)	(0.069)	(0.064)
Log Prior Launched Drugs (Firm)	-0.432	-0.441	-0.334			
	(0.529)	(0.457)	(0.395)			
Log Prior Launched Drugs (Firm-Disease)				-0.051	-0.056	-0.039
				(0.061)	(0.063)	(0.059)
Sponsor FE	Yes	Yes	Yes			
Country-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Sponsor-Category FE				Yes	Yes	Yes
Category-Year FE				Yes	Yes	Yes
First-stage F	9.98	12.36	13.45	23.76	26.39	28.93
Observations	21427	21427	21427	18301	18301	18301

Notes: Columns 1–3 repeat the control-function IV-PPML specification in Table 2, column 2, while columns 4–6 repeat the 2SLS specification in column 4. Columns 1 and 4 use the baseline university-exposure instrument constructed with CS PhD graduates rather than AI graduates. The remaining columns interact the share of a firm's early employees trained at the top 200 US and European AI departments in CSRankings (based on 2000–2009 AI publications) with those departments' cumulative AI graduates (A-CSR200) or CS PhD graduates (C-CSR200) since 2010. First-stage F-statistics follow Table 2.

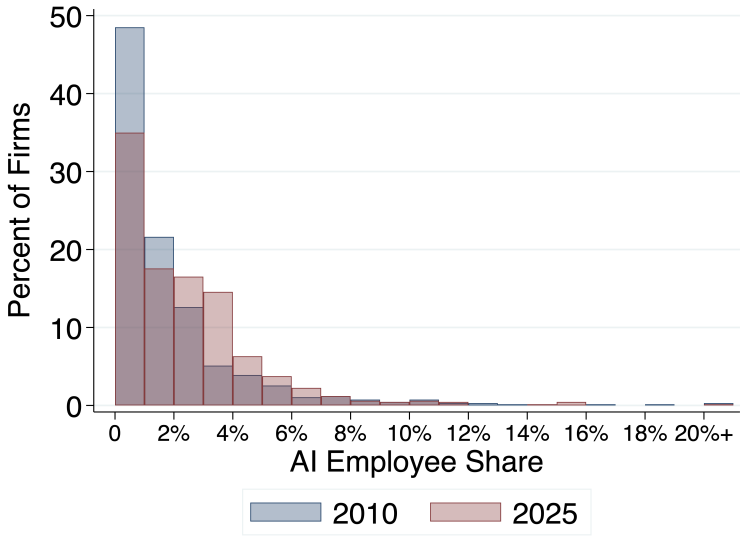
Table A.3: Robustness: Log New Drug Projects as the Dependent Variable

	Number of New Drugs		
	(1) Baseline	(2) Young Firms	(3) US Firms
Log AI Employees	0.029***	0.034***	0.047***
	(0.010)	(0.011)	(0.017)
Log AI Employees × Young Firm		-0.025**	
		(0.012)	
Log AI Employees × US Firm			-0.026
			(0.019)
Log Total Employees	0.029***	0.029***	0.030***
	(0.008)	(0.008)	(0.008)
Log Prior Launched Drugs (Firm)	-0.012	-0.021	-0.012
	(0.024)	(0.024)	(0.024)
Constant	0.224***	0.225***	0.222***
	(0.032)	(0.033)	(0.032)
Sponsor FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Observations	21427	21427	21427

Notes: The table reports firm-year OLS regressions of log(1 + new drug projects). Column 1 reports the baseline specification; columns 2 and 3 interact log AI employees with indicators for young firms (five years old or younger) and US-headquartered firms, respectively. All regressions control for log total employment and log prior launched drugs and include firm and year fixed effects. Standard errors are clustered by firm.

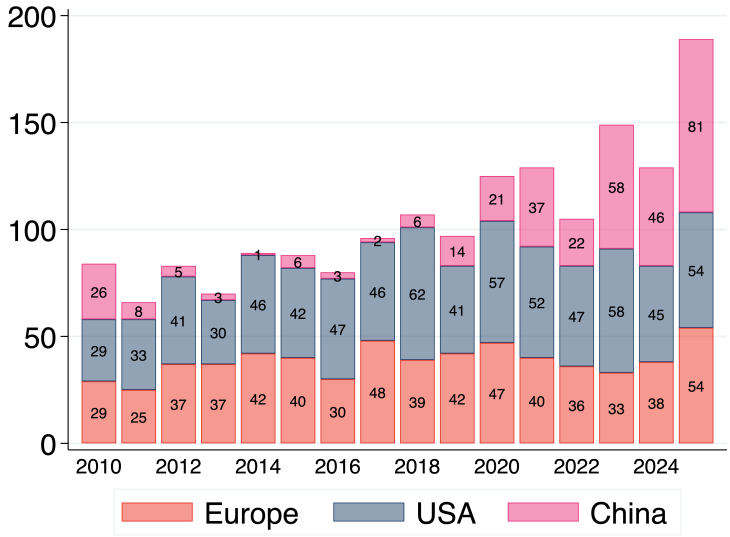
B Additional Figures

Figure B.1: Distribution of AI Employee Share



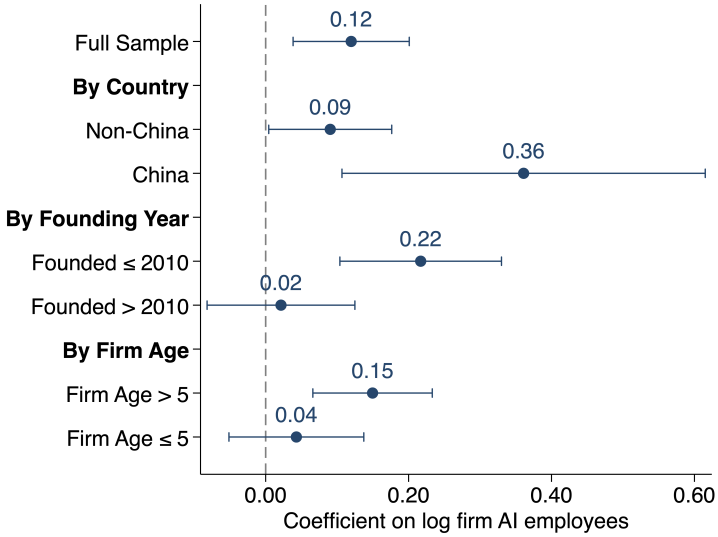
Notes: AI employees are workers in Revelio Labs with at least one skill in the AI skill cluster, including skills such as machine learning and deep learning. The figure plots the distribution of AI employees as a share of firm employment in 2010 and 2025; the final bin pools shares of 20 percent and above.

Figure B.2: Newly Approved Drugs



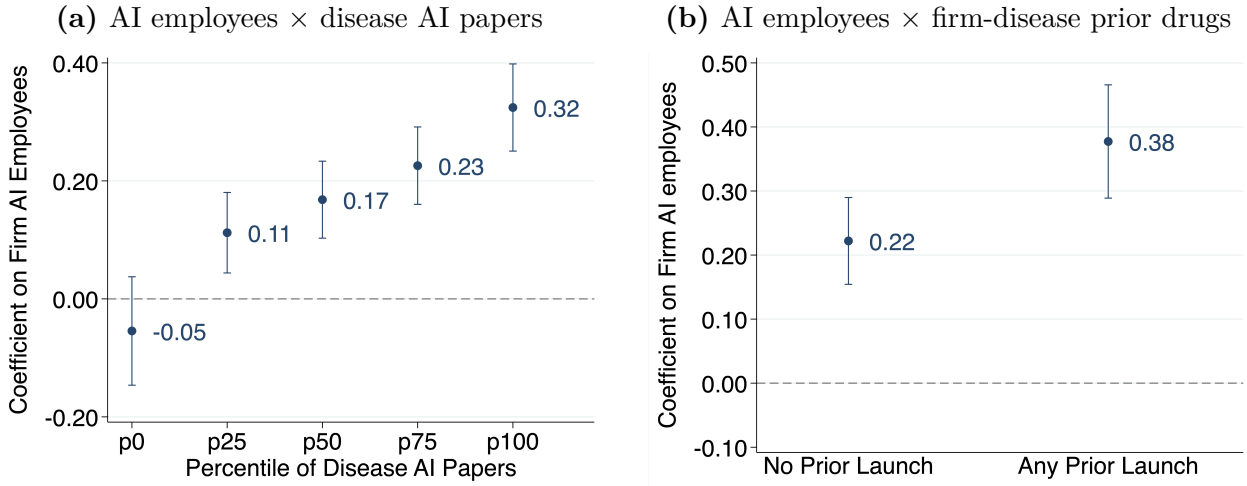
Notes: The figure counts newly approved drugs by year and country. Europe aggregates European countries.

Figure B.3: AI Employees and New Drug Projects: Heterogeneity



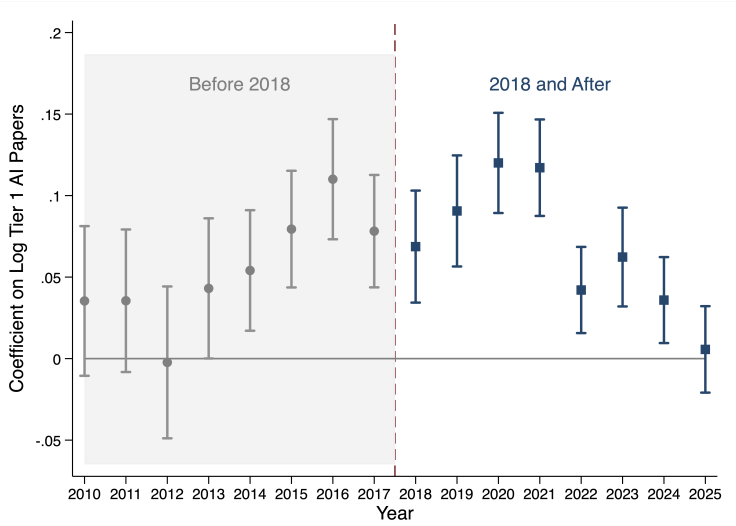
Notes: This figure summarizes the heterogeneity estimates. It reports Poisson coefficients on log firm AI employees across the firm subsamples shown in the figure. The dependent variable is the number of new drug projects initiated by the firm. All specifications include firm and year fixed effects and control for log total employment and log prior launched drugs. Standard errors are clustered at the firm level; bars report 95 percent confidence intervals.

Figure B.4: AI Employees and New Drug Projects: Complementary Knowledge



Notes: This figure complements Table 5 by showing how the coefficient on log firm AI employees varies across levels of complementary knowledge. Panel (a) groups observations by the number of AI-related publications in the disease over the preceding three years, while panel (b) groups observations by the firm's prior launched drugs in the disease. Regressions are estimated at the firm–disease–country–year level and include firm-by-disease, disease-by-year, and country-by-year fixed effects and the same controls as Table 5. Standard errors are clustered at the firm-by-disease level; bars report 95 percent confidence intervals.

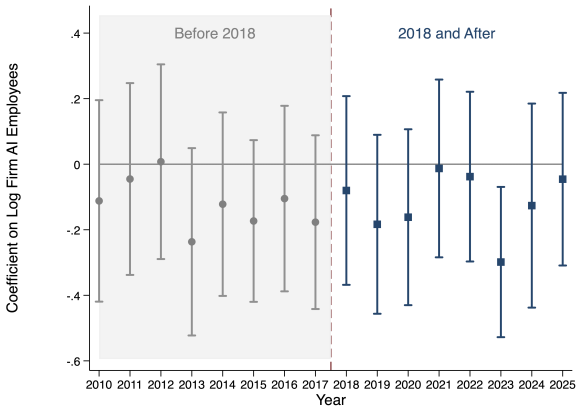
Figure B.5: Disease AI Science and Firm Entry: Event Study



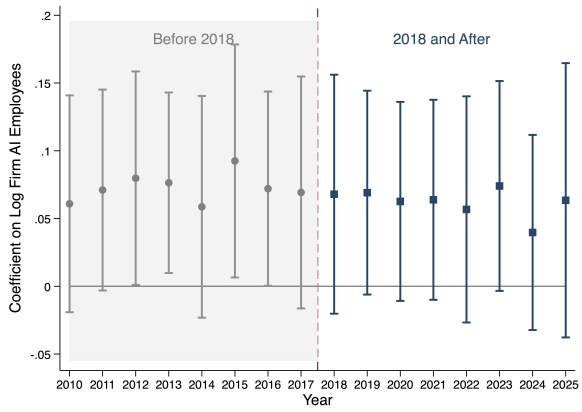
Notes: This figure reports year-specific Poisson coefficients on the log number of AI-related publications in a disease and country over the preceding three years. The dependent variable is the number of firms entering the disease in the country and year. The regression includes disease-by-country and year fixed effects and controls for log total biomedical publications over the preceding three years and prior drug development in the disease. Standard errors are clustered at the disease-by-country level; bars report 95 percent confidence intervals.

Figure B.6: AI Employees and Development Duration: Event Study

(a) Days from preclinical to Phase I



(b) Phase I trial duration



Notes: This figure reports year-specific coefficients on log firm AI employees from regressions at the drug-candidate level. The dependent variable is log days from preclinical development to Phase I among candidates that reach Phase I in panel (a), and log Phase I trial duration in panel (b). All specifications include firm-by-disease, disease-by-year, and country-by-year fixed effects and the same controls as the drug-level analysis in Table 2. Standard errors are clustered at the firm-by-disease level; bars report 95 percent confidence intervals.

C Measurement

This appendix documents how we measure AI in papers, patents, and clinical trials, how we place documents on a common disease axis, and how we assign papers to countries. Every measure is versioned so that results can be tied to a specific release of the classifier and crosswalk.

C.1 Data Layers

Table C.1 summarizes the four data layers used in the disease-level analysis. The unit of analysis for these layers is the disease-country-year, and outcomes run through 2023 to limit right-censoring. The firm-level analysis in the main text additionally uses Revelio Labs (Appendix C.5) and OpenTargets.

Table C.1: Data Layers

Layer	Source	Country assignment	Period
Paper	PubMed; AI flag from the PharmAI ontology	First and last author, fractional	2000–2025
Patent	Patents filed in China, the US, and Europe	Filing jurisdiction (applicant country)	2010–2025
Trial	Trials conducted in China, the US, and Europe	Sponsor country	2010–2024
Drug	Projects initiated by firms in China, the US, and Europe	Firm country	2010–2024

Notes: Each layer is aggregated to the disease-country-year level using the 332 Citeline disease indications (Appendix C.3). Papers are assigned to countries fractionally across first and last authors (Appendix C.4).

C.2 Classifying AI in Papers

A paper is classified as AI-related if its title or abstract reports the application of an AI method or names an AI system. Classification in this analysis relies on PharmAI, a structured ontology of AI methods and named systems in pharmaceutical research and development that we developed for this purpose. PharmAI maps alternative labels and synonyms to common concepts and records the relationships among them, which permits the identification of technologies referenced by name even when a publication does not explicitly mention arti-

ficial intelligence or machine learning. The release used in this paper, v0.3.46, comprises 654 entities: 442 classes inherited from the Artificial Intelligence Ontology (AIO), 49 additional method classes, and 163 named systems (Figure C.1). All results reported in this subsection, and all paper-based measures in the main text, are based on this release. The remainder of the subsection motivates the ontology, describes its composition, explains how it maps into a paper-level indicator, and compares the resulting measure with the keyword-based indicator used in earlier versions of this project.

The PharmAI ontology

Measuring AI adoption in the scientific literature requires a definition of AI that is both precise and applicable at scale to millions of abstracts. Existing approaches fall short of this standard in pharmaceutical research. Patent classification codes, the conventional instrument for identifying AI inventions, have no counterpart for scientific publications, and even among patents they systematically over- and under-identify AI.

The most comprehensive of these, AIO (Joachimiak et al., 2024), is an open, machine-readable ontology in which each concept carries a persistent identifier, a definition, synonyms, and hierarchical relations to broader concepts; a convolutional network, for instance, is recorded as a type of neural network, which is in turn a type of model. AIO is, however, domain-agnostic and heavily weighted toward deep learning. Of its 442 classes, 167 describe neural-network layers, 79 describe model and language-model variants, and 61 describe forms of statistical bias, whereas the machine-learning tasks that predominate in applied pharmaceutical research account for only 48. AIO also omits many classical and soft-computing methods that the pharmaceutical literature regards as AI, and it contains no named systems. PharmAI retains the structure of AIO and extends its coverage to pharmaceutical research and development.

PharmAI draws on three sources (Figure C.1, upper bar). The first is AIO itself: all 442 classes are inherited with their original identifiers, which preserves interoperability with AIO and with research that builds on it. The second is a set of 49 method classes that AIO lacks. To construct it, we synthesized fourteen review articles on AI in pharmaceutical research into a single inventory of methods and compared this inventory with the AIO class list.

The 26 methods that the reviews treat as AI but AIO omits span classical machine learning (gradient boosting, Gaussian processes, k-means, UMAP, and hidden Markov models), soft computing (genetic algorithms, swarm intelligence, fuzzy logic, and expert systems), and methods specific to the domain (QSAR modeling, machine-learning-augmented docking, knowledge graphs, diffusion models, and physics-informed neural networks). A further 23 classical methods, predominantly clustering algorithms such as DBSCAN and spectral clustering, were added from the methodological literature. The third source comprises 163 named systems that pharmaceutical research cites by name, compiled from the same reviews and from curated software inventories: 89 models (e.g., AlphaFold, RoseTTAFold, REINVENT, and DiffDock), 52 software tools (e.g., DeepChem, AutoDock Vina, and RDKit), and 22 company platforms (e.g., Insilico Medicine, Atomwise, and Recursion). Each named system is linked to the methods it implements.

Each of the 654 entities records a persistent identifier, a preferred label, synonyms and acronyms, the source in which it appears, and, where applicable, the stages of drug development it serves. Because the additions carry a distinct identifier prefix, their contribution beyond AIO can be measured directly. The complete ontology is available at https://pharm.jing-yi.org/pharmai_ontology.html.

AI paper classifier

The classifier examines the title and abstract of each paper for the entities in PharmAI. Each entity is matched through its preferred label, its synonyms, and its acronyms, and a match requires a word boundary on both sides of the term, so that an entity is not detected inside a longer, unrelated word.

Not every entity in the ontology counts toward the indicator. Each entity carries a tier (Figure C.1, lower bar), and the tier determines whether a match classifies the paper. Tier 1 comprises 230 entities, namely statistical and classical machine-learning methods, such as principal component analysis and Bayesian optimization. Tier 2 comprises 261 advanced AI methods, including neural networks, transformers and large language models, generative adversarial networks, and diffusion models. Tier 3 comprises the 163 named systems. A paper is classified as AI-related if its title or abstract matches at least one of the 424 entities in

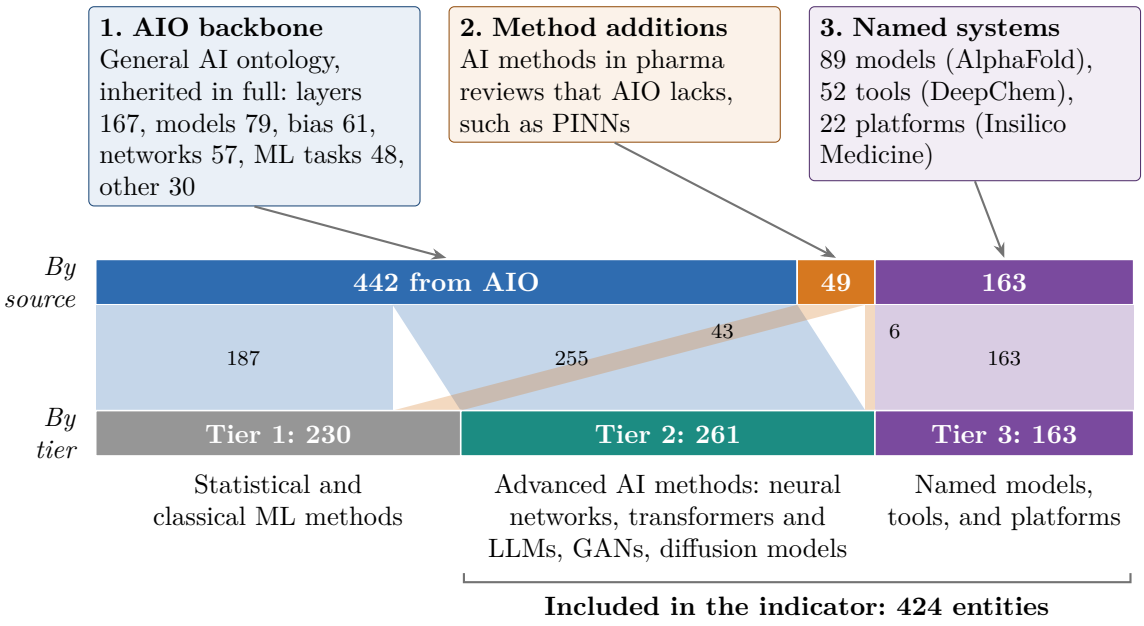


Figure C.1: Composition of the PharmAI Ontology, Release v0.3.46

Notes: Bar lengths are proportional to the number of entities. The top bar splits the 654 entities by source; the bottom bar splits them by tier, which determines whether a match contributes to the paper-level indicator. Shaded bands show how many entities from each source fall in each tier. Tier 1 includes 187 AIO classes (among them all 61 bias classes) and 43 of the 49 method additions; the 6 additions in Tier 2 are diffusion models, physics-informed neural networks, tool-augmented language-model agents, quantum kernel methods, tabular foundation models, and flow matching.

Tiers 2 and 3. As a robustness check, we also construct a broader indicator that classifies a paper as AI-related if its title or abstract matches at least one of the 654 entities in any of the three tiers.

Two sets of rules guard against false positives. The first concerns ambiguous acronyms. Short forms such as AI, ML, DL, GP, RL, and KG remain in the ontology but never classify a paper on their own: they count only when a Tier 2 or Tier 3 term, or the developer of the named system, also appears in the title or abstract. The second concerns polysemy, where an AI term shares its spelling with a non-AI concept. "Neural network" counts when it refers to a trained network with hidden layers, but not when it refers to networks of cultured neurons; chemical prefixes such as the DL- in DL-methionine do not trigger deep learning. Every match that passes through one of these rules is logged together with the reason, so that each classification can be traced back to the term and the rule that produced it. Table C.2 illustrates these rules.

Table C.2: Examples of Matching Decisions

		Classified as AI	
Term in title or abstract	Reason	Strict	Broad
convolutional neural network	Tier 2 method	Yes	Yes
AlphaFold	Tier 3 named system	Yes	Yes
principal component analysis	Tier 1 method	No	Yes
ML, with no other AI term	Ambiguous acronym; cannot trigger alone	No	No
neural network of cultured neurons	Polysemy guard	No	No
DL-methionine	Chemical prefix, not deep learning	No	No

Notes: Illustrative cases under PharmAI v0.3.46. *Strict* is the main indicator, under which a paper is AI-related if its title or abstract matches at least one entity in Tier 2 or 3 that passes the acronym and polysemy rules. *Broader* is the indicator used as a robustness check, which counts matches in all three tiers (654 entities) under the same rules.

C.3 Disease Crosswalk: PACT to Citeline-332

The PACT ontology

Papers, patents, and trials are placed on a common disease axis using PACT (the Patent

Table C.3: Diseases with the Highest AI Share of Papers, by Classifier

Legacy classifier		PharmAI v0.3.46	
Disease	AI / all papers	Disease	AI / all papers
Adrenal insufficiency	211/1,656 = 12.7%	Monkeypox	32/943 = 3.4%
Diabetic retinopathy	658/7,509 = 8.8%	CNS, meningioma	40/1,327 = 3.0%
CNS, meningioma	110/1,312 = 8.4%	Tremor	67/2,350 = 2.9%
Progressive supranuclear palsy	62/934 = 6.6%	Epilepsy	712/26,746 = 2.7%
Age-related macular degeneration	377/5,976 = 6.3%	Movement disorders	67/2,613 = 2.6%

Notes: The five Citeline disease indications with the highest share of AI-related papers under each classifier. Entries are the number of AI papers over the number of all papers in the disease.

And Clinical-Trial disease ontology), a coded disease ontology that we developed for this purpose. PACT assigns each disease a permanent identifier and maps the disease vocabularies of the three data sources onto it, so that a patent, a paper, and a trial concerning the same disease carry the same code. For the analysis, PACT codes are converted into the 332 disease indications of Citeline, the taxonomy of the trial and drug data. The results in this paper use PACT release v9 and the V8 release of the Citeline conversion. The remainder of the subsection motivates the ontology, describes its composition and structure, explains how documents are labeled, and documents the conversion to the Citeline taxonomy.

Each data source describes diseases in its own vocabulary, and none is adequate on its own. Table C.4 summarizes the four sources used to construct PACT, together with their strengths and limitations. Pharmaceutical patents carry codes in class A61P of the International Patent Classification, which record the therapeutic activity of an invention. These codes are assigned by examiners and are independent of language, an important property because many Chinese patents contain little English text. They are, however, coarse: the 22 main groups divide into only 149 subgroups, many of which name a mechanism of action or a target organ rather than a disease. PubMed papers are indexed with Medical Subject Headings (MeSH), which are specific but have no counterpart in patents. Trials and drug projects are coded to Citeline's 332 indications, grouped into nine therapeutic areas, which define the unit of analysis but are not attached to patents or papers. The same disease therefore appears under a different name in each source. HIV infection, for example, is A61P31/18 ("for HIV") in patents, descriptor D015658 in MeSH, "HIV" in Citeline, and MONDO:0005109 in the Mondo Disease Ontology (MONDO), an open ontology that assigns each disease a persistent

identifier, cross-references to other disease vocabularies, and a hierarchy of subtypes. PACT gathers these names under a single identifier.

Table C.4: Data Used to Construct PACT

Source	Content	Strengths	Limitations	Role in PACT
A61P (IPC subclass)	Examiner-assigned classification of the therapeutic activity of a patent; 22 main groups and 149 subgroups	Attached to every pharmaceutical patent; independent of language	Coarse; many subgroups name a mechanism, organ, or procedure (e.g., A61P23, anaesthetics) rather than a disease	Backbone of therapeutic areas; code layer for patents
Citeline (Pharma-projects)	Commercial taxonomy of drug development; 332 indications in nine therapeutic areas	Defines the disease scope of the trial and drug data	Not attached to patents or papers; includes non-disease indications (e.g., contraception, smoking cessation)	Disease level of PACT; link to the trial and drug data
MeSH	Controlled vocabulary of the U.S. National Library of Medicine for indexing the biomedical literature	Attached to indexed PubMed papers; specific	No counterpart in patents; indexing lags publication by one to two years	Code layer for papers; attached to 207 active classes
MONDO	Open disease ontology integrating DOID, OMIM, Orphanet, and other vocabularies	Persistent identifiers, cross-references, and a subtype hierarchy	Covers diseases only, not procedures or prevention; not attached to documents	Identifiers for 273 active classes; 35,800 subtype and reference concepts

Notes: Counts of active classes refer to PACT release v9, which contains 536 active classes. Indications that MONDO does not name, such as procedures and prevention, retain a PACT code but carry no MONDO or MeSH cross-reference.

PACT release v9 contains 536 active disease classes, which are the units used to label documents (Figure C.2). They derive from five reviewed components: 149 therapeutic-activity classes defined by A61P subgroups; 35 diseases seeded from A61P subgroups that have no Citeline counterpart; 42 diseases in which an A61P subgroup and a Citeline indication are unified; 290 Citeline indications without an A61P counterpart; and 20 residual classes for documents that fall within a therapeutic area but name no specific disease in it. Of the 536

classes, 332 are linked to a Citeline indication and can therefore be joined to the trial and drug data. MONDO identifiers are attached to 273 of the active classes and MeSH descriptors to 207. Beneath the active classes, PACT retains 35,800 subtype and reference concepts derived from MONDO, for a total of 36,336 concepts. These reference concepts supply the subtype hierarchy and cross-references to other vocabularies but are not used as labels. The complete ontology is available at https://pharm.jing-yi.org/pact_complete_v9.html.

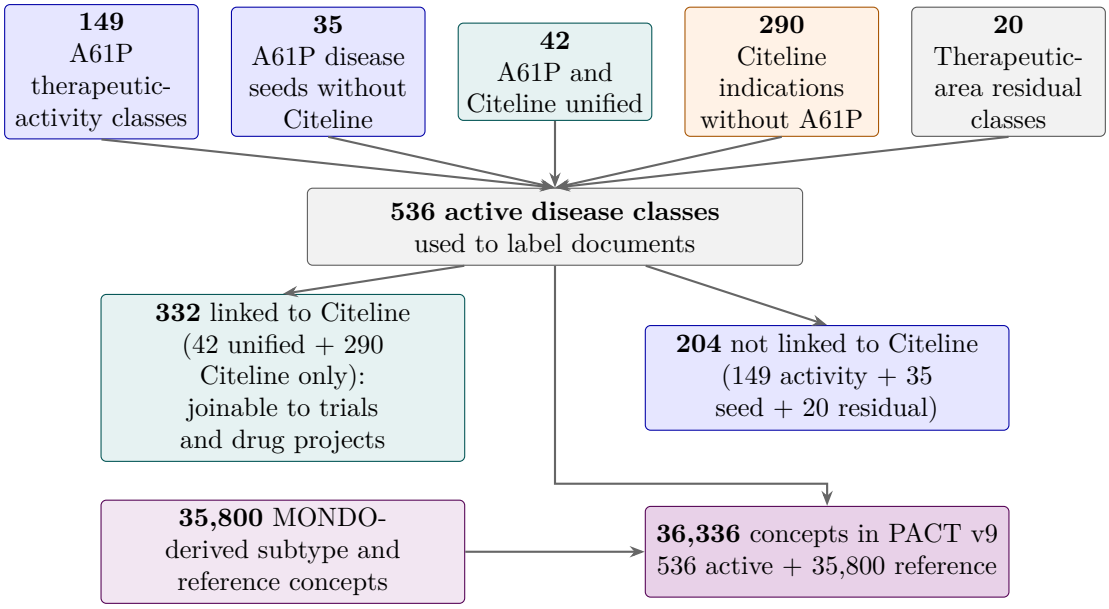


Figure C.2: Composition of the PACT Disease Ontology, Release v9

Notes: Five reviewed components combine into the 536 active disease classes used to label documents. Of these, 332 carry a link to a Citeline indication and can be joined to the trial and drug data. The complete ontology additionally contains 35,800 subtype and reference concepts derived from MONDO, which provide hierarchy and cross-references but are not used as labels.

PACT is organized in three levels. The first level comprises broad therapeutic areas, such as oncology, derived from the A61P main groups. The second level comprises the 536 active disease classes, such as breast cancer, which are the units to which documents are labeled. The third level comprises disease subtypes, such as triple-negative breast carcinoma, inherited from the MONDO hierarchy; subtypes are reached mainly by papers indexed to specific MeSH descriptors and roll up to their parent disease. A document is labeled at the most specific level that its text supports: a patent that mentions only "cancer" is assigned to the oncology area, whereas one that names breast cancer is assigned to that disease. Identifiers are opaque and permanent. The therapeutic area is recorded as a parent field rather than

encoded in the identifier, so renaming a disease or moving it to another area leaves its code unchanged, consistent with standard practice for released disease ontologies such as MONDO and the Human Disease Ontology.

Two layers are applied to every document. The code layer maps structured codes to PACT classes: A61P codes for patents and MeSH descriptors for papers. The text layer applies, for each active class, curated lists of included and excluded terms, compiled into word-boundary regular expressions and matched against titles, abstracts, and, for patents, claims. The two layers are complementary rather than sequential; each label records whether it was produced by the code layer, the text layer, or both. At least one disease label is assigned to 41.7 percent of U.S. patents, 54.7 percent of Chinese patents, 46.5 percent of European patents, and 45.0 percent of papers.

PACT and Citeline differ in granularity. PACT contains therapeutic-activity classes, broad parent diseases, and residual classes, whereas Citeline often enumerates disease subtypes; leukemia, for example, is a single PACT class but several Citeline indications. A forced one-to-one mapping would therefore create false precision. Conversion therefore proceeds as a cascade. Eight duplicate or superseded PACT codes are first collapsed, leaving 528 canonical codes. Codes with a reviewed one-to-one Citeline equivalent (348) are converted directly, without reading the document again. For codes broader than any single Citeline indication (163), the Citeline indication is recovered from the document itself: the text is searched first against a family of candidate indications specific to the code and then against all 332 indications, and an indication is assigned only if the text names it. A document labeled only "leukemia," for instance, is not assigned to acute myeloid leukemia unless its text says so. The remaining 25 codes (six area residuals and 19 classes defined by an agent or use rather than a disease, such as antioxidants or local anesthetics) are searched directly against all 332 indications in two further reviewed rounds. Documents for which no indication is supported retain their PACT code as unresolved rather than receiving a guessed label.

Resolution. Each PACT code is resolved by a cascade. Codes with a direct, curated mapping (348 codes) are assigned first. For the remaining codes with recovery enabled (163 codes), we try the code's shortlist of candidate diseases and, failing that, scan the text of

Table C.5: Source Vocabularies in the PACT Disease Ontology

Source	Contribution	Classifier layer	Comprehensive layer
A61P	Patent therapeutic-use codes (backbone)	536	536
Citeline	Clinical-trial indications	332	332
MONDO	Disease ontology graph (coverage)	273	36,073
MeSH	Biomedical vocabulary cross-reference	207	8,126

Notes: The classifier layer is the set of curated diseases used to label documents; the comprehensive layer is the set of reference concepts nested under those diseases. Entries are the number of concepts contributed by each source vocabulary to each layer.

all 332 Citeline diseases. Codes that yield no match (25 codes) remain unresolved. Recovery was built in reviewed rounds: a semantic patch with four corrections, a round re-enabling recovery for six disease residual categories, and a round covering 19 codes defined by agent or mechanism rather than disease. Precision guards remove "Supportive Care" from the text scan, because the phrase is common in cancer patents, and leave generic "cancer" with no specified type unresolved.

An independent explainer reproduces the mapping from the 332 Citeline indications to PACT: 151 indications merge onto curated diseases and 181 are added as new diseases. The reproduction matches the relation for all indications and the label for all single-label indications; only two multi-label ties differ. In the trial data, exact Citeline-332 names match 95.0 percent of disease tokens for China, 89.8 percent for the United States, and 89.7 percent for Europe (Table C.6). The most common unmatched tokens are "Unassigned" and vaccine categories, which are reported rather than counted as matched.

Table C.6: Trial Disease Mapping Audit

Region	Matched tokens	All tokens	Share matched
China	62,729	65,998	95.0%
United States	86,067	95,824	89.8%
Europe	69,516	77,471	89.7%

Notes: A token is one disease label attached to a trial record in Citeline TrialTrove. A token is matched if its name exactly matches one of the 332 Citeline indications. Unmatched tokens are mostly "Unassigned" and vaccine categories.

C.4 Assigning Papers to Countries

We assign each paper to countries using its first and last authors, with weights that sum to one. The first and last authors each receive a weight of 0.5. A single-author paper gives its author a weight of one, and if only one of the two authors' countries can be determined, that author receives the full weight. Papers for which neither country can be determined are excluded. An author with multiple affiliations splits the weight equally across them, and weights are then summed by country; for example, a last author affiliated with institutions in the United States and the United Kingdom contributes 0.25 to each. When an affiliation does not state a country, we infer it from a country name, a US state, or a demonym, and record the evidence used; about 7 percent of papers rely on an inferred country. The resulting allocation covers 11.2 million of 12.3 million papers.

C.5 Firm AI Human Capital

We measure firm AI human capital using Revelio Labs profiles, which are constructed from public professional profiles and contain information on job titles, skills, employers, and employment start and end dates. We define AI employees as workers whose job titles or skills are related to AI, machine learning, or data science, or who hold a PhD in computer science. We aggregate these workers to the firm-year level and also construct measures of total employment and firm age.

We match Citeline sponsors to Revelio company identifiers, accounting for subsidiaries and firm name changes, and exclude unmatched sponsors. The underlying Citeline data contain 8,011 firms and 44,310 drug projects; after matching and applying our sample restrictions, the regression sample contains 3,633 firms and 25,251 drug projects. We restrict the sample to firms with at least five employees on average, observed from their founding year through 2025, and located in China, the United States, or Europe. As alternative measures of AI human capital, we use the share of employees classified as AI workers and an indicator for AI-intensive firms, defined as firms with an AI employee share of at least 5 percent. Revelio coverage is relatively thinner in China, and AI-related job titles are noisier before 2015.