From Discovery to Treatment: AI and the Economics of Pharmaceutical Innovation

Kristopher J. Hult[1]

November 2026

Abstract

Artificial intelligence promises to transform drug discovery. But will an explosion of discoveries produce a comparable increase in treatments? We argue that the answer depends on a second, often overlooked form of innovation: discovering how to use new technologies. A novel drug molecule is a platform whose value depends on developing therapeutic applications. The same logic applies recursively to AI itself. AI creates scientific capabilities, but exploiting them requires additional investments in knowledge, validation, and application development. Improvements in discovery can therefore expand the frontier of pharmaceutical opportunities without proportionately increasing the production of treatments.

We develop an economic framework in which the incentives to create new molecular platforms and develop their applications depend on appropriability, complementary research capabilities, and competition for scarce R&D resources. The central issue for a medical innovation system is not only how many scientific opportunities arise, but who has incentives to undertake the costly work of realizing them. Different AI advances can therefore change the composition of innovation differently, and gains in scientific productivity need not translate directly into therapeutic output.

We investigate these ideas using clinical-development data and two distinct empirical approaches. Therapeutic areas with richer historical research opportunities exhibit stronger growth in novel molecular experimentation, without comparable growth in additional-use development. Exploiting AlphaFold's 2021 expansion of protein structural information, we also find stronger relative development of new indications for existing drugs associated with previously unsolved proteins. This differential is smaller for outcomes incorporating more advanced clinical-stage documentation, although we cannot directly measure subsequent progression to approved treatments.

Together, the evidence highlights a central economic challenge for the medical innovation system: creating new possibilities for treatment and developing treatments from those possibilities are separate innovation problems, shaped by different incentives and complementary capabilities.

Keywords: artificial intelligence; pharmaceutical innovation; novel molecules; incremental innovation; molecular platforms; complementary innovation; clinical development. JEL: O31, O33, I11, L65.

[1] NERA Economic Consulting

Introduction

AI may make scientific discoveries dramatically cheaper without making their applications equally profitable to develop. A promising molecule is not a medicine: someone must identify its uses, validate its biology, and finance clinical trials. If AI multiplies the discoveries available to researchers, who will invest in converting them into treatments? The answer depends on incentives to develop complementary knowledge and competition for scarce research resources—not simply on how quickly AI improves.

Pharmaceutical innovation offers a particularly revealing way to understand this distinction. A new drug molecule is more than a candidate for a single disease: it can become the foundation for therapies in several conditions. But those further uses are not contained in the molecule in any economically useful sense. They must be found, tested, financed, and developed through additional research. Hult (2015) describes this relationship as one between novel innovation, which creates molecular platforms, and incremental innovation, which develops their therapeutic applications. The incentive to discover a new molecule consequently depends in part on the value of uses that may be discovered later. Creating a platform and realizing the options embedded in that platform are complementary, but distinct, investment decisions.

Our central insight is that the same problem recurs at the level of the technology used to make discoveries. AI is itself a platform for pharmaceutical innovation. More capable models create opportunities to search different targets, evaluate new molecular designs, and generate hypotheses that would otherwise go unexplored. Those opportunities become productive only when researchers learn how to use the technology: which predictions to trust, how to connect them to biological experiments, how to identify commercially and clinically worthwhile indications, and how to incorporate results into development decisions. The relevant applications of AI must themselves be invented and refined. Thus there are two levels of platform innovation: new molecules and the research that develops their medical uses; and AI capabilities and the research that discovers their productive applications. This recursive connection is the organizing idea of the paper.

The allocation of R&D effort, rather than the speed of discovery alone, determines which scientific opportunities become clinical projects. Firms choose among originating molecules, developing their uses, and learning to exploit new research tools. The returns to those investments depend on rights to follow-on applications, product competition, and access to specialized scientists and clinical-development capacity.

Application knowledge is often costly to produce and difficult to appropriate. A validated assay or reliable AI-assisted workflow can benefit projects beyond the one that paid for it. Developers may capture only part of those returns while competing for the same scientists, experiments, and clinical resources needed to test a growing pool of candidates. Discovery can therefore accelerate even when investment in exploiting discoveries remains privately unattractive. This is an investment distortion, not necessarily a temporary engineering delay.

We develop a model that makes these choices explicit. At the molecular level, research effort can create new chemical platforms or discover further uses for existing platforms. A new molecule earns not only the return to an initial project but also an option on future, appropriable applications. Better research into additional uses therefore has two competing effects on novel discovery. It raises the expected value of owning a new molecule, encouraging origination; but it also draws on scarce research inputs and can

displace rents from competing products. The model identifies when the platform-value effect dominates the competition effect. This matters for AI because an advance in one research task need not shift resources toward the same type of innovation as an advance in another. The composition of pharmaceutical research is an economic outcome, not a direct readout of technological capability.

At the AI level, we distinguish the stock of available capability from the accumulated knowledge required to deploy that capability productively. Firms allocate research effort among originating molecules, developing additional therapeutic uses, and learning how to apply AI. Improvements in AI may directly increase the productivity of candidate generation, but investment in learning applications responds to privately captured returns and the cost of shared research inputs. The model allows a striking possibility: even when better AI increases the value of productive applications, firms may choose not to make the complementary investment if the private payoff remains below its marginal cost. When learning occurs, downstream development catches up only if the resulting improvement in application productivity is sufficiently large relative to the expansion of early experimentation. The model thus explains both a widening discovery-to-development gap and the conditions under which that gap can narrow.

Early experiments can generate information and valuable options even when they do not yet produce treatments. But a growing candidate pool cannot reveal whether firms have adequate incentives to validate predictions, pursue additional uses, and build clinical-development capacity. Those complementary investments must be made; time alone does not supply them.

Evidence from an important advance in AI-enabled science makes this question especially salient. Hill and Stein (2026) study the release of AlphaFold2, which suddenly made predicted protein structures widely available. They document increased basic research on proteins that had previously lacked structural information, while finding a much more muted response in applied research and early drug-development activity over the period they examine. The lesson is not that AI-generated scientific information lacks value; it is that an expansion of the scientific opportunity set and a response in commercial development are different outcomes. Their study identifies a specific information shock at the protein level. Our paper asks a complementary economic question: what incentives determine whether a discovery platform induces investment in its downstream applications? We examine that question through the relationship between novel molecular experimentation and later stages of pharmaceutical research.

Our empirical analysis separates three observable research decisions: first-observed clinical development of a molecule, first study of an additional therapeutic indication for a previously observed molecule, and documented activity in more advanced clinical phases. A new drug–indication start is a clinical-development project, not an approved indication or a new molecule. In RePairDB we observe 4,272 first appearances of parent molecules and 53,490 subsequent molecule–disease research starts across 33 therapeutic categories from 2005 through 2025. These records make it possible to compare the margins of innovation emphasized by the model rather than infer treatment output from an aggregate count of AI-related drug candidates.

Our first empirical comparison uses the pre-2016 AI Research Opportunity Index (AIRO), constructed from historical clinical-research networks across therapeutic categories. These networks may facilitate evaluation of AI-generated hypotheses, but AIRO does not measure whether researchers adopted a particular AI system. A one-standard-deviation higher index is associated with 18.2% greater differential entry of first-observed molecules in 2016–2020, versus a 0.4% differential in new

therapeutic-use starts. To put the novel result in units, a category-year otherwise predicted to have ten first-observed molecule starts would have roughly 1.8 more under an 18.2% relative difference. That is an illustrative translation of the regression coefficient, not 1.8 molecules causally created by AI. The evidence is strongest for early molecular experimentation, not for a broad increase across every clinical-development margin.

The downstream distinction is equally important. In a separate mature-cohort analysis, the AIRO association with documented Phase II-or-higher activity is −7.4%, with a confidence interval spanning economically meaningful positive and negative effects. The data do not establish a reduction in the probability that a given molecule subsequently advances. Indeed, the decline in the share of recent entrants with Phase II+ records is concentrated among programs already at that stage at their first clinical observation; the share with Phase II+ documented strictly later is approximately unchanged in the mature cohorts. Thus, the evidence shows a clearer novel-entry response than later-stage expansion, not that AI-generated molecules fail to become treatments.

Our second empirical design exploits a different kind of AI-related opportunity. Unlike AIRO's broad measure of historical research-network richness, the 2021 AlphaFold release made predicted protein structures publicly available at a specific date, with especially large potential information gains for proteins lacking experimental structures. We use that change to study additional-indication development among drugs associated with those proteins. The two designs deliberately measure different exposures, research decisions, populations, and periods; comparing their coefficients cannot identify the relative effectiveness of different AI technologies. Together they ask whether expanding a research opportunity translates into proportionate investment farther along the pharmaceutical-development process.

The AlphaFold results make the economic scale tangible. In the unadjusted data, annual new-indication starts declined from 204 to 189 among 352 drugs with at least one previously unsolved target, compared with a decline from 2,006 to 1,232 among 1,964 other drugs. The central result is therefore a smaller relative decline, not an absolute increase. With drug-class-by-year effects, the estimate is +40.2% for a drug whose mapped targets are all dark compared with one whose targets are all bright; the fitted post-period differential corresponds to roughly 48 indication starts per year, or 2.1 per 100 sample drugs. The model-implied number is not a causal count of new medicines, and the coefficient falls sharply when the most active drugs are excluded. In a separate short-window comparison, the association is larger for all new indications than for those with Phase II+ documented, but those records are not verified phase transitions.

The paper connects these empirical margins to a common allocation problem. At the molecular level, firms decide whether to create platforms or develop their uses; at the AI level, they decide whether to build capabilities or invest in learning their applications. Appropriability, platform option value, and scarce R&D inputs govern both choices. AIRO captures historical research opportunity rather than observed AI adoption. AlphaFold supplies a dated structural-information shock, although its drug–target exposure is imperfectly measured. The evidence identifies changes in clinical research activity, not the number of treatments delivered to patients.

We first develop the model, then examine novel molecule starts, additional indications, and advanced clinical documentation using AIRO and AlphaFold. We conclude with implications for incentives to develop applications of scientific discoveries.

1. A Recursive Model of Pharmaceutical Platforms and Applications

1.1 Molecular platforms and the value of discovering their uses

A newly originated molecule can support several therapies, and its value can therefore exceed the return to the first indication studied. In Hult (2015), novel innovation expands the stock of molecular platforms and incremental innovation develops uses for inherited molecules. This is a complementarity across research activities, not merely a classification of drug approvals: an expected future indication can affect the incentive to originate the molecule today. Whether the option is realized depends on the costs and returns to follow-on research. Budish, Durvasula, Roin, and Williams (2025, revised 2026) show that enforceability of intellectual-property rights can have large consequences for investment in new uses of existing drugs. The model begins from this economic connection between present origination and future applications.

The implication is subtle. Cheaper additional-use discovery can raise the value of discovering an entirely new molecule because the new platform will support more valuable follow-on opportunities. At the same time, pursuing those uses can absorb research inputs and replace profitable products. Whether the two forms of innovation complement or compete at the margin is an equilibrium result rather than a technological assumption.

Consider a therapeutic research market with a stock N_t of distinct molecular platforms and a stock I_t of incremental innovations developed from those platforms. These stocks need not correspond to the number of unique firms or approved products. The key structural assumption inherited from Hult (2015) is that novel innovation creates future platforms and incremental innovation draws on, and partially exhausts, opportunities in the pre-existing platform stock.

Let z_t^N and z_t^I denote real R&D inputs devoted to novel and incremental activity. Hult's baseline flow technology is

$$n_t = \delta z_t^N, \qquad i_t = \delta \zeta(N_t, I_t) z_t^I, \qquad \zeta(N_t, I_t) = \theta \frac{N_t}{I_t},$$

with $\delta, \theta > 0$. The derivative of incremental productivity with respect to N is positive and with respect to I negative. In Hult's original model, firms choose R&D according to the expected returns of new drug products under quality-ladder competition. Equalized returns and balanced accumulation yield $i/n = \theta$ and $I/N = \theta$ in the standard reference normalization. This is a model of approval flows in the original dissertation; applying its economic production structure to early clinical starts requires an explicit change of observation stage, not a claim that the objects are identical.

To represent AI-assisted prediction, let $m(A)$ be the effective productivity of novel-molecule discovery and $d(A)$ that of incremental hypothesis generation. At the predetermined stock ratio $x = I/N$, the marginal research costs of generating an innovation flow are

$$a_N(m) = \frac{1}{\delta m}, \qquad a_I(d; x) = \nu(d; x) = \frac{x}{\delta \theta d}.$$

Both technologies may change as AI improves, but they need not change proportionally. A reduction in experimental screening cost, better structure information, or more effective hypothesis ranking can raise m or d only to the extent that the particular task contributes to project selection. An information-gap representation such as $\phi_j(A) = 1 + \lambda_j g_j a_j(A)$ is useful for motivating differences in task relevance, but it does not sign relative productivity unless task coverage, the quality of unexplored opportunities, and implementation are specified. We therefore use m and d as **effective**, economically relevant productivities rather than interpreting the fraction of hypothetical untested drug–disease pairs as measured AI exposure.

The productivity parameters therefore represent distinct tasks rather than a single undifferentiated "AI shock." A system that improves candidate screening need not improve indication selection to the same degree; even a model that improves both tasks can change the optimal research mix. The relative response depends on the option value of new platforms, competition among drug products, and shared development inputs. These are the economic margins that connect AI's technological capabilities to the composition of pharmaceutical experimentation.

This distinction gives a disciplined interpretation of the two empirical settings. AIRO is an ex ante measure of research-network opportunity, not an observed shift in either task-productivity parameter; its associations cannot be read as causal estimates of a particular AI capability. The AlphaFold release, by contrast, is a specific information event, but its effect on a_N versus a_I is not known a priori. Predicted protein structures can inform molecular origination, target validation, and the investigation of additional indications, depending on which decisions the information actually changes. Consequently, the model does not infer a structural productivity shift solely from a positive clinical-development coefficient. A useful empirical test must establish both the relevant outcome margin and why the technological information should change its expected return.

1.2 Rents, replacement, and scarce R&D inputs

A successful project earns a flow of operating surplus $\pi > 0$ until displaced by subsequent innovation in its economically relevant market. Let n and i denote contemporaneous rates of novel and incremental entry. The market's displacement hazard is

$$H = n + \gamma i, \qquad 0 \le \gamma \le 1, \qquad R(H) = \frac{\pi}{r + H}, \quad r > 0.$$

The parameter γ distinguishes market expansion from replacement. A new indication for a drug already used to treat another disease need not displace the highest-quality treatment in either market to the same degree as a novel rival. The equal-displacement assumption $\gamma = 1$ is a special case, not a neutral baseline. The function $R(H)$ is a tractable quality-ladder closure with endogenous profit duration. It is not a full indication-level demand system, and the appropriate empirical market definition remains an economic question.

A new molecular platform may earn future rents from additional uses developed internally or by external researchers under licensing arrangements. Instead of assigning an arbitrary linear coefficient to the option, let $s \in [0,1)$ denote the **ex ante share of the project's expected private value attributable to appropriable follow-on options at a reference incremental productivity** d_0. A parsimonious normalized value factor is

$$c(d;s) = 1 + \frac{s}{1-s}\left(\frac{d}{d_0}\right)^{\eta}, \qquad \eta \geq 0,$$

so that $c(d_0;s) = 1/(1-s)$ and the continuation option is exactly the fraction s of total project value at the reference state. The normalization is transparent but does not estimate s: future cash flows should formally be weighted by probabilities of development, date-specific discounting, remaining protection and rights, and the risk of failure. The common $R(H)$ applied to direct and future rents is a reduced form. Our quantitative examples are conditional on that approximation.

For a developer of an existing-molecule incremental project, let $\rho \in (0,1]$ represent its appropriable share of the contemporaneous project return. The integrated originator case takes $\rho = 1$ while allowing positive s. An external licensing arrangement can set $\rho < 1$ and a separate originator continuation claim; we do not impose a universal $\rho = 1 - \omega$ rule or treat sponsor identity as direct evidence of royalties. Consistent accounting requires a prospective platform option only if additional projects can in fact be financed and developed.

R&D resources can expand, but at increasing marginal cost. Project-specific costs have slopes $k_N, k_I > 0$, and shared specialized-resource congestion contributes $\kappa \geq 0$ per unit of total entry. Atomistic entry then generates the locally interior equilibrium conditions

$$\begin{aligned} c(d;s)R(n+\gamma i) &= a_N(m) + k_N n + \kappa(n+i), \\ \rho R(n+\gamma i) &= v(d;x) + k_I i + \kappa(n+i). \end{aligned}$$

This system allows total R&D to expand. The shared-input parameter can represent scarce experienced discovery chemists, validation teams, or clinical operations capabilities; the data do not directly quantify it. Distinguishing κ from purely type-specific expansion costs matters because resource competition need not dominate if inputs can respond elastically.

The shared-input term is central to the extension. A validation scientist, a disease model, or an experienced clinical-development team can help originate a molecule, evaluate a new indication, or discover how to use an AI system effectively. Expanding one activity raises the opportunity cost of the others unless the capacity of these inputs grows sufficiently. This is a tractable way to make complementary innovation an allocation decision rather than assume that every downstream capacity adjusts automatically.

1.3 Why better indication research can increase molecular origination

There are two opposing consequences of improving the discovery of additional therapeutic uses. New uses increase the future option value of owning a novel molecule. But they also bid for scarce R&D resources and can displace existing product rents. Proposition 1 isolates the economic condition under which the first effect dominates the second.

For fixed inherited stocks (N, I) and market parameters, equation (5) determines an interior flow pair. Conditional on a candidate displacement hazard H, both flow quantities decline as $R(H)$ declines, while $H = n + \gamma i$ increases in the candidate innovation flows. With nonnegative participation and positive own-input adjustment costs, the equilibrium can be computed from a scalar fixed-point condition; uniqueness holds for the maintained monotone-cost closure. Its comparative statics are especially transparent.

Define $h = -R'(H) = \pi/(r+H)^2 > 0$ and

$$A = k_N + \kappa + ch, \quad B = \kappa + c\gamma h, \quad C = \kappa + \rho h, \quad E = k_I + \kappa + \rho\gamma h.$$

Under $c \geq \rho$, $k_N, k_I > 0$, $0 \leq \gamma \leq 1$, and $\kappa \geq 0$, the Jacobian determinant is strictly positive:

$$D = AE - BC = k_N k_I + \kappa(k_N + k_I) + h\{ck_I + \rho\gamma k_N + \kappa(1-\gamma)(c-\rho)\} > 0.$$

Let $c_d = \partial c/\partial d$, $G = c_d R$, and $V = v/d > 0$. The derivative of novel and incremental flows with respect to an **incremental-specific** discovery gain is

$$\boxed{n_d = \frac{EG - BV}{D}, \qquad i_d = \frac{AV - CG}{D}.}$$

The two parts of n_d have distinct economic meanings. The term EG captures the higher expected value of owning a newly originated molecular platform. The term BV captures competition from additional-use entry through shared R&D costs κ and product-market displacement γh. The stronger productivity of incremental search does not automatically imply a negative novel response. It does so when the platform-value term is absent or insufficient.

Proposition 1 (platform value versus research competition). In the regular interior equilibrium satisfying (3)–(7), an incremental-specific productivity increase changes novel entry with sign(EG − BV) and incremental entry with sign(AV − CG), with terms defined above. If appropriable follow-on option value is absent (G=0), the additional-use productivity improvement reduces novel entry locally. Sufficiently strong platform option value reverses that response.

The result captures a distinctive feature of pharmaceutical innovation. Better methods for discovering new drug uses need not draw innovators away from originating new drugs: they can make new drug platforms more profitable to create. Which force prevails depends on how prospective indications are valued and appropriated and on how strongly additional-use research competes for scientists and product rents. AI can therefore affect molecular origination even when the immediate technical advance is in indication discovery rather than molecule design.

$$n_m = \frac{E\, a_N/m}{D} > 0, \qquad i_m = -\frac{C\, a_N/m}{D} < 0.$$

Thus novel-specific AI raises novel entry and draws some investment away from incremental development at a locally interior equilibrium, while incremental-specific AI may do the opposite. The net effect of broad AI on composition depends on both m_A and d_A, not on an area-level label alone. Even an effective improvement in both tasks need not increase both entry types simultaneously, because competition and platform returns operate at the margin.

1.4 The long-run composition of innovation

The connection to Hult is substantive rather than a rhetorical citation. Set $s = 0$, $\rho = 1$, $\gamma = 1$, $m = d = 1$, and let $k_N, k_I, \kappa \to 0$, preserving positive entry. Both entry conditions become $R = 1/\delta$ and $x = \theta$. Balanced stock accumulation requires $i/n = I/N = x$, yielding Hult's familiar shares $n/(n+i) = 1/(1+\theta)$ and $i/(n+i) = \theta/(1+\theta)$. The total innovation rate is determined by the quality-ladder value condition. This is the relevant *limiting equilibrium*, not an identity between the enriched model and every welfare or demand assumption in the dissertation.

Proposition 2 (relative task productivity in the frictionless balanced-stock limit). With zero platform option, fully appropriable current development, equalized private returns and zero convex expansion costs, an interior balanced-stock equilibrium with constant m, d satisfies

$$\frac{I}{N} = \frac{i}{n} = \theta \frac{d}{m}.$$

In that limit, a proportional gain in incremental discovery relative to novel discovery increases the long-run incremental share; a common proportional gain has no effect on the ratio. With $\gamma = 1$ the total innovation hazard is $H = \delta m\pi - r$ when positive. With $\gamma \neq 1$, this hazard is apportioned across the two innovation types according to (10). The composition result follows from equalized costs and balanced stock accumulation; the formal proof and conditions are in Appendix A.2.

The frictionless benchmark gives a transparent comparative static: a relative improvement in finding indications raises the long-run share of additional-use innovation. The richer equilibrium explains why that prediction is conditional. Product-market displacement, appropriability, shared capacity and the evolving stock of usable molecular platforms all affect the research mix. This distinction is important for AI, which can improve different tasks at different rates and can change the returns to future applications at the same time as it changes direct discovery costs.

The inherited stocks change as

$$N_{t+1} = N_t + n_t, \quad I_{t+1} = I_t + i_t, \quad x_{t+1} - x_t = \frac{i_t - x_t n_t}{N_t + n_t}.$$

Novel entry expands the future base for follow-on research; incremental entry uses up some of the stock-adjusted opportunity set. A shock that initially tilts investment toward incremental development may reduce its relative attractiveness later as the stock I rises. Conversely, novel entry today can increase future incremental options. We do not extrapolate the local derivative (8) into a complete long-run welfare result without solving firms' forward-looking investment in both stages and evolving rights. The model's welfare content is also limited: the number of additional projects is distinct from patient health impact, market access, incremental therapeutic benefit, and avoidance of low-quality projects.

1.5 AI is a second platform: productive applications must be invented

The molecular platform logic now recurs at the level of the discovery technology. A molecule makes new therapies possible; AI capability makes new research activities possible. Let K denote available AI capability, and let L denote the stock of knowledge, validated procedures, and organizational routines that make that capability useful in pharmaceutical work. Increasing K is not the same decision as investing in L. A model may become much better at generating hypotheses while firms still need to determine which predictions merit experiments, which data support clinical inference, and which workflows improve development decisions. As with molecular indications, useful applications are economic innovations: costly to find, partly reusable, and imperfectly appropriable.

$$L_{t+1} - L_t = \delta_l \varphi(K_t/L_t)\, z_{l,t}, \quad \varphi' > 0; \quad benchmark\ \varphi(x) = x.$$

The recursive analogy produces an allocation problem with a clear private-investment margin. The opportunity to learn AI applications grows with the capability-to-knowledge gap K/L, but application R&D competes for scientists and development capacity with molecule origination and indication

discovery. Better capability can increase the gross return to learning while simultaneously raising congestion costs through stronger candidate generation. The general-purpose-technology literature emphasizes co-invention and delayed intangible returns (Bresnahan and Trajtenberg 1995; Brynjolfsson, Rock, and Syverson 2021); the distinctive contribution here is to put that second investment decision inside an economy where molecular platforms themselves have endogenous follow-on value.

The recursive platform analogy yields an economic result that does not depend on the clinical pattern we subsequently observe: firms must choose whether to invest in the knowledge that makes AI capability productive. We derive that investment condition first. We then show how an initial difference between entry and documented advanced development can arise as a conditional implication when complementary capacity adjusts slowly.

Proposition 3 (appropriability and research costs govern investment in AI applications). In the model's reduced two-effort case, a zero-learning corner persists when the private marginal return to learning productive AI applications, U_L, does not exceed direct marginal cost plus congestion from molecular discovery: U_L ≤ 1+κz_N. This outcome can arise even when AI capability has improved and some application benefits are privately appropriable. When the inequality reverses, application learning becomes worthwhile; the full molecular equilibrium in Appendix A.5 retains platform rents, displacement, and competition with additional-indication research.

The learning threshold is the central second-level economic mechanism. Greater AI capability raises the set of technically feasible research opportunities, but it need not generate privately profitable investment in reusable assays, validation methods or research routines. The returns depend on appropriability, specialized capacity and competition with other research. The model consequently admits both expansion of complementary knowledge and persistent underinvestment in productive applications, independently of any assumed clinical-stage conversion function.

Task-specific AI shocks can be represented without adding a new technology to the equilibrium. Let a particular capability j shift the existing effective task-productivity pair (a_N, a_I) by (Delta a_N(j), Delta a_I(j)), with the economically relevant changes mediated by application knowledge L. Molecular candidate generation may primarily affect a_N; target-structure information may alter both components; indication-ranking tools may primarily affect a_I; and clinical-development tools may chiefly improve the conversion margin rather than either entry flow. These are hypotheses about task relevance, not empirical labels assigned to AIRO or guaranteed predictions for AlphaFold. Because Proposition 1 includes both platform-option value and shared-input competition, even the sign of a cross-task response must be derived from the specified productivity shift and market parameters. The model therefore predicts conditional heterogeneity across capabilities rather than a universal ranking of novel and incremental responses.

Proposition 4 (a conditional short-run development gap). In an interior equilibrium, improved AI capability K affects novel molecular entry through task productivity and endogenous research allocation; the sign of the entry response depends on economic parameters. Conditional on an increase in early entry, temporarily fixed AI-application knowledge L, and diminishing returns in the specified conversion technology, documented advanced-stage activity expands less than proportionately on impact. This stage implication does not imply that individual candidates become less likely to succeed over their full development histories.

$$E = \delta\, m(K, L)\, z_{\mathrm{n}}; \qquad dE/dK = \delta\,[m_{\mathrm{k}}\, z_{\mathrm{n}} + m \cdot dz_{\mathrm{n}}/dK].$$

The stage comparison is a conditional implication of the model, rather than the source of its investment mechanism. In the uncapped region, advanced activity is A=Eq with q=ψ(L)E^(α−1), 0<α<1, and ψ′>0. Consequently, d ln A=α d ln E+d ln ψ(L). Holding L fixed on impact, documented advanced-stage activity grows by only α times the proportional change in early entry. The learning investment characterized in Proposition 3 can subsequently raise ψ(L). The identity applies to matched entry cohorts and outcomes, not to the separately estimated empirical columns.

$$q \;=\; min\{1, \psi(L)\; E\text{^}(\alpha - 1)\}, \quad A \;=\; E\; q.$$

The two levels of application innovation must remain analytically separate. An additional disease indication is an application of a molecule, and can be approximated by an observed molecule–disease start. A validated procedure for using AI in assay design, patient selection, or candidate triage is an application of the discovery technology, represented by L and not observed in RePairDB. These innovations interact, but they are not the same variable. That distinction is the paper's recursive contribution and the foundation for interpreting clinical outcomes.

A clinically observed development stage adds a third, distinct margin: conversion of research opportunities into treatments. Let E denote the flow of eligible early clinical projects and A a consistently defined advanced-stage flow for comparable cohorts. The conversion relation in Proposition 4 is a reduced-form representation of how complementary knowledge and capacity shape A conditional on E; it is not an approval production function and does not equate an observed Phase II+ record with a therapy reaching patients. In particular, if a data source records only the highest phase ever documented for a program, comparing early starts with Phase II+ documentation does not reveal within-program transition probabilities. A claim about persistence requires dated stage entries and common follow-up horizons. Ultimately patient-facing treatment output also depends on efficacy, regulatory approval, and access, none of which is identified by the current stage-documentation exercise.

A forward-looking extension of this recursive mechanism (Appendix A.5) makes the private return to AI-application learning depend on the discounted pipeline of molecular opportunities awaiting development. Conditional on the existing molecular-market equilibrium, the same improvement in AI capability can produce different investment in application knowledge and downstream development when appropriability and research-input congestion differ. The extension does not predict a universal delay or identify clinical progression rates; it makes complementary investment an intertemporal economic decision.

The theoretical message is therefore about investment incentives, not an automatic development lag. At both levels of the recursive framework, creating a platform expands the set of potential applications, but realizing their value depends on who can appropriate future returns and how scarce research resources are allocated. The forward-looking extension adds that the expected pipeline of molecular opportunities affects the value of learning AI applications. Neither more capable AI nor a larger pipeline guarantees that complementary investment will occur.

2. From Economic Mechanisms to Observable Development Decisions

2.1 The three observable margins of pharmaceutical investment

Three investment decisions are visible in clinical records: initiating research on a molecule not previously observed, testing an additional indication for an existing molecule, and undertaking more advanced clinical work. The model permits different responses at each margin. Molecular search primarily affects candidate origination; indication research affects the value of existing platforms; application knowledge and clinical capacity affect the ability to move projects further.

These outcomes also correspond to different objects in the theory. First-observed parent-molecule starts proxy the molecular origination margin, while additional molecule–disease starts proxy research into new uses. Dated Phase II+ observations measure documented advanced clinical activity, not treatment approvals or a clean Phase I-to-II transition probability. The learned stock L is a fourth, conceptually separate investment margin that the clinical records do not directly measure. The distinction matters because economic value depends on how these decisions interact, not merely on how many candidates appear.

We estimate the margins separately. Their relative responses are economically meaningful only when counting conventions, populations, and observation windows are comparable.

Early starts can also have option and information value. Even candidates that do not appear in a later-phase record may reveal target biology or improve the next experiment. Conversely, shifts toward registering earlier work can change measured stage shares without changing the probability that a genuinely comparable project advances. This is why the data require a careful distinction between entry-stage composition and subsequent development.

2.2 Why complementary capacity matters for translation

For innovation margin j, let $\Delta D_j(t)$ be a change in scientific discoveries induced by a productivity advance, $p_j(u)$ the probability that a discovered opportunity ultimately reaches a measured stage after delay u, and $g_j(u)$ the conditional delay density. The time profile of observable incremental starts can be represented by

$$\Delta Y_j(t) = \int_0^t \Delta D_j(t-u) p_j(u) g_j(u)\, du.$$

A new prediction changes treatment development only when it improves a decision made with real evidence and scarce resources. Translationally useful assays, disease models, indication-specific validation, development finance, and clinical operations remain costly. This is the practical counterpart to the model's learned stock L. As Scannell et al. (2022) emphasize, predictive validity matters for whether early information forecasts clinically relevant outcomes; Bender et al. (2026) similarly argue that drug-discovery AI should be judged by improvements in downstream decisions rather than benchmark performance alone. The economics adds that validated uses of AI are not simply waiting to be adopted: they must be discovered, organized, and financed.

An illustrative development-lag calculation is retained in Appendix C rather than used as evidence of AI-driven progression. It shows why later discovery cohorts are mechanically less mature by the end of the observation window; its assumed delay distribution is not estimated from AI-exposed programs. The main empirical comparison therefore focuses on mature entry cohorts and separates phase at first observation from advancement documented strictly afterward.

Independent evidence makes the implementation margin concrete. In a survey of 302 respondents across pharmaceutical companies, clinical research organizations and technology providers, Lamberti et al. (2025) find that, across 36 clinical-development activities, only 10.7% of respondents on average reported fully implemented AI/ML use, with much larger shares not implementing, piloting or partially implementing it. These survey responses do not validate AIRO or identify effects on drug progression. They do document a distinction central to the theory: having AI capabilities available and embedding them in repeatable clinical-development workflows are different economic activities. Our clinical analysis therefore separates early entry from later activity rather than treating either as an adoption measure.

2.3 Predictions, interpretation, and the AlphaFold comparison

Hill and Stein (2026) offer an especially useful parallel. Their AlphaFold2 research design studies a sudden improvement in protein-structure information and finds a substantially larger response in basic research on previously less-characterized proteins than in measured downstream applied R&D. Our setting examines an earlier period and different units—therapeutic categories, clinical molecules, and new disease uses—so it does not identify the AlphaFold effect. The papers instead meet at an economic question: why can an expansion of scientifically attractive opportunities fail to produce a commensurate immediate expansion in the development activities that could exploit them?

New molecular platforms are worth more when developers can capture returns from follow-on uses. AI capabilities face an analogous constraint: they become valuable in practice when firms invest in learning how to apply them. Cheaper candidate generation can raise early starts before validation and clinical capacity expand. Whether complementary investment subsequently closes that gap depends on its private return; our clinical panel does not observe that investment directly.

Three empirical implications follow, each conditional on an observable source of variation. First, if an AI advance primarily raises effective novel-task productivity while follow-on capabilities remain relatively fixed, novel starts should respond more strongly than additional-use starts in a comparable population and window; a robust opposite differential would challenge that restriction. Second, if structural information chiefly lowers the cost of identifying new molecular candidates, the same novel-leading prediction applies to AlphaFold, but only insofar as the information is usable for that task. A positive additional-use response would therefore be a finding to explain, not confirmation of that initial prediction. Third, holding the initial opportunity shock comparable, greater independently measured application knowledge or clinical capacity should predict stronger subsequent translation into advanced development; absent such heterogeneity, that particular complementary-capacity prediction receives no support. These are testable restrictions, not unconditional sign predictions for every AI technology.

Stage composition cannot substitute for observed progression. A program first documented at Phase II is different from one shown to advance there after an earlier trial. Identifying clinical conversion requires comparable cohorts and dated phase histories.

The empirical exercises in this paper probe different portions of this mechanism, not interchangeable estimates of one AI treatment effect. The AI Research Opportunity Index (AIRO) measures historical research environments potentially positioned to benefit from computational advances; it does not identify adoption of molecular-generation systems or any specific AI technology. AlphaFold instead represents a dated expansion in predicted protein-structure information. Its drug-level comparison concerns additional indication starts for established molecules, not the complete universe of novel molecular entry. Differences between the AIRO and AlphaFold coefficients therefore cannot, by themselves, identify technology-specific causal effects. A positive incremental association after AlphaFold would challenge a simple prediction of novel entry rising while additional uses remain flat; it is not automatically evidence for the complementary-capabilities mechanism. The test of that mechanism requires independent measures of complementary capacity and informative downstream outcomes.

3. Relation to the Economics of Discovery and Complementary Innovation

The closest empirical conversation is with Hill and Stein (2026). The AlphaFold2 release sharply increased access to predicted protein structures, and their study documents a redirection of basic research toward previously less-characterized proteins. Downstream applied-R&D responses are considerably more muted in their observation window. Their contribution is to identify responses to a particular AI-enabled scientific information shock. The present paper asks a complementary economic question: what determines whether the opportunities created by such a shock attract investment in pharmaceutical applications? Our evidence covers different clinical margins and an earlier historical comparison, while the model characterizes the missing allocation decisions.

Hult (2015) is the key pharmaceutical starting point, because it explains why this question occurs twice. A new molecule does not embody all the value of the indications that might later be discovered; it creates a platform for further inventions. Expected future uses feed back into incentives to originate molecules. AI similarly creates a menu of research opportunities whose value depends on future validated applications. The recursive extension makes this feedback a second investment problem and allows the two levels to interact through rents, research capacity, and competition.

The theory of general-purpose technologies provides the wider economic perspective. Bresnahan and Trajtenberg (1995) emphasize innovational complementarities between a core technology and the inventions made by its users. Brynjolfsson, Rock, and Syverson (2021) explain how intangible complementary investment can postpone measured productivity gains. We take the importance of complements as established; the new element is their allocation within pharmaceutical research, where the value of discovering a molecule depends on downstream use development and where learning AI applications competes with both activities for specialized inputs.

Prediction and experimentation are another essential pair. Cockburn, Henderson, and Stern (2019) characterize AI as a general-purpose method of invention; Agrawal, McHale, and Oettl (2024) develop a model of AI-prioritized hypothesis search in which testing capacity can constrain realized gains. Our distinction goes beyond an exogenous stock of tests. It allows research into the productive use of AI itself to be an appropriable, resource-consuming investment that interacts with molecule origination and indication development. This gives an economic reason why the returns to the same AI advance can differ across stages and over time.

The pharmaceutical innovation literature disciplines the private-return mechanism. Acemoglu and Linn (2004) show how market size affects drug-entry incentives; Krieger, Li, and Papanikolaou (2022) study incentives for chemical novelty; and Frankel, Krieger, Li, and Papanikolaou (2023, revised 2024) emphasize learning and evaluation in R&D. Most closely related, Budish, Durvasula, Roin, and Williams (2025, revised 2026) identify how the enforceability of intellectual-property rights changes investment in new uses for existing drugs. Their empirical contribution isolates a specific appropriability mechanism. Ours examines a different question: how opportunities created by advances in the technology of discovery interact with incentives to originate molecular platforms, develop their uses, and learn productive AI applications. We do not estimate the causal effect of patent enforceability. In our data, 'novel' denotes first-observed parent-molecule clinical entry, not the chemical novelty measure of Krieger and coauthors.

The paper thus connects three literatures that are often studied separately: discovery technologies, the option value of molecular platforms, and the institutions that govern complementary research. The recursive extension is economically consequential because the innovation required to exploit AI can compete for the same scarce research inputs as molecular origination and indication development. The empirical contribution is to examine those distinct clinical-development margins rather than infer medical progress from one count of discoveries. This approach speaks to the rate and direction of medical innovation, while leaving the effects of specific policy instruments to designs that identify them directly.

4. Clinical Evidence: Molecular Entry, New Uses, and Advanced Development

4.1 Clinical research records and the innovation margins

We distinguish novel and incremental pharmaceutical innovation using first-observed clinical research starts in RePairDB. Novel molecular entry is the first qualifying clinical start recorded for a parent molecule, counted once in its first observed year. Incremental innovation is the first subsequent clinical start for a new molecule–disease pair involving a molecule already observed in research for another disease. Repeated trials for the same molecule–disease pair do not count again; diseases first observed simultaneously with initial molecular entry are not later incremental starts. The measures capture the start of clinical experimentation, not verified molecular invention, FDA approval, or new formulations of an existing drug.

The underlying source contains 4,272 first-observed parent molecules and 53,490 subsequent molecule–disease starts during 2005–2025, classified into 33 therapeutic categories in 14 families. For the novel-entry regression, we use 2,948 molecules whose initial clinical-start dates are explicitly recorded as ACTUAL, and assign each initial molecule once to a therapeutic category. This prevents a molecule associated with several first-date indications from being counted as multiple distinct molecular platforms; precise tie-resolution and fractional allocation rules are documented in Appendices B and C. The additional-use outcome uses the full qualifying clinical-pair histories. Documented Phase II+ trial activity is examined as a third, distinct development margin.

The therapeutic classification and the historical network inputs are dated before the comparison window, although the RePairDB compilation itself is retrospective. We keep that distinction explicit.

The project-level observables map imperfectly but transparently onto the model's three clinical research decisions; they do not measure the unobserved AI-application knowledge stock.

4.2 The AI Research Opportunity Index

The AI Research Opportunity Index (AIRO) summarizes pre-2016 therapeutic-area clinical research networks. We use it as a theory-motivated proxy for opportunities to benefit from AI, not as a direct measure of AI adoption. It averages standardized measures of research breadth, observed molecule–disease links, and experience investigating multiple uses of a molecule. We remove the component explained by baseline clinical-pair volume and standardize the residual across 33 categories. Higher scores indicate richer historical research connections conditional on scale, which may make AI-enabled hypotheses easier to evaluate. Appendix B documents the components and category scores.

The index is motivated by complementary innovation: existing molecule–disease relationships and clinical research networks can raise the returns to AI-assisted search by supporting hypothesis assessment and experimental validation. Independent inventories show substantial AI-involved drug research, particularly in oncology and neurology (Druedahl et al., 2024; Haque et al., 2026). Appendix B compares those broad groupings and a five-family FDA submission check with AIRO; neither independently validates exposure across our 33 therapeutic categories.

The separate pre-2016 FDA molecular-platform-density proxy is essentially uncorrelated with AIRO (r=0.04), consistent with the two measures capturing different dimensions of research opportunity. Its construction and estimates are in Appendix B.

We compare changes after 2015 across categories with different historical AIRO scores, controlling for category and year effects. The regressions quantify the distinct clinical-development margins highlighted by the model.

4.3 Empirical specification and interpretation

For novel molecule starts and new therapeutic-use starts, indexed by outcome j, therapeutic category c, and calendar year t, we estimate the following conditional-mean count model:

$$\mathbb{E}\left[Y_{ct}^{j} \mid X_{ct}\right] = \exp\left(\alpha_c^j + \lambda_t^j + \beta_E^j R_c D_t^E + \beta_L^j R_c D_t^L\right).$$

Y is the number of first-observed novel molecule starts or new therapeutic-use starts in category c and year t. R is the standardized AI Research Opportunity Index (AIRO), fixed from the pre-2016 research network. D^E and D^L equal one in 2016–2020 and 2021–2025, respectively, with 2005–2015 the omitted period. Category fixed effects (alpha) absorb time-invariant research differences, and year fixed effects (lambda) absorb common changes in clinical activity. Each period coefficient measures the differential change associated with a one-standard-deviation higher AIRO score.

We estimate the conditional mean by Poisson pseudo-maximum likelihood (PPML), which accommodates nonnegative counts and zero-count category-years. All reported standard errors are clustered at the 14-parent-therapeutic-family level; the significance stars use two-sided t(13) reference values. Both entry regressions use 33 categories over 2005–2025 (693 category-year observations).

The two entry outcomes share the same estimating equation but have distinct clinical definitions: novel entry counts 2,948 first-observed molecules with ACTUAL first-start dates, each assigned to one

therapeutic category, whereas new therapeutic-use starts count 53,490 subsequent molecule–disease research starts. The five-year Phase II+ analysis has a separate mature-cohort specification in Section 4.6. Appendix C reports the stacked PPML contrast between novel and additional-use research.

Table 1 reports both post-2015 interaction coefficients and their clustered standard errors for the novel and incremental research margins. It also reports a pooled 2016–2025 specification for each outcome, replacing the two period interactions with a single AIRO-by-postperiod term. Percentage effects equal 100[exp(coefficient)−1]. The 2016–2020 novel-entry coefficient of 0.167 implies 18.2% greater entry per standard deviation of historical AI research opportunity.

4.4 Novel molecular experimentation responds most strongly

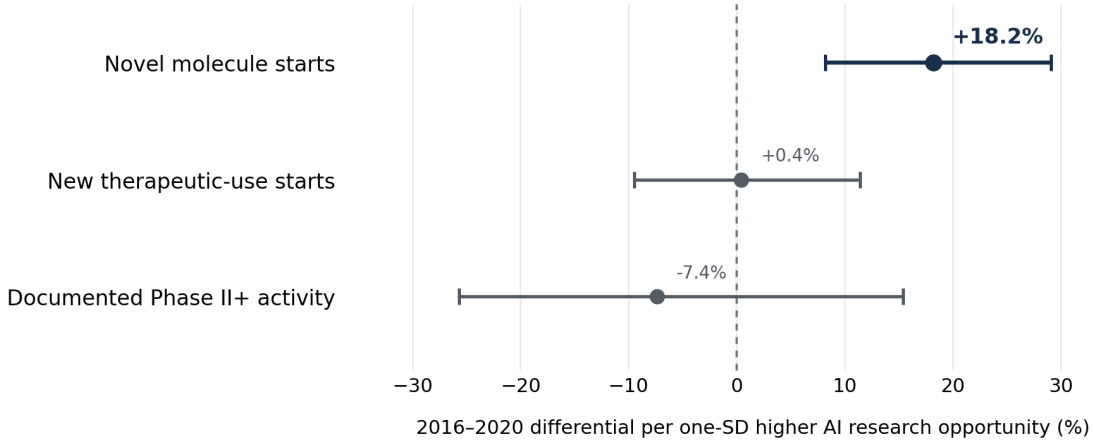
Table 1. AI research opportunity and novel versus incremental clinical innovation

Regressor / specification	(1) Novel starts Two periods	(2) New uses Two periods	(3) Novel starts Pooled	(4) New uses Pooled
AIRO × 2016–2020	0.167*** (0.041)	0.004 (0.047)	—	—
AIRO × 2021–2025	0.104 (0.083)	0.006 (0.029)	—	—
AIRO × 2016–2025	—	—	0.149*** (0.047)	0.005 (0.038)
2016–2020 effect (%)	+18.2%	+0.4%	—	—
2021–2025 effect (%)	+10.9%	+0.6%	—	—
2016–2025 effect (%)	—	—	+16.1%	+0.5%
Category fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Observations	693	693	693	693
Parent-family clusters	14	14	14	14

Notes: PPML count regressions with therapeutic-category and calendar-year fixed effects; parent-family-clustered standard errors in parentheses (14 clusters). Columns (1) and (2) estimate separate 2016–2020 and 2021–2025 interactions, with 2005–2015 omitted. Columns (3) and (4) instead include one 2016–2025 interaction. Novel entry counts 2,948 ACTUAL-date molecules once each; if simultaneous first-date indications span categories, assignment is to the alphabetically first category. New therapeutic uses count 53,490 later first-observed molecule–disease starts. There are 693 category-year observations in every column. Percentage effects equal 100[exp(beta)−1]. Two-sided t(13) inference: *** p<0.01, ** p<0.05, * p<0.10. Appendix C reports fractional-attribution sensitivities.

The central finding is a positive, precisely estimated novel-entry response: in 2016–2020, a one-standard-deviation higher AIRO score is associated with 18.2% greater first-observed novel molecular experimentation (p=0.0012). The later-period coefficient also remains positive (+10.9%), though less precisely estimated, and the pooled 2016–2025 response is +16.1% (p=0.0076). New therapeutic-use starts show no comparable expansion in either period (+0.4% and +0.6%, respectively). Figure 1 focuses on the 2016–2020 estimates for the three distinct development margins.

Figure 1. Differential clinical-development responses across the three innovation margins



Notes: Points are 2016–2020 PPML differentials per one-SD higher pre-2016 AIRO; bars are 95% family-clustered confidence intervals. The outcomes are first-observed novel molecule starts (+18.2%), first-observed additional therapeutic-use starts (+0.4%), and documented Phase II+ activity within five years (−7.4%). They use separate, explicitly defined research populations and should not be interpreted as a matched clinical-survival funnel. Appendix C reports alternative therapeutic-category allocations.

Novel clinical molecule entry is 18.2% higher per one-standard-deviation AIRO difference during 2016–2020 (95% CI: +8.2% to +29.1%; p=0.0012). For scale, if a therapeutic category-year would otherwise have ten fitted first-observed molecule starts, an 18.2% relative difference corresponds to about 1.8 more starts in that category-year. This is a model-based illustration of an association, not an estimate of molecules created by AI or of approvals. As a robustness check, we also fractionally distribute first-date molecules across therapeutic categories. On common fractional definitions, the novel-versus-incremental response differential remains positive at +14.1% and is statistically significant at the 10% level (p=0.085; Appendix C). Molecular origination is thus the clearest expanding clinical-research margin in areas with stronger historical research opportunity.

4.5 Additional indications and the value of applications

New therapeutic-use starts have a 2016–2020 differential of +0.4% per standard deviation of AIRO (95% CI: −9.3% to +11.2%; p=0.932). The confidence interval is informative about economically large responses: it excludes, within this additional-use equation, an increase as large as the +18.2% novel-molecule entry estimate. The conditional family-clustered power calculation in the statistical power calculation, evaluated at the estimated full-sample additional-use standard error, has approximately 91% power to detect a +20% increase at the 5% significance level. A modest incremental response remains possible, but the evidence does not indicate expansion of new-use research comparable in magnitude to new-molecule entry. Economically, the difference matters because originating a platform and developing additional therapeutic applications involve distinct expected returns and demands on research resources.

The incentives are first order. Budish, Durvasula, Roin, and Williams (2025, revised 2026) document how enforceability of rights can strongly affect investment in drug repurposing. In our model, weaker appropriability of additional indications reduces the return to developing molecular applications and can also diminish the prospective option value of originating molecules. The RePairDB regressions do not

isolate this channel, but the model clarifies why a larger menu of candidates does not guarantee proportionately more indication investment.

The distinct pre-2016 FDA platform-density variable provides a useful comparison of research maturity, retained in Appendix B. It is not interchangeable with AIRO and does not measure the stock ratio in the structural model.

4.6 Development stage at entry versus progress after entry

We examine advanced clinical development in a separate PPML regression using five-year Phase II+ activity for mature entry cohorts through 2020. Unlike first clinical entry, this outcome requires a fixed follow-up horizon; 2021–2025 entrants cannot contribute complete five-year histories in a dataset ending in 2025. Let H be the recoverably documented number of Phase II+ programs associated with first-observed molecule cohorts in category c and entry year t. We estimate:

$$E[H_{ct}^{(5)} \mid X] = exp(\alpha_c + \lambda_t + \gamma R_c D^E), \; t \leq 2020$$

Table 2. AI research opportunity and documented advanced clinical development

Panel A. PPML: Five-year Phase II+ programs

Outcome	Five-year Phase II+
AIRO × 2016–2020	−0.077
95% CI, coefficient	[−0.297, 0.143]
Implied change (%)	−7.4%
Two-sided p-value	0.465
Category fixed effects	Yes
Entry-year fixed effects	Yes
Parent-family clusters	14

Panel B. Timing of documented Phase II+ (descriptive cohort shares)

First-observed cohort	Any Phase II+ by year 5	Already Phase II+ at first observation	Phase II+ documented strictly later
2005–2015 (N=2,648)	29.7%	20.3%	10.1%
2016–2020 (N=1,157)	24.2%	15.1%	10.1%
Difference (pp)	−5.5	−5.2	0.0

Notes: Panel A reports the existing PPML regression for recoverably dated Phase II+ activity within five years of first clinical observation, comparing 2016–2020 entry cohorts with 2005–2015. Its 95% confidence interval is given on the log-coefficient scale (converted from the reported percent-effect interval); inference uses 14 therapeutic-family clusters and t(13). A complete five-year horizon is unavailable for the 2021–2025 cohorts. Panel B shows descriptive shares from recoverably dated RePairDB trials, with 2,648 earlier and 1,157 later first-observed molecules. Initial and strictly later Phase II+ records can both occur, so the columns do not sum. Differences are percentage points; initial molecules are fractionally allocated across simultaneous categories for this cohort decomposition to retain the weight of a molecule spanning categories; Table 1 instead assigns each newly observed molecule to a single category for the origination outcome. An unrecorded trial is not evidence of clinical failure; Appendix D gives source-recovery details.

Table 2, Panel A reports the coefficient on AIRO × 2016–2020 from this mature-cohort regression: −0.077, an implied −7.4% differential (95% CI for the percentage effect: −25.7% to +15.4%; p=0.465). Panel B distinguishes whether Phase II+ activity is already present at a molecule's first recorded clinical date or is documented strictly afterward. The distinction matters because a newly recorded molecule may already be in Phase II or III, while another may first appear in an early trial and later advance.

The regression decomposition supports the same distinction. The 2016–2020 AIRO differential is −4.49 percentage points for any dated Phase II+ record within five years (nominal p=0.0013), compared with −0.76 points for Phase II+ activity documented strictly after first observation (p=0.41). A bounded source-coverage diagnostic provides context: completeness of recoverable date-and-phase information for all source-listed trials rises from 48.0% in the earlier cohort to 54.7% in the later cohort (the clinical-source completeness audit in Appendix D.3). Thus the decline in the documented higher-stage share is not mechanically explained by a uniform loss of recoverable metadata, although broader capture of early trials and nonrandom missing histories remain possible. The economic conclusion concerns the changing observed stage mix, not a demonstrated fall in clinical success.

Later entrants cannot yet resolve the eventual adjustment path. On a common three-year horizon, strictly subsequent Phase II+ AIRO differentials are −0.81 percentage points for 2016–2020 cohorts and −0.30 for 2021–2022 cohorts; the +0.50-point change is imprecise (p=0.659). The paper's theory explains the incentives determining later catch-up; the observed cohort histories have not yet established such catch-up. Appendix D documents the date-recovery rules and the associated sensitivity bounds.

This decomposition sharpens the paper's central distinction. First clinical entry, the recorded stage at entry, and advancement after entry measure different investments and information events. An enlarged early research pipeline need not yield an immediate increase in advanced programs, but the same records do not imply that new molecules are becoming less likely to advance. The model identifies the complementary investments that can bridge this gap; the data identify why the relevant stages must be measured separately.

4.7 Attribution and historical-pattern checks

Each novel molecular platform is counted once in the main specification, consistent with the economic distinction between originating a molecule and pursuing multiple indications. The fractional allocation check distributes simultaneous first-date molecule entries across their therapeutic categories rather than assigning an entire molecule to one category. On common fractional definitions, the novel-versus-additional-use differential is positive and statistically significant at the 10% level (+14.1%; p=0.085). Appendix C reports the individual fractional estimates, oncology exclusions, and historical checks. In 500 randomized resolutions of simultaneous cross-category ties, the early novel-entry association is positive in every draw but smaller on average and less precisely estimated than under alphabetical assignment; statistical significance therefore depends on attribution.

The clinical-stage results distinguish first observed phase from development documented strictly after entry. Source-recovery and historical-pattern diagnostics, including changes in attributable-link quality, are reported in Appendices B and D. These checks clarify what the registry histories measure without equating early clinical entry with eventual clinical success. The raw novel-entry series by AIRO tercile also suggest that differential growth was visible before the comparison window, so the regression is not interpreted as a causal AI-shock estimate.

4.8 An emerging second wave: learning to use AI in drug development

The model distinguishes an improvement in AI capability from the costly accumulation of knowledge needed to use that capability productively. Our clinical-entry panel does not observe this second stock directly. Evidence on AI deployment in clinical development nevertheless provides a way to locate the industry along that adjustment path: researchers are beginning to develop, validate, and routinize

applications of AI beyond candidate generation. Table 3 brings together three independent indicators of that transition.

Table 3. Early evidence of AI application learning in pharmaceutical development

Margin	Observed evidence	Economic interpretation
Regulatory use	AI/ML-related FDA CDER submissions rose from 1 in 2016 to 132 in 2021.	Growing use in development and regulatory workflows, not only molecular discovery.
Implementation maturity	Tufts survey (302 responses, 36 clinical-development tasks): 30.3% piloting, 22.1% partially implementing, and 10.7% fully implementing on average.	Complementary know-how is being accumulated; routine deployment remains limited.
Clinical translation	Among 129 AI-enabled assets reaching clinical trials by March 2026, 24 had completed Phase II and five had entered or completed Phase III.	Some applications are reaching advanced evaluation; treatment impact is not yet established.

Notes: Source populations and units differ and are not pooled or used as a causal estimate. FDA submissions are regulatory documents, not distinct drugs or proven efficiencies (Liu et al., 2023). Survey shares are average implementation states across 36 activities, not a longitudinal transition rate (Lamberti et al., 2025). The therapeutic-asset inventory is based on reported AI involvement and includes candidates at different degrees of maturity (Haque et al., 2026).

The substance of the emerging complement is visible in specific development tasks. Liu et al. (2023) document AI/ML applications in clinical trial design, dose optimization, endpoints, and other regulatory activities. In December 2025, the FDA qualified AIM-NASH, an AI-assisted liver-biopsy assessment tool for MASH clinical trials, following validation against expert assessments (FDA, 2025). Unlike the generation of a new molecule, the development of a usable clinical assessment tool requires measurement standards, validation, and a reliable decision process. This is the kind of application-specific investment represented by the learning stock in our model.

The organizational counterpart to these technical milestones is visible in implementation evidence. Drawing on industry workshops, an executive survey, and statistical analysis of pharmaceutical AI/ML use cases, Makarov et al. (2026) find that reported business success is associated with project maturity and organizational support, whereas model type and data source show no comparable independent association. Read alongside Lamberti et al. (2025), who document that most clinical-development applications have not reached repeatable full implementation, these findings point to an investment margin beyond acquiring AI models: developing validation routines, organizational coordination, and reliable workflows that make the technology useful. The implementation outcomes in these studies are distinct from drug approvals.

The therapeutic pipeline shows movement along a different, longer margin. Haque et al. (2026) identify 129 AI-enabled candidates tested in humans across 69 companies by March 2026, but only five had entered or completed Phase III. Rentosertib, an AI-discovered candidate for idiopathic pulmonary fibrosis, generated exploratory results in a randomized Phase IIa study in 2025 and began Phase III testing in September 2026 (Xu et al., 2025; ClinicalTrials.gov, 2026). These are concrete milestones in the development of AI-enabled drugs; neither the trial inventory nor an individual Phase III start establishes higher overall success rates.

Related evidence from clinical-development decisions underscores the role of validation. Rathee (2026) examines 55,527 industry-sponsored interventional trials and measures sponsors' AI capability using cumulative prior AI-related trial experience. In sponsor fixed-effects analyses, greater AI experience is associated with lower novelty of drug–disease combinations on average, particularly in later-stage trials; when more AI-experienced sponsors pursue novel projects, their trial designs also tend to include more primary outcomes. The pattern is consistent with stage-specific proof requirements shaping clinical-development choices, not with an automatic increase in downstream novelty. It does not establish a causal effect of AI experience on clinical success.

Oncology highlights a further economic possibility. It accounts for 74 of the 129 AI-enabled clinical assets in Haque et al. (2026), yet in our broader RePairDB oncology cohorts the fraction with a recoverably documented Phase II+ trial within five years is 20.3% in 2005–2015 and 12.5% in 2016–2020 (Appendix D). The latter comparison is sensitive to how early trials enter the records and is not evidence that individual oncology programs have become less likely to progress. Nor is oncology uniformly ranked by AIRO: the index varies considerably across its ten subcategories (Appendix Table B.2). The juxtaposition is nevertheless informative. Concentrated discovery effort may raise demand for scarce clinical capacity, validation resources, and experienced researchers, so the research-input competition in our model could limit downstream expansion. The present evidence motivates that crowding-out mechanism but does not identify it.

Taken together, adoption surveys, regulatory applications, validated clinical tools, and advancing AI-enabled candidates suggest that the industry has begun to invent the complementary capabilities needed to exploit AI. Most applications remain at an early stage of implementation or clinical evaluation. The distinction between the growing stock of AI capability and the still-developing stock of its productive applications is therefore observable in the industry, even though the eventual effect of that learning on approved treatments remains to be measured.

These developments give the theoretical mechanism a concrete industry interpretation. The diffusion of AI capabilities is changing what pharmaceutical researchers can attempt; validated workflows, organizational know-how, and clinical-development methods are the complementary innovations through which those opportunities may become treatments. The evidence documents the beginnings of that investment process, while the model explains why its timing and scale depend on economic incentives rather than technological progress alone.

5. AlphaFold: A Specific AI Information Shock and Clinical Development

5.1 Motivation, historical exposure, and clinical-development outcomes

Our first empirical design links pre-2016 clinical-research opportunity to later changes across therapeutic categories. The AlphaFold setting examines a different margin: a dated expansion of protein structural information in July 2021. Following Hill and Stein (2026), we distinguish proteins that lacked a publicly released experimental structure before the breakthrough (“dark”) from those with such a structure (“bright”). AlphaFold predicted structures for both groups, but the addition to directly observed structural information was potentially greater for dark proteins. Structural-information gains

are not randomly assigned: experimentally studied targets differed in scientific attention, biological tractability, and pharmaceutical investment well before 2021.

We construct historical dark/bright status by combining Protein Data Bank release dates and SIFTS links to UniProt with the original AlphaFold human-protein list. RePairDB clinical histories identify drug–disease programs, while Open Targets drug-mechanism annotations connect ChEMBL compounds to protein targets. Our drug-level exposure is the fraction of a drug's distinct mapped protein targets classified dark before AlphaFold. Thus a drug with one dark target among two mapped targets has exposure 0.5. Unlike a protein-year outcome that repeats a new indication for every mapped protein, our drug-year outcome counts a drug–indication start exactly once.

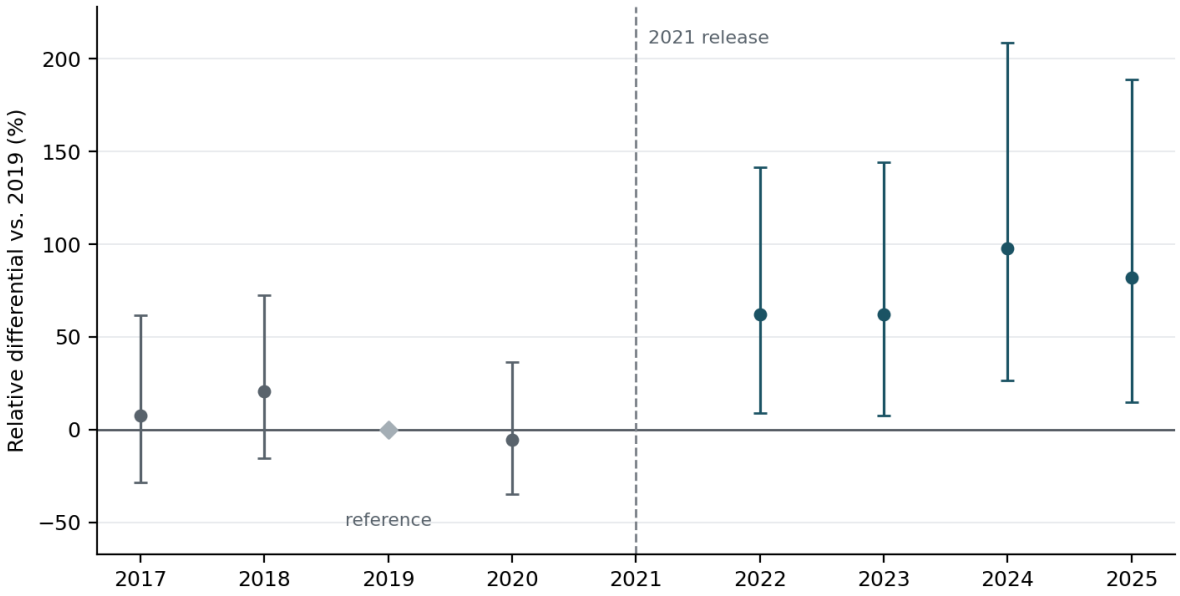
The principal analysis follows 2,316 drugs observed clinically no later than 2016. Of these, 1,964 have no mapped dark target, 284 have only dark targets, and 68 have a mixture; 352 have at least one dark target. The average share of dark targets across drugs is 0.135. The full 0-to-1 exposure contrast is therefore supported by observed drugs but is not the sample-average exposure shift. Crucially, the drug–protein annotations were assembled retrospectively in the 2026 data release: we observe which proteins a drug is mapped to today, but have not verified that each relationship was known to its developer before AlphaFold. This historical-provenance issue qualifies the treatment interpretation.

5.2 Drug-level difference-in-differences and dynamic patterns

For each eligible drug d and year t, the outcome is the number of distinct additional drug–indication programs first observed in year t, excluding indications in the drug's first observed clinical year and repeated trials for the same drug–disease pair. We estimate a conditional Poisson model with drug fixed effects and year fixed effects, interacting pre-2021 dark-target share with a post-2021 indicator. The drug-class-adjusted model replaces common calendar-year effects with effects specific to a drug's modal target class and year. Percentage coefficients equal 100[exp(beta) − 1] for the complete 0-to-1 dark-share contrast. Standard errors in Table 4 are clustered by drug.

Observed counts clarify the relative-change estimand. Among 352 drugs with at least one mapped dark target, annual new-indication starts declined from about 204 in 2017–2020 to 189 in 2022–2025, or from approximately 58 to 54 starts per 100 drugs per year (−7%). Among 1,964 drugs without dark targets, annual starts fell from about 2,006 to 1,232, or from approximately 102 to 63 per 100 drugs per year (−39%). Both groups declined, and their starting activity levels differed substantially. These unadjusted any-dark comparisons illustrate the data; the regressions instead use each drug's fractional dark-target exposure and account for drug and calendar-time effects.

Figure 2. New-indication development by historical AlphaFold exposure



Notes: Each dot is the exponentiated dark-target-share-by-year interaction relative to 2019; error bars are nominal drug-clustered 95% intervals. The event study uses drug and calendar-year effects but not drug-class-by-year effects, and includes 2020 as a pre-release diagnostic. The 2021 release year is omitted. The vertical scale is the full 0-to-1 dark-share contrast; the figure is not a count of additional indications or a causal estimate.

Figure 2 shows no statistically detectable joint proportional pretrend in this original drug-level specification (joint $p = 0.572$), followed by positive relative differentials in 2022–2025. The pre-release intervals are wide, so the diagnostic does not establish parallel trends. The event-study specification differs from the cleaner class-by-year model reported below; the two must not be conflated.

Table 4. AlphaFold exposure and additional-indication starts: drug-level estimates

Specification	Relative difference	95% CI	Drugs
Original: 2017–20 vs. 2022–25	+67.0%	+29.3 to +115.5%	2,316
Exclude 2020–21	+59.9%	+21.1 to +111.1%	2,316
Plus drug-class × year effects	+40.2%	+6.8 to +84.0%	2,316
Plus exclude top 5% active drugs	+5.9%	−22.1 to +43.9%	2,206

Notes: Conditional Poisson with drug fixed effects; errors clustered by drug. "Exclude 2020–21" uses 2017–2019 pre and 2022–2025 post. Drug class is the modal class of mapped targets. The top-5% exclusion ranks drugs by 2017–2019 indication starts. Models differ in sample and controls; estimates are associations, not causal effects. Source: Authors' calculations using RePairDB drug–indication histories, drug–target mechanisms, and historical PDB/SIFTS structural information.

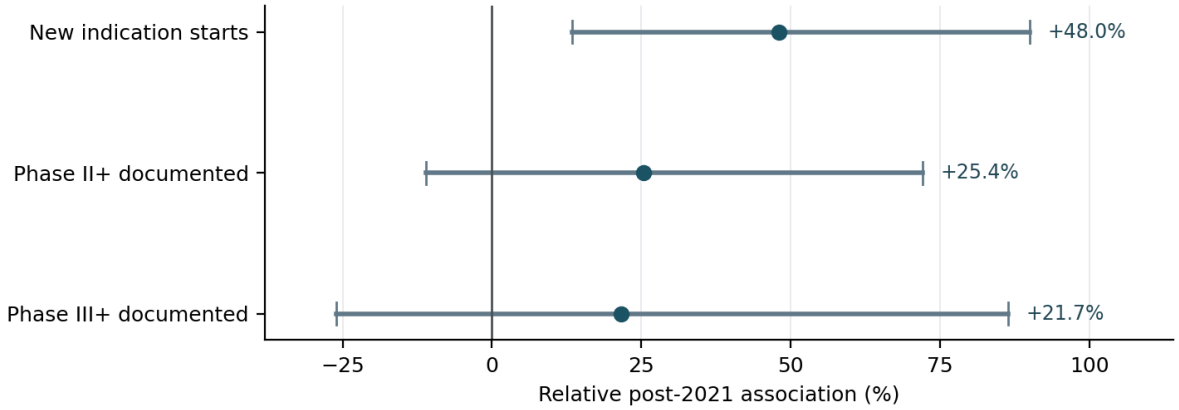
The preferred class-adjusted model estimates a +40.2% relative post-2021 indication-development differential (95% drug-cluster CI +6.8% to +84.0%) when comparing a drug with all mapped targets dark with an otherwise comparable drug with none dark. For a 10-percentage-point difference in exposure, the implied relative differential is approximately 3.4%; the exposure distribution includes 1,964 drugs at zero, 284 at one, and 68 between those endpoints. To express the fitted magnitude in actual research activity, summing predictions over 2,316 drugs in 2022–2025 yields an average of 47.8 more drug–indication starts per year with the observed post-treatment exposure interaction than with that interaction set to zero: approximately 2.1 starts per 100 drugs annually. This is a model-implied comparison of clinical research starts, not 48 new molecules, 48 approved indications, or 48 treatments caused by AlphaFold.

The result has an economically consequential concentration. Excluding the 5% of drugs with greatest pre-2020 additional-indication activity reduces the class-adjusted estimate to +5.9% (95% CI −22.1% to +43.9%). Those drugs account for a substantial part of activity in established molecular platforms, so excluding them changes the population being studied; nevertheless the reduction prevents us from claiming a uniform effect across drugs. The earlier protein-year comparison yielded a much smaller +11.1% adjusted association. Different risk sets, binary versus fractional exposure, and repeated drug-to-protein attribution make the two estimates different estimands, not interchangeable confirmations. Their sample-attribution reconciliation and other robustness results appear in Appendix E.

5.3 Does the differential extend to more advanced clinical outcomes?

We next ask whether the relative association appears only in all new-indication starts or also in programs with later-stage clinical documentation. On a shorter window—2017–2019 versus 2022–2023, excluding 2020–2021—we classify each new drug–indication start by whether RePairDB documents Phase II+ or Phase III+ at any point by the data snapshot. This provides more post-release follow-up than 2024–2025 entrants, though the earlier cohort still has a longer observation horizon. The source records highest attributable phase, not dates of first entry into Phase II or III; consequently, these are nested documentation outcomes rather than verified transitions of comparable initial-phase projects.

Figure 3. AlphaFold-associated differences are larger for indication starts than for outcomes with advanced-stage documentation



Notes: Points show drug-level conditional Poisson relative associations for 2017–2019 vs. 2022–2023; intervals are percentile 95% CIs from 1,000 paired drug-bootstrap replications (no failed fits). Programs with Phase II+ or III+ documented are subsets of all starts, not observed sequential advancement cohorts. The early-minus-Phase II+ contrast is 0.166 log points (95% bootstrap CI 0.003 to 0.335).

The short-window relative association is +48.0% for all new-indication starts (95% bootstrap CI +13.4% to +90.0%), +25.4% for starts with Phase II+ documented (−11.1% to +72.0%), and +21.7% for starts with Phase III+ documented (−26.1% to +86.3%). A paired 1,000-replication drug bootstrap estimates a 0.166-log-point early-minus-Phase II+ differential (95% CI 0.003 to 0.335; approximate two-sided p = 0.040). The stage-documentation association is therefore smaller than the initial-start association in this specification. These are relative effects for different outcome counts, not changes in the probability that an individual project progresses through clinical phases; nor do they establish the number of new treatments ultimately approved.

This is not an estimate of clinical survival or the probability of advancing from Phase I to Phase II. For 2022–2023 first-observed indication programs, 227 of 370 associated with drugs with any dark target (61.4%) have Phase II+ documented by the snapshot, compared with 1,776 of 2,621 (67.8%) for drugs without dark targets. These are selected and unequally followed programs. The coefficient contrast is borderline, and both later-stage estimates are imprecise. We therefore interpret the pattern as suggestive attenuation across stage-documentation outcomes—not as a finding that AlphaFold-generated discoveries failed to become treatments.

5.4 Novel molecular entry: the limits of a common-target comparison

The intended complementary margin is first-observed molecular entry. RePairDB contains 473 first-observed molecules for 2022–2025, but only 38 (8.0%) receive a reviewed protein mechanism in our strict direct mapping. Broader ChEMBL activity screening yields possible targets for additional molecules; only a small subset have independently corroborated clinical mechanisms, and some recent first appearances correspond to historical drugs or diagnostic agents rather than newly developed therapeutics. We do not promote unvalidated bioactivity associations into AlphaFold exposure. First-ever targeting of an entirely new protein is likewise very rare in the current observed clinical panel (six post-2021 entries).

Protein-year novel-link PPML point estimates are positive but extremely imprecise, and alternative binary or pooled estimators do not solve the selective target mapping. The available data therefore do

not identify whether AlphaFold changed novel molecular entry more or less than new-indication development. We report the novel counts and provenance audit in Appendix E, rather than using the coefficient's sign to support our preferred theory.

5.5 Interpretation and relation to the economic framework

The AlphaFold result presents a useful empirical challenge to a simple version of our initial hypothesis. If structural information primarily reduced the cost of originating new molecular candidates, one might expect novel entry to respond more strongly than additional-use research. Instead, among established drugs with mapped targets, the clearer association is with additional indications. This is not proof of a complementary-learning mechanism: the heterogeneity contrast by historical target-development experience is statistically inconclusive, and the drug-level relationship is concentrated among historically active drugs.

Our model nonetheless identifies why this margin is economically meaningful. A new molecule creates a platform whose prospective uses matter for its value; similarly, a predicted structure provides information that researchers may deploy in studying existing platforms. Different AI capabilities change the productivity of different tasks. But a positive association with additional clinical experimentation is not equivalent to a proportionate increase in clinical-stage advancement or treatments delivered to patients. The two empirical exercises therefore examine distinct opportunities and development decisions rather than supplying two causal estimates of a universal "AI effect."

Identification remains qualified by historically nonrandom experimental-structure availability, retrospective drug-target mapping, class-specific development patterns, influential drugs, and incomplete recent clinical histories. The section does not claim that AlphaFold caused the measured divergence or that the relative novel-versus-incremental response is identified. Its contribution is an economically informative extension of Hill and Stein (2026) into drug–indication development and available clinical-stage documentation, evaluated alongside—not substituted for—the broader AIRO evidence.

6. The Economics of Learning to Harness AI

6.1 Why better AI creates options, not finished treatments

Molecular-search tools may increase candidate origination, while predicted protein structures may inform work on established drugs. Neither delivers a treatment by itself. Further experimentation, application-specific knowledge, and clinical investment determine which opportunities are pursued. Appropriability and competition for scarce research inputs affect those decisions.

The empirical exercises capture different parts of this allocation problem. Across therapeutic categories, AIRO is associated with 18.2% greater first-observed molecular entry and a 0.4% differential in additional-use starts in 2016–2020. Among established drugs, the class-adjusted AlphaFold association with additional indications is +40.2%, but falls sharply when historically active drugs are excluded. The shorter-window association is smaller for Phase II+ documentation than for all indication starts. These comparisons use different exposures, samples, and outcomes: AIRO does not measure adoption, and some AlphaFold target mappings are retrospective. Their coefficients should not be read as effects of competing AI technologies.

6.2 Why the complement may be underproduced

Application knowledge has a classic appropriability problem. A validated assay, a transferable workflow, or a reliable clinical-triage procedure may improve projects beyond the one that finances its creation. Its social return can therefore exceed its private return. The learning threshold in Proposition 3 formalizes this mechanism: a valuable AI capability can remain underexploited if the expected private return to building applications does not cover specialized research costs and congestion. At the molecular level, Budish, Durvasula, Roin, and Williams (2025, revised 2026) identify a specific missing market arising from weakly enforceable rights to new drug uses. Our argument is complementary, not a claim to replicate that policy experiment: related appropriability and capacity constraints may also govern the costly development of AI's productive research applications, but the present data do not identify their individual causal effects.

These investment margins interact. Better indication research raises the prospective value of a new molecule but also draws on scarce scientists and clinical resources. Better AI can increase the returns to discovery and application development while intensifying competition for those same inputs. The resulting allocation depends on rights to downstream uses, validation capacity, and the organization of research across firms.

6.3 What the mechanism implies for future pharmaceutical innovation

Early experimentation can create information and options before clinical success is observable. Application learning may eventually improve validation and selection, but firms will invest only where expected returns cover the costs. Scarce specialized capacity and weak appropriability can prolong the gap between discovery and development. Our AIRO and AlphaFold stage measures do not establish long-run progression or treatment output; they show why delays in observation must be distinguished from insufficient investment in the complements to discovery.

A direct test of this mechanism would measure AI deployment, investment in its applications, and dated progression of comparable drug–indication cohorts. That evidence would distinguish learning and incentive effects from changes in the composition of early research.

For medical innovation policy, the implication is to evaluate discovery and its complements separately. Funding science or improving access to data can expand opportunities, while rights to follow-on uses, knowledge sharing, validation infrastructure, and clinical capacity influence which opportunities are developed. We do not estimate the effect of a particular U.S. policy. The framework instead identifies the investment margins such policies may affect.

Conclusion: The Value of a Discovery Platform Depends on Its Uses

A novel molecule is valuable partly because it can support uses not yet discovered. AI has a similar economic property: its productive applications must be developed. Extending Hult (2015), we model the decisions to originate molecular platforms, develop additional indications, and learn to exploit AI as distinct investments. Platform option value, appropriability, product competition, and scarce research inputs shape their allocation. A discovery breakthrough therefore need not produce a proportional increase in therapeutic development.

The clinical evidence illustrates these distinct margins. Across therapeutic categories, one standard deviation higher pre-2016 AIRO is associated with 18.2% greater first-observed molecule entry in 2016–2020, compared with 0.4% for new-use starts; the −7.4% Phase II+ association comes from a separate mature-cohort sample. After AlphaFold's 2021 release, drugs with more previously unsolved targets exhibit stronger relative additional-indication activity. The class-adjusted full-exposure estimate is +40.2% (about 3.4% per ten percentage points of exposure), corresponding to approximately 48 model-implied drug–indication starts annually in the sample—not approved medicines. The estimate is sensitive to historically active drugs, and incomplete protein mapping limits the novel-entry comparison. These are distinct observational designs, not comparable causal estimates of AI adoption.

In the shorter AlphaFold window, the differential is +48.0% for indication starts and +25.4% for starts with Phase II+ documentation, with a borderline paired-bootstrap difference. Because the source does not date phase transitions, this is not evidence that the same programs failed to advance. Neither analysis measures additional treatments reaching patients. The economic question remains why investment in exploiting scientific opportunities may lag investment in creating them—and which institutions could change that balance.

Acknowledgements and Disclosures

The author received no funding or other research support for this study. The author is affiliated with NERA Economic Consulting. The views expressed are those of the authors and do not necessarily reflect those of any supporting institution.

References

Acemoglu, Daron, and Joshua Linn. 2004. “Market Size in Innovation: Theory and Evidence from the Pharmaceutical Industry.” *Quarterly Journal of Economics* 119 (3): 1049–1090. https://doi.org/10.1162/0033553041502144.

Agrawal, Ajay, John McHale, and Alexander Oettl. 2024. “Artificial Intelligence and Scientific Discovery: A Model of Prioritized Search.” *Research Policy* 53 (5): 104989. https://doi.org/10.1016/j.respol.2024.104989.

Bender, Andreas, Morgan C. Thomas, Jack W. Scannell, et al. 2026. “Artificial Intelligence in Drug Discovery—What It Is, Where We Stand and the Path Forward.” Nature Reviews Drug Discovery 25: 855–876. https://doi.org/10.1038/s41573-026-01496-2.

Bresnahan, Timothy F., and Manuel Trajtenberg. 1995. “General Purpose Technologies: ‘Engines of Growth’?” Journal of Econometrics 65 (1): 83–108. https://doi.org/10.1016/0304-4076(94)01598-T.

Budish, Eric, Maya M. Durvasula, Benjamin N. Roin, and Heidi L. Williams. 2025, revised 2026. “Missing Markets for Innovation: Evidence from New Uses for Existing Drugs.” NBER Working Paper 34222. https://doi.org/10.3386/w34222.

Brynjolfsson, Erik, Danielle Li, and Lindsey Raymond. 2025. “Generative AI at Work.” Quarterly Journal of Economics 140 (2): 889–942. https://doi.org/10.1093/qje/qjae044.

Brynjolfsson, Erik, Daniel Rock, and Chad Syverson. 2021. “The Productivity J-Curve: How Intangibles Complement General Purpose Technologies.” American Economic Journal: Macroeconomics 13 (1): 333–372. https://doi.org/10.1257/mac.20180386.

ClinicalTrials.gov. 2026. "Study Evaluation Rentosertib (INS018_055) Administered Orally in Patients With Idiopathic Pulmonary Fibrosis (IPF)." NCT07687459, trial start September 9, 2026. https://clinicaltrials.gov/study/NCT07687459.

Cockburn, Iain M., Rebecca Henderson, and Scott Stern. 2019. “The Impact of Artificial Intelligence on Innovation: An Exploratory Analysis.” In The Economics of Artificial Intelligence: An Agenda, edited by Ajay Agrawal, Joshua Gans, and Avi Goldfarb, 115–146. Chicago: University of Chicago Press. https://www.nber.org/chapters/c14006.

ChEMBL Consortium. 2015. “ChEMBL 20 Released.” ChEMBL Blog, February 3, 2015. https://chembl.blogspot.com/2015/02/chembl-20-released.html.

Congressional Budget Office. 2021. *Research and Development in the Pharmaceutical Industry.* Washington, DC. https://www.cbo.gov/publication/57126.

Druedahl, Louise C., W. Nicholson Price II, Timo Minssen, and Ameet Sarpatwari. 2024. "Use of Artificial Intelligence in Drug Development." JAMA Network Open 7(5): e2414139. doi:10.1001/jamanetworkopen.2024.14139.

Fakhouri, Tala H. 2024. “Responsive Regulation of Artificial Intelligence in Drug Development.” U.S. Food and Drug Administration, Center for Drug Evaluation and Research, presentation, May 29, slide 6. https://www.fda.gov/media/184256/download.

Ferrero, Enrico, Ian Dunham, and Philippe Sanseau. 2017. “In Silico Prediction of Novel Therapeutic Targets Using Gene-Disease Association Data.” *Journal of Translational Medicine* 15:182. https://doi.org/10.1186/s12967-017-1285-6.

Frankel, Alexander P., Joshua L. Krieger, Danielle Li, and Dimitris Papanikolaou. 2023, revised 2024. “Evaluation and Learning in R&D Investment.” NBER Working Paper 31290. https://doi.org/10.3386/w31290.

Haque, Waqas, et al. 2026. "Artificial Intelligence-Enabled Drug Discovery: A Growing but Largely Early-Stage Clinical Pipeline." Cancer Treatment Reviews 151: 103223. doi:10.1016/j.ctrv.2026.103223.

Hill, Ryan R., and Carolyn Stein. 2026. “How Artificial Intelligence Shapes Science: Evidence from AlphaFold.” NBER Working Paper 35143. https://doi.org/10.3386/w35143.

Hult, Kristopher J. 2015. *Incremental Innovation and Pharmaceutical Productivity.* PhD dissertation, University of Chicago.

Krieger, Joshua L., Danielle Li, and Dimitris Papanikolaou. 2022. “Missing Novelty in Drug Development.” *Review of Financial Studies* 35 (2): 636–679. https://doi.org/10.1093/rfs/hhab024. Earlier circulated as “Developing Novel Drugs.”

Lamberti, Mary Jo, Maria I. Florez, Hana Do, et al. 2025. “The Adoption and Use of Artificial Intelligence and Machine Learning in Clinical Development.” Therapeutic Innovation & Regulatory Science 59 (5): 1074–1086. https://doi.org/10.1007/s43441-025-00803-0.

Liu, Qi, Ruihao Huang, Julie Hsieh, et al. 2023. "Landscape Analysis of the Application of Artificial Intelligence and Machine Learning in Regulatory Submissions for Drug Development From 2016 to 2021." Clinical Pharmacology & Therapeutics 113 (4): 771–774. https://doi.org/10.1002/cpt.2668.

Makarov, Vladimir A., Mario Sänger, Roman Saiz, Matthias Hartung, and Nick Lynch. 2026. “Why Do AI Projects Fail in Drug Development and Pharma? Insights from Multi-Year Pistoia Alliance Studies.” Journal of Computer-Aided Molecular Design 40: 197. https://doi.org/10.1007/s10822-026-00894-3.

Mendeley Data. 2026. *RePairDB: An Interventional Clinical Trial Drug–Disease Association Database*, version 1. https://doi.org/10.17632/8wjx8yb55n.1.

Mullen, Joseph, Simon J. Cockell, Peter Woollard, and Anil Wipat. 2016. “An Integrated Data Driven Approach to Drug Repositioning Using Gene-Disease Associations.” *PLOS ONE* 11 (5): e0155811. https://doi.org/10.1371/journal.pone.0155811.

Pletscher-Frankild, Sune, Albert Pallejà, Kalliopi Tsafou, Janos X. Binder, and Lars Juhl Jensen. 2015. “DISEASES: Text Mining and Data Integration of Disease-Gene Associations.” *Methods* 74: 83–89. https://doi.org/10.1016/j.ymeth.2014.11.020.

Rathee, Chandrika. 2026. “AI Capability and Innovation Novelty: Evidence from Clinical Trials.” Academy of Management Proceedings 2026 (1), conference abstract. https://doi.org/10.5465/AMPROC.2026.21458abstract.

Scannell, Jack W., James Bosley, John A. Hickman, et al. 2022. “Predictive Validity in Drug Discovery: What It Is, Why It Matters and How to Improve It.” Nature Reviews Drug Discovery 21: 915–931. https://doi.org/10.1038/s41573-022-00552-x.

U.S. Food and Drug Administration. 2010. *Oravig: Prescribing Information*, NDA 022404. https://www.accessdata.fda.gov/drugsatfda_docs/label/2010/022404lbl.pdf.

U.S. Food and Drug Administration. 2012. *QNASL: Prescribing Information*, NDA 202813. https://www.accessdata.fda.gov/drugsatfda_docs/label/2012/202813s000lbl.pdf.

U.S. Food and Drug Administration. 2018. *Banzel Pediatric Utilization Review*, NDA 201367. https://www.fda.gov/media/111570/download.

U.S. Food and Drug Administration. 2019. *Vumerity Regulatory Action Summary*, NDA 211855. https://www.accessdata.fda.gov/drugsatfda_docs/nda/2019/211855Orig1s000SumR.pdf.

U.S. Food and Drug Administration. 2022. *NDA Classification Codes*, Manual of Policies and Procedures 5018.2. https://www.fda.gov/media/94381/download.

U.S. Food and Drug Administration. 2025. "FDA Qualifies First AI Drug Development Tool, Will Be Used in MASH Clinical Trials." December 8. https://www.fda.gov/drugs/drug-safety-and-availability/fda-qualifies-first-ai-drug-development-tool-will-be-used-mash-clinical-trials.

U.S. Food and Drug Administration. 2026. *Drugs@FDA Database and Original NDA Chemical-Type Classifications.* FDA public data export accessed October 2026. https://www.fda.gov/drugs/drug-approvals-and-databases/drugsfda-data-files.

Woven Health Collective. 2025. "The Current State of the Life Sciences." State of Life Sciences Report, Figure C.1 (2024 Pharmaprojects pipeline by therapy area).

Xu, Zuojun, et al. 2025. "A Generative AI-Discovered TNIK Inhibitor for Idiopathic Pulmonary Fibrosis: A Randomized Phase 2a Trial." Nature Medicine 31: 2602–2610. https://doi.org/10.1038/s41591-025-03743-2.

Technical Appendix A. Model derivations and verification

A.1 Interior equilibrium, uniqueness and the incremental-specific comparative static

Consider the system (5) at fixed inherited stocks and positive interior flows. Define

$$F_N(n,i,d) = c(d;s)R(n+\gamma i) - a_N(m) - k_N n - \kappa(n+i),$$

$$F_I(n,i,d) = \rho R(n+\gamma i) - \nu(d;x) - k_I i - \kappa(n+i).$$

The Jacobian with respect to (n,i) is

$$J = -\begin{pmatrix} A & B \\ C & E \end{pmatrix},$$

where A, B, C, E are defined in equation (6). Direct multiplication gives

$$AE - BC \quad = k_N k_I + \kappa(k_N + k_I) + h\{ck_I + \rho\gamma k_N + \kappa(1-\gamma)(c-\rho)\}.$$

For $c \geq \rho > 0$, $0 \leq \gamma \leq 1$, and $k_N, k_I > 0$, the determinant is strictly positive. Conditional on a candidate H, the linear system has a unique solution because the own-cost block is strictly increasing. Substitution into $H = n + \gamma i$ yields a scalar fixed point. More directly, since each marginal-return equation is strictly decreasing in its own flow and cross-decreasing in the other flow, the strictly positive determinant makes the implicit solution locally unique; under the maintained linear costs and decreasing rational $R(H)$, the relevant positive-flow branch has no turning point. The resulting local comparative statics are sufficient for the proposition; global existence requires that the resulting interior flows be nonnegative, otherwise the appropriate corner conditions apply.

Implicit differentiation with respect to d gives

$$\begin{pmatrix} A & B \\ C & E \end{pmatrix}\begin{pmatrix} n_d \\ i_d \end{pmatrix} = \begin{pmatrix} c_d R \\ -\nu_d \end{pmatrix} = \begin{pmatrix} G \\ V \end{pmatrix},$$

because $\nu = x/(\delta\theta d)$ and $-\nu_d = \nu/d = V$. Inverting the 2×2 block yields equation (8). With $s = 0$, $c_d = 0$, and thus $G = 0$:

$$n_d = -\frac{BV}{D}, \qquad i_d = \frac{AV}{D}, \qquad (n+i)_d = \frac{(A-B)V}{D} = \frac{[k_N + c(1-\gamma)h]V}{D} > 0.$$

Since $B = \kappa + c\gamma h$, $n_d < 0$ if $\kappa > 0$ or $\gamma > 0$, and $n_d = 0$ if both vanish. For positive platform returns, $G = c_d R \geq 0$ and the sign of n_d is determined exactly by $EG - BV$. The sign of i_d is determined by $AV - CG$; the proposition does *not* assert $i_d > 0$ for arbitrary large continuation value. Under $s > 0$, the aggregate derivative is

$$(n+i)_d = \frac{(E-C)G + (A-B)V}{D} = \frac{[k_I - \rho(1-\gamma)h]G + [k_N + c(1-\gamma)h]V}{D},$$

which is not mechanically positive when $\gamma < 1$. The sign restrictions in Proposition 1 follow. Analogously, differentiating with respect to m gives the right-hand vector $(a_N/m, 0)$ and establishes equation (9), since $a_N = 1/(\delta m)$ and $-a_N'(m) = a_N/m$.

A.2 Balanced-stock composition and the restricted baseline equilibrium

Let $x = I/N$ and use $a_N = 1/(\delta m)$, $\nu = x/(\delta\theta d)$. A frictionless interior benchmark with $s = 0, \rho = 1, k_N, k_I, \kappa \to 0$ equalizes the two returns, so

$$\frac{1}{\delta m} = \frac{x}{\delta\theta d} \quad \Rightarrow \quad x = \theta\frac{d}{m}.$$

Balanced stock accumulation requires $x_{t+1} = x_t$, implying $(i - xn)/(N + n) = 0$ and thus $i/n = x$. Hence $I/N = i/n = \theta d/m$. If $m = d = 1$, the original Hult ratio $I/N = i/n = \theta$ is recovered; with equal displacement $\gamma = 1$, $R(n+i) = 1/(\delta m)$ determines the total innovation hazard $n + i = \delta m\pi - r$ where this is positive. This is the no-friction steady-state benchmark; it does not prove that a positive continuation-claim share, finite resource adjustment, differentiated product markets or firm-level patent rents reproduce the original Hult equilibrium identically.

A.3 Innovation-stock dynamics

From $N_{t+1} = N_t + n_t$ and $I_{t+1} = I_t + i_t$ with $x_t = I_t/N_t$, algebra gives

$$x_{t+1} - x_t = \frac{I_t + i_t}{N_t + n_t} - \frac{I_t}{N_t} = \frac{i_t - x_t n_t}{N_t + n_t}.$$

When a shock increases i_t holding n_t fixed, the next stock ratio rises. Given $\nu = x/(\delta\theta d)$, a higher ratio reduces later incremental productivity holding technology fixed. This mechanism creates dynamic adjustment, but the direction of the full forward-looking innovation response additionally depends on the evolution of d, m, market values, input costs and the ownership option. A numerical local comparative static at fixed x is not automatically a unique global dynamic path.

A.4 A recursive AI capability/application layer with restored molecular platform rights and displacement

At fixed molecular stocks N and I, let x=I/N. The original molecular equilibrium takes market value R(H)=π/(r+H) with H=n+γi, and the appropriable novel-platform option multiplier c(d;s)=1+[s/(1−s)](d/d₀)^η. The incremental developer captures fraction ρ of a contemporaneous market return. Effective discovery productivities m(K,L) and d(K,L;x) reflect raw AI capability and learned applications, while application-learning effort zL competes for shared R&D inputs with n and i.

$$c(d;s)R(H) \; - \; aN/m \; = \; kN\,n \; + \; \kappa(n + i + z_l),$$

$$\rho R(H) - vx/d = kI\, i + \kappa(n + i + z_l),$$

$$\sigma_l W \delta_l \varphi(K/L) - 1 = kL\, z_l + \kappa(n + i + z_l).$$

These are exact local interior return equations under the stated reduced-form W, and explicitly restore the inherited molecular product-market displacement and platform-rent closure. Nonnegative corners require complementary-slackness inequalities, not clipping a negative unconstrained equilibrium coordinate.

Define h=−R′(H)>0 and c_d=∂c/∂d. The negative Jacobian of the three return equations with respect to y=(n,i,zL) is the matrix M below; its inverse exists at a regular interior equilibrium. The equation for learning is conditional on W being a locally fixed marginal private value, not a forward-looking Bellman condition.

$$M = [[kN + \kappa + c\, h, \kappa + c\, \gamma\, h, \kappa], [\kappa + \rho\, h, kI + \kappa + \rho\, \gamma\, h, \kappa], [\kappa, \kappa, kL + \kappa]].$$

At fixed L and inherited x, the partial K-shock vector is rK=(c_d dK R+aN mK/m², νx dK/d², σL W δL φ′(K/L)/L). If det(M)≠0 then dy/dK=M⁻¹rK. This is the correctly restored local comparative static: platform-option value, product-market displacement, molecular task productivity, and application-learning diversion all enter. No unconditional sign for novel entry or incremental entry follows. The numerical verifier solves the full three-equation system and confirms these derivatives against finite differences for hundreds of admissible local equilibria.

$$(dn/dK, di/dK, dz_l/dK)' = M^{-1} rK.$$

In the transparent reduced two-effort version, let U_N=δVm and U_L=σL W δL φ(K/L). At an interior optimum, D=kN kL+κ(kN+kL)>0, zN=[(kL+κ)(U_N−1)−κ(U_L−1)]/D. Unlike the preliminary draft, the derivative of entry E=δ m zN includes both the direct mK zN term and the equilibrium zN response.

$$dE/dK = \delta\{m_k z_n + m[(kL + \kappa)\delta V\, m_k - \kappa\sigma_l W\delta_l\varphi'/L]/D\}.$$

Thus entry rises only if (kL+κ)δV mK + (D mK/m)zN exceeds κσL WδLφ′/L. With L predetermined, q in its uncapped region has a derivative opposite in sign to dE/dK because α<1. But if initial application knowledge, advanced capacity, or molecular stocks also move, the simple sign opposition need not hold. The included counterexample demonstrates declining discovery effort and rising entry, so the earlier effort-only sign condition is invalid.

At a no-learning corner, zL=0 and zN=(U_N−1)/(kN+κ)>0. The corner persists if U_L ≤ 1+κ zN. Consequently σL>0 does not guarantee learning. For φ=K/L, fixed W and positive mL, the stopping inequality has at most one threshold above the initial L; conditional on a positive-investment trajectory and the regularity restrictions described in the verification note, the final L rises with the appropriable marginal return. However, a higher L can expand E more than ψ(L), so q does not necessarily rise. At a common impact and later state, log recovery requires Δln ψ(L)>(1−α)Δln E. Numerical examples show both improvement and deterioration in q as L rises. A globally optimal dynamic solution would require specifying a Bellman value W and forward-looking research choices, which are outside this bounded extension.

$$z_l = 0\ is\ optimal\ if\ \sigma_l W\delta_l\varphi(K/L) \le 1 + \kappa(U_N - 1)/(kN + \kappa).$$

The corrected software checks the nonnegative quadratic optimum (4,000 parameter cases), direct entry and q derivatives (1,763 admissible interior examples), and local full-molecular equations (428

admissible examples). It reports numerical parameter-sensitive learning paths separately from the analytic conditional propositions. This exercise does not estimate welfare, the value of σL, or clinical-stage success probabilities.

A.5 Forward-looking AI application learning and pipeline value

This extension retains the molecular platform rents, displacement, and shared research-input costs in Appendix A.4. Conditional on inherited AI capability K and learned application knowledge L, the novel-origination and additional-use flows n(K,L,z) and i(K,L,z) solve the existing competitive molecular-return equations, allowing application-learning effort z to bid for the same specialized inputs. The added choice is forward-looking; the molecular block remains locally competitive. The resulting conditional subgame is not a jointly forward-looking equilibrium for all drug developers.

A.5.1 Stocks, timing, and the Bellman problem

Let B denote molecular opportunities awaiting additional development and L the reusable stock of validated AI-application knowledge. Advanced-development flow A is generated from the inherited backlog; unprocessed projects may be abandoned or become obsolete at hazard δ_B. These quantities are theoretical flows, not the paper's five-year clinical-progression outcomes:

$$A_t = \mu(L_t)B_t, \quad B_{t+1} = [1 - \mu(L_t) - \delta_B]B_t + n_t.$$

Investment z accumulates application knowledge, which depreciates at rate δ_L:

$$L_{t+1} = (1 - \delta_L)L_t + \eta z_t.$$

The application developer captures share σ of development-service surplus p and pays the same congestion cost that affects novel- and additional-use research. The cost of learning is:

$$C_L(z; n, i) = c_L z + \frac{k_L + \kappa}{2} z^2 + \kappa(n + i)z.$$

For a fixed AI capability regime K, the developer chooses z while anticipating the contemporaneous competitive molecular response. Its value function is:

$$V_K(B, L) = \sigma p \mu(L)B + \max_{z \geq 0}\{-C_L(z; n, i) + \beta V_K(B', L')\}.$$

At an interior optimum, assuming differentiability, the discounted marginal return G equals the full marginal cost:

$$G = \beta[\eta V_L(B', L') + V_B(B', L')n_z],$$

$$G = c_L + (k_L + \kappa)z + \kappa(n + i) + \kappa z(n_z + i_z).$$

Here V_L is the discounted private value of additional application knowledge, while V_B n_z captures how learning effort changes the future molecular pipeline through resource competition. At the zero-learning corner, the left side cannot exceed the right side evaluated at z=0; the corner condition alone does not establish global optimality without checking the entire choice set.

A.5.2 A backlog-dependent shadow value

An exact restricted result clarifies the mechanism. Hold the novel inflow n fixed, set future learning to zero, and let current knowledge equal zero. Write $\mu_0 = \mu(0)$, $\mu'_0 = \mu'(0)$, q = 1 - μ_0 - δ_B, and let V^0 denote the private continuation value under perpetual noninvestment. Direct differentiation of its Bellman recursion, including depreciation of L, gives:

$$V_L^0(B,0) = aB + b, \qquad q = 1 - \mu_0 - \delta_B.$$

$$v_B = \frac{\sigma p \mu_0}{1 - \beta q}, \quad a = \frac{\mu_0'(\sigma p - \beta v_B)}{1 - \beta(1 - \delta_L)q}.$$

$$b = \frac{\beta(1 - \delta_L)an}{1 - \beta(1 - \delta_L)}.$$

For $\sigma > 0$, $\mu'_0 > 0$, and the maintained discount and transition restrictions, a > 0. Thus the marginal value of application knowledge rises with the pipeline backlog B and with expected future molecular arrivals n (through b). A one-period marginal deviation into learning, followed by no additional learning, is attractive under the restricted constant-inflow benchmark when:

$$\beta\eta(aB' + b) > c_L + \kappa(n + i), \qquad B' = qB + n.$$

Proof. Along the zero-learning policy, V^0(B,L) = σ p μ(L)B + β V^0([1-μ(L)-δ_B]B+n,(1-δ_L)L). Differentiating once with respect to B yields v_B; differentiating with respect to L at L=0 and equating coefficients on B and the constant gives a and b. The displayed deviation comparison follows because z increases next period's knowledge by η z and costs [c_L + κ(n+i)]z to first order when n and i are held constant. This is an exact restricted value derivative and a one-step incentive test, not a theorem locating the global policy threshold of the full Bellman problem.

Technical Appendix B. Data definitions and historical measurement

B.1 Clinical-first-start definitions

The main RePairDB outcome construction uses core/attributable RePairDB drug–disease histories with 4,272 first-observed attributable parent-molecule starts and 53,490 subsequent first-observed molecule–disease starts during 2005–2025. The ACTUAL-date variant includes 2,948 molecules whose first clinical start dates have a recorded ACTUAL flag rather than an ESTIMATED or unspecified flag. The single-category rule records one start per molecule and chooses the alphabetically first therapeutic category in simultaneous-date ties. This is a reproducible assignment convention, not an inferred medical primary indication. Fractional attribution instead assigns each tied first-date category an equal share of one molecule, preserving the same total molecule count. The original all-date fractional construction, the ACTUAL-only fractional alternative, and the older v4.4 exclusive-allocation panel (4,566 novel and 56,403 additional-use records) are documented separately. No definition equates first registry observation with verified molecular discovery or approval. ACTUAL-date quality flags are not uniformly recorded across calendar years, making the full-sample checks important.

B.2 Pre-period clinical AIRO measure

The frozen clinical-network index is reconstructed from pre-2016 RePairDB evidence breadth, molecule–disease links and multi-indication reuse, with three standardized components averaged and residualized against baseline log clinical-pair volume. See the archived pre-2016 AIRO index for each component, normalization, and category score. Pre-2016 start dates restrict the temporal input, but a retrospectively curated 2026 database is not an immutable 2015 archive; index provenance and contemporaneous availability therefore remain distinct issues.

B.3 Differential observed-link-quality diagnostics

AIRO is associated with a +2.31-percentage-point differential in observed attributable-link share in the later period (95% CI +0.15 to +4.47). This establishes that source attribution changes with the preperiod index and can affect measured starts. It does not isolate whether such changes reflect new projects or altered registry capture; category and year fixed effects do not remove differential measurement. We retain the diagnostic because it constrains the interpretation of early-entry counts.

Technical Appendix C. Estimation and statistical resolution

C.1 Outcome-specific PPML and family-clustered inference

For each outcome and index, estimate a separate Poisson GLM with category and year fixed effects and interactions of the fixed 2005–2015 index with 2016–2020 and 2021–2025 indicators. Report exponentiated percentage effects, 95% intervals using parent-family clustered sandwich standard errors with 13 reference degrees of freedom, and corresponding t_{13} tail probabilities. To avoid artificial precision, neither the two outcomes nor the two exposure indexes are combined into a stacked equation. The research question does not require a cross-equation coefficient-equality test.

The full base panel has 693 category-years. For the original clinical-network index, 2016–2020 coefficients are 0.202808 (SE 0.092973) for novel and 0.004067 (SE 0.051036) for incremental clinical starts; 2021–2025 coefficients are 0.125677 (SE 0.129851) and −0.004670 (SE 0.033247). For the FDA molecular-platform proxy, early coefficients are −0.120948 (SE 0.095338) and −0.151709 (SE 0.044141); advanced coefficients are −0.013228 (SE 0.134908) and −0.133479 (SE 0.033183). Stored regressions include 14 leave-one-family-out estimates for every outcome-index pair, preperiod placebo cutoffs, category-specific trend sensitivity and a COVID-year omission. The historical clinical-network index is constructed independently of the postperiod estimated coefficients.

C.2 Scope of causal claims

A causal interpretation as an AI productivity effect would require a predetermined measure that predicts task-specific AI opportunity and credible comparisons of untreated trends. AIRO and the FDA platform-density index capture different aspects of research structure, while external asset classifications do not validate the primary 33-category ranking. Category attribution, oncology concentration, source reporting and historical trends therefore limit attribution of observed clinical changes to AI. The empirical comparisons remain useful descriptions of economically distinct development activities.

C.3 Supplementary specifications and inference diagnostics

Table C.2. Original-outcome and AIRO-tier sensitivities

Outcome	Index representation	Early-period differential (95% CI)	Advanced-period differential (95% CI)
Novel molecule starts	Per 1-SD index	+22.5% (+0.2%, +49.7%)	+13.4% (-14.3%, +50.1%)
Novel molecule starts	High vs. all other tiers	+53.8% (+8.9%, +117.2%)	+43.7% (-0.7%, +107.9%)
Novel molecule starts	High vs. low tier	+56.5% (+18.6%, +106.6%)	+27.0% (-24.8%, +114.6%)

Outcome	Index representation	Early-period differential (95% CI)	Advanced-period differential (95% CI)
Novel molecule starts	Middle vs. low tier	+23.1% (-7.4%, +63.6%)	+36.4% (-19.6%, +131.6%)
Additional-use starts	Per 1-SD index	+0.4% (-10.1%, +12.1%)	-0.5% (-7.4%, +6.9%)
Additional-use starts	High vs. all other tiers	+4.0% (-17.9%, +31.9%)	-5.5% (-15.7%, +6.1%)
Additional-use starts	High vs. low tier	-4.3% (-19.8%, +14.2%)	-8.1% (-19.6%, +5.1%)
Additional-use starts	Middle vs. low tier	+11.2% (-23.6%, +62.0%)	+12.3% (-11.7%, +42.8%)

Notes: Alternative historical outcome definition (4,566 novel and 56,403 additional-use observations), retained only as a secondary outcome-definition check; the main results use the preferred RePairDB outcome construction.

Figure C.1. Illustrative development lag assumptions, not estimated from AI-exposed cohorts.

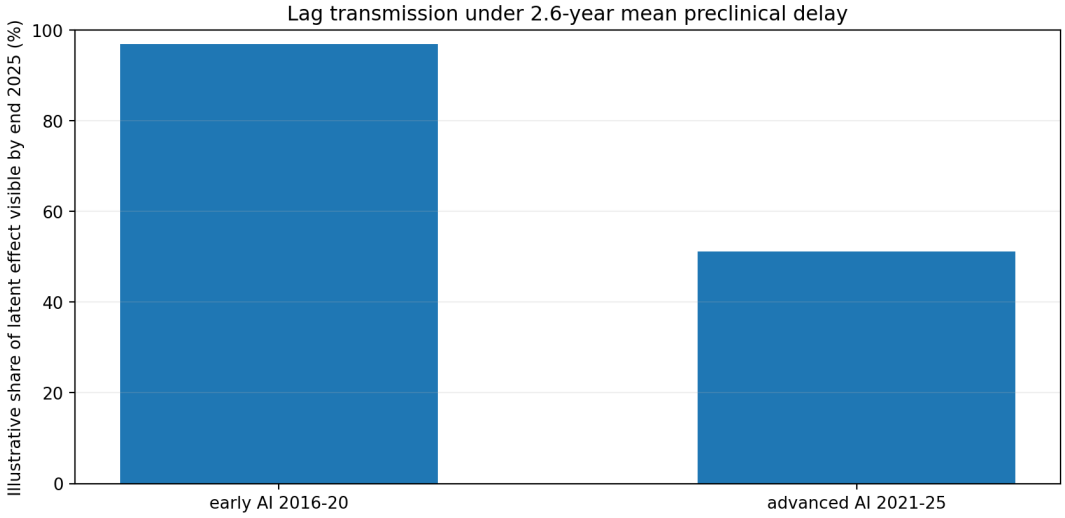


Figure C.2. Leave-one-family-out early novel-AIRO estimate.

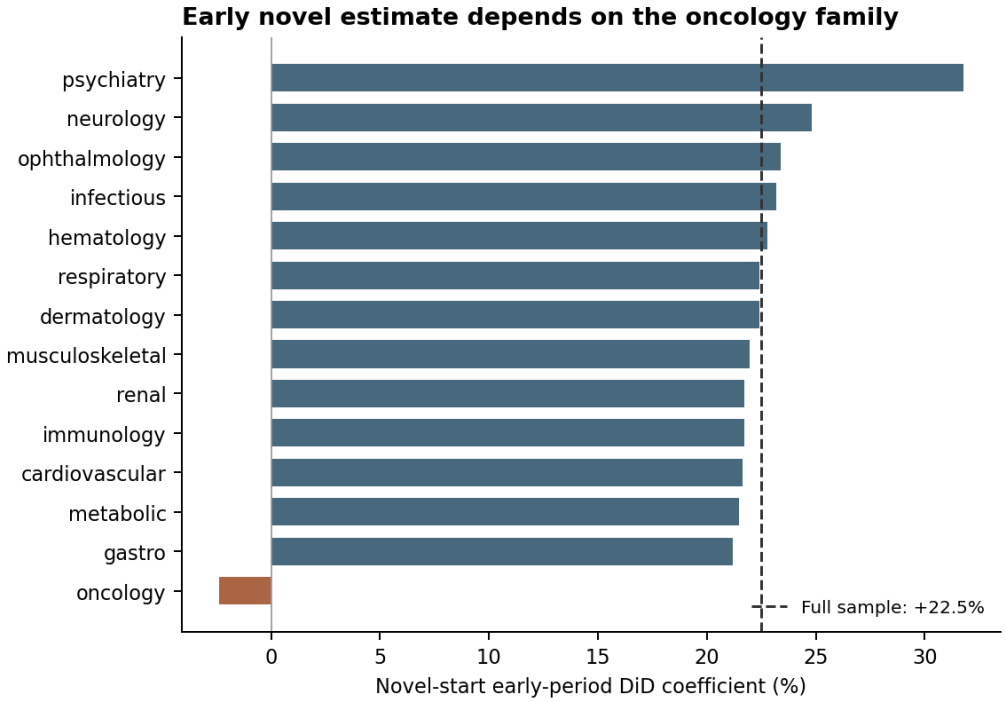


Table C.3. Direct stacked-PPML comparison of the two innovation margins

Period	Novel / additional-use relative differential	95% family-clustered CI	t(13) p-value
2016–2020	+14.1%	[-2.1%, +33.0%]	0.085
2021–2025	+9.7%	[-10.3%, +34.2%]	0.340

Notes: Joint stacked PPML with separate category and year fixed effects by outcome; same frozen AIRO score, 14 family clusters. The estimate compares log interaction coefficients across outcomes and does not identify task-specific AI productivity.

C.4 Alternative attribution rules and preperiod diagnostics

Table C.4. Novel-entry attribution and AIRO sensitivity

Outcome and predetermined opportunity measure	2016–2020 effect (95% CI)	2021–2025 effect (95% CI)
Novel entry: ACTUAL date, one category	**+18.2% (+8.2%, +29.1%); p=.0012**	—
Novel entry: all dates, fractional (full-sample benchmark)	+14.6% (−2.9%, +35.2%)	+10.3% (−10.4%, +35.8%)
Novel entry: ACTUAL date, fractional (attribution check)	+8.9% (−2.3%, +21.3%); p=.114	—
Additional-use entry: all dates (full sample)	+0.4% (−9.3%, +11.2%)	+0.6% (−5.6%, +7.1%)
Novel minus additional-use (full-sample fractional, stacked PPML)	+14.1% (−2.1%, +33.0%); p=.085	+9.7% (−10.3%, +34.2%); p=.340

Outcome and predetermined opportunity measure	2016–2020 effect (95% CI)	2021–2025 effect (95% CI)
Novel, excluding oncology (full-sample fractional)	−1.9% (−20.4%, +20.8%)	—
Novel, category-specific trends (full-sample fractional)	+4.6% (−25.9%, +47.4%)	—

Notes: All estimates are associations per one standard deviation of pre-2016 AIRO. The ACTUAL-date first-start cohort includes 2,948 molecules. The unique-category variant resolves simultaneous ties alphabetically; fractional attribution distributes the same molecule across tied categories with weights summing to one. The ACTUAL-date single-category estimate is +18.2% (p=.0012), versus +8.9% (p=.114) with fractional assignment; the broader all-date fractional estimate is +14.6%. Oncology exclusion reduces the single-category estimate to +6.9% (p=.391), and the broad fractional estimate to −1.9%; including category-specific trends in the broad sample yields +4.6%. The randomized-allocation distribution is now computed directly from the 2,948 molecule-by-first-date-category records, using the exact same PPML specification and 14-family clustering (Appendix Table C.6).

Table C.5. AIRO: preperiod trends and placebo tests

Measure and outcome	Pre-2016 exposure × annual trend	Pre-2016 placebo treatment, 2011–2015	Main interpretation
Clinical readiness / novel	+1.3% per year (p=.365)	+6.8% (p=.385)	No strong rejection; imprecise
Clinical readiness / additional-use	−0.4% per year (p=.210)	−3.2% (p=.060)	Suggestive preperiod difference

Notes: The placebo uses 2005–2015 with fictitious treatment beginning in 2011. The novel-entry placebo is +6.8% (p=.385), and the additional-use placebo is −3.2% (p=.060). Separate preperiod AIRO-by-linear-trend estimates are noisy and do not validate parallel untreated trajectories. Appendix Table C.4 provides the main novel-entry attribution sensitivities.

Technical Appendix D. Recovering dated clinical-stage evidence from molecule–disease summaries

D.1 Source structure and recoverability

A RePairDB pair combines all linked NCT trials and reports the earliest and latest start dates and highest phase. These cannot generally be joined into a trial-specific phase/date sequence. For any pair with one total trial and one attributable trial, the sole NCT necessarily has the pair's first start date and attributable phase. Pooling those unique NCT observations across every drug–disease row and joining by NCT identifiers produces 34,231 inferable NCT start dates, 31,330 with nonmissing phases. Consistency audits verify no conflicting start or nonmissing phase among these singleton rows. The resulting NCT trial-phase and start-date pairs are joined back to the frozen first-observed disease indications of 4,272 molecules. The method makes no claims about NCT histories that cannot be recovered in this way. These trial-level values are inferred from RePairDB's NCT-linked summaries; they have not been verified by downloading the complete ClinicalTrials.gov or AACT trial history.

D.2 Source-observed selection, fixed-horizon definitions, and censoring

Molecule cohort t0 equals the globally earliest core/attributable date in RePairDB, and simultaneous first-date therapeutic categories receive fractional allocations summing to one molecule. Events are dated NCT starts at or after t0 and on/before the end of the 3- or 5-year interval, with recorded start no later than December 31, 2025. Phase II+ admits 2, 2/3, or 3; Phase III admits 3; Phase IV alone is excluded. The observed NCT registry date may be marked ACTUAL, ESTIMATED, or unspecified. A stricter ACTUAL-only subset is reported for auditing but is not comparable across older and newer years because older dates frequently lack type labels. The source's maximum attributable phase is used only as an undated outer envelope, never as an assumed date of transition.

D.3 Missing histories and interpretation

Date-and-phase data are recoverable for every source-listed original-indication NCT for 48.0% of 2005–2015 entrants and 54.7% of 2016–2020 entrants (see the clinical-source completeness audit in Appendix D.3), a +6.7-percentage-point increase. Within the auxiliary high-documented-AI families, completeness rises from 53.3% to 60.6%; in other families it rises from 42.9% to 45.1%. Thus the lower documented Phase II+ share among recent entrants does not coincide with a general decline in completeness of source-listed phase/date information. This diagnostic does not test whether earlier-stage trials were increasingly registered or whether unlisted trials changed over time. The documented higher-phase shares remain minimum recorded activity, not estimated clinical survival rates; undated source-maximum phases are only loose envelopes.

D.4 Inference and clinical-economic meaning

Conditional-on-entry estimates and counted later-stage programs are distinct estimands. A descriptive high-AI-family Phase II-or-III count contrast of +5.0% has wide conventional intervals and relies on only two treated families. Relative observed rates can move for compositional reasons even without a change in underlying advancement. We do not interpret few-cluster t(13) intervals or permuted labels as independent AI-causal evidence. Appendix D.5 adds explicit source-timing envelopes and within-oncology standardization.

D.5 Source-missingness envelopes and fixed oncology-category composition

The source evidence supports two different partial measures: a conservative lower rate with a recoverably dated Phase II/III NCT and a loose upper envelope in which the source pair ever reports a sufficiently high maximum phase, regardless of when it occurred. The upper envelope does not cover external or absent trials and does not imply advancement within five years. Values below apply to source-listed original-indication trials and the 2005–2015 / 2016–2020 five-year-mature cohorts.

Table D.1. Early-cohort minimum dated and maximum phase-ever evidence within the source

Group	Phase	Dated minimum	Source-ever upper envelope	Full NCT recovery
High documented	II+	15.1%	35.5%	60.6%
High documented	III	2.3%	15.3%	60.6%
Other	II+	38.8%	77.5%	45.1%
Other	III	8.8%	29.6%	45.1%

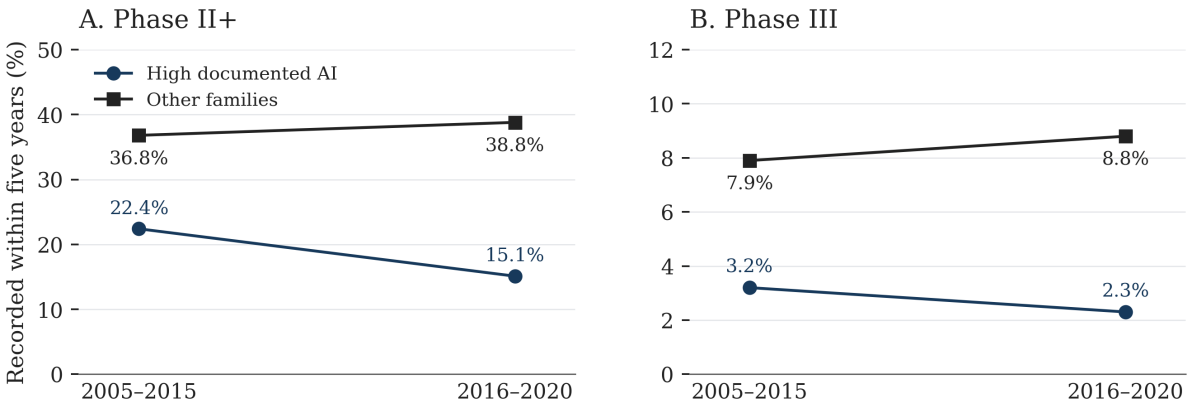
The high-to-other difference in five-year documented rates is not rescued by fine oncology category mixing. However, the envelope allows arbitrary preperiod/postperiod and high/other missing timing: the extremal relative early-versus-pre Phase II+ change ranges from −85.1% to +213.6%, and the Phase III range is wider still. A common unobserved-timing fraction applied across cells would give an illustrative Phase II+ relative change between −35.9% and −27.1%, but its equality assumption is unverified. These calculations concern recorded phase timing, not the probability of a complete drug program advancing.

Table D.2. Within-oncology recorded-stage rates: fixed subcategory composition

Outcome	2005–2015 raw	2016–2020 raw	2016–2020 at preperiod category weights
Phase II-or-III within five years	20.3%	12.5%	12.4%
Phase III within five years	2.1%	1.6%	1.7%
Source NCT dates/phases complete	55.3%	62.7%	64.0%

Fixed-mix calculations hold the distribution of first-observed oncology cohorts over ten oncology subcategories at its 2005–2015 weighted shares. Standardization shifts the early Phase II+ recorded rate from 12.5% to 12.4%, so reallocations across these ten subcategories explain little of that observed decline. They do not control for within-category changes in molecule quality, entry stage, reporting, or AI uptake. With only two non-randomly selected high-AI families and demonstrably different pretrends, the exchangeability assumptions for a causal family-label randomization-inference p-value are not credible; no such p-value is reported.

Figure D.1. Dated higher-phase activity by entry cohort and auxiliary family grouping.



Notes: The lines plot minimum documented Phase II+ and Phase III shares within five years of first-observed entry. The binary high-documented-AI grouping differs from the continuous AIRO measure. Incomplete NCT-date recovery differs across the groups; these are not underlying progression probabilities.

Technical Appendix E. AlphaFold exposure, drug-year estimation, and clinical-stage documentation

E.1 Linked sample, exposure and outcome conventions

The drug-level population is 2,316 drugs first observed clinically by 2016, with reviewed UniProt target annotations in the Open Targets 26.06 mechanism data and historical dark/bright status inferred from PDB public release dates via SIFTS. A drug with multiple targets receives the simple unweighted fraction of dark targets; alternative biological mechanisms are not weighted by contribution to efficacy. Exposure uses pre-2021 structure dates but 2026 target annotations. The latter do not establish contemporaneous developer knowledge and may bias the exposed/control comparison if related annotation improves differentially after AlphaFold.

The original sample uses 2017–2020 as pre and 2022–2025 as post, omitting 2021. The cleaner sample removes 2020 as well, leaving 2017–2019 as pre. A genuine 2016–2019 preperiod additionally restricts to drugs first clinically observed by 2015 (2,224 drugs); its drug-year estimate is +50.3% (95% CI +12.9% to +100.3%) without class-year controls. The corresponding high-dimensional class-year fit did not converge to stated tolerance and is not reported. Event studies use 2019 as the omitted reference year and include 2020 as a pretrend diagnostic rather than as the preferred pooled pre-period.

Table E.1. Exposure composition in the drug-level estimation sample

Dark-target share	Drugs	Interpretation
0	1,964	All mapped targets bright
Between 0 and 1	68	Mixed dark and bright targets
1	284	All mapped targets dark
Total	2,316	Mean dark-share exposure 0.135

Notes: Authors' calculations from the linked drug–protein exposure panel. These group counts do not reflect the relative economic importance of particular drugs.

E.2 Drug versus protein units and inference

In the harmonized clean-window drug sample, 11,807 distinct drug–indication starts become 15,573 attributed starts if every drug–protein mapping receives full event credit (1.32 attributed starts per unique start). This explains why protein-year counts cannot be treated as drug-level counts. The adjusted +11.1% protein-level association uses 1,168 class-overlap proteins and held-out balancing, and is not a commensurable regression coefficient with the unbalanced drug-year result. The class-adjusted drug-year model yields +40.2% and the same model after excluding the top 5% of historically most active drugs yields +5.9%; the latter finding is an economically important influence check.

Table 4 confidence intervals use drug-clustered model covariance. The original drug-year model was also evaluated using 1,000 successful drug bootstrap draws, producing a percentile interval from +29.0% to +112.2% around the original +67.0% specification. This bootstrap does not automatically validate the later class-adjusted model. The protein-level nested 120-draw bootstrap re-estimates balancing weights but does not resample shared drug programs, and its tail quantiles are imprecise. We therefore do not claim comprehensive dependence-robust causal inference.

E.3 Novel clinical entry and direct protein-target coverage

Table E.2. Post-2021 first-observed molecule starts with reviewed protein mappings

Year	First-observed molecules	Reviewed mappings	Coverage
2022	171	13	7.6%
2023	147	11	7.5%
2024	82	7	8.5%
2025	73	7	9.6%
Total	473	38	8.0%

Notes: First observation in RePairDB is not verified initial clinical development worldwide. Mapping means a reviewed direct target relationship, not a biochemical screen hit. Later ChEMBL screening yields candidate links without establishing intended clinical mechanisms.

A target-level first-ever-entry design yields only six post-2021 first entries in the broad observed panel. A new molecule–target program is conceptually different from first-ever molecular innovation and can include a previously studied molecule. Sparse and selectively mapped novel starts do not support a credible formal comparison of novel and incremental AlphaFold treatment effects.

E.4 Stage-documentation populations and paired uncertainty

Table E.3. New indications and highest phase documented by snapshot

Cohort	Exposure	Starts	Phase II+	Phase III+
2017–19	Any dark	577	374	182
2017–19	No dark	5,548	3,534	1,186
2022–23	Any dark	370	227	109
2022–23	No dark	2,621	1,776	618

Notes: Phase II+ and III+ columns represent highest attributable phase in the dataset snapshot; they are not verified phase entry after first observation. Cohorts have different follow-up durations. Source: Authors' calculations from RePairDB drug–indication cohort and phase documentation records.

The paired drug bootstrap resamples the same drugs for all three outcome regressions, preserving covariance among nested stage documentation measures. In 1,000 successful fits (zero failures), the difference between all-start and Phase II+ log interactions equals 0.1658 (95% percentile interval 0.0026–0.3350, approximate sign p=0.040). The Phase II+-minus-Phase III+ log contrast equals 0.0303 (95% CI −0.287–0.352). The available data have max_phase_attributable and first_start_date but no dated Phase II/III entry variables. The results are not conditional clinical advancement hazards or evidence on approvals.

E.5 Economic interpretation and limits of the AlphaFold comparison

The positive additional-indication association concerns established drugs rather than first-ever targeting or molecular invention. Its concentration among historically active drugs may be economically meaningful, but the protein-level high- versus low-capability interaction is imprecise (p=0.595) and does not confirm that mechanism. The estimates remain observational, subject to retrospective target assignments, unmeasured biological or drug-class shocks, selective mappings, and clinical-development lags.