Essay

Where value is made in drug R&D

Human genetics, the critical value creation period, and why proof of concept is the inflection point that matters most

Author	Robert Plenge, MD, PhD, Executive Vice President and Chief Research Officer, Bristol Myers Squibb
Date	October 2026

About this essay

This post expands on a talk I gave at the 5 Prime Sciences Symposium at the annual meeting of the American Society of Human Genetics (ASHG) in Montréal on October 20, 2026, titled *Managing an R&D portfolio through the critical value creation period (CVCP)*. The slides and a companion guide with full references are available here **[LINK TO BE INSERTED]**.

A paradox at the heart of drug R&D

Biopharmaceutical R&D does not lack investment. Investor funding for AI-enabled drug discovery roughly doubled between 2023 and 2025, to about $8.4 billion, and the headline value of AI partnerships signed by the 20 largest pharmaceutical companies reached about $30 billion in 2025 (McKinsey & Company, 2026). The number of active pipeline programs nearly doubled between 2015 and 2024. Yet R&D remains inefficient and expensive, and productivity has not kept pace with spending. **The question is not how much capital the industry invests, but how effectively that capital is deployed to improve R&D productivity.**

Too often, that capital returns to the same biology. Over the same decade, the number of novel targets entering the pipeline each year fell by about 70% (L.E.K. Consulting, 2025). About 92% of human protein-coding genes have never been the subject of a clinical-stage drug program, while just 81 genes, 0.4% of the genome, account for 44% of all clinical programs started since 2000 (Citeline data released with Minikel et al., 2024).

Even the newest money follows familiar paths. Of the 15 largest pharma–AI discovery deals from 2023 to mid-2026, 14 included molecule design, four included target or disease biology, and none addressed the translation from preclinical models to patients. As McKinsey put it, novel, causally validated biology "remains biopharma's scarcest commodity."

I believe this paradox has a structural explanation and a structural answer. Value in drug R&D is not created evenly along the pipeline. It is concentrated in a narrow window, the path from first-in-human dosing to clinical proof of concept, which I call the **critical value creation period (CVCP)**, a concept discussed previously on this blog. Human genetics is the most scalable source of evidence for winning in that window, and AI is most valuable when it is pointed there too.

The argument in brief

- **Value is created at proof of concept (PoC).** Probability of success rises from about 10% at the start of Phase 1 to 30–60% after a well-designed PoC, and only then can a mega-blockbuster be told apart from a stable product.
- **Large companies need mega-blockbusters to grow.** Because returns are so skewed, roughly 10–20% of launches must become mega-blockbusters for a large company to sustain growth through loss of exclusivity.
- **Human genetics changes the odds twice.** It raises the starting probability of success, and it supplies the biomarkers, doses and patient selection that make PoC decisive.
- **Programs and portfolios should be managed by dollars spent per change in probability of success.** Genetic evidence moves programs toward the efficient frontier; AI can scale this, but only if it learns the integrative judgments that connect analyses.

Three problems, two solutions

Drug R&D is failure-prone, expensive and inefficient, and the numbers are familiar to anyone who works in the field. Only about 10–14% of programs entering Phase 1 reach approval (Wong, Siah and Lo, 2019). Most late-stage failures occur due to lack of efficacy, and about a quarter due to safety (Harrison, 2016). The average cost to take a single asset from discovery to launch at the largest companies reached about $2.67 billion in 2025 (Deloitte, 2025). And target crowding means that much of this effort duplicates work already under way elsewhere.

Two solutions address all three problems at once. The first is **disciplined adherence to R&D principles, combined with managing every program through clinical PoC**. The second is **integrating AI and automation** to improve efficiency, shorten decision cycles and raise the quality of what enters the clinic.

R&D principles matter more than they might seem. At Bristol Myers Squibb we apply five, spanning the full continuum: causal human biology, matching modality to mechanism, path to clinical PoC, accelerated full development, and maximizing market access. The first three govern research and early development. Their value is not in any single principle but in creating a common language for decisions across Research, Development and Commercial teams. There is good evidence that such frameworks work: after AstraZeneca adopted its "5R" framework (right target, tissue, safety, patient and commercial potential), the share of its projects completing Phase 3 rose from 4% to 19% (Morgan et al., 2018).

Where value is created: the critical value creation period

Ask where along the R&D value chain the most value is created, and the answer is not discovery, where programs are cheap but almost always fail, nor late development, where programs are de-risked but expensive. **The answer is the path to clinical proof of concept.**

The arithmetic is simple. A program entering Phase 1 has roughly a 10% chance of becoming an approved medicine. After a well-designed PoC study succeeds, that probability rises to somewhere between 30% and 60%, depending on the strength of the therapeutic hypothesis and the quality of the PoC. At the start of Phase 3 it is about 60–70% (plengegen.com/blog/cvcp). The largest single jump in value happens in early development. This is consistent with a well-known finding from R&D productivity modeling: Phase 2 has the lowest transition probability of any phase, and reducing Phase 2 attrition by stopping weak programs early and cheaply improves the cost per launch more than

almost any other lever (Paul et al., 2010).

PoC here has a specific meaning: **the earliest point at which data suggest a reasonable likelihood of achieving the aspirational target product profile (aTPP)**. That definition carries a commercial insight. Before PoC, it is not possible to know whether a program will become a stable product or a mega-blockbuster. Commercial potential can and should be estimated up front through the aTPP, but it remains a hypothesis until human data test it.

Four actions follow from this framing:

1. **Minimize the time and capital needed to reach PoC.**
2. **Increase the probability of success at PoC and, ultimately, at approval.**
3. **Increase the ratio of mega-blockbusters to stable products** in the portfolio.
4. **Increase speed to high-value markets** once PoC is achieved, including running indications in parallel.

Why mega-blockbusters matter

The third action deserves explanation, because it is where R&D strategy and commercial reality meet. As defined here, a mega-blockbuster is a product with more than $5 billion in annual revenue, and a stable product is one with $3 billion or less. Many stable products are important medicines; the distinction is financial, not clinical. But three features of biopharma economics make mega-blockbusters essential for a large company.

First, returns are extremely skewed. In the cohort of new drugs introduced in the United States between 1988 and 1992, the top decile accounted for 56% of total sales revenue (Grabowski and Vernon, 2000). Earlier work found that only the top 30 of 100 drugs introduced in the 1970s covered average R&D costs. The pattern persists. In the 2025 late-stage pipelines of the top 20 companies, 54 blockbuster assets, only 9% of drugs in advanced development, are forecast to deliver about 70% of projected sales, and about half of pipeline asset-indications have forecast peak sales below $250 million (Deloitte, 2025).

Second, average returns barely clear the cost of capital. Deloitte's projected internal rate of return on late-stage pipelines rose to 7% in 2025. Excluding a single drug class, the GLP-1 and GIP agonists for obesity and diabetes, it falls to 2.9%. Average forecast peak sales per asset were about $600 million. Without a few very large products, the average program does not earn back what it costs.

Third, revenue must be replaced before it can grow. Between 2025 and 2030, estimates of branded revenue facing loss of exclusivity range from about $180 billion to more than $300 billion, including about 70 blockbusters. Meanwhile, a top 20 company averaged roughly one new FDA approval per year between 2014 and 2023, counting acquisitions (Schuhmacher et al., 2024).

Put these together and a simple calculation makes the point. Consider a company with $50 billion in annual revenue that aims to grow at 3–5% a year for a decade while a third to half of its current revenue loses exclusivity. Assume generously that stable products reach $1 billion in peak sales and mega-blockbusters $7 billion. **A company launching two to three new medicines a year needs roughly 10–20% of them to become mega-blockbusters to sustain mid-single-digit growth. At one to two launches a year, the required share rises to 20–45%, which is not realistic.** The model is deliberately simplified, but the conclusion is robust: large companies grow on a small number of very large products, and they cannot tell which programs those will be until PoC.

That is why the CVCP matters commercially, not just scientifically. A portfolio must carry enough programs through PoC, quickly and cheaply, to find the few that matter.

Human genetics changes the odds twice

If PoC is where value is created, the question becomes what raises the odds of reaching it successfully. The strongest evidence points to human genetics. Drug mechanisms with human genetic support were about twice as likely to reach approval in an early analysis (Nelson et al., 2015), with stronger effects when the causal gene is assigned confidently (King, Davis and Degner, 2019). The most recent analysis puts the advantage at 2.6-fold from Phase 1 to launch. Strikingly, the benefit is largely independent of genetic effect size or allele frequency (Minikel et al., 2024). A small-effect common variant can be as informative about a target as a rare variant of large effect.

Genetics acts on the CVCP in two distinct ways, and it helps to separate them.

- **It raises the starting probability of success.** Causal evidence from human genetics is the strongest single input to the first R&D principle, causal human biology, and it moves a program higher on the probability axis before it ever enters the clinic.
- **It makes the PoC readout more decisive.** Genetics specifies the direction of effect (inhibit, degrade or activate), the level of target modulation needed (from allelic series), the biomarkers that should move (from protein and expression quantitative trait loci), and the patients most likely to respond. A decisive PoC sends good programs sharply up and poor programs toward zero.

Figure 1 shows why both effects matter, using an illustrative model. Genetic support alone raises the starting probability, but a weak PoC leaves even a genetically supported program stranded between 17% and 37%, having consumed PoC capital without producing a clear decision. **Only genetic support combined with a decisive PoC reaches the range above 50% that justifies acceleration and concentrated investment.**

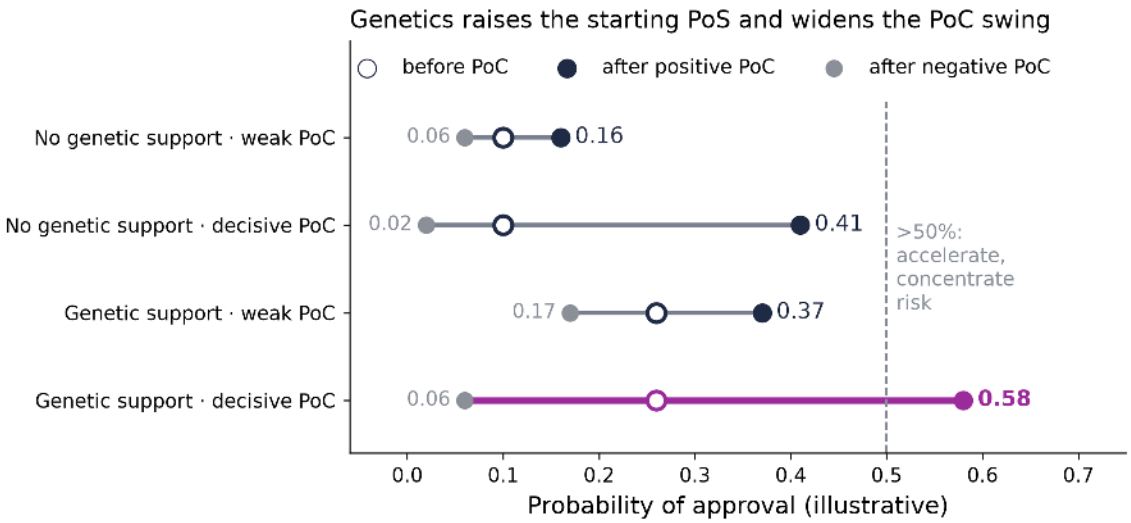


Figure 1. *Genetics raises the starting probability of success and widens the swing at proof of concept. Open circles: before PoC; filled circles: after a positive PoC (dark) or a negative PoC (grey). Illustrative model: probability of approval given a true hypothesis = 0.7; genetic support about 2.6-fold; decisive PoC sensitivity 0.85 and specificity 0.90; weak PoC 0.65 and 0.65.*

This framing has a practical consequence for how programs are managed. Every program passes through a series of inflection points: development candidate, toxicology, IND, safety and tolerability, pharmacokinetics and pharmacodynamics, early clinical data and PoC. Each raises the probability of success by a different amount, at a different cost. The useful measure is not the probability at any one point but **the expected change in probability of success per dollar and per month spent to reach the next inflection point.**

Programs differ on all three dimensions: starting probability, the inflection points they pass through, and the time and capital needed to reach PoC. A fast path to PoC is attractive regardless of commercial potential. A medium path

is acceptable for medium and large commercial opportunities. A slow, expensive path is acceptable only when the commercial opportunity is very large, and only when strong causal biology provides the conviction to invest.

Three programs, three kinds of human evidence

"Genetic support" is not a single thing. Three programs at Bristol Myers Squibb illustrate how different sources of human evidence shape a program's path through the CVCP (Table 1).

An HbF-inducing degrader for sickle cell disease rests on germline genetics. Sickle cell disease is caused by a single amino-acid change in adult beta-globin. Fetal hemoglobin (HbF) does not participate in sickle polymerization, and higher HbF levels are associated with longer survival (Platt et al., 1994). People who inherit hereditary persistence of fetal hemoglobin alongside the sickle mutation often have mild disease, a natural experiment showing that lifelong high HbF is protective and well tolerated. Genome-wide association studies found common variants at BCL11A that explain much of the population variation in HbF (Uda et al., 2008), fine-mapping located them in an erythroid-specific enhancer (Bauer et al., 2013), and editing that enhancer eliminated vaso-occlusive crises in early clinical studies (Frangoul et al., 2021). That path led to an approved gene-editing therapy: clinical proof that genetically guided HbF induction works.

The two targets of BMS-986470 have different sources of validation. ZBTB7A, master transcriptional repressor of fetal hemoglobin, has direct germline support: naturally occurring mutations near the fetal globin genes disrupt its binding and raise fetal globin expression (Martyn et al., 2018). WIZ, another transcriptional repressor of fetal hemoglobin, was discovered through chemical genetics, in a phenotypic screen of molecular glue degraders (Ting et al., 2024). Degrading both repressors raises HbF synergistically, and the oral degrader is now in a Phase 1/2 PoC study (NCT06481306). Because HbF is a quantitative biomarker with known clinical thresholds, PoC can be read early. **Genetics fixed the target, the direction, the biomarker and the target level, which places this program on a fast path.**

A BCL6 degrader for lymphoma rests on somatic genetics. BCL6 is a transcriptional repressor that is essential for germinal center B cells (Ye et al., 1997). In lymphoma it is deregulated by chromosomal translocations and promoter mutations, and BCL6 alterations help define genetic subtypes of diffuse large B-cell lymphoma (DLBCL) (Schmitz et al., 2018). Mice engineered to express BCL6 constitutively develop a disease resembling human DLBCL (Cattoretti et al., 2005). BCL6 was long considered hard to drug, but small molecules that bind it can trigger its degradation (Slabicki et al., 2020). BMS-986458 is a highly selective ligand-directed degrader of BCL6 (J Med Chem, 2026), now in a Phase 1/2 study (NCT06090539). **Because the genetic lesion is in the tumor, it defines both causality and the patients most likely to respond.** About 30% of DLBCL cases show high BCL6 expression, and tumor degradation provides a direct measure of target engagement.

CD19 cell therapy for autoimmune disease rests on human pharmacology. Here the decisive human evidence came not from genetics but from treating patients. CD19 CAR T cells produced drug-free remission in refractory systemic lupus erythematosus (Mackensen et al., 2022), and in a series of 15 patients with lupus, myositis or systemic sclerosis, all stopped immunosuppression, with B cells returning predominantly as naive cells (Müller et al., 2024). Genetics plays a supporting role. People with inherited CD19 deficiency have an antibody-deficiency syndrome but otherwise develop normally, a human knockout that bounds the expected on-target effect (van Zelm et al., 2006), and lupus genetics implicates many B-cell genes (Bentham et al., 2015). **Immune reset visible in a handful of patients is early evidence of mega-blockbuster potential, and it justifies investing to accelerate.** The genetics opportunity now lies downstream, in ranking concurrent indications and identifying patients likely to achieve durable reset.

Table 1. Three programs, three sources of causal human evidence.

Program	Primary source of causal evidence	What the evidence informs	Path to PoC
HbF-inducing degrader (sickle cell disease)	Germline genetics: HbF persistence, common HbF variants, rare variants disrupting ZBTB7A binding	Target, direction of effect, biomarker and target level	Fast
BCL6 degrader (B-cell lymphoma)	Somatic genetics: translocations and mutations that deregulate BCL6	Causality, patient selection, target engagement	Medium
CD19 cell therapy (autoimmune disease)	Human pharmacology, supported by a human knockout and lupus genetics	Mechanism, on-target safety, indication breadth	Fast

Knowing the source of a program's causal evidence tells a portfolio team which inflection point carries the most risk, and therefore where the next dollar should go.

From programs to portfolios

The same logic scales from one program to many. Each program in a portfolio can be placed on two axes: **scientific conviction**, scored against the first three R&D principles, and **CVCP value**, the expected change in probability of success from first-in-human to PoC per dollar invested. The aTPP sets each program's commercial weight. Each region of that map implies a default action. Programs with high conviction and high CVCP value should be funded fully, with concurrent indications planned for after PoC. Programs with high conviction but an expensive path should have that path optimized, often with genetic patient segmentation or better biomarkers. Programs with an efficient path but weaker conviction need decisive experiments before PoC money is spent. And programs that are weak on both should be stopped or redesigned, freeing capital for the rest.

Genetic evidence moves programs across this map: up through causal human biology, and toward efficiency through more decisive and cheaper PoCs. It also has a portfolio-level benefit that is easy to miss. Crowded targets are correlated bets. When many programs pursue the same target, a single clinical readout can move all of them at once. Genetically supported novel targets offer both a higher starting probability and genuine diversification.

For commercial teams, the framework changes the conversation. The aTPP is set before the program enters the clinic and refined through PoC, rather than fixed at the start of Phase 3. Indication sequencing can draw on genetic evidence about which diseases share a mechanism. And the fourth CVCP action, speed to market after PoC, depends on having made those decisions before the data arrive.

Pointing AI at the bottleneck

AI enters this story in a specific place. Today, most AI investment goes to molecule design, where data are structured and experimental feedback is fast. **AI for target identification is not yet as mature as other applications, such as molecule design and clinical modeling.** Models still struggle to separate causation from correlation, and feedback on whether a target was right arrives only years later, in the clinic (McKinsey & Company, 2026). Yet target identification is where the CVCP begins, because it sets a program's starting probability of success. The CVCP framework therefore points AI at the part of the value chain where it is least mature and where most of the money is not going.

Our aspiration, the "dream", is an AI-enabled research organization in which every target nominated in early discovery comes with a prediction of its safety and efficacy profile from a causal foundation model, development candidates are generated by a fully automated closed-loop laboratory, and every clinical experiment is managed dynamically through the CVCP with predictive modeling. Clinical results would flow back to update the models. The "build" is more prosaic: five interdependent capabilities spanning disease biology insights from human data, an automated lab of the future, a clinical data feedback loop, AI-ready data infrastructure, and a co-scientist layer of agents for experiment design and knowledge extraction.

How close are we? A recent preprint from Stanford University offers a useful window. **DrugTargetWorld** generates synthetic biobanks, or "worlds", each with genotypes, proteins, health records, cardiac imaging and outcomes for 54,000 simulated participants, governed by a hidden causal model (Margolis et al., 2026). Because the true causal drivers are known by construction, AI agents can be scored against ground truth on the full target-discovery workflow. They must build a disease phenotype from raw data, nominate causal proteins, state whether each should be inhibited or activated, reject proteins made to look causal by planted biases such as confounding, reverse causation or pleiotropy, and decide whether to spend a limited budget on virtual experiments.

The results are instructive. Across 540 episodes, the two leading agents recovered about 64% of causal drivers on average, and both relied mainly on cis Mendelian randomization with appropriate checks on instrument strength. But no agent reliably rejected misleading non-causal proteins, and the best overall scores were about 35 to 40 of 100. A larger experimental budget did not detectably improve the leading agents' scores. In the authors' words, current agents can perform most of the individual analyses in a biobank study but do not yet consistently make the integrative judgments that connect them: **how to measure disease, how to separate causal drivers from non-causal proteins, and when evidence is sufficient to support a claim.**

Those judgments are precisely the ones that set a program's starting probability and shape its path to PoC. Environments with known ground truth and verifiable rewards could help train and test the causal models and agents that the dream requires, before they are trusted with real portfolio decisions.

What this asks of us

The CVCP framework makes specific demands on each of the communities it touches.

- **Of human geneticists**, it asks for outputs that change the shape of a program's path: effect sizes reported with direction and uncertainty, not just P values; allelic series that indicate how much target modulation is needed; quantitative, genetically anchored biomarkers that read out early; genetic patient subgroups and endophenotypes; phenome-wide maps that rank concurrent indications; and genotyped clinical trials that feed results back into genetic priors.
- **Of R&D leaders**, it asks for investment in the decisiveness of PoC, not just in reaching it; for programs to be judged by the expected change in probability of success per dollar; and for discipline in avoiding strategic de-prioritization of programs before they read out.
- **Of commercial teams**, it asks for an aspirational product profile defined before the clinic, refined through PoC, and connected to an indication strategy informed by genetics.

Conclusion

Drug R&D remains failure-prone, expensive and inefficient. The most value is created on the path to clinical proof of concept, and that is where evidence, capital and AI should be concentrated. **High-conviction targets, paired with the right modality and a decisive path to PoC, unlock novel biology and clinical differentiation.** Human genetics is the most scalable source of that conviction. AI and automation can scale these principles across the pipeline, but only if they learn to make the judgments that turn analyses into decisions.

Further reading

The companion guide to the ASHG talk contains a fuller reference list. Key sources cited in this post:

- **R&D productivity:** Paul et al., Nat Rev Drug Discov 2010; Wong, Siah and Lo, Biostatistics 2019; Harrison, Nat Rev Drug Discov 2016; Morgan et al., Nat Rev Drug Discov 2018.
- **Genetic support and clinical success:** Nelson et al., Nat Genet 2015; King, Davis and Degner, PLoS Genet 2019; Minikel et al., Nature 2024.
- **Returns and the economics of mega-blockbusters:** Grabowski and Vernon, PharmacoEconomics 2000; Deloitte, Measuring the return from pharmaceutical innovation, 2025; Schuhmacher et al., 2024; L.E.K. Consulting, *Is biopharma doing enough to advance novel targets?*, 2025.
- **The CVCP framework:** plengegen.com/blog/cvcp.
- **HbF induction:** Platt et al., NEJM 1994; Uda et al., PNAS 2008; Bauer et al., Science 2013; Frangoul et al., NEJM 2021; Martyn et al., Nat Genet 2018; Ting et al., Science 2024.
- **BCL6:** Ye et al., Nat Genet 1997; Cattoretti et al., Cancer Cell 2005; Schmitz et al., NEJM 2018; Slabicki et al., Nature 2020; BMS-986458, J Med Chem 2026.
- **CD19 cell therapy:** Mackensen et al., Nat Med 2022; Müller et al., NEJM 2024; van Zelm et al., NEJM 2006; Bentham et al., Nat Genet 2015.
- **AI in drug discovery:** Margolis et al., DrugTargetWorld, arXiv 2026; McKinsey & Company, *AI drug discovery: focusing on what matters most*, September 2026.

Figure 1 and the growth calculation are illustrative models, not forecasts for any company or program.