Title: Financial rewards to scientific novelty in small-molecule drug development

Author: Boshen Jiao[1,2]

Affiliation:
1. Department of Pharmaceutical and Health Economics, Alfred E. Mann School of Pharmacy and Pharmaceutical Sciences, University of Southern California, Los Angeles, California, USA
2. The Leonard D. Schaeffer Center for Health Policy & Economics, University of Southern California, Los Angeles, California, USA

*** Corresponding author:**
Boshen Jiao
Department of Pharmaceutical and Health Economics
University of Southern California
Los Angeles, CA, USA
Email: boshenji@usc.edu

Working Draft — November 2026

Abstract

Drug developers must choose which candidates to advance before knowing whether they will reach the market. We estimated expected financial value at Phase I entry for small-molecule drugs by combining chemical and target novelty with US approval outcomes, clinical development time and costs, and subsequent revenue. Structurally distinct candidates had estimated values of US$165 million, compared with US$126 million for analogues and US$76 million for candidates based on previously approved structures. Greater revenue after approval was partly offset by a lower probability of approval and longer development. Candidates acting on targets with no earlier approved drugs also had higher estimated value than those with more predecessors, although this association depended on model specification. Estimated values for several novelty groups were higher when later revenue years were used, but differences in data coverage limit interpretation as a change in development incentives.

Introduction

Whether scientific novelty is financially rewarded depends on the decisions drug developers face before they know which candidates will succeed. A structurally distinct drug may earn more if it reaches the market, but developing it can also involve a lower probability of approval, a longer development period, and higher costs. The question is therefore not simply whether novel drugs generate greater revenue after approval, but whether that revenue advantage translates into higher expected financial value when clinical development begins.

Existing research highlights this tension. Krieger et al. found that chemically novel drugs face greater development risk but have higher estimated value[1]. In our earlier study, we used measures of chemical similarity, biological targets, and delivery-related properties to examine revenues among approved small-molecule drugs, finding that more novel products earned higher gross and net revenues, particularly in later years[2]. Those findings establish a commercial advantage among successful drugs, but revenue conditional on approval does not reveal the value of pursuing a candidate before the outcome is known. Evidence on clinical attrition[3,4], R&D costs[5–7], uneven drug revenues[8–10], and the consequences of long development periods[11] underscores that distinction. The returns to innovation also depend on market size and policy incentives[12–16], competition and market exclusivity[17–20], and other dimensions of private innovation value[21].

Building on our previous study of the relationship between scientific novelty and revenues among approved drugs, we examine whether pursuing novelty is financially attractive at the start of clinical development. We assembled a cohort of small-molecule drugs entering Phase I and linked their chemical structures and biological targets to subsequent regulatory outcomes, development timelines, clinical trial costs, and US revenues. By integrating these components, we estimated expected financial value at Phase I entry and examined how it varies with chemical and target novelty. We further assessed how these relationships differ across revenue periods and whether the novelty measures are associated with independently assessed therapeutic benefit.

Methods

Study cohort and data sources

We identified 1,141 small-molecule drug candidates that entered Phase I between 1996 and 2008 in Pharmaprojects, allowing at least 12 years of follow-up. We matched candidates to US Food and Drug Administration (FDA) approval records using drug identities and regulatory histories. An approval counted only if it occurred during or after the candidate's Phase I entry year; earlier approvals of the same molecule did not count as a new success. Drugs marketed only outside the United States were not counted as US approvals.

We used the Medical Expenditure Panel Survey (MEPS) to estimate US outpatient prescription spending and SSR Health to obtain branded-drug sales and discounts. MEPS data covered 2000–2019 and SSR Health covered 2007–2019; our primary revenue window was 2007–2019. We linked sales information to approved drugs using product and regulatory records, accounted for periods of market exclusivity, and expressed financial amounts in 2020 US dollars. MEPS methods are described elsewhere[22]. Pharmaprojects and SSR Health data were obtained under commercial licenses.

Measuring novelty and delivery properties

We measured chemical novelty by comparing each candidate's molecular structure at Phase I entry with those of previously approved drugs. Extended-connectivity fingerprints (ECFP4) summarized molecular features, and a Tanimoto similarity score identified the closest earlier structure on a scale from zero to one[23]. We classified candidates as structurally distinct (less than 0.4), analogues (0.4 to less than 1), or based on a previously

approved structure (equal to 1). We obtained structures and targets from ChEMBL and DrugBank[24,25] and calculated molecular fingerprints using the Chemistry Development Kit[26].

Target novelty captured how many previously approved drugs acted on the same biological targets. We grouped candidates as first in target, with one to nine predecessors, or with ten or more, and also examined novelty relative to earlier drugs within the same therapeutic class using the World Health Organization's Anatomical Therapeutic Chemical classification[27]. Separately, we counted departures from five conventional molecular properties: molecular weight, lipophilicity, hydrogen-bond acceptors, hydrogen-bond donors, and molecular flexibility[28,29]. We grouped drugs according to whether they had zero or one, two or three, or four or five departures. This measure describes physicochemical properties, not a new delivery technology.

Estimating expected financial value

We estimated the expected net present value (ENPV) of each drug candidate at Phase I entry by combining its probability of US approval with the revenue expected if approved, accounting for market exclusivity and the delay before approval, and subtracting expected clinical-trial costs:

$$\mathrm{ENPV} = \mathrm{P(US\ approval)} \times R_{net} \times A \times D - C$$

Here, P(US approval) is the probability of receiving US approval, R_{net} is expected annual US revenue after rebates, A represents the discounted period of market exclusivity, D accounts for the time from Phase I entry to approval, and C is the expected discounted cost of clinical trials. We discounted future revenue and costs at 10% annually. Because other drugs may reach the market during development, the revenue estimate reflects each candidate’s projected relative novelty at approval.

We estimated approval, development time, trial costs, and revenue separately while accounting for drug type, therapeutic area, and relevant time periods. Trial expenditures covered Phases I–III and drew on observed enrollment and published per-patient costs[30]. The primary revenue analysis used identifiable MEPS prescriptions supplemented by selected SSR Health records for drugs that could have appeared in MEPS without suppression of their identities. These 1,240 additional drug-years represented 312 drugs. We assigned them weights reflecting their estimated chance of appearing in the survey and adjusted their revenue accordingly. This approach approximates missing drug-level observations; it does not establish which drugs were actually masked.

Results

Structural differentiation is associated with higher expected financial value

We examined 1,141 small-molecule drug candidates entering Phase I between 1996 and 2008, of which 130 (11.4%) subsequently received US approval. Among candidates with a chemical similarity score, 807 were structurally distinct, 257 were analogues, and 76 involved a previously approved structure. We compared their expected financial value at the start of clinical development, rather than considering only drugs that ultimately reached the market.

Expected financial value declined as chemical similarity to previously approved drugs increased. For the 2007–2019 revenue window, structurally distinct candidates had an estimated value of US$165 million, compared with US$126 million for analogues and US$76 million for candidates based on a previously approved structure (Fig. 1a). The reported difference between the extreme groups was US$88 million using unrounded estimates. The same pattern appeared when similarity was measured only against drugs in the same therapeutic class (US$164 million, US$126 million, and US$77 million, respectively; Fig. 1b). Structural differentiation

therefore remained associated with higher expected value after accounting for approval risk, development time, and clinical-trial costs.

Target novelty has a smaller, model-sensitive association with value

Target novelty showed a smaller association with expected financial value than chemical novelty. Among the drug candidates, 313 were first in target, 370 had one to nine earlier drugs acting on the same targets, and 256 had at least ten. Their estimated values declined from US$186 million for first-in-target candidates to US$170 million and US$148 million for candidates with more predecessors (Fig. 1c). The corresponding values within therapeutic classes were US$181 million, US$156 million, and US$124 million (Fig. 1d), showing a similar pattern within the same therapeutic area.

The size of the target-novelty association depended on how earlier approvals were represented in the models. A continuous count of predecessors produced a smaller difference for first-in-target drugs than a model using separate target groups. Thus, chemical structure and target novelty provided related but different information about expected financial value, with the target results more sensitive to model specification.

Estimated financial value is higher in later revenue windows

Estimated financial value was higher for several drug groups when we used revenues from later years, particularly among candidates with greater chemical or target novelty (Fig. 1). For structurally distinct candidates, expected value was US$133 million using 2007–2012 revenues, US$205 million using 2013–2019 revenues, and US$232 million using 2015–2019 revenues; estimates for candidates based on previously approved structures changed much less. First-in-target candidates showed a similar pattern (US$149 million, US$233 million, and US$275 million, respectively), while drugs with at least ten target predecessors showed smaller increases. These differences describe how the same development cohort is valued using different revenue periods, not changes in returns to successive groups entering Phase I.

Post-approval revenue advantages narrow at Phase I entry

After approval, structurally distinct drugs were predicted to earn substantially more than drugs based on previously approved structures. Estimated annual net revenue was US$539 million for structurally distinct drugs, versus US$146 million for drugs using an earlier structure (Fig. 2a). Comparisons based on gross revenue also favored structural novelty, with ratios of 3.52 and 2.08 across the two revenue specifications. These comparisons capture commercial rewards among successful drugs, but not the financial risk taken when clinical development begins.

The greater revenue potential of structurally distinct drugs was partly offset by a lower likelihood of approval and slower development. Their estimated probability of US approval was 11.1%, compared with 14.1% for candidates based on a previously approved structure (Fig. 2b). They also had a lower time-to-approval discount factor (0.435 versus approximately 0.71; Fig. 2c), reflecting the longer wait for potential revenue. The approval and development-time factors were 0.78 and 0.61 in the component analysis. After allowing for these differences, rebates, and trial costs, the estimated value ratios at Phase I entry were 1.96 and 1.11. Those final ratios include adjustments beyond the gross-revenue, approval, and time factors shown here.

More departures from conventional physicochemical ranges are associated with lower expected value

Expected financial value also varied with the number of departures from conventional physicochemical-property ranges. Drugs with zero or one departure from this Lipinski/Veber-type index[28,29] had an estimated value of US$216 million, compared with US$148 million for two or three departures and US$103 million for

four or five (Fig. 3). The index describes molecular properties rather than delivery technology or novelty relative to previously approved drugs.

Both novelty measures are associated with assessed therapeutic benefit

Both novelty measures were also associated with assessed therapeutic benefit among approved drugs. In French ASMR assessments, greater chemical similarity to earlier drugs was associated with lower odds of receiving a more favorable rating (adjusted odds ratio, 0.53; 95% confidence interval, 0.38–0.73, per interquartile-range increase). Greater target crowding showed a similar association (odds ratio, 0.59; 95% confidence interval, 0.43–0.81). Because these assessments include only approved drugs, they do not establish that novelty itself causes better therapeutic outcomes.

Discussion

The financial rewards to scientific novelty look different when assessed at the start of clinical development rather than after a drug reaches the market. Structurally distinct candidates retained higher expected value at Phase I entry, although their larger potential revenues were partly offset by lower approval rates and longer development. This result extends earlier work linking chemical novelty to development risk and commercial value[1], as well as our analysis of revenues among marketed drugs[2], by considering the costs and uncertainty faced before approval. It shows why the performance of successful drugs alone cannot fully describe the financial choices developers face.

Chemical and target novelty capture different aspects of drug innovation. A new chemical structure may act on an established target, while a drug directed at a less-explored target need not have a markedly different structure. Their different associations with expected value suggest that scientific novelty cannot be treated as a single economic characteristic. The more model-sensitive target results do not rule out substantial rewards for first-in-target drugs, but they provide less clear evidence about the size of that reward. Similarly, drugs based on established structures should not automatically be viewed as imitations: competing drugs are often developed in parallel, before the commercial success of a class leader is known[32].

Estimated values were higher for several drug groups when later revenue years were used, raising the possibility that the commercial opportunities available to developers have changed. However, the comparisons concern the same cohort of drug candidates, while revenue coverage and the contribution of reconstructed observations differ across periods. The estimates therefore should not be read as proof that incentives for newly developed drugs increased over time. The broader importance of development time is consistent with earlier research on how delayed returns can discourage pharmaceutical investment[11].

For drug developers, the findings suggest that potential sales should be weighed against the likelihood, duration, and costs of reaching approval. A drug with substantial commercial potential may also face considerable uncertainty, whereas a more familiar structure may offer a smaller but more predictable return. Policymakers face a related distinction when considering incentives for exploratory and follow-on research. Our estimates describe patterns in private financial value; they do not establish whether industry investment is socially optimal or whether one type of drug should receive greater policy support.

Several limitations affect these interpretations. This observational study cannot determine whether novelty itself causes differences in revenue or approval. The estimated financial values use US revenue and clinical-trial costs but exclude discovery, preclinical development, manufacturing, marketing, and other expenses, so they are not complete estimates of manufacturer profit. MEPS and SSR Health cover different groups of marketed drugs, and reconstructing unidentified MEPS drugs requires assumptions about survey inclusion and spending. The

preferred weighting method performed worse than simpler approaches in a historical validation exercise, limiting confidence in the precise dollar differences. The cohort also represents one period of small-molecule development, and uncertainty around the full financial-value estimates has not been quantified.

In conclusion, evaluating drug innovation at the start of clinical development provides a different picture from comparing sales among approved medicines. Structural novelty was associated with higher expected financial value despite greater development risk, while target novelty showed a smaller and less certain association. These findings indicate that the financial rewards to drug innovation depend on what is novel, the risks required to bring it to market, and the period used to value future sales. This decision-stage perspective can inform how drug developers and policymakers think about incentives for different approaches to pharmaceutical research.

References

1. Krieger, J., Li, D. & Papanikolaou, D. Missing novelty in drug development. Rev. Financ. Stud. 35, 636–679 (2022). https://doi.org/10.1093/rfs/hhab024
2. Jiao, B., Haworth, I. S. & Lakdawalla, D. N. Highly novel drugs outperformed less novel drugs in gross and net revenues in the US, driven primarily by utilization, 2013–19. Health Aff. (Millwood) 45, 846–854 (2026). https://doi.org/10.1377/hlthaff.2026.00054
3. Hay, M., Thomas, D. W., Craighead, J. L., Economides, C. & Rosenthal, J. Clinical development success rates for investigational drugs. Nat. Biotechnol. 32, 40–51 (2014). https://doi.org/10.1038/nbt.2786
4. Wong, C. H., Siah, K. W. & Lo, A. W. Estimation of clinical trial success rates and related parameters. Biostatistics 20, 273–286 (2019). https://doi.org/10.1093/biostatistics/kxx069
5. Paul, S. M. et al. How to improve R&D productivity: the pharmaceutical industry's grand challenge. Nat. Rev. Drug Discov. 9, 203–214 (2010). https://doi.org/10.1038/nrd3078
6. Scannell, J. W., Blanckley, A., Boldon, H. & Warrington, B. Diagnosing the decline in pharmaceutical R&D efficiency. Nat. Rev. Drug Discov. 11, 191–200 (2012). https://doi.org/10.1038/nrd3681
7. DiMasi, J. A., Grabowski, H. G. & Hansen, R. W. Innovation in the pharmaceutical industry: new estimates of R&D costs. J. Health Econ. 47, 20–33 (2016). https://doi.org/10.1016/j.jhealeco.2016.01.012
8. Grabowski, H. G. & Vernon, J. The distribution of sales revenues from pharmaceutical innovation. PharmacoEconomics 18 (Suppl. 1), 21–32 (2000). https://doi.org/10.2165/00019053-200018001-00005
9. Grabowski, H. G. & Vernon, J. A new look at the returns and risks to pharmaceutical R&D. Manage. Sci. 36, 804–821 (1990). https://doi.org/10.1287/mnsc.36.7.804
10. Pammolli, F., Magazzini, L. & Riccaboni, M. The productivity crisis in pharmaceutical R&D. Nat. Rev. Drug Discov. 10, 428–438 (2011). https://doi.org/10.1038/nrd3405
11. Budish, E., Roin, B. N. & Williams, H. Do firms underinvest in long-term research? Evidence from cancer clinical trials. Am. Econ. Rev. 105, 2044–2085 (2015). https://doi.org/10.1257/aer.20131176
12. Dubois, P., de Mouzon, O., Scott Morton, F. & Seabright, P. Market size and pharmaceutical innovation. RAND J. Econ. 46, 844–871 (2015). https://doi.org/10.1111/1756-2171.12113
13. Acemoglu, D. & Linn, J. Market size in innovation: theory and evidence from the pharmaceutical industry. Q. J. Econ. 119, 1049–1090 (2004). https://doi.org/10.1162/0033553041502144
14. Blume-Kohout, M. E. & Sood, N. Market size and innovation: effects of Medicare Part D on pharmaceutical research and development. J. Public Econ. 97, 327–336 (2013). https://doi.org/10.1016/j.jpubeco.2012.10.003
15. Finkelstein, A. Static and dynamic effects of health policy: evidence from the vaccine industry. Q. J. Econ. 119, 527–564 (2004). https://doi.org/10.1162/0033553041382166

16. Scherer, F. M. The link between gross profitability and pharmaceutical R&D spending. Health Aff. 20, 216–220 (2001). https://doi.org/10.1377/hlthaff.20.5.216
17. Caves, R. E., Whinston, M. D. & Hurwitz, M. A. Patent expiration, entry, and competition in the U.S. pharmaceutical industry. Brookings Pap. Econ. Act. Microecon. 1991, 1–66 (1991).
18. Hemphill, C. S. & Sampat, B. N. Evergreening, patent challenges, and effective market life in pharmaceuticals. J. Health Econ. 31, 327–339 (2012). https://doi.org/10.1016/j.jhealeco.2012.01.004
19. Kesselheim, A. S., Avorn, J. & Sarpatwari, A. The high cost of prescription drugs in the United States: origins and prospects for reform. JAMA 316, 858–871 (2016). https://doi.org/10.1001/jama.2016.11237
20. Grabowski, H. G., DiMasi, J. A. & Long, G. The roles of patents and research and development incentives in biopharmaceutical innovation. Health Aff. 34, 302–310 (2015). https://doi.org/10.1377/hlthaff.2014.1047
21. Kogan, L., Papanikolaou, D., Seru, A. & Stoffman, N. Technological innovation, resource allocation, and growth. Q. J. Econ. 132, 665–712 (2017). https://doi.org/10.1093/qje/qjw040
22. Abdus, S., Hill, S. C. & Ahrnsbrak, R. Outpatient prescription drugs: data collection and editing in the 2021 Medical Expenditure Panel Survey. Methodology Report No. 37 (Agency for Healthcare Research and Quality, 2024). https://meps.ahrq.gov/data_files/publications/mr37/mr37.shtml
23. Rogers, D. & Hahn, M. Extended-connectivity fingerprints. J. Chem. Inf. Model. 50, 742–754 (2010). https://doi.org/10.1021/ci100050t
24. Gaulton, A. et al. The ChEMBL database in 2017. Nucleic Acids Res. 45, D945–D954 (2017). https://doi.org/10.1093/nar/gkw1074
25. Wishart, D. S. et al. DrugBank 5.0: a major update to the DrugBank database for 2018. Nucleic Acids Res. 46, D1074–D1082 (2018). https://doi.org/10.1093/nar/gkx1037
26. Willighagen, E. L. et al. The Chemistry Development Kit (CDK) v2.0: atom typing, depiction, molecular formulas, and substructure searching. J. Cheminform. 9, 33 (2017). https://doi.org/10.1186/s13321-017-0220-4
27. WHO Collaborating Centre for Drug Statistics Methodology. ATC classification index with DDDs, 2026 (WHO Collaborating Centre for Drug Statistics Methodology, Oslo, 2026). https://atcddd.fhi.no/atc_ddd_index_and_guidelines/atc_ddd_index/
28. Lipinski, C. A., Lombardo, F., Dominy, B. W. & Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv. Drug Deliv. Rev. 23, 3–25 (1997). https://doi.org/10.1016/S0169-409X(96)00423-1
29. Veber, D. F. et al. Molecular properties that influence the oral bioavailability of drug candidates. J. Med. Chem. 45, 2615–2623 (2002). https://doi.org/10.1021/jm020017n
30. Sertkaya, A., Beleche, T., Jessup, A. & Sommers, B. D. Costs of drug development and research and development intensity in the US, 2000–2018. JAMA Netw. Open 7, e2415445 (2024). https://doi.org/10.1001/jamanetworkopen.2024.15445
31. Ministère chargé de la Santé (France). ASMR: amélioration du service médical rendu. https://sante.gouv.fr/soins-et-maladies/medicaments/glossaire/article/asmr (accessed 2 November 2026).
32. DiMasi, J. A. & Faden, L. B. Competitiveness in follow-on drug R&D: a race or imitation? Nat. Rev. Drug Discov. 10, 23–27 (2011). https://doi.org/10.1038/nrd3296

Figure legends

Fig. 1 | Expected financial value at Phase I entry by novelty group. Chemical similarity to previously approved drugs is shown across all therapeutic classes (a) and within therapeutic class (b); target novelty is measured by the number of earlier approved drugs acting on the same targets across classes (c) and within class (d). Lines show estimates based on revenues from 2007–2019, 2007–2012, 2013–2019, and 2015–2019. The primary revenue model supplements identifiable MEPS observations with estimates for potentially unidentified drug-years based on SSR Health (Methods). Values are point estimates in 2020 US$ millions per drug candidate at Phase I entry; the same development cohort is used in every period.

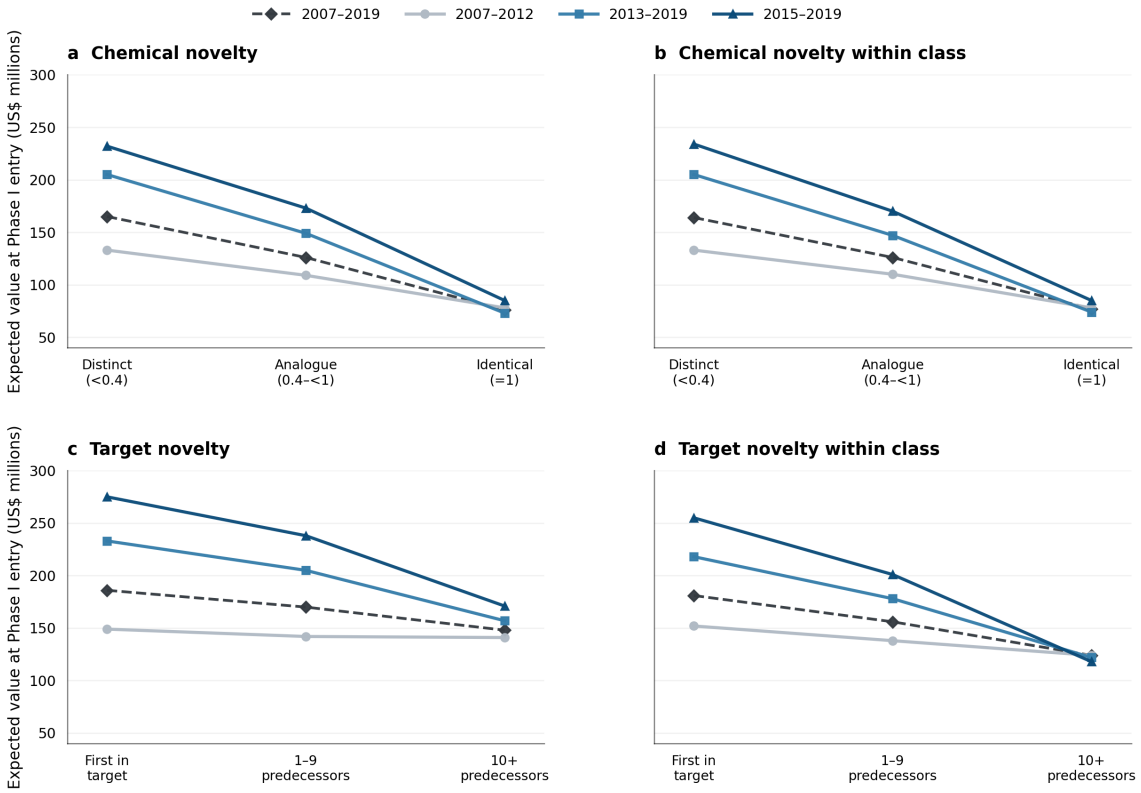


Fig. 2 | Revenue and clinical development by chemical novelty. The panels compare structurally distinct drug candidates with candidates based on a previously approved structure. (a) Predicted annual US net revenue conditional on approval. (b) Estimated probability of US approval. (c) Discount factor reflecting time from Phase I entry to approval. Revenue estimates in panel a come from the identifiable-MEPS model rather than the supplemented primary model, and the gross-revenue ratios in the text refer to a different measure of revenue. The complete valuation also accounts for rebates and expected clinical-trial costs.

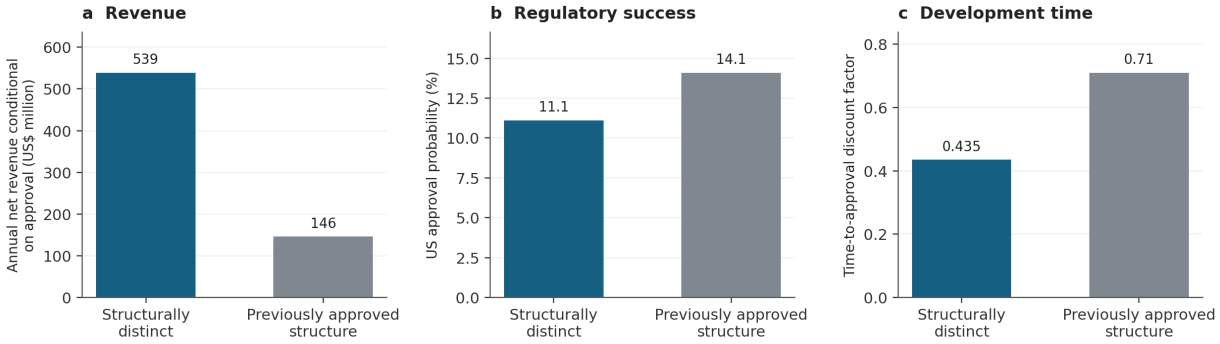


Fig. 3 | Delivery-related physicochemical properties and expected financial value. The index counts departures from five study-defined molecular-property ranges and groups drug candidates by zero or one, two or three, or four or five departures. Lines show estimates based on 2007–2019, 2007–2012, 2013–2019, and 2015–2019 revenue. Values are point estimates in 2020 US$ millions per drug candidate at Phase I entry. The index does not measure delivery technology or novelty relative to earlier approved drugs.

