Contribution of NIH funding to new drug approvals 2010–2016

Ekaterina Galkina Cleary*, Jennifer M. Beierleinab, Navleen Surjit Khanuja, Laura M. McNameea,b, and Fred D. Ledleya,b,c,1

*Center for Integration of Science and Industry, Bentley University, Waltham, MA 02452; bDepartment of Natural and Applied Sciences, Bentley University, Waltham, MA 02452; and cDepartment of Management, Bentley University, Waltham, MA 02452

Edited by Solomon H. Snyder, Johns Hopkins University School of Medicine, Baltimore, MD, and approved December 27, 2017 (received for review September 1, 2017)

This work examines the contribution of NIH funding to published research associated with 210 new molecular entities (NMEs) approved by the Food and Drug Administration from 2010–2016. We identified >2 million publications in PubMed related to the 210 NMEs (n = 131,092) or their 151 known biological targets (n = 1,966,281). Of these, >600,000 (29%) were associated with NIH-funded projects in RePORTER. This funding included >200,000 fiscal years of NIH project support (1985–2016) and project costs >$100 billion (2000–2016), representing ~20% of the NIH budget over this period. NIH funding contributed to every one of the NMEs approved from 2010–2016 and was focused primarily on the drug targets rather than on the NMEs themselves. There were 84 first-in-class products approved in this interval, associated with >$64 billion of NIH-funded projects. The percentage of fiscal years of project funding identified through target searches, but not drug searches, was greater for NMEs discovered through targeted screening than through phenotypic methods (95% versus 82%). For targeted NMEs, funding related to targets preceded funding related to the NMEs, consistent with the expectation that basic research provides validated targets for targeted screening. This analysis, which captures basic research on biological targets as well as applied research on NMEs, suggests that the NIH contribution to research associated with new drug approvals is greater than previously appreciated and highlights the risk of reducing federal funding for basic biomedical research.

NIH funding | drug development | basic science | translational science

Ongoing public debate concerning budget allocations for the NIH and the propriety of pharmaceutical pricing have raised questions about the roles of the public and private sectors in drug discovery and development. The classic linear model of innovation in drug development posits that basic research, or use-inspired basic research (1), provides a scientific foundation for drug discovery by elucidating mechanisms of disease and strategies for therapy, validating drug targets, and, sometimes, identifying prototype compounds (2). This research is funded largely by the public sector, primarily by government (3), and is performed principally in academic institutions or government laboratories. The insights and intellectual property arising from this basic research are then transferred to the private sector for development. Biopharmaceutical companies are responsible for conducting applied preclinical research and clinical research, obtaining regulatory approval, and establishing the manufacturing, control, distribution, and marketing required to commercialize a new molecular entity (NME). This development is funded primarily from the profits generated by earlier products as well as by capital investments. This simplified model does not account for dynamic interactions that occur at the boundary between the academic and commercial sectors (4), contributions made to use-inspired basic research by the biotechnology and pharmaceutical industries (5–7), or efforts to promote translational science in the public domain (2, 8, 9). Nevertheless, this model is commonly invoked by scientists and policy makers alike to justify government support for basic biomedical research.

The timelines and costs involved in the commercial development of an NME have been extensively characterized (10–15). Recent data suggest that companies invest an average of $1.4 billion in out-of-pocket expenses for each NME launched, with the total cost of capital exceeding $2.5 billion (10).

Quantifying the contribution of public-sector funding to the emergence of new drugs is less well characterized. Stevens et al. (16) assessed how many NMEs arise from public-sector research institutions by identifying patents licensed by biopharmaceutical companies from academic institutions. Their analysis showed that 9.3% of NMEs approved from 1990–2007 were patented by public-sector institutions and subsequently licensed to commercial entities for development. These results are consistent with studies examining patents cited in the Food and Drug Administration’s (FDA’s) Orange Book (17), which suggest that 7.6% of drugs approved 1981–1990 (18) and 6.7% of new drugs approved 1990–1999 (11) originated in academia.

Using an analogous approach, Sampat (19) has estimated that 7.7% of all FDA approvals and 10.6% of NMEs are based on academic patents. Expanding on this analysis, Sampat and Lichtenberg (20) explored the indirect contributions of public-sector funding to drug patents by characterizing the prior art referenced in these patents. Their studies show that, of 379 drugs approved from 1988–2007, 48% were associated with a patent that cited prior art generated in the public sector. Kneller (21) also examined this indirect contribution by examining patents that described novel biological targets and prototype compounds, concluding that nonprofit research organizations made major contribution to the patent estate for 14% of NMEs approved from

Significance

This report shows that NIH funding contributed to published research associated with every one of the 210 new drugs approved by the Food and Drug Administration from 2010–2016. Collectively, this research involved >200,000 years of grant funding totaling more than $100 billion. The analysis shows that >90% of this funding represents basic research related to the biological targets for drug action rather than the drugs themselves. The role of NIH funding thus complements industry research and development, which focuses predominantly on applied research. This work underscores the breadth and significance of public investment in the development of new therapeutics and the risk that reduced research funding would slow the pipeline for treating morbid disease.


The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

This open access article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

1To whom correspondence should be addressed. Email: fledley@bentley.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1715368115/-/DCSupplemental.

Published online February 12, 2018.
1998–2007 and some contribution to 35% of NMEs. This fraction is consistent with the results of Patridge et al. (22), who found that for 38% of FDA-approved NMEs the first synthesis or purification of the molecular entity was reported from an academic institution. Categorically, analysis suggests the public-sector contribution to NME discovery and development may be even higher. Cockburn and Henderson (23) examined the development of 21 drugs that had the “most impact” on practice from 1965–1992, observing that 76% were associated with some input from the public sector. Chakravarty et al. (6) studied the discovery and development of the 19 “most transformative” drugs of the past 25 y, concluding that the public sector contributed to the basic science underlying 54% of these NMEs but contributed directly to the discovery of only 15%. Similar results were observed in the overlap of NMEs by Zych (24).

We have previously explored the relationship between the advance of biomedical research and the emergence of NMEs from this research (24, 25). We used an analytical model for the growth and maturation of research that quantifies technology growth based on the rate of accumulation of publications in PubMed. These studies show that the growth of research on novel drug targets follows a characteristic S-curve pattern, with a point of initiation leading to a phase of exponential growth, which slows as the technology becomes established. Studies on more than 400 NMEs show that few NMEs discovered through targeted screening or biological products are approved when the underlying research is in the exponential growth phase and that NME approval occurs on average of 14 y after the established point (26–28). These data are consistent with the expectation that targeted drug discovery, including biologicals, is enabled by a body of basic research, which identifies and validates drug targets as well as potential mechanisms for therapeutic action (29, 30). In contrast, our analysis shows no relationship between metrics of technology growth and approval of NMEs discovered through phenotypic methods (26, 28), consistent with the expectation that phenotypic discovery is not predicated on knowledge of drug targets or mechanisms of action (31–34).

In the present study, we examined the scope of NIH support for published research associated with the 210 NMEs approved by the FDA from 2010–2016. Specifically, we used PubMed to identify publications related to each of these NMEs as well as the 151 known molecular targets for these drugs. We then used the NIH RePORTER database to identify publications that cited NIH funding, the core projects (i.e., grants) that supported this research, and the fiscal-year costs of those projects. We specifically focused on the 84 first-in-class drugs approved from 2010–2016, identifying publications, projects, and project costs leading to the first approval of these innovative therapeutics.

We identified NIH-funded research associated with every one of the 210 NMEs approved from 2010–2016, most of which was focused on the biological targets rather than on the drugs themselves. This research comprises over 2 million publications and was supported by more than 200,000 fiscal years of federal (primarily NIH) funding totaling more than $100 billion. These results demonstrate the scale of basic research involved in bringing a novel product to market and the magnitude of public sector support for this research.

### Methods

NMEs approved by the FDA from 2010–2016 were identified from FDA reports (35) and designated “first in class” or “follow-on” based on assessment by the FDA (36). NMEs were designated “phenotypic” or “targeted” based on the criteria of Swinney et al. (29, 30). Known molecular targets for each NME and approved clinical indications were determined from FDA labels (37) and other sources as described in SI Methods. For biological products comprising a naturally occurring protein, the target is considered to be the normal counterpart of the biological product.

PubMed searches were performed for each drug (“drug search”) using an ontology of drug name synonyms in ChEMBL (38) and the National Center for Biotechnology Information (NCBI) Query Translation. PubMed searches for molecular targets (“target searches”) were performed using Boolean search terms and NCBI Query Translation. The PubMed identifier (PMID) was recorded for each publication identified in the search.

Data associating publications with specific NIH-funded projects were obtained from the RePORTER/ExPORTER format files catalog (39). The “Link Tables for Project to Publication Associations” (hereafter, “Link Table”) associates PMIDs from 1980–present with projects that provided research funding and the PMID may be associated with multiple publications identified in the search project number and year in the Link Table. The Project Data Table provides the fiscal year cost for each project (2000–present). Costs were assigned for each funding year corresponding to the program cost in the year associated with the PMID in the Link Table. For publications with dates 1–4 y after the end of the project, costs for the final year of the project were used. The activity code associated with the core project number indicates the grant type.

We identified funding and funding years associated with a publication if every PMID associated with that project searched was associated with the same grant type. We categorized projects in the RePORTER database (Table 1). This includes 17% of the publications identified from target searches and 30% of the publications identified from target searches. These fractions are consistent with data showing that 29–35% of entries in PubMed originate from US institutions (3, 41, 42). Significantly, 94% of all publications and 96% of the publications associated with NIH funding were identified through target searches rather than drug searches. Data from individual drug and target searches are described in Tables S2 and S3.

NIH-supported PMIDs were identified from the 210 drug searches and in all 151 target searches. NIH funding was directly or indirectly associated with every one of the 210 NMEs approved from 2010–2016. We identified 221,891 funding years associated with the corpus of published research. Of these, 14,292 funding years were associated with publications identified in one or more drug searches and represent applied research related to the NME. The other 207,599 funding years were associated with publications identified...
in target searches but not drug searches (target only). These publications and funding years represent more basic research, which contributes to the body of knowledge related to the drug target without explicit reference to the NME (Table 1).

Analysis of the grant activity codes showed that the largest fraction of grants were R01 research project grants (53%), followed by T32 training programs (6%), P01 research program projects (4%), and ZIA intramural research programs (4%) (Fig. S4 and Table S3).

The time course of publications and funding years is shown in Fig. 1. Publications identified through target searches increased through the 1970s and then slowed after 2010. Publications identified through drug searches showed little growth through the 2000s and then accelerated after 2007 (Fig. 1A). The growth of NIH-funded publications exhibited a parallel pattern (Fig. 1A and B). The time course of funding years mirrors the growth of NIH-funded publications, with the number of funding years related to molecular targets increasing from the 1980s and the number of funding years directly related to the NMEs increasing only in the late 2000s (Fig. 1C).

We associated the cost of one fiscal year of project funding with 160,309 funding years. The inflation-adjusted costs for these funding years totaled $115.3 billion, with $12.5 billion for funding years associated with publications related to the NMEs and $102.8 billion for funding years associated with publications on molecular targets (Table 1). The average cost of each funding year associated with publications identified in a drug search was significantly greater than the cost of the funding years associated with publications identified in target-only searches ($1.20 million versus $684,000). The time course of costs paralleled the growth of publications and funding years, with the exception of 2009–2010 when the costs were significantly higher (Fig. 1D). This discrepancy is consistent with the increased NIH funding provided by the American Recovery and Reinvestment Act (ARRA), which provided a supplemental allocation of $10.4 billion from 2009–2010, of which $8.2 billion was invested in research (43).

Therapeutic Areas. The largest fraction of NMEs approved from 2010–2016 were antineoplastic agents, followed by antiinfectives (primarily for HIV and hepatitis) and metabolic, cardiovascular, immunologic, and central nervous system therapies. Costs associated with NMEs for major therapeutic indications are shown in Fig. S5. The number of NMEs in each therapeutic area was correlated with both drug costs and target costs ($r^2 > 0.86, d.f. = 9$). For each indication, costs associated with research on molecular targets constituted >85% of the total costs.

NIH Contribution to First-in-Class NMEs. The 210 NMEs approved from 2010–2016 include 84 first-in-class NMEs, defined as NMEs that work through a novel mechanism of action or molecular target. These 84 first-in-class NMEs are associated with 77 unique molecular targets. To assess the NIH funding that contributed to the emergence of first-in-class compounds, we clustered PMID, funding year, and cost data for each of these 77 targets with data for each of the drugs associated with these targets. Clusters were characterized as “first in class” if the first-in-class NME against that target was discovered using targeted screening methods (including biological products) or phenotypic methods (29, 30). There were 67 clusters with first-in-class NMEs discovered by targeted screening and 10 clusters with first-in-class NMEs discovered through phenotypic methods. The total number of funding years associated with these clusters is shown in Table 2.

The time course of funding years leading to approval of a first-in-class NME discovered through targeted or phenotypic methods is shown in Fig. 24. The relative proportion of funding years directly related to targets is shown in Fig. 2B for products discovered through targeted screening and phenotypic methods. These data show that for first-in-class products, a larger fraction of the funding years was associated with target only and that this fraction dropped in the years immediately before approval. Overall, for first-in-class NMEs discovered through targeted screening, 95% of

---

Fig. 1. PMIDs, NIH-funded PMIDs, NIH funding year, and costs associated with 210 NMEs approved from 2010–2016 or the 151 known molecular targets for these NMEs. (A) PMIDs identified searching for NMEs (drug search) or their molecular targets (target search). (B) PMIDs associated with NIH funding in RePORTER (1980–present). (C) Funding years associated with NIH-funded PMIDs directly related to NMEs (drug) or their targets (target only). (D) Project costs (2000–2016) associated with funding years (all), costs directly related to the NMEs (drug), or costs related to the molecular target only (target only). Open circles indicate years of supplemental funding from the ARRA. Dashed line shows trend without ARRA data. Shaded areas show the years of drug approvals.

Galkina Cleary et al.
Table 2. Funding years contributing to first-in-class products discovered through targeted or phenotypic methods

<table>
<thead>
<tr>
<th>Metrics</th>
<th>Drug</th>
<th>Target only</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Funding years</td>
<td>12,840</td>
<td>184,015</td>
<td>196,855</td>
</tr>
<tr>
<td>Targeted</td>
<td>9,170</td>
<td>166,970</td>
<td>176,140</td>
</tr>
<tr>
<td>Phenotypic</td>
<td>3,670</td>
<td>17,045</td>
<td>20,715</td>
</tr>
</tbody>
</table>

Funding years associated with multiple NMEs are not unique.

Discussion

This study provides a perspective on the scale of the public-sector contribution to the discovery and development of new drugs. We identified NIH-funded publications and projects directly related to all the 210 NMEs approved by the FDA from 2010–2016 or their molecular targets. This research comprised more than 200,000 funding years and project costs totaling more than $100 billion.

This analysis paints a more expansive picture of the public sector’s contribution to new drug discovery and development than previous studies. Previous studies showed that 6–10% of NMEs were first patented by the public sector and academic institutions (11, 16, 18–20), that as many as half of the patents on new drugs cite prior art produced in the public sector (20, 21), and that up to 40% of the new molecular entities were first synthesized or purified in academic institutions (22). Case studies have identified a public-sector contribution to the basic or applied science underlying 50–75% of new drugs (5, 6, 23).

The differences in our findings likely relate to the use of a method designed to capture the contribution of basic research or use-inspired basic research as well as applied research. Both the NIH and National Science Foundation (NSF) define basic research as being undertaken “without specific applications towards processes or products in mind” (44, 45). Our analysis focuses on published research, which is the primary deliverable for NIH-funded basic research.

In contrast, studies based on drug patents are explicitly limited to contributions that meet legal standards for patentability. In the United States, these standards require inventors to establish that their inventions are both “new and useful” according to 35 USC 101 (46, 47). In the European Union, European Patent Office standards...
require inventors to establish that their invention is “susceptible of industrial application” according to EPC Article 52 and Article 57 (48). Thus, while there is often considerable latitude in the interpretation of these standards, patent-based analyses are implicitly biased in favor of applied research with specific processes or product applications. Studies have shown, in fact, that patent analysis is a poor measure of knowledge flow from public institutions or funding (49, 50). Case-study methods are more likely to recognize the contributions of basic science or “fundamental scientific knowledge” (23) but may be similarly biased in favor of research that is linked to a specific process or product. Thus, the methods used in previous studies may have systematically underestimated the contribution of basic research to new drug approvals.

In underestimating the scope of basic research, previous studies may have also underestimated the contribution of public funding, specifically work funding from the NIH. The NIH has traditionally prioritized funding for investigator-initiated basic research (51, 52) and continues to direct more than half of its funding to basic, as opposed to applied, research (53). The finding that more than 90% of the publications citing NIH support appeared in PubMed searches for the targets of newly approved drugs, but not in searches for the drugs themselves, suggests that this research was not focused on specific processes or products to the extent that this association would be identified in the title, abstract, or PubMed medical subject headings (MeSH) terms. Thus, most of these publications likely represent basic research. This work does not attempt to identify publications that describe seminal inventions or milestones in drug discovery or development. Rather, this work posits that the corpus of published research, taken as a whole, reflects the advancing forefront of knowledge from which new drugs are discovered and developed. This approach builds on an increasing body of scholarship that views innovation not as the end product of a discrete sequence of insights or inventions but as an outgrowth of extensive knowledge networks, social networks, and intersecting communities in an ecosystem that both generates and applies new knowledge (4, 54, 55). Consistent with this view, our previous work has shown that increasingly the growth in the accumulation of publications are strongly associated with successful development of targeted and biological NMEs (26, 28).

These data demonstrate that a sizable public-sector investment occurs before the approval of first-in-class NMEs, particularly those discovered using targeted discovery methods (including recombinant biologicals). The scale of this investment can be estimated from the costs associated with first-in-class NMEs approved in 2010–2016 and their molecular targets (Table 3). These data suggest that the public-sector investment in research underlying each first-in-class drug is as high as $839 million, with 89% of this cost associated with target research and 11% of the cost associated with the first-in-class compound or follow-on compounds approved from 2010–2016.

This estimate of public-sector contribution to first-in-class products should not be interpreted as a per-drug investment. Since most of the public-sector funding we identified is associated with research on molecular targets, rather than specific NMEs, this funding may contribute not only to the first-in-class NME approved from 2010–2016 but also to follow-on products in the same class. It is also likely that basic research on drug targets has spillover effects that could lead to new classes of products that are not yet anticipated as well as to new diagnostics, devices, or approaches to disease management.

Overall, this analysis suggests that as much as 20% of the NIH budget allocation from 2000–2016 was associated with published research that directly or indirectly contributed to NMEs approved from 2010–2016. This fraction does not include research that contributed to drugs approved before 2010 or drugs currently in development that may emerge in years to come. Thus, while the NIH may continue to emphasize basic or use-inspired basic research arising from investigator-driven initiatives (56), this analysis shows that a substantial fraction of NIH spending is contributing directly or indirectly to new therapies for disease.

There are several significant limitations to the methods utilized in this analysis. First, PubMed searches identify papers by the content of the title, abstract, and metadata, including MeSH terms, but may fail to identify research published before its relevance to a specific target or NME is recognized. Our method also does not identify research on enabling technologies, such as new analytical or genomic technologies, with broad applications in biomedical research or research focused generally on mechanisms or outcomes of disease. Second, we encountered numerous mismatches between the years inferred from the Link database in RePORTER and the actual fiscal years of funded projects. Related problems have been noted by others who describe false-positive matches between publications and patents as well as false negatives in which grant information was not indexed within articles (42). Many mismatches involved publications with dates after the last year of grant funding. Based on the observations of Boyack and Jordan (42), who estimated a 3- to 4-year lag between publication dates and entries in RePORTER, we associated the costs associated with the final year of the project with research published 1–4 years after that point. With NIH data, we were able to associate costs with 86.3% of funding years from 2000–2016. Possible explanations for mismatches between the other 13.7% of funding years and fiscal years of projects include errors in acknowledging specific projects or errors in data curation. These experimental limitations would tend to underestimate the magnitude of NIH funding for research related these products and should be considered lower bounds (42). Third, this analysis does not account for research funding from government agencies, such as the Department of Defense or NSF, or other nations, which are not represented in the RePORTER database, and does not include research funding from non-governmental, nonindustrial medical and health research and development (R&D) institutions or philanthropies. A recent report estimated that 30% of nonindustrial medical and health research and development (R&D) was funded by universities, research institutes, or foundations (57). Thus, these results do not reflect the full public-sector contribution to research underlying NME approvals over this interval but predominantly reflect the contribution made by NIH funding.

This work demonstrates that NIH funding was associated directly or indirectly with every drug approved from 2010–2016 and suggests that the scale of this contribution is larger than generally appreciated. This contribution primarily involved funding for research on enabling technologies, such as new analytical or genomic targets or NME is recognized. Our method also does not identify research published before its relevance to a specific target or NME is recognized. This work underscores the breadth and significance of public investment in the development of new therapeutics and the risk that reduced research funding would slow the pipeline for treating diseases.
The science community has long recognized the importance of basic research as the engine of innovation and development. We are currently witnessing the emergence of ex- tending the phenomenon. A PubMed-based assessment, F1000 Res 3:292.


