Increasing disparities between resource inputs and outcomes, as measured by certain health deliverables, in biomedical research

Anthony Bowen[a] and Arturo Casadevall[a,b,1]

[a]Department of Microbiology and Immunology, Albert Einstein College of Medicine, Bronx, NY 10461; and [b]Department of Molecular Microbiology and Immunology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD 21205

Society makes substantial investments in biomedical research, searching for ways to better human health. The product of this research is principally information published in scientific journals. Continued investment in science relies on society's confidence in the accuracy, honesty, and utility of research results. A recent focus on productivity has dominated the competitive evaluation of scientists, creating incentives to maximize publication numbers, citation counts, and publications in high-impact journals. Some studies have also suggested a decreasing quality in the published literature. The efficiency of society's investments in biomedical research, in terms of improved health outcomes, has not been studied. We show that biomedical research outcomes over the last five decades, as estimated by both life expectancy and New Molecular Entities approved by the Food and Drug Administration, have remained relatively constant despite rising resource inputs and scientific knowledge. Research investments by the National Institutes of Health over this time correlate with publication and author numbers but not with the numerical development of novel therapeutics. We consider several possibilities for the growing input-outcome disparity including the prior elimination of easier research questions, increasing specialization, overreliance on reductionism, a disproportionate emphasis on scientific outputs, and other negative pressures on the scientific enterprise. Monitoring the efficiency of research investments in producing positive societal outcomes may be a useful mechanism for weighing the efficacy of reforms to the scientific enterprise. Understanding the causes of the increasing input-outcome disparity in biomedical research may improve society's confidence in science and provide support for growing future research investments.

biomedical research | health outcomes | innovation | research efficiency

Few forces have transformed society more than science. Knowledge gained through the scientific method has lifted billions out of poverty, fueled industrialization and mass communication, eradicated smallpox, and placed human footprints on the moon. During the 20th century, science was increasingly funded by governments and corporations vying for military and economic advantage. Furthermore, the realization that investments in biomedical research translated into medical advances garnered strong societal support for the expenditure of public funds to support science. Today, science is a vast industry producing new knowledge, usually to address particular problems or questions facing humankind. Like any industry, the scientific enterprise uses tools and resources—scientists, money, and time—to produce an output: scientific knowledge, which can be represented by publications in the scientific literature.

To remain competitive, corporations have long sought to maximize production efficiency, defined here as the ratio of output to input, by trimming waste and producing more (1). With increasing competition for research grants and jobs, funders and employers have turned to measures of efficiency and productivity to evaluate scientists (2). Such pressures have created well-documented incentives for scientists to maximize apparent productivity through

publication numbers, citation counts, and publishing in high-impact journals (3, 4). Although this approach is designed to reward those who contribute most to the knowledge base, recent studies have raised questions about the quality of the biomedical literature (5–8). One such study found that only 11% of findings could be confirmed in 53 "landmark" hematology and oncology publications (6). Another study found that 43 of 67 published cardiovascular, oncological, and women's health findings could not be reproduced (7). Recently, it was estimated that more than $28 billion is spent each year in the United States on irreproducible preclinical research and that the prevalence of these studies in the literature exceeds 50% (8). We note that this lack of reproducibility was identified by the National Institutes of Health (NIH) as a major problem and has led to initiatives for enhancing reproducibility and transparency (9). The irreproducibility of published data could potentially waste limited funding, years of work, and threatens to undermine public confidence in the scientific enterprise.

A pure focus on scientific outputs can ignore the quality of those outputs. This is partially due to the fact that the quality and importance of scientific publications, unlike most human products, is subjective and difficult to assess in the present (10, 11). The true measure of a study's quality would involve time-consuming and costly independent replication and an analysis of the work's outcomes, including assessing its downstream utility to other applications and its effects on society. However, given the length of time needed and low probability of any given study

---

Significance

The general public funds the vast majority of biomedical research and is also the major intended beneficiary of biomedical breakthroughs. We show that increasing research investments, resulting in an increasing knowledge base, have not yielded comparative gains in certain health outcomes over the last five decades. We demonstrate that monitoring scientific inputs, outputs, and outcomes can be used to estimate the productivity of the biomedical research enterprise and may be useful in assessing future reforms and policy changes. A wide variety of negative pressures on the scientific enterprise may be contributing to a relative slowing of biomedical therapeutic innovation. Slowed biomedical research outcomes have the potential to undermine confidence in science, with widespread implications for research funding and public health.

MEDICAL SCIENCES

generating substantial societal impact, it is very difficult to judge individual scientists or their work using simple outcome measures. Valid scientific research can have tremendous intrinsic societal value in producing information without the need for tangible outcomes. However, we also assume that all valid scientific research has some probability of generating tangible outcomes, even though such outcomes can be unpredictable and distant in both time and subject matter. For example, the theory of general relativity proposed in 1916 did not produce a useful outcome to the public until the late 20th century, when it provided a means to account for differences in clock rates under different gravitational influences, thereby enabling the global positioning systems found in a majority of modern cell phones and vehicles (12). Despite the intrinsic social value of biomedical knowledge, it is evident that public agencies, such as the NIH, support biomedical research with outcome-oriented goals, including benefitting health, preventing disease, and increasing return on investment (13). Although it is impractical to measure the outcomes generated by individual researchers, we believe that it is possible to estimate the outcomes of the biomedical research enterprise. In contrast to the production-based definition of efficiency (outputs ÷ inputs), we believe that evaluating research outcomes relative to inputs can be used to monitor the efficiency of biomedical innovation and the impact of research investments. Without studying outcomes, there is no way of knowing if society's investment in research is paying off, and there is no way of evaluating the efficacy of systemic modifications to the scientific enterprise.

Results

We used the annual number of PubMed publications to estimate total biomedical research output, whereas the number of uniquely named authors was used to estimate the number of contributing scientists. Both publication number and author number have risen exponentially since 1965 (Fig. 1A). PubMed now adds more than

Fig. 1. Changing inputs, outputs, and outcomes in biomedical research from 1965 to 2014. (A) Yearly publication number and author number. (B) Yearly received and approved NDAs and approved NMEs. (C) The annual NIH budget, inflation-adjusted to 2014 dollars. (D) Yearly inputs, outputs, and outcomes in biomedical research are represented by NIH budget, author number, publication number, and number of approved NMEs. All categories are normalized to 1.0 for their respective levels in 1965. (E) Cumulative knowledge in the form of scientific publications since 1965 is compared with cumulative NMEs since 1965. Both publications and NMEs are normalized to 1.0 for their respective levels in 1965.

1 million new biomedical publications annually, a 527% increase over 1965. The number of authors has risen even faster—809% since 1965—with more than 1.63 million individuals contributing in 2014. The pace of increasing authors is an underestimate, because many investigators share common names. Furthermore, the average number of authors per published manuscript in the PubMed database has more than doubled since 1965 (Fig. S1A). The number of retractions found in the literature has also risen rapidly since the late 1990s (Fig. S1B), as documented previously (14, 15). The number of scientific journals represented in PubMed has tripled since 1965 (Fig. S1C). The growth in US-affiliated publications and authors from 1995 to 2012 was substantial, but less than worldwide increases (Fig. S2). Furthermore, total biomedical knowledge is better approximated by the global increase in publications rather than publications affiliated with any particular nation.

The number of New Molecular Entities (NMEs) approved by the Food and Drug Administration (FDA) was used to estimate the outcomes of the biomedical research enterprise (Fig. 1B). NMEs are defined as drugs containing active moieties, as either a single ingredient or as a combination product, that have not previously been approved by the FDA. Biologic products may be defined as NMEs for the purposes of FDA review even if a closely related moiety has previously been approved (16). NMEs are an attractive outcome measure because they are unequivocal indicators of progress in global biomedical research. NME approvals grew steadily each decade from 1965 to 2004 but regressed over the last 10 y (Fig. S3 A and B). Furthermore, the magnitude of growth observed in annual NME approvals was much smaller than the changes in author or publication number. The numbers of received and approved New Drug Applications (NDAs) have remained relatively constant (Fig. 1B), suggesting that the FDA's ability to review applications is not a major limiting factor on new therapeutic approvals.

As the largest contributor to global biomedical research funds (17), the NIH budget (inflation-adjusted to 2014 dollars) was used as a proxy for total biomedical research funds (Fig. 1C). The budget rose exponentially from 1965 to 1999. Over the next 4 y, the budget doubled before a steady decrease occurred during 2003–2014, which is larger than apparent because of the rapidly rising cost of scientific experiments. The cost per NME, in millions of dollars of NIH budget, was compared for each decade from 1965 to 2014, and has grown rapidly since the mid-1980s (Fig. S3 C and D). When normalized to their 1965 values, NIH budget, author number, publication number, and number of approved NMEs enjoyed a combined 5.5-fold increase over the last 50 y (Fig. 1D). The greatest share of that growth occurred in the number of scientists (9.1-fold) and the number of publications (6.3-fold). NIH budget increased to a lesser extent (4.2-fold), whereas approved NMEs increased least of all (2.3-fold). Consequently, society's acquisition of new knowledge through biomedical publications drastically outpaced the production of novel therapeutics by the biomedical research enterprise (Fig. 1E).

Biomedical research efficiency (outcomes ÷ input) was investigated by calculating approved NMEs each year per relative units of new scientific knowledge, yearly scientific authors, and yearly NIH funding (Fig. 2 A–C). Decreasing research efficiency from 2000 to 2014 is demonstrated by the drop in approved NMEs for each unit of personnel and financial input. Despite 2014 being the most productive year for novel therapeutics since 1996, the number of NMEs developed per billion dollars of NIH funding was 56% lower in 2000–2014 compared with 1965–1999. The same drop is evident in approved NMEs for each unit of published scientific knowledge, illustrating that increasing scientific outputs have not translated into new therapeutics in a proportional manner. Cumulative NMEs since 1965 divided by cumulative biomedical knowledge remained relatively constant from 1976 to
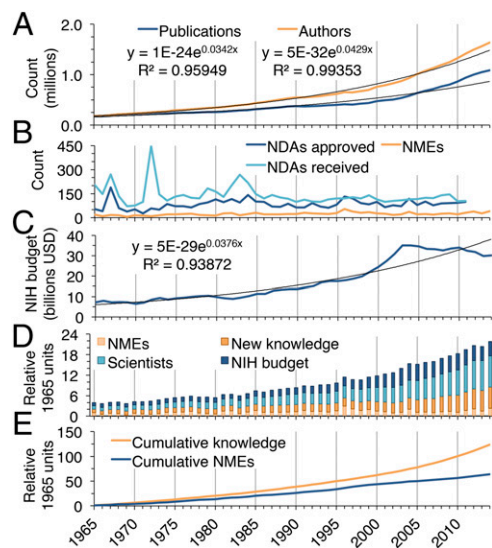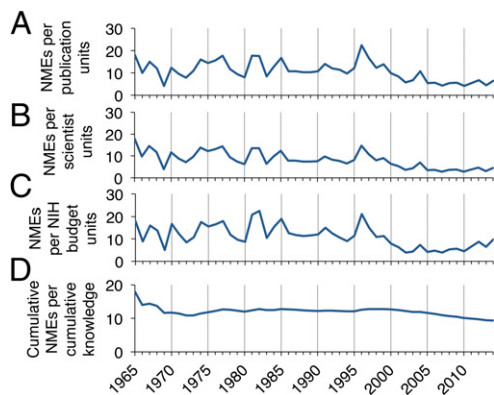
Fig. 2. Biomedical research efficiency measured by outcomes ÷ outputs and outcomes ÷ inputs. (A–C) NMEs produced each year divided by yearly units of new knowledge (A), scientific authors (B), and NIH funding (C) relative to 1965 levels. (D) The cumulative number of NMEs produced since 1965 divided by cumulative biomedical knowledge generated since 1965.

1999, followed by a marked decline through 2014 (Fig. 2D). As might be expected, the number of authors was highly correlated with the number of publications in a year (Fig. 3A). The number of NMEs produced in a given year, however, correlated very poorly with the resources and knowledge needed to develop them (Fig. 3 B and C). Finally, both author number and publication number had closer relationships to the NIH budget over the last 50 y than to NMEs (Fig. 3D).

Because discovery takes time to translate into products, we investigated whether the observed effect was due to a growing time lag between these two events. To estimate the time between discovery and product, we analyzed 12 examples of medical breakthroughs resulting from 11 key scientific discoveries during the last century (Table S1). The translation lag time between discovery and implementation was between 10 and 14 y for most discoveries and did not increase with more recent examples. Although this retrospective analysis suffers from selection bias in choosing ultimately productive scientific discoveries, it does provide an estimate for bench-to-bedside translation time. Other more systematic studies have found longer translational lag times, but did not indicate that these lags have increased (18, 19). The observed declines in research efficiency did not change even when approved NMEs per units of new scientific knowledge and yearly scientific authors were adjusted with a discovery to product translation lag of 15 y (Fig. S4 A, B, and D). Research efficiency followed an even sharper decline when the NIH budget was adjusted with a 15-y translation lag (Fig. S4C).

Despite being affected by many factors other than biomedical research, we explored life expectancy and patent applications as alternative outcome and output measures. US life expectancy at birth rose linearly from 1965 to 2012 with an average yearly growth of slightly more than 2 mo (Fig. 4A). Patent applications in the United States increased 504% since 1965, similar to the 527% growth that was observed in publication number (Fig. S5). We used these alternative outcome and output measures to track biomedical research efficiency over time. Dividing life expectancy (LE) and the yearly change in life expectancy (ΔLE) by the output measure of biomedical publication number (Fig. 4B) showed a steady decline. A similar drop in research efficiency is evident when LE and ΔLE per unit of NIH funding is followed over time (Fig. 4C). Many major medical breakthroughs in recent decades have been targeted to diseases affecting older patient populations. To determine whether life expectancy has risen faster in older populations, we examined the life expectancy at age 70 (LE_{70}) of 11 European nations from 1965 to 2012 and the European Union from 2002 to 2012. All LE_{70} measurements

increased linearly, following the pattern of US LE at birth (Fig. S6A). Furthermore, the 11 analyzed European nations experienced an average yearly increase in LE_{70} of less than 1.5 mo from 1965 to 2012 (Fig. S6B). This result indicates that life expectancy gains among older populations have also not kept pace with rising biomedical research inputs and outputs.

Discussion

The data presented above suggest an increasing disparity between resource inputs and outcomes in biomedical research. To visualize the relationship between scientific inputs, outputs, and outcomes, we constructed a scheme showing the flow of resources within the biomedical research enterprise (Fig. S7). Increasing resource investments have led to an explosion in scientific knowledge, but the resulting gains in new therapies and improved human health have been proportionally smaller, as measured by NMEs and prolongation of life expectancy. Several possible causes for slowed growth in biomedical research outcomes are considered.

First, the problems could have become more difficult through prior elimination of easier research questions or the challenge of producing therapies that are better than current standards of care (20). This possibility could account for a decreasing marginal impact of new biomedical innovations. Increasing regulatory requirements over the last half-century—generally in the interest of public safety—have undoubtedly placed larger financial and temporal burdens on the development of new therapies. Furthermore, as scientists now spend nearly half of their time on administrative tasks, they are able to devote less time to research and training responsibilities (21). Other economic factors, such as the recent proliferation of patent trolls (22), may have also contributed to a stifling of biomedical innovation. Additionally, society has seen a shift from acute to chronic illnesses, and the development of therapies for chronic illnesses may be inherently more difficult. We note that improving health outcomes in chronic patients often requires prolonged and coordinated medical care. It is possible that the
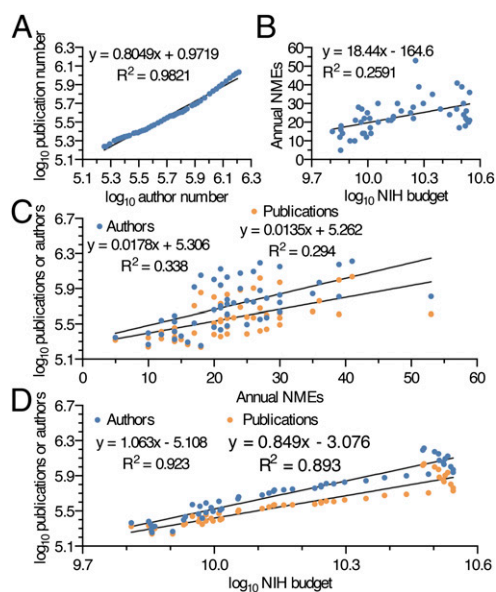

Fig. 3. Correlations between authors, NMEs, NIH budget, and publication numbers from 1965 to 2014. NIH budget, author number, and publication number are plotted in log base 10. (A) Log-log plot of publication vs. author number. (B) Plot of NMEs vs. NIH budget. (C) Plot of publication number and author number vs. NMEs. (D) Plot of publication and author number vs. the NIH budget.
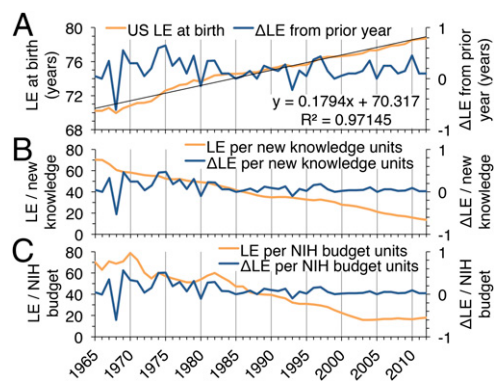
Fig. 4. Alternative measures of biomedical research efficiency from 1965 to 2012. (*A*) Mean US life expectancy at birth (LE, blue line) and the yearly change in life expectancy (ΔLE, orange line). (*B* and *C*) LE and ΔLE for each divided by that year's biomedical research publications (*B*) and NIH budget (*C*). ΔLE curves in *B* and *C* are not identical, but they appear similar at this scale due to the small differences between them.

current health care delivery system and public health policies are not optimally aligned to today's foremost medical problems.

Second, biomedical research findings may be accurate, but many have little immediate clinical relevance. In this regard, mice are widely used for biomedical discovery, but this species may not accurately predict human outcomes (23). Similarly, reliance on cell lines for basic discovery has many limitations including the facts that experiments may not reflect tissue conditions and that many cell lines are not authenticated or are contaminated (24, 25). A "brute-force bias" in pharmaceutical research and development has been attributed as a likely reason for stagnant NME approvals (20). This bias is manifested by an increasing dependence on high-throughput screening in highly simplified, automated systems for very specific molecular effects (i.e., high affinity binding to a target). A corollary to this bias is the assumption that more funding will directly translate into more drug candidates. Although such reductionist approaches have facilitated the analysis of enormous candidate drug libraries, there is evidence that techniques used decades ago may have been more effective at identifying clinically relevant molecules (20). While a focus on reductionism has greatly expanded our knowledge of biology's most basic components, the extent to which this applies to complex biological networks is not clear. Specialization in science has also grown with the expansion of the scientific knowledgebase, increasing the necessity of technical collaboration (26). The mean number of authors per PubMed publication has grown from 1.94 to 5.61 since 1965, suggesting that more collaboration is now necessary ([Fig. S1*A*](#)). We note that there is some evidence that increasing authorship may reflect inflationary practices more so than research complexity or collaboration (27). Specialization may also raise barriers to interfield collaboration and cause pertinent findings or techniques to be overlooked by scientists in disparate fields.

A third possibility is that the quality of the scientific literature has decreased. Several articles illustrate a recent surge in retracted publications and widespread irreproducibility in the literature (5–8, 14, 15). The high prevalence of irreproducible research likely reflects the use of highly specialized techniques, the incomplete publication of methods, the failure to share necessary reagents or software, poor experimental design, and the bias toward publication of positive results. A recent study of error in retracted publications suggests that many erroneous findings remain uncorrected in the literature, creating the possibility for friction in the discovery process as scientists are distracted by invalid leads and/or spend time investigating and correcting

published errors (28). Common commercial products such as research antibodies suffer from batch-to-batch variability, uncharacterized cross-reactivity, and protocol-specific changes in binding specificity (29). We also note that the documented growth of retracted and irreproducible research likely reflects a combination of increased prevalence and surveillance of problematic manuscripts and science (15). Irreproducible, incorrect, and falsified results in the literature can all contribute to inefficiencies in the process of scientific discovery. The incidence of these publications may be linked to the use of scientific outputs as metrics for the competitive evaluation of scientists (3, 4). We note that certain metrics are easier to game than others, and thus the use of less-pliable scientific output measures may lead to a decline in inappropriate incentives (30). It is also not known whether the use of such metrics might concurrently affect the creativity and work ethic of scientists. Behavioral research has indicated that task-based incentives tend to increase productivity, but with little impact on creative work (31). Finally, the current mania for publishing science in high-impact factor journals delays the publication of key findings as scientists shop for high-impact venues and may be creating perverse incentives that affect outcomes because high-impact science does not necessarily imply high importance (4).

We suspect that slowed biomedical outcomes growth is the result of some combination of the problems detailed above. Suggested remedies to these problems include reforms that could ease the regulatory hurdles for new therapeutics (11, 32), reduce administrative burdens (32), improve the clinical relevance of reductionist biomedical techniques (20), shift incentives away from scientific outputs (3, 4), encourage the publication of negative results (33), mandate public access to government-funded research results and methods (34), facilitate independent reproduction of important studies (35), and/or create a framework for measuring the outcomes of biomedical research (11, 36). A simple focus on boosting any given biomedical research metric, such as the number of approved NMEs, is not likely to solve the problem of slowed outcomes, but may instead contribute to misdirected incentives. To address the systemic problems detailed in this manuscript, it will be essential to consider each of the possible problems and their respective solutions in relation to the entire biomedical research enterprise and the culture of science. However, for policy changes to have the desired effect, they will need to be tested experimentally and that may require the development of new tools to measure the efficiency of science. Efforts to reduce the disparity between biomedical research inputs and outputs and their eventual societal outcomes will likely require a variety of reforms on the part of individual scientists, academic institutions, private corporations, and governmental entities (37, 38).

We acknowledge that any link between biomedical innovations and overall life expectancy is controversial, with many other factors—including political stability, food production, hygiene, education, income, transportation, and communication—potentially having greater impact. However, these factors are more likely to affect developing regions, and for both the United States and Europe, political stability, abundant food, adequate hygiene, and developed communication and transportation networks were available and stable during the time of our study. In this regard, long-term linear mortality reduction may be the result of increasing public health and biomedical resources combined with their decreasing marginal effectiveness (39). Other studies have attributed decreases in disease-specific mortality to certain biomedical innovations (40, 41). A key limitation of these analyses is the difficulty in appropriating patient outcomes that result from a given clinical innovation to the initial funding of many underlying basic research advances. Even in the face of sluggish therapeutic development and constant life expectancy gains, biomedical innovations over the last 50 y have transformed many individual fields, including imaging, criminology, the

treatment of some cancers, and HIV/AIDS (40, 42–44). It is also possible that the exponential growth in research investment and scientific knowledge over the previous five decades has simply not yet grown fruit and that a deluge of medical cures are right around the corner, a development that would make us extremely happy. One explanation for such a possibility is humanity's relatively recent development of the necessary technology to generate and manipulate vast quantities of data coupled with inadequate tools to analyze and make use this of data.

Our results are best interpreted as a cautionary tale that will hopefully motivate new efforts to understand the parameters that influence the efficiency of science and its ability to translate discovery into practical applications. Considering that humanity is currently facing numerous challenges such as climate change, emerging infectious diseases, environmental degradation, and a faltering green revolution, each of which requires scientific advances to provide new solutions, it is essential for science to function at its best. Science's power to influence society has historically resided in its ability to reliably predict phenomena and improve human lives. A worrisome consequence of the growing input-outcome disparity in biomedical research may be an erosion of society's trust in the scientific enterprise at a time when science and its findings are often under attack (45, 46), with wide-ranging consequences for public health and future research funding. In support of this concern we note that the number of NMEs flattened from 2004 to 2014, a decade when the NIH budget decreased in real terms. Nevertheless, we are hopeful that the scientific community can use this information to adopt measures that improve the efficiency of the scientific enterprise. The ongoing motivation of scientific research should continually refocus to mirror its historical objectives: making discoveries that expand human knowledge and generating outcomes that positively impact humanity.

Methods

PubMed Data Retrieval. The PubMed search engine can be used to access the most comprehensive aggregation of global biomedical literature, which contains more than 23 million records and is maintained by the NIH and the US National Library of Medicine (47). An algorithm was written in Perl to query PubMed for all citations per year from 1809 to 2014. All citations with "journal article" publication type were downloaded in MEDLINE format and parsed according to documented field descriptors (48). Unique journal names were determined from the "journal title" data field. Unique author names were determined from the "author" data field, which includes last name, first and middle initials, and a suffix, if applicable. Corporate or group authors were excluded from this analysis. The author count based on distinct names in the database represents a low estimate of the true number of authors because scientists that share a common name are counted as one. Additionally, PubMed limited the maximum author number in its MEDLINE records to 10 per publication from 1984 to 1995 and then to 25 per publication from 1996 to 1999, which further contributes to an undercounting of scientists (48). Yearly publication number was calculated as the number of PubMed citations with at least one individual author. Retracted publications were counted by original year of publication and also by year of retraction. Retraction data were gathered by automatically querying PubMed each year for "retracted publication" and "retraction of publication" publication types.

To gather data on publications and authors arising from the United States, we performed the same analysis as described above but added "USA" affiliation as a search constraint. Affiliation details were only recorded in PubMed citations for first authors from 1988 to 2013. The designation USA was added to author affiliations arising from the United States starting in 1995, and quality control of this field ceased during 2013 to facilitate affiliation details being added for all authors (48). Therefore, our analysis was limited to 1995–2012 and underestimates both the number of US scientists and publications with contributions from US scientists. We compared the number of USA-affiliated first authors to the global number of first authors during the same time period (Fig. S2). PubMed data presented in this manuscript was gathered on March 7, 2015 and reflects the PubMed database as of this date. PubMed's records do not comprehensively cover the biomedical literature before 1965, so this was chosen as the starting year for most of our analyses. The Perl script and raw data used in this analysis are available and annotated at https://github.com/bowenanthony/biomedical-outcomes.

NIH Investment Data. For the purposes of this study, the NIH budget was used as an estimate for the annual investment of public funds in the biomedical research enterprise. NIH budget appropriations from 1938 to 2014 were obtained (49, 50) and inflation adjusted to 2014 dollars using the annual average Consumer Price Index (51). Although the NIH is responsible for a fraction of worldwide biomedical research funds, it is the largest public sector contributor by far (17) and has an easily accessible and comprehensive budget history. Including private sector investments and other national research budgets would make year-to-year comparisons over the last five decades prohibitively difficult. Furthermore, including additional sources of private or public biomedical research funds would only increase the input-outcome disparities presented in this article because the NIH funds are just a portion of worldwide investment. It can also be argued that much of the innovation in the private sector is based on the discoveries of government-funded research, presumably because governments are able to take on greater risk than corporations (52).

Research Outcomes Data. The long-term outcomes of biomedical research on human health and the economy are difficult to quantify due to numerous confounding variables. Therefore, NMEs approved by the FDA were used to estimate the relative yearly outcomes of the biomedical research enterprise. The number of received and approved NDAs and approved NMEs from 1938 to 2014 were obtained from the FDA (53, 54). From 2004 to 2014, these data include Biologic License Applications and approved biological products. Because the United States is the largest market for drug manufacturers (55), most new therapeutics eventually make their way to the FDA's approval pipeline. We acknowledge that some NMEs, known as "me too drugs" are generated by making minor changes to existing therapeutics, creating compounds with little to no benefit over the original medication (56). Removing such "me too drugs" from our analysis, however, would be somewhat subjective and would only worsen the problem of slowed biomedical research outcomes. Although some NMEs have a much greater impact than others on health outcomes, it is very difficult to quantify the disease-specific outcomes of each NME. Additionally, a comparison of NME-related health outcomes over time would likely overvalue older therapeutics due to the time required for drugs to impact population health and also because new uses for existing drugs are often discovered over time. We recognize that much of the yearly NIH investment in biomedical research does not go directly to creating novel therapeutics. Funds spent on the development of new research tools and new uses for existing therapeutics are not represented by NMEs, but quantifying outcomes of this research is difficult. Additionally, a large part of the NIH budget funds basic research with the potential for long-term utility but no immediate impact on human health. For the purpose of this analysis we assume that all valid contributions to the cumulative scientific knowledgebase have some probability of future therapeutic utility. We also assume that the rate of NME approval is proportional to the rate of overall outcomes in biomedical science.

Alternative Measures. The products of biomedical research are embodied by more than new pharmaceuticals on the market and peer-reviewed scientific publications. To address this, we examined the average US life expectancy at birth as an outcome measure and the number of utility patent applications in the US as an output measure. Data for these measures from 1965 to 2012 were obtained from the World Bank and the US Patent and Trademark Office (57, 58). Patent application numbers reflect utility patents and exclude design patents. We also recognize that many medical advances throughout the last several decades have targeted illnesses present in older patient populations and therefore might be better represented by measuring life expectancy at age 70. Because these data were not readily available for the United States, life expectancy at age 70 was obtained from Eurostat for 11 European nations with comprehensive data from 1965 to 2012 and for the European Union from 2002 to 2012 (59).

1. Witzel M (2002) A short history of efficiency. *Bus Strateg Rev* 13(4):38–47.
2. Hirsch JE (2005) An index to quantify an individual's scientific research output. *Proc Natl Acad Sci USA* 102(46):16569–16572.
3. Brembs B, Button K, Munafò M (2013) Deep impact: Unintended consequences of journal rank. *Front Hum Neurosci* 7:291.
4. Casadevall A, Fang FC (2014) Causes for the persistence of impact factor mania. *MBio* 5(2):e00064–e14.
5. Ioannidis JPA (2005) Why most published research findings are false. *PLoS Med* 2(8):e124.
6. Begley CG, Ellis LM (2012) Drug development: Raise standards for preclinical cancer research. *Nature* 483(7391):531–533.
7. Prinz F, Schlange T, Asadullah K (2011) Believe it or not: How much can we rely on published data on potential drug targets? *Nat Rev Drug Discov* 10(9):712.
8. Freedman LP, Cockburn IM, Simcoe TS (2015) The economics of reproducibility in preclinical research. *PLoS Biol* 13(6):e1002165.
9. Collins FS, Tabak LA (2014) Policy: NIH plans to enhance reproducibility. *Nature* 505(7485):612–613.
10. Casadevall A, Fang FC (2009) Important science: It's all about the SPIN. *Infect Immun* 77(10):4177–4180.
11. Bornmann L (2012) Measuring the societal impact of research: Research is less and less assessed on scientific impact alone—We should aim to quantify the increasingly important contributions of science to society. *EMBO Rep* 13(8):673–676.
12. Ashby N (2003) Relativity in the global positioning system. *Living Rev Relativity* 6(1):1–42.
13. National Institutes of Health (2013) Mission. Available at perma.cc/H7EP-F42S. Accessed June 7, 2015.
14. Fang FC, Steen RG, Casadevall A (2012) Misconduct accounts for the majority of retracted scientific publications. *Proc Natl Acad Sci USA* 109(42):17028–17033.
15. Steen RG, Casadevall A, Fang FC (2013) Why has the number of scientific retractions increased? *PLoS One* 8(7):e68397.
16. US Food and Drug Administration (2015) New drugs at FDA: CDER's New Molecular Entities and New Therapeutic Biological Products. Available at perma.cc/Z8LX-3Q9E. Accessed March 6, 2015.
17. Chakma J, Sun GH, Steinberg JD, Sammut SM, Jagsi R (2014) Asia's ascent–global trends in biomedical R&D expenditures. *N Engl J Med* 370(1):3–6.
18. Contopoulos-Ioannidis DG, Alexiou GA, Gouvias TC, Ioannidis JPA (2008) Medicine. Life cycle of translational research for medical interventions. *Science* 321(5894):1298–1299.
19. Contopoulos-Ioannidis DG, Ntzani E, Ioannidis JPA (2003) Translation of highly promising basic science research into clinical applications. *Am J Med* 114(6):477–484.
20. Scannell JW, Blanckley A, Boldon H, Warrington B (2012) Diagnosing the decline in pharmaceutical R&D efficiency. *Nat Rev Drug Discov* 11(3):191–200.
21. Barham BL, Foltz JD, Prager DL (2014) Making time for science. *Res Policy* 43(1):21–31.
22. Ledford H (2015) Congress seeks to quash patent trolls. *Nature* 521(7552):270–271.
23. Seok J, et al.; Inflammation and Host Response to Injury, Large Scale Collaborative Research Program (2013) Genomic responses in mouse models poorly mimic human inflammatory diseases. *Proc Natl Acad Sci USA* 110(9):3507–3512.
24. Domcke S, Sinha R, Levine DA, Sander C, Schultz N (2013) Evaluating cell lines as tumour models by comparison of genomic profiles. *Nat Commun* 4:2126.
25. Geraghty RJ, et al.; Cancer Research UK (2014) Guidelines for the use of cell lines in biomedical research. *Br J Cancer* 111(6):1021–1046.
26. Casadevall A, Fang FC (2014) Specialized science. *Infect Immun* 82(4):1355–1360.
27. Papatheodorou SI, Trikalinos TA, Ioannidis JPA (2008) Inflated numbers of authors over time have not been just due to increasing research complexity. *J Clin Epidemiol* 61(6):546–551.
28. Casadevall A, Steen RG, Fang FC (2014) Sources of error in the retracted scientific literature. *FASEB J* 28(9):3847–3855.
29. Baker M (2015) Reproducibility crisis: Blame it on the antibodies. *Nature* 521(7552):274–276.
30. Ioannidis JPA (2015) A generalized view of self-citation: Direct, co-author, collaborative, and coercive induced self-citation. *J Psychosom Res* 78(1):7–11.
31. Eckartz K, Kirchkamp O, Schunk D (2012) How do incentives affect creativity? CESifo Working Paper Series No. 4049. Available at perma.cc/ET8J-6KWB. Accessed June 19, 2015.
32. Al-Shahi Salman R, et al. (2014) Increasing value and reducing waste in biomedical research regulation and management. *Lancet* 383(9912):176–185.
33. Young NS, Ioannidis JPA, Al-Ubaydli O (2008) Why current publication practices may distort science. *PLoS Med* 5(10):e201.
34. Chan A-W, et al. (2014) Increasing value and reducing waste: Addressing inaccessible research. *Lancet* 383(9913):257–266.
35. Anonymous (2012) Further confirmation needed. *Nat Biotechnol* 30(9):806.
36. Crow MM (2011) Time to rethink the NIH. *Nature* 471(7340):569–571.
37. Casadevall A, Fang FC (2012) Reforming science: Methodological and cultural reforms. *Infect Immun* 80(3):891–896.
38. Fang FC, Casadevall A (2012) Reforming science: Structural reforms. *Infect Immun* 80(3):897–901.
39. Tuljapurkar S, Li N, Boe C (2000) A universal pattern of mortality decline in the G7 countries. *Nature* 405(6788):789–792.
40. Mocroft A, et al. (1998) Changing patterns of mortality across Europe in patients infected with HIV-1. EuroSIDA Study Group. *Lancet* 352(9142):1725–1730.
41. Goldman L, Cook EF (1984) The decline in ischemic heart disease mortality rates. An analysis of the comparative effects of medical interventions and changes in lifestyle. *Ann Intern Med* 101(6):825–836.
42. Collins FS (2010) Revealing the body's deepest secrets. *Parade*. Available at perma.cc/RWJ4-C456. Accessed June 18, 2015.
43. Jobling MA, Gill P (2004) Encoded evidence: DNA in forensic analysis. *Nat Rev Genet* 5(10):739–751.
44. Chabner BA, Roberts TG, Jr (2005) Timeline: Chemotherapy and the war on cancer. *Nat Rev Cancer* 5(1):65–72.
45. Rosenau J (2012) Science denial: A guide for scientists. *Trends Microbiol* 20(12):567–569.
46. Achenbach J (2015) Why science is so hard to believe. *The Washington Post.* Available at perma.cc/7N96-DV7E. Accessed March 6, 2015.
47. US National Library of Medicine PubMed search engine. Available at www.ncbi.nlm.nih.gov/pubmed/. Accessed March 7, 2015.
48. US National Library of Medicine (2014) MEDLINE/PubMed data element (field) descriptions. Available at perma.cc/Q9XL-CNW2. Accessed January 22, 2015.
49. National Institutes of Health (2014) The NIH almanac: Appropriations (section 2). Available at perma.cc/62PJ-F3FS. Accessed January 21, 2015.
50. National Institutes of Health (2013) National Institutes of Health: Operating plan for FY 2014. Available at perma.cc/F3Z5-XEDS. Accessed January 21, 2015.
51. US Bureau of Labor Statistics (2015) CPI detailed report: Data for December 2014. Available at perma.cc/H8SJ-CT2C. Accessed January 21, 2015.
52. Madrick J (2014) Innovation: The government was crucial after all. *The New York Review of Books.* Available at perma.cc/8J6H-4XXE. Accessed June 18, 2015.
53. US Food and Drug Administration (2013) Summary of NDA approvals & receipts, 1938 to the present. Available at perma.cc/6SSM-W4B7. Accessed January 21, 2015.
54. US Food and Drug Administration (2015) Novel new drugs: 2014 summary. Available at perma.cc/ZN5J-LLGG. Accessed January 29, 2015.
55. The World Health Organization (2004) The World medicines situation. Available at perma.cc/FBH3-XS5X. Accessed February 2, 2015.
56. Hollis A (2004) Me-too drugs: Is there a problem? Available at perma.cc/U4W6-T6VA. Accessed June 18, 2015.
57. The World Bank (2014) World Bank open data. Available at data.worldbank.org. Accessed February 2, 2015.
58. US Patent and Trademark Office (2014) US patent statistics chart. Available at perma.cc/LN2Q-JFGT. Accessed March 7, 2015.
59. European Commission Eurostat (2014) Life expectancy by age and sex. Available at appsso.eurostat.ec.europa.eu/nui/show.do?dataset=demo_mlexpec&lang=en. Accessed October 2, 2014.
60. Wilson D (2005) The top pharmaceuticals that changed the world: Thyroxine. *Chemical & Engineering News.* Available at perma.cc/4Y2Q-U9TT. Accessed February 2, 2015.
61. Santella TM, Wertheimer AI (2005) The levothyroxine spectrum: Bioequivalence and cost considerations. *Formulary* 40(8):258–271.
62. Fleming A (1980) Classics in infectious diseases: On the antibacterial action of cultures of a penicillium, with special reference to their use in the isolation of B. influenzae by Alexander Fleming, Reprinted from the British Journal of Experimental Pathology 10:226-236, 1929. *Rev Infect Dis* 2(1):129–139.
63. American Chemical Society (1999) Discovery and development of penicillin. Available at perma.cc/9NRJ-PZXB. Accessed January 21, 2015.
64. Ahlquist RP (1948) A study of the adrenotropic receptors. *Am J Physiol* 153(3):586–600.
65. Quirke V (2006) Putting theory into practice: James Black, receptor theory and the development of the beta-blockers at ICI, 1958-1978. *Med Hist* 50(1):69–92.
66. Carr H (1952) Free precession techniques in nuclear magnetic resonance. PhD thesis (Harvard Univ, Cambridge, MA).
67. Mallard J (2003) The evolution of medical imaging: From Geiger counters to MRI–a personal saga. *Perspect Biol Med* 46(3):349–370.
68. Prince AM (1968) Relation of Australia and SH antigens. *Lancet* 2(7565):462–463.
69. Gerlich WH (2013) Medical virology of hepatitis B: How it began and where we are now. *Virol J* 10:239.
70. Erdös EG (2006) The ACE and I: How ACE inhibitors came to be. *FASEB J* 20(8):1034–1038.
71. Smith CG, Vane JR (2003) The discovery of captopril. *FASEB J* 17(8):788–789.
72. Jackson DA, Symons RH, Berg P (1972) Biochemical method for inserting new genetic information into DNA of Simian Virus 40: Circular SV40 DNA molecules containing lambda phage genes and the galactose operon of *Escherichia coli. Proc Natl Acad Sci USA* 69(10):2904–2909.
73. Junod SW (2009) Celebrating a milestone: FDA's approval of first genetically-engineered product. Available at perma.cc/BTN6-LZBZ. Accessed February 2, 2015.
74. Köhler G, Milstein C (1975) Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature* 256(5517):495–497.
75. Smith SL (1996) Ten years of Orthoclone OKT3 (muromonab-CD3): A review. *J Transpl Coord* 6(3):109–119, quiz 120–121.
76. Tobert JA (2003) Lovastatin and beyond: The history of the HMG-CoA reductase inhibitors. *Nat Rev Drug Discov* 2(7):517–526.
77. US Food and Drug Administration (2014) HIV/AIDS historical time line 1981-1990. Available at perma.cc/D7JM-9ZQE. Accessed January 21, 2015.
78. Bartlett JG (2006) Ten years of HAART: Foundations for the future. *Medscape.* Available at perma.cc/XQS4-FD24. Accessed February 2, 2015.
79. Okazaki T, Honjo T (2007) PD-1 and PD-1 ligands: From discovery to clinical application. *Int Immunol* 19(7):813–824.
80. Iwai Y, et al. (2002) Involvement of PD-L1 on tumor cells in the escape from host immune system and tumor immunotherapy by PD-L1 blockade. *Proc Natl Acad Sci USA* 99(19):12293–12297.