Innovation in the pharmaceutical industry: New estimates of R&D costs

Joseph A. DiMasi a,*, Henry G. Grabowski b, Ronald W. Hansen c

a Tufts Center for the Study of Drug Development, Tufts University, United States
b Department of Economics, Duke University, United States
c Simon Business School, University of Rochester, United States

A R T I C L E   I N F O

Article history:
Received 15 August 2014
Received in revised form 28 January 2016
Accepted 29 January 2016
Available online 12 February 2016

JEL classification:
L65
O31

Keywords:
Innovation
R&D cost
Pharmaceutical industry
Discount rate
Technical success rates

A B S T R A C T

The research and development costs of 106 randomly selected new drugs were obtained from a survey of 10 pharmaceutical firms. These data were used to estimate the average pre-tax cost of new drug and biologics development. The costs of compounds abandoned during testing were linked to the costs of compounds that obtained marketing approval. The estimated average out-of-pocket cost per approved new compound is $1395 million (2013 dollars). Capitalizing out-of-pocket costs to the point of marketing approval at a real discount rate of 10.5% yields a total pre-approval cost estimate of $2588 million (2013 dollars). When compared to the results of the previous study in this series, total capitalized costs were shown to have increased at an annual rate of 8.5% above general price inflation. Adding an estimate of post-approval R&D costs increases the cost estimate to $2870 million (2013 dollars).

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

We provide an updated assessment of the value of the resources expended by industry to discover and develop new drugs and biologics, and the extent to which these private sector costs have changed over time. The costs required to develop these new products clearly play a role in the incentives to invest in the innovative activities that can generate medical innovation. Our prior studies also have been used by other researchers, including government agencies, to analyze various policy questions (US Congressional Budget Office, 1998, 2006).

The full social costs of discovering and developing new compounds will include these private sector costs, but will also include government-funded and non-profit expenditures on basic and clinical research that can result in leads and targets which drug companies can explore. These additional costs can be substantial. However, it is difficult to identify and measure non-private expenditures that can be linked to specific new therapies. Thus, we focus here on the private sector costs.

The methodological approach used in this paper follows that used for our previous studies, although we apply additional statistical tests to the data (Hansen, 1979; DiMasi et al., 1991, 1995a,b, 2003, 2004; DiMasi and Grabowski, 2007). Because the methodologies are consistent, we can confidently make comparisons of the results in this study to the estimates we found for the earlier studies, which covered earlier periods, to examine and illustrate trends

http://dx.doi.org/10.1016/j.jhealeco.2016.01.012
0167-6296/© 2016 Elsevier B.V. All rights reserved.
in development costs. These studies used compound-level data on the cost and timing of development for a random sample of new drugs first investigated in humans and annual company pharmaceutical R&D expenditures obtained through surveys of a number of pharmaceutical firms.

We analyze private sector R&D activities as long-term investments. The industrial R&D process is marked by substantial financial risks, with expenditures incurred for many development projects that fail to result in a marketed product. Thus, our approach explicitly links the costs of unsuccessful projects to those that are successful in obtaining marketing approval from regulatory authorities. In addition, the pharmaceutical R&D process is very lengthy, often lasting a decade or more (DiMasi et al., 2003). This makes it essential to model accurately how development expenses are spread over time.

Given our focus on resource costs and how they have changed over time, we develop estimates of the average pre-tax cost of new drug development and compare them to estimates covering prior periods. We corroborated the basic R&D cost results in this study by examining the representativeness of our sample firms and our study data, and by incorporating a number of independently derived results and data relating to the industry and the drug development process into analyses that provide rough comparators for at least components of our cost results. The details of those analyses are provided in our online supplement.

The remainder of this paper is organized as follows. We briefly discuss the literature on pharmaceutical industry R&D costs since our 2003 study in Section 2. Section 3 briefly outlines the standard paradigm for the drug development process. In Section 4 we describe the survey sample data and the population from which they were drawn, and briefly outline the methodology used to derive full R&D cost estimates from data on various elements of the drug development process. We present base case pre- and post-marketing approval R&D cost estimates in Section 5. Sensitivity analyses are presented in Section 6. We describe the representativeness of our data, various approaches to validating our results, and responses to various critiques in Section 7. Finally, we summarize our findings in Section 8.

2. Previous studies of the cost of pharmaceutical innovation

Much of the literature on the cost of pharmaceutical innovation dating back decades has already been described by the authors in their previous two studies (DiMasi et al., 1991, 2003). The interested reader can find references and discussions about the prior research in those studies. The earliest studies often involved a case study of a single drug (typically without accounting for the cost of failed projects) or they analyzed aggregate data. We will focus here on studies and reports that have emerged since DiMasi et al. (2003) that involve the use of new data for at least some parts of the R&D process. The basic elements of these analyses are shown in Table 1.

Adams and Brantner (2006, 2010) sought to assess the validity of the results in DiMasi et al. (2003) with some alternative data. Specifically, in their 2006 article, they used a commercial pipeline database to separately estimate clinical approval and phase attrition rates, as well as phase development times. They found a similar overall cost estimate ($868 million versus $802 million in year 2000 dollars). The authors followed that study with another

study that featured clinical phase out-of-pocket cost estimates derived from regressions based on publicly available data on company R&D expenditures (Adams and Brantner, 2010). They found a somewhat higher overall cost estimate ($1.2 billion in year 2000 dollars).

In a paper authored by two of the authors of this study (DiMasi and Grabowski, 2007), we provided a first look at the costs of developing biotech products (specifically, recombinant proteins and monoclonal antibodies). The methodological approach was the same as that used for our studies of traditional drug development. We used some data from DiMasi et al. (2003) combined with new data on the costs of a set of biotech compounds from a single large biopharmaceutical company. Biotech drugs were observed to have a higher average clinical success rate than small molecule drugs, but this was largely offset by other cost components. We found that the full capitalized cost per approved new compound was similar for traditional and biotech development ($1.3 billion for biotech and $1.2 billion for traditional development in year 2005 dollars), after adjustments to compare similar periods for R&D expenditures.

The other studies shown in Table 1 are discussed in detail in the online supplement. One important finding emerging from the survey of cost studies in Table 1 is that clinical success rates are substantially lower for the studies focused on more recent periods. This observed trend is consistent with other analyses of success probabilities (DiMasi et al., 2010; DiMasi et al., 2013; Hay et al., 2014; Paul et al., 2010) and our analysis below. Average R&D (inflation-adjusted) cost estimates are also higher for studies focused on more recent periods, suggesting a growth in real R&D costs. While suggestive, these studies are not strictly comparable to our earlier analyses of R&D costs given methodological differences and data omissions that are discussed in the online supplement (Appendix A).

3. The new drug development process

The new drug development process need not follow a fixed pattern, but a standard paradigm has evolved that fits the process well in general. We have described the process in some detail in previous studies, and the FDA’s website contains a schematic explaining the usual set of steps along the way from test tube to new compound approval (http://www.fda.gov/Drugs/DevelopmentApprovalProcess/SmallBusinessAssistance/ucm053131.htm). Marketing approval applications for investigational compounds submitted to the FDA for review by manufacturers are referred to as new drug applications (NDAs) or biologic license applications (BLAs), depending on the type of product.

In basic form, the paradigm portrays new drug discovery and development as proceeding along a sequence of stages and activities (some of which often overlap). Basic and applied research initiate the process with discovery programs that result in the synthesis or isolation of compounds that are tested in assays and animal models in preclinical development. We do not have the level pipeline databases are also thin prior to the mid-1990s. The DiMasi et al. (2003) study covered new drugs that had first entered clinical testing anywhere in the world from 1983 to 1994 and followed the same set of drugs through time.

4 However, the authors interpreted their estimate as a marginal, as opposed to an average, drug cost. The concept, though, of marginal cost has an unclear meaning here. With high fixed costs and a development process that varies by drug, it is difficult to understand what marginal pharmaceutical R&D cost means in this context. It seems that the relevant marginal concept here is marginal profitability. The marginal profitability drug could have a very high or a very low cost. What’s more, marginal profitability may only have meaning at the firm, not the industry, level. The cost of a marginally profitable drug in the pipeline of a firm may be high for one firm and low for another firm.

---

2 For mean out-of-pocket phase costs, they used the estimates in DiMasi et al. (2003).

3 The Adams and Brantner (2006) study used records in the pipeline database that were reported to have entered some clinical testing phase from 1989 to 2002. Thus, they did not follow the same set of drugs through time. The data for the commercial
of granularity to disaggregate R&D expenditure data into discovery and preclinical development testing costs, so for the purposes of this study, as in prior studies, discovery and preclinical development costs are grouped and referred to as pre-human costs.\(^5\)

Clinical (human) testing typically proceeds through three successive, sometimes overlapping phases. Historically, human testing has often been initiated first outside the United States (DiMasi, 2001). For any of these clinical phases, pharmaceutical companies may pursue development of their investigational compounds in multiple indications prior to and/or after the initial indication approval.

### 4. Data and methods

Ten multinational pharmaceutical firms of varying sizes provided data through a confidential survey of their new drug and biologics R&D costs.\(^6\) Data were collected on clinical phase expenditures and development phase times for a randomly selected sample of the investigational drugs and biologics of the firms participating in the survey.\(^7\) The sample was taken from a Tufts Center for the Study of Drug Development (CSDD) database of the investigational compounds of top 50 firms. Tufts CSDD gathered information on the investigational compounds in development and their development status from commercial pipeline intelligence databases (IMS R&D Focus and Thomson Reuters Cortellis database [formerly the IDea database]), published company pipelines, clinicaltrials.gov, and web searches. Cost and time data were also collected for expenditures on the kind of animal testing that often occurs concurrently with clinical trials.\(^8\) The compounds chosen were self-originated in the following sense. Their development from synthesis up to initial regulatory marketing approval was conducted under the auspices of the surveyed firm. This inclusion criterion is broader than it might at first seem since it includes compounds of firms that were acquired or merged with the survey firm during development and drugs that originated with the survey firm and were co-developed (and for which full cost data were available).\(^9\) Licensed-in and co-developed compounds without partner clinical cost data were excluded because non-survey firms would have conducted significant portions of the R&D.\(^10\)

We also collected data from the cost survey participants on their aggregate annual pharmaceutical R&D expenditures for the period 1990–2010. The firms reported on total annual R&D expenditures broken down by expenditures on self-originated new drugs, biologics, diagnostics, and vaccines. Data were also provided on annual R&D expenditures for licensed-in or otherwise acquired new drugs, and on already-approved drugs. Annual expenditures on self-originated new drugs were further decomposed into expenditures during the pre-human and clinical periods.

The survey firms accounted for 35% of both top 50 firm pharmaceutical sales and pharmaceutical R&D expenditures. Of the 106 investigational compounds included in the project dataset, 87 are small molecule chemical entities (including three synthetic peptides), and 19 are large molecule biologics (10 monoclonal antibodies and nine recombinant proteins). For ease of exposition, we will refer to all compounds below as new drugs, unless otherwise indicated. Initial human testing anywhere in the world for these compounds occurred during the period 1995–2007. Development costs were obtained through 2013.

We selected a stratified random sample of investigational compounds.\(^11\) Stratification was based on the status of testing as of the end of 2013. Reported costs were weighted to reflect the development status of compounds in the population relative to those in the cost survey sample, so that knowledge of the distribution of development status in the population from which the sample was drawn was needed. The population is composed of all investigational compounds in the Tufts CSDD investigational drug database that met study criteria: the compounds were self-originated and first tested in humans anywhere in the world from 1995 to 2007. We found 1442 investigational drugs that met these criteria. Of these compounds, 103 (7.1%) have been approved for marketing, 13 (0.9%) had NDAs or BLAs that were submitted and are still active, 11 (0.8%) had NDAs or BLAs submitted but abandoned, 576 (39.9%) were abandoned in phase I, 19 (1.3%) were still active in phase I, 492 (34.1%) were abandoned in phase II, 84 (5.8%) were still active in phase II, 78 (5.4%) were abandoned in phase III, and 66 (4.6%) were still active in phase III. For both the population and the cost survey sample, we estimated approval and discontinuation shares for the active compounds by phase so that the population and sample distributions consisted of shares of compounds that were approved or discontinued in phase I, phase II, phase III, or regulatory review.

---

\(^5\) We capture out-of-pocket discovery costs with our data, but the pre-synthesis discovery period is highly variable with no clear starting point. For our analyses we began our representative discovery and development timeline at the point of compound synthesis or isolation. Thus, our estimates of time costs are somewhat conservative.

\(^6\) Using pharmaceutical sales in 2006 to measure firm size, 5 of the survey firms are top 10 companies, 7 are top 25 firms, and 3 are outside the top 25 (Pharmaceutical Executive, May 2007).

\(^7\) A copy of the survey instrument can be found in our online supplement (Appendix C).

\(^8\) Long-term teratogenicity and carcinogenicity testing may be conducted after the initiation of clinical trials, and is often concurrent with phase I and phase II testing.

\(^9\) The criterion also does not preclude situations in which the firm sponsors trials that are conducted by or in collaboration with a government agency, an individual or group in academia, a non-profit institute, or another firm.

\(^10\) Large and mid-sized pharmaceutical firms much more often license-in than license-out new drug candidates. Firms that license-in compounds for further development pay for the perceived value of the prior R&D typically through up-front fees, development and regulatory milestone payments, and royalty fees if the compound should be approved for marketing. For a breakdown of new drugs and biologics approved in the United States in the 2000s by business arrangements among firms initiated during clinical development, see DiMasi et al. (2014).

\(^11\) To ease the burden of reporting and increase the likelihood that firms would respond, we limited the number of compounds to be reported on to a maximum of 15 for any firm (with fewer compounds for smaller firms).
cost survey sample was purposely weighted toward compounds that lasted longer in development to increase the amount of information on drugs that reached late-stage clinical testing. Weights, determined as described above, were then applied to the compounds in the cost dataset so that the results would reflect the development status distribution for the population from which the sample was drawn.

Some firms were not able to provide full phase cost data for every new drug sampled. For example, phase I cost data were available for 97 of the 106 new drugs in the dataset (92%). Of the 82 compounds in the dataset that had entered phase II, cost data were available for 78 (95%). For phase III, cost data were available for 42 of the 43 compounds that entered the phase (98%). However, we had cost data for at least one phase for each of the 106 drugs in the sample. In aggregate, we had cost data for all phases entered for 94 of the 106 compounds (89%). In addition, five compounds were still active in a phase at the time that data were reported. For these drugs it is likely that there will be some additional future costs for the drug’s most recent phase. Thus, for this reason our cost estimates are likely to be somewhat conservative. However, given the small number of drugs in this category and the fact that the impact would be on only one phase for each of these drugs, our overall cost estimates are not likely to be substantially affected.

The methodology that we use to estimate development costs is the same as the approach used in our earlier studies (Hansen, 1979; DiMasi et al., 1991, 2003). We refer the reader to the earlier studies and to our online supplement (Appendix A) for details. The methodology results in a full risk-adjusted cost per approved new compound that also takes into account time costs. That is, we link the cost of compound failures to the cost of the successes (investigational compounds that attain regulatory marketing approval), and we utilize a representative time profile along with an industry cost of capital to monetize the cost of the delay between when R&D expenditures are incurred and when returns to the success can first be realized (date of marketing approval). We refer to the sum of out-of-pocket cost (actual cash outlays) and time cost per approved new compound as the capitalized cost per approved new compound. The full capitalized cost estimate is built through a number of estimates of various components of the drug development process. These individual component estimates are interesting as objects of analysis in their own right, and we provide estimates for those components.

5. Base case R&D cost estimates

5.1. Out-of-pocket clinical cost per investigational drug

To determine expected costs, we need estimates of the clinical development risk profile. We examined the dataset of 1442 self-originated compounds of top 50 pharmaceutical firms described above and estimated the phase transition probabilities shown in Fig. 1. The overall probability of clinical success (i.e., the likelihood that a drug that enters clinical testing will eventually be approved) was estimated to be 11.83%. This success rate is substantially lower than the rate of 21.50% estimated for the previous study, but consistent with several recent studies of clinical success rates. Such an increase in overall risk will contribute greatly to an increase in costs per approved new drug, other things equal.

Cost survey sample was purposely weighted toward compounds that lasted longer in development to increase the amount of information on drugs that reached late-stage clinical testing. Weights, determined as described above, were then applied to the compounds in the cost dataset so that the results would reflect the development status distribution for the population from which the sample was drawn.

Some firms were not able to provide full phase cost data for every new drug sampled. For example, phase I cost data were available for 97 of the 106 new drugs in the dataset (92%). Of the 82 compounds in the dataset that had entered phase II, cost data were available for 78 (95%). For phase III, cost data were available for 42 of the 43 compounds that entered the phase (98%). However, we had cost data for at least one phase for each of the 106 drugs in the sample. In aggregate, we had cost data for all phases entered for 94 of the 106 compounds (89%). In addition, five compounds were still active in a phase at the time that data were reported. For these drugs it is likely that there will be some additional future costs for the drug’s most recent phase. Thus, for this reason our cost estimates are likely to be somewhat conservative. However, given the small number of drugs in this category and the fact that the impact would be on only one phase for each of these drugs, our overall cost estimates are not likely to be substantially affected.

The methodology that we use to estimate development costs is the same as the approach used in our earlier studies (Hansen, 1979; DiMasi et al., 1991, 2003). We refer the reader to the earlier studies and to our online supplement (Appendix A) for details. The methodology results in a full risk-adjusted cost per approved new compound that also takes into account time costs. That is, we link the cost of compound failures to the cost of the successes (investigational compounds that attain regulatory marketing approval), and we utilize a representative time profile along with an industry cost of capital to monetize the cost of the delay between when R&D expenditures are incurred and when returns to the success can first be realized (date of marketing approval). We refer to the sum of out-of-pocket cost (actual cash outlays) and time cost per approved new compound as the capitalized cost per approved new compound. The full capitalized cost estimate is built through a number of estimates of various components of the drug development process. These individual component estimates are interesting as objects of analysis in their own right, and we provide estimates for those components.

5. Base case R&D cost estimates

5.1. Out-of-pocket clinical cost per investigational drug

To determine expected costs, we need estimates of the clinical development risk profile. We examined the dataset of 1442 self-originated compounds of top 50 pharmaceutical firms described above and estimated the phase transition probabilities shown in Fig. 1. The overall probability of clinical success (i.e., the likelihood that a drug that enters clinical testing will eventually be approved) was estimated to be 11.83%. This success rate is substantially lower than the rate of 21.50% estimated for the previous study, but consistent with several recent studies of clinical success rates. Such an increase in overall risk will contribute greatly to an increase in costs per approved new drug, other things equal.

As described above, we calculated weighted means, medians, standard deviations, and standard errors for clinical phase costs. Some of the firms could not separate out long-term animal testing costs during clinical development, and instead, included these costs in their phase cost estimates by year. To be consistent, therefore, for those compounds where animal costs were separately reported, we allocated those costs to the clinical phases according to when the animal testing costs were incurred. Thus, the clinical phase costs presented in Table 2 are inclusive of long-term animal testing costs. Weighted mean and median costs per investigational drug entering a phase increase for later clinical phases, particularly for phase III (which typically includes a number of large-scale trials). In comparison to our previous study (DiMasi et al., 2003), both mean and median phase III cost are notably higher relative to the earlier phases. While the ratio of mean phase III cost to mean phase I cost was 5.7 for the previous study, it was 10.1 here. Similarly, the ratio of mean phase III to phase II cost was 3.7 for the earlier study, but was 4.4 for this study. Mean phase II cost was also higher relative to phase I cost in the current study compared to the previous one (2.3 times as high compared to 1.5 times as high). Thus, while mean cost in real dollars for phase I increased 28% relative to the previous study, phase II costs were notably lower relative to both phase II and phase III for the current study.

As we will see below, the differential in cost per approved new drug between the two studies will be much greater than cost per investigational drug because of the much lower overall clinical approval success rate. However, our results do show that the impact is mitigated to some degree by firms failing the drugs that they do abandon faster for the current study period. The distribution of clinical period failures for this study were 45.9% for phase I, 43.5% for phase II, and 10.6% for phase III/regulatory review. The

---

12 Phase cost correlation results presented in the online supplement, together with an examination of relative phase costs for drugs that had some missing phase cost data, suggest that our phase cost averages (exclusive of missing data) are conservative.

13 See, for example, Paul et al. (2010), DiMasi et al. (2013), and Hay et al. (2014).

---

Fig. 1. Estimated phase transition probability and overall clinical approval success rates for self-originated new molecular entity (NME) and new therapeutically significant biologic entity (NBE) investigational compounds first tested in humans anywhere from 1995 to 2007.

As described above, we calculated weighted means, medians, standard deviations, and standard errors for clinical phase costs. Some of the firms could not separate out long-term animal testing costs during clinical development, and instead, included these costs in their phase cost estimates by year. To be consistent, therefore, for those compounds where animal costs were separately reported, we allocated those costs to the clinical phases according to when the animal testing costs were incurred. Thus, the clinical phase costs presented in Table 2 are inclusive of long-term animal testing costs.

Weighted mean and median costs per investigational drug entering a phase increase for later clinical phases, particularly for phase III (which typically includes a number of large-scale trials). In comparison to our previous study (DiMasi et al., 2003), both mean and median phase III cost are notably higher relative to the earlier phases. While the ratio of mean phase III cost to mean phase I cost was 5.7 for the previous study, it was 10.1 here. Similarly, the ratio of mean phase III to phase II cost was 3.7 for the earlier study, but was 4.4 for this study. Mean phase II cost was also higher relative to phase I cost in the current study compared to the previous one (2.3 times as high compared to 1.5 times as high). Thus, while mean cost in real dollars for phase I increased 28% relative to the previous study, phase II costs were notably lower relative to both phase II and phase III for the current study.

As we will see below, the differential in cost per approved new drug between the two studies will be much greater than cost per investigational drug because of the much lower overall clinical approval success rate. However, our results do show that the impact is mitigated to some degree by firms failing the drugs that they do abandon faster for the current study period. The distribution of clinical period failures for this study were 45.9% for phase I, 43.5% for phase II, and 10.6% for phase III/regulatory review. The
corresponding figures for the previous study were 36.9% for phase I, 50.4% for phase II, and 12.6% for phase III/regulatory review.

5.2. Cost of capital estimates

To account for the time value of money in our previous paper (DiMasi et al., 2003), we utilized an 11% real after-tax weighted average cost of capital (WACC). In particular, we employed the capital asset pricing model (CAPM) to estimate the cost of equity capital. This was combined with the cost of debt, appropriately weighted with the cost of equity, to yield a representative, pharmaceutical industry weighted after-tax cost of capital. The resultant parameters were estimated at regular intervals from the mid-1980s to the year 2000, given the time period spanned by our sample of R&D projects.

In the present paper, we follow the same methodology to compute WACC. In the current R&D cost analysis, we have a sample of new drugs that began clinical trials in 1995 through 2007 and which have an average introduction period in the latter part of the 2000 decade. Hence, a relevant time period for our cost of capital is the mid-1990s through 2010. Our analysis yielded an after-tax weighted cost of capital of 10.5%, moderately lower than in our last paper. This reflects the fact that the cost of equity capital has declined in pharmaceuticals since 2000 (as well as for other industrial sectors). Research intensive industries, including the pharmaceutical industry, generally finance most of their investments through equity, rather than through debt. This is the case even when the cost of debt is significantly below the cost of equity (Hall, 2002; Vernon, 2004). One of the primary reasons is that servicing debt requires a stable source of cash flows, while the returns to R&D activities are skewed and highly variable (Scherer and Harhoff, 2000; Berndt et al., 2015). Given the low debt-to-equity ratios that exist for pharmaceutical firms, the cost of equity component dominates the computed WACC values in Table 3.

To obtain a real cost of capital, we first compute the nominal values and then subtract the expected rate of inflation. The nominal cost of capital in 1994 is from a CAPM study by Myers and Howe (1997). The estimates for 2000, 2005, and 2010 are based on our own analysis, utilizing a comparable approach, with a large sample of pharmaceutical firms.14 As this table shows, the estimated nominal cost of capital for pharmaceuticals was fairly stable during the period 1994–2000 (14.2–14.9%). However, it decreased during the decade of 2000s, particularly after the global recession occurred (with a value of 11.4% observed in 2010).

As discussed in DiMasi et al. (2003), the rate of inflation was above historical values during the first part of the 1980s, but then receded back to or below historical levels throughout most of the 1990s. Hence, we utilized the long run historical value for inflation for the expected inflation level in 1994 and 2000 (3.1%), as in our prior work. For the 2000s decade, inflation was significantly below historical values. In this case, we employed a 5-year lagged moving average to compute the expected rate of inflation in 2005 and 2010 (calculated as 2.5% and 2.0%, respectively).

As shown in Table 3, our estimates for the real cost of capital varied between 9.4% and 11.8% for pharmaceutical firms over the 1994–2010 period. We elected to use the midpoint of this range, or approximately 10.5%, as the representative COC to capitalize our R&D cost estimates.

The focus of our analysis is R&D investment expenditures and privately financed resources for new drugs undertaken by the biopharmaceutical industry. Accordingly we capitalized these expenditures utilizing a cost of capital estimate based on financial data from publicly listed firms. Drug development is also sponsored and funded by government and non-profit agencies (e.g., public–private partnerships devoted to developing medicines for neglected diseases). To the extent that our cost estimates are applicable to these ventures, a social rate of discount would be appropriate to capitalize R&D outlays. We provide a sensitivity analysis in Section 6 with respect to a wide spectrum of alternative cost of capital values.

5.3. Capitalized clinical cost per investigational drug

Opportunity cost calculations for clinical period expenditures require estimates of average phase lengths and average gaps or overlaps between successive clinical phases to generate an average clinical development and regulatory review timeline. Mean phase lengths and the mean lengths of time between successive phases are shown in Table 4, along with the associated capitalized mean phase costs and capitalized expected phase costs by phase for investigational compounds. The time between the start of clinical testing and submission of an NDA or BLA with the FDA was estimated to be 80.8 months, which is 12% longer (8.7 months) than the same period estimated for the previous study. The average time from the start of clinical testing to marketing approval for our timeline was 96.8 months for the current study, 7% (6.5 months) longer than for the earlier study. The difference is accounted for by shorter FDA approval times. The period for the previous study included, in part, a period prior to the implementation of the Prescription Drug Use Fee Act of 1992 (PDUFA), and, in part, the early user fee era for which approval times were somewhat higher than for later user fee periods (Berndt et al., 2005).15 While the approval

---

14 The sample is composed of all publicly traded drug firms in the Value Line Survey which also provides beta values and the other pharma-specific parameters used in the CAPM calculations for the relevant years. The long-term horizon equity risk premium, and the yield on long-term government bonds employed in the CAPM analysis, are from Ibbotson Valuation yearbooks for 2000, 2005, and 2010.

15 The user fee legislation sunsets every 5 years. It has been renewed every 5 years since its original enactment. Performance goals for FDA review of marketing
phase averaged 18.2 months for the earlier paper’s study period, that phase averaged 16.0 months for drugs covered by the current study. Other things being equal, the observed longer times from clinical testing to approval yielded higher capitalized costs relative to out-of-pocket costs. However, the discount rate that we used for the current study is also lower than for the previous study (10.5% versus 11.0%). The two effects work in offsetting ways. In addition, capitalized clinical cost per investigational compound will also depend on the gaps and overlaps between phases. On net, the ratio of mean capitalized to out-of-pocket cost per investigational compound was slightly lower for the current study compared to the previous one (1.5 versus 1.7).20

5.4. Clinical cost per approved new drug

Average cost estimates for investigational drugs are useful, but we are primarily interested in estimates of cost per approved new drug. As noted above, our analysis of drugs in development for the relevant period yielded a predicted overall clinical success rate of 11.83%. Applying this success rate to our estimates of out-of-pocket and capitalized costs per investigational drug results in estimates of cost per approved new drug that link the cost of drug failures to the successes.

Aggregating across phases, we found an out-of-pocket clinical period cost per approved new drug estimate of $965 million and a capitalized clinical period cost per approved new drug estimate of $1460 million. In constant dollars, these costs are 2.6 and 2.4 times higher than those we found in our previous study, respectively.

5.5. Pre-human out-of-pocket and capitalized costs per approved drug

The pre-human period, as defined here, includes discovery research as well as preclinical development. Some costs incurred during this period cannot be associated with specific compounds. To deal with this issue, we analyzed reported aggregate annual firm expenditures on self-originated new drugs by the pre-human and clinical periods. We gathered data on aggregate expenditures for these periods from survey firms for 1990–2010. Both times series tended to increase over time in real terms. Given this outcome, and the fact that the clinical expenditures in 1 year will be associated with pre-human expenditures that occurred years earlier, the ratio of total pre-human expenditures to total R&D (pre-human plus clinical) expenditures over the entire study period would yield an underestimate of the share of total cost per new drug that is accounted for by the pre-human period. To accurately estimate

---

### Table 4
Average phase times and clinical period capitalized costs for investigational compounds (in millions of 2013 dollars).

<table>
<thead>
<tr>
<th>Testing phase</th>
<th>Mean phase length</th>
<th>Mean time to next phase</th>
<th>Capitalized mean phase cost (^a)</th>
<th>Capitalized expected phase cost (^b,c)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phase I</td>
<td>33.1</td>
<td>19.8</td>
<td>49.6</td>
<td>49.6</td>
</tr>
<tr>
<td>Phase II</td>
<td>37.9</td>
<td>30.3</td>
<td>95.3</td>
<td>56.7</td>
</tr>
<tr>
<td>Phase III</td>
<td>45.1</td>
<td>30.7</td>
<td>314.0</td>
<td>66.4</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>172.7</td>
<td></td>
</tr>
</tbody>
</table>

\(^a\) All costs were deflated using the GDP implicit price deflator. Weighted values were used in calculating means for costs and phase times. Phase times are given in months.
\(^b\) The NDA/BLA approval phase was estimated to be 16.0 months on average (2000–2012).
\(^c\) Costs were capitalized at an 10.5% real discount rate.

---

![Fig. 2. Pre-human phase, clinical phase, and total out-of-pocket and capitalized costs per approved new compound.](image)

this share we built in a lag structure that associates pre-human expenditures with clinical expenditures incurred some time later.

The survey firms reported on dates of synthesis or isolation for compounds for which we sought cost data, as well as dates of first human testing. We had data for the period from synthesis to first human testing for 78 of the compounds. The average time from synthesis to initial human testing for these compounds was 31.2 months, down considerably from 52.0 months for the previous study.21 Our analyses of clinical phase lengths and phase gaps and overlaps indicated a period of 95.2 months over which clinical period development costs are incurred. We approximated the lag between pre-human and clinical expenditures for a representative new drug as the time between the midpoints of each period. This yields a lag of 63.2 months, or approximately 5 years. Thus, we used a 5-year lag in analyzing the aggregate expenditure data, although we also examined 4-year and 6-year lags. A 5-year lag applied to the aggregate expenditure data resulted in a pre-human to total R&D expenditure ratio of 30.8%, which was only slightly different from the corresponding ratio used in our previous study (30.0%). The share was applied to our clinical cost estimates to determine associated pre-human cost estimates.

Given the estimates of out-of-pocket and capitalized clinical cost per approved new drug noted in Section 5.4 and the pre-human expenditure to total R&D expenditure ratio, we can infer pre-human out-of-pocket and capitalized costs per approved new drug of $430 million and $1098 million, respectively (Fig. 2). The results are very robust to different values for the length of the lag structure. For example, if we assume a lag of 4 years instead of 5 years, then out-of-pocket pre-human costs would be 6.8% higher. Alternatively, if we assume a 6-year lag, then out-of-pocket pre-human costs would be 8.5% lower.22

---

\(^20\) The results for the current study are consistent with data for a small number of compounds reported in a recently published study (Stergiopoulos and Getz, 2012). The mean time from synthesis to human testing there was 37.9 months for 17 compounds.

\(^21\) The pre-human to total R&D expenditure ratios for four- and six-year lags were 32.2% and 28.9%, respectively.

\(^22\) The differences in the ratios of capitalized to out-of-pocket cost for the individual phases were also small. For the current study they were 2.0, 1.6, and 1.2 for phase I, phase II, and phase III, respectively. For the earlier study, we found the ratios to be 2.0, 1.8, and 1.3 for phase I, phase II, and phase III, respectively.

---

applications under PDUFA were tightened somewhat for some applications after the initial 5-year period.
5.6. Total capitalized cost per approved drug

Total cost estimates are the sum of pre-human and clinical period cost estimates. Our base case total out-of-pocket cost per approved new drug is $1395 million, while our fully capitalized total cost estimate is $2558 million (Fig. 2). Time costs (differences between capitalized cost and out-of-pocket cost) account for 45% of total cost. This share is down from the share in our previous study (50%) and that for the study that preceded it (51%). This is due in part to a shorter pre-human period and a lower discount rate.

5.7. Trends in R&D costs

Fig. 3 presents capitalized pre-human, clinical, and total cost per approved new drug for the previous three studies in this series and for our current study. In constant dollars, total capitalized cost increased 2.31 times for the second study in comparison to the first, 2.53 times for the third study in comparison to the second study, and 2.45 times for the current study in comparison to the third study. However, the samples for these studies include drugs that entered clinical testing over periods that are not uniformly distributed. In addition, while the samples were chosen on the basis of when drugs entered clinical testing, changes over time in the average length of the development process make ascribing differences in the study periods according to the year of first human testing problematic. An alternative is to determine an average approval date for drugs in each study sample and use the differences in these dates to define the time differences between the studies. Our previous study described this approach and presented the corresponding annual growth rates between successive studies for the first three studies.

Drugs in the current study sample obtained FDA marketing approval from 2005 to 2013. The mean and median approval dates for drugs in the current study sample were both in 2008. For the previous study, we reported that the average approval date was in 1997. Thus, we used 11 years as the relevant time span between the studies and calculated compound annual rates of growth between the two studies accordingly.

Using the period differences described here and in our previous study, we determined the compound annual growth rates between the studies for out-of-pocket and capitalized cost per approved drug for pre-human, clinical, and total costs (Table 5). Compared to the growth rate for the results in the previous study, the growth rates for total out-of-pocket and capitalized costs for the current study are somewhat higher (9.3% and 8.5% per year). The results for the current study in comparison to those for the previous study are also noteworthy in that, after a substantial decline in the growth rate for real pre-human costs described in the previous study and presented in Table 5, pre-human costs for the current study resumed a much higher rate of growth. Conversely, the growth rates for clinical period expenditures declined from the very high rates for the previous study, although they are still substantial.

5.8. Cost of post-approval R&D

As we did for our most recent study, we develop indirect estimates of post-approval R&D costs. Post-approval R&D consists of efforts subsequent to original marketing approval to develop the active ingredient for new indications and patient populations, new dosage forms and strengths, and to conduct post-approval (phase IV) research required by regulatory authorities as a condition of original approval. We follow the methodology that we used in previous study.23 We utilize our pre-approval estimates together with aggregate pharmaceutical industry data regarding the drug development process to construct an estimate of the cost of post-approval R&D, which together with our pre-approval estimates, provide estimates of average total R&D cost per new drug covering the entire development and product life-cycle. The data that we collected from the survey firms on company annual aggregate expenditures on biopharmaceutical R&D show that over the study period these firms spent 73.1% of their prescription biopharmaceutical R&D expenditures on investigational self-originated new compounds.24 10.2% on investigational compounds that were licensed-in or otherwise acquired, and 16.5% on improvements to drugs that have already been approved.25

We cannot, however, use the percentage of aggregate R&D expenditures spent on post-approval R&D on a current basis and apply it to a pre-approval cost estimate to obtain an appropriate estimate of the cost of post-approval R&D per approved compound. The reason is that pre-approval costs occur years before post-approval costs. We used our aggregate annual firm R&D data to obtain an appropriate ratio by building in a reasonable lag structure between pre-approval and post-approval costs.

For our base results we used, as we did for the previous study, a 10-year lag for the aggregate data (which is the approximate time between median pre-approval development costs and median post-approval costs, given an 8-year post-approval expenditure period), we assumed that post-approval R&D cost per approval is the same, on average, for licensed-in and self-originated compounds, and we determined the percentage of approvals for the cost survey firms that are self-originated to estimate the ratio of post-approval R&D cost per approved compound to pre-approval cost per approved compound. The data indicated that this share was 33.4%. Applying this ratio, we estimated the out-of-pocket per approved compound for post-approval R&D to be $466 million (Fig. 4). Since these costs occur after approval and we are capitalizing all costs to the point of marketing approval, our discounted cost estimate is lower ($312 million). Thus, out-of-pocket cost per approved compound for post-approval R&D is 25.0% of...

---

23 We refer to the discussion in DiMasi et al. (2003) and an accompanying Appendix A for more detail on the method.

24 This figure includes expenditures on biologics, vaccines, and diagnostics. The self-originated share for therapeutic investigational drugs and biologics was 71.2%.

25 These expenditure shares are similar to those found for the previous study for the 1980 to 1999 period. The results here are also similar to figures that the trade association Pharmaceutical Research and Manufacturers of America (PhRMA) has published for its member firms for the years 2003 and 2005 to 2010. Those data do not separate out expenditures on existing products, but they do distinguish between self-originated and licensed products. Aggregating across those years, the shares for self-originated, licensed, and uncategorized were 74.3%, 17.6%, and 8.1%, respectively.
total R&D cost (pre- and post-approval), while capitalized cost for post-approval R&D is 10.9% of total cost.

5.9. Extensions to the base case

We can extend the base case results on drug development costs prior to original approval in a number of interesting ways. The sample dataset includes information on compound-level costs for both chemical compounds (small molecules) and biologics (large molecules). As reported in the online supplement (Appendix B), we examined investigational compounds by molecule size for differences in individual clinical phase costs. Since the distributions of compounds across therapeutic classes differ for large and small molecules, we conducted a regression analysis of phase costs for investigational compounds for each of the three clinical phases, while controlling for molecules size and therapeutic class. Sample sizes were somewhat limited when cut by both sample size and therapeutic class, but we found statistically significant higher phase II costs for large molecules. However, we found that clinical approval success rates for large molecules are substantially higher than for small molecules. As a result, clinical period cost per approved compound was appreciably higher for small molecules, with the ratio of costs nearly the same as we had estimated in a previous paper for an earlier period (DiMasi and Grabowski, 2007). Compete results are given and discussed in the online supplement (Appendix B).

The base case results on full R&D costs link expenditures on drug failures to the costs of drugs that attain regulatory success. We can also estimate the clinical period cost of taking a successful drug all the way to approval by examining the data for just the approved drugs in the sample. Focusing on that subsample also allowed us to examine evidence on the costs for the more therapeutically significant drugs (according to what is known at the time of approval) by using an FDA prioritization system for reviewing drugs submitted to the agency for marketing approval. We found that clinical period costs were substantially higher for the approved compounds in the sample relative to our results for the sample as a whole, and that costs were lower (although not at a statistically significant level) for compounds that the FDA had designated for a priority review (compounds thought to represent a significant gain over existing therapy). These results are presented in full and discussed in the online supplement (Appendix B).

6. Sensitivity analysis

We examined how sensitive the results were to extreme values in the data and to changes in certain critical parameters. In particular, we focus in detail in this section on variation in the discount rate used to calculate capitalized costs. We also determine the extent to which key cost drivers (cash outlays, risks, time, and the cost of capital) explain the increase in total cost per approved drug found for this study relative to our previous study.

In addition, since all of the parameters are subject to sampling error, we conducted Monte Carlo simulations, reported on in detail in the online supplement (Appendix C), allowing all parameters to vary according to their sampling distributions (using Crystal Ball™ software). For the full capitalized pre-approval cost estimate, 80% of the simulation forecasts (set of 1000) varied between $2.3 billion and $2.8 billion. All of the forecasts varied between $1.9 billion and $3.2 billion.

Finally, we also conducted an outlier analysis to determine the impact of the most extreme values in the dataset. The results show that drugs with high and low costs have a fairly small impact on cost estimates. For example, if all cost data for the drugs with the highest and lowest aggregate clinical costs are dropped from the analysis, then the full capitalized cost estimate falls by only 3.0% (3.5% if only the drug with the highest aggregate cost is dropped). The online supplement (Appendix D) further describes in detail various outlier analyses, including those that examine results when a number of high and/or low values for each clinical phase are excluded even though no one drug has uniformly high or low values across all clinical phases.

### Table 5

<table>
<thead>
<tr>
<th>Approval periods</th>
<th>Out-of-pocket</th>
<th>Capitalized</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre-human</td>
<td>Clinical</td>
</tr>
<tr>
<td>1970s to 1980s</td>
<td>7.8%</td>
<td>6.1%</td>
</tr>
<tr>
<td>1980s to 1990s</td>
<td>2.3%</td>
<td>11.8%</td>
</tr>
<tr>
<td>1990s to early 2010s</td>
<td>9.6%</td>
<td>9.2%</td>
</tr>
</tbody>
</table>

* Costs for 1970s approvals are from Hansen (1979), costs for 1980s approvals are from DiMasi et al. (1991), costs for the 1990s to the early 2000s are from DiMasi et al. (2003), and costs for the 2000s to early 2010s are from the current study.

### Table 6

<table>
<thead>
<tr>
<th>Discount rate</th>
<th>Pre-human</th>
<th>Clinical</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0%</td>
<td>472</td>
<td>1012</td>
<td>1476</td>
</tr>
<tr>
<td>2.0%</td>
<td>517</td>
<td>1044</td>
<td>1561</td>
</tr>
<tr>
<td>3.0%</td>
<td>567</td>
<td>1086</td>
<td>1653</td>
</tr>
<tr>
<td>4.0%</td>
<td>621</td>
<td>1129</td>
<td>1750</td>
</tr>
<tr>
<td>5.0%</td>
<td>679</td>
<td>1175</td>
<td>1854</td>
</tr>
<tr>
<td>6.0%</td>
<td>742</td>
<td>1222</td>
<td>1964</td>
</tr>
<tr>
<td>7.0%</td>
<td>811</td>
<td>1271</td>
<td>2082</td>
</tr>
<tr>
<td>8.0%</td>
<td>885</td>
<td>1322</td>
<td>2207</td>
</tr>
<tr>
<td>9.0%</td>
<td>965</td>
<td>1376</td>
<td>2341</td>
</tr>
<tr>
<td>10.0%</td>
<td>1052</td>
<td>1431</td>
<td>2483</td>
</tr>
<tr>
<td>11.0%</td>
<td>1145</td>
<td>1489</td>
<td>2634</td>
</tr>
<tr>
<td>12.0%</td>
<td>1246</td>
<td>1549</td>
<td>2795</td>
</tr>
<tr>
<td>13.0%</td>
<td>1355</td>
<td>1612</td>
<td>2967</td>
</tr>
<tr>
<td>14.0%</td>
<td>1473</td>
<td>1677</td>
<td>3150</td>
</tr>
<tr>
<td>15.0%</td>
<td>1600</td>
<td>1744</td>
<td>3344</td>
</tr>
</tbody>
</table>

Fig. 4. Out-of-pocket and capitalized total cost per approved new drug for new drugs and for improvements to existing drugs.
6.1. Effects of variation in the discount rate

Table 6 shows how pre-human, clinical, and total capitalized costs would vary by discount rate at one percentage point intervals. The values for a zero percent discount rate are out-of-pocket costs. In the neighborhood of our base case discount rate (10.5%), clinical cost changed by approximately $30 million, pre-human cost changed by approximately $45 million, and total cost changed by approximately $75 million for every half of one percent shift in the discount rate. In our previous study, the base case discount rate was 11.0%. At an 11.0% discount rate, total capitalized cost here was $2634 million or 3% higher than our base case result. At more extreme values for the discount rate, Table 6 indicates that total capitalized cost with a 15% discount rate was $3334 million, or 30% higher than our base case result. Similarly, a 3% discount rate (a figure often used as a social discount rate) yielded a total capitalized cost per approved new drug of $1561 million, or 39% lower than the base case result.26

6.2. Impact of cost drivers

As noted in the previous section, the full cost estimate is a function of numerous parameters that interact in a non-linear (often multiplicative) manner. That makes it difficult to isolate the extent to which changes in individual parameters alone drive changes in total costs. However, we can get a sense for which parameters had the greatest impacts, in either direction, on the change in total R&D cost between the previous study and the current one by calculating what R&D costs would have been if only a single parameter (or a set of related parameters) had changed from what it was for the previous study to what we found it to be for the current study period.

Table 7 shows our results for these thought experiments for the major parameters categorized into four groupings (direct pre-human and clinical average phase cost outlays, technical risks, average development and approval times, and the cost of capital). The base result is total cost per approved new compound for the DiMasi et al. (2003) study in year 2013 dollars ($1044 million). The current study full cost estimate is 145% higher than the base result. That change reflects the cumulative effect of all parameter changes. For the table, we examined parameter-by-parameter changes from the parameter values for the DiMasi et al. (2003) study to those values found for the current study.

The largest impact on the change in costs between the studies was driven by changes in average out-of-pocket clinical phase costs, which resulted in an 82.5% increase in full cost.27 Considering also the small difference between the studies in the estimated ratio of pre-human to clinical costs, the impact of the change in direct out-of-pocket phase costs was an increase in total cost of 85.5%. The increase in total cost was also driven to a substantial extent by much higher development risks. The overall clinical approval success rate declined from approximately one-in-five to approximately one-in-eight. That change alone accounts for a 57.3% increase in total cost. However, the impact of a lower clinical approval success rate was mitigated to a small extent by a shift in the distribution of failures to earlier in development. Taking both effects into account resulted in an increase in total cost of 47.3%. Changes in the development and approval timeline had a relatively small depressing effect on total cost. This impact was driven by a shorter pre-human testing phase and a shorter average approval phase. Average clinical development time increased modestly, and this had a relatively small impact on total cost. Overall, the effect of changes in the development and approval timeline was a 5.6% decrease in total cost. Finally, the small change in the cost of capital had a 3.1% depressing effect on total cost. The aggregation of the direct impacts across the four cost factor groupings accounted for a 124% increase in costs between the two studies. We attribute the residual increase (21%) to interaction effects.

7. Critiques, sample representativeness, and validation

Our prior study results have been questioned on a number of methodological and data grounds (Angell, 2005; Goozner, 2004; Light and Warburton, 2005a,b; Love, 2003; Young and Surrusco, 2001). We have rebutted each of these criticisms in detail in a number of venues (e.g., DiMasi et al., 2004, 2005a,b). We review the critics’ main arguments only briefly here.

Goozner (2004) and Angell (2005) reject opportunity cost calculations because they, in essence, deny that industrial pharmaceutical R&D expenditures can be viewed as investments at risk.28 These points are addressed more fully in DiMasi et al. (2004). Clearly, industrial pharmaceutical R&D meets the criteria for being considered investments that have opportunity costs. In any event, an estimate with no opportunity costs is simply the out-of-pocket cost estimate.

26 The appropriate social rate of discount for government backed expenditures has been analyzed and debated extensively in the economics literature. See for example, Moore et al., 2013 and Burgess and Zerbe, 2013. A standard reference in the cost-effectiveness literature (Gold et al., 1996) recommends 3% as the base case rate in comparing alternative medical therapies (“Therefore, we recommend that the base rate of 3% and an alternate rate of 5% be retained for a period of at least 10 years,” p.233).

27 Given the methodology, higher out-of-pocket clinical phase costs also get associated with higher out-of-pocket pre-human phase costs.

28 In the case of Goozner (2004), the claim is made that R&D expenditures are expenses rather than investments, because accountants have traditionally treated them as such for tax purposes (failing to recognize practical measurement problems underlying why this has been the practice, such as great uncertainty regarding future regulatory and commercial success). The basis offered for rejecting opportunity costs in Angell (2005, p.45) is simply the claim that pharmaceutical firms “have no choice but to spend money on R&D if they wish to be in the pharmaceutical business.”
A number of the critiques question how representative the data were for prior studies, whether tax deductions and credits must be included, and whether any FDA application for product marketing approval (as opposed to the active ingredient that is at the core of all such applications) should be taken as the unit of observation. As noted, we have addressed all of these issues in earlier publications as they relate to our prior studies. In this section we examine the representativeness of the survey firms and data used for this study, what the level of tax credits has been in relation to R&D expenditures in recent years, an analysis of molecules that have been approved for orphan drug indications recently, and we outline a variety of methods using independent data that can be used to validate our results (full details of the methods and analysis can be found in our online supplement).

7.1. Representativeness of the survey firm data

Questions about data representativeness should be framed in terms of the population from which the sample was selected. In particular, it is relevant to compare characteristics of the investigational drugs in our cost survey sample and for our cost survey firms generally to those of all drugs in our database of top 50 pharmaceutical firms, which is the relevant population. This is the main focus of the analysis in this section.

Smaller research-oriented firms may have a comparative advantage in the discovery and pre-human stages because they often have scientific researchers with close ties to the basic research underlying new classes of therapies and technology platforms. Even if this is the case, the literature indicates that smaller firms also tend to have significantly higher costs of capital, especially when they are start-ups financed by venture firms. The literature also indicates that firms with larger R&D pipelines and greater R&D experience have a higher probability of success during the costly clinical stages of drug R&D. It is not evident, therefore, that the R&D costs for compounds originating in smaller firms, whether developed internally or in alliances would be systematically lower than those originating in mid-sized and large firms. We discuss what is known about R&D metrics for small firms in Appendix E of the online supplement.

As noted, the appropriate comparator dataset for our cost survey sample is the population of investigational compounds of the top 50 pharmaceutical firms over the relevant period. There are 1442 compounds in the top 50 firm database that met our study inclusion criteria. Of these, 510, or 35.4% belonged to nine of our 10 cost survey firms. Thus, the cost survey sample (n = 106) constitutes 20.8% of the survey firm compounds and 7.4% of the population compounds.

We determined the therapeutic class distribution for the drugs in the larger dataset for the four largest therapeutic classes and one miscellaneous class (with a wide variety of drug types) for drugs in the dataset that met our study inclusion criteria and compared it to the therapeutic class distribution for our cost sample. The population shares for antineoplastic, cardiovascular, central nervous system (CNS), and systemic anti-infective drugs were 21.5%, 8.7%, 19.0%, and 8.5%, respectively. The corresponding shares for the cost survey sample were 19.8%, 9.4%, 24.5%, and 8.5%, respectively. We used a chi-squared goodness-of-fit test to compare the therapeutic class distributions for cost survey firm drugs and for the drugs of the entire set of 50 firms in the database, and found no statistically significant differences in the class shares ($\chi^2 = 2.4257, df = 4$).

We also examined the degree to which the top 50 firms in aggregate and the sample of cost survey firms agreed in terms of how molecule type (biologic versus small molecule) and the sourcing of compounds are distributed. For the set of top 50 firms, 14.6% of their self-originated investigational compounds over the study period are large molecules, compared to 13.7% for the survey firms ($p = 0.3933$). In terms of the share of investigational compounds for the study period that are self-originated (as broadly defined here), we found the share to be 74.1% for the cost survey firms and 71.1% for all top 50 firms ($p = 0.1039$).

Finally, we also examined the phase transition and overall approval success rates for the cost survey firms and compared them to the corresponding estimates for the larger dataset. The phase transition rates for just the cost survey firms were 58.0% for phase I to phase II, 36.0% for phase II to phase III, 58.2% for phase III to regulatory review, and 89.5% for regulatory review to approval. The corresponding figures for the population, as shown in Fig. 1, are 59.5%, 35.5%, 62.0%, and 90.4%. The overall clinical approval success rate for just the cost survey firms implied by the phase transition rates is 10.9%, which compares to 11.8% for the entire dataset.

7.2. Orphan drug development

Some past critiques have focused to some extent on orphan tax credits, which can provide incentives to develop some drugs for a class of indications. We examine the extent to which these tax credits and other tax issues are empirically significant in the context of drug development as a whole in the next section. Here we briefly discuss the nature of development of molecules that are approved for orphan indications and the distinction between costs for orphan drug indications and the full development costs for molecules with orphan drug indication approvals.

Compounds developed for orphan indications may well have lower clinical development costs for those indications, as trial sizes tend to be lower. The share of U.S. original new drug approvals from 2000 to 2014 for drugs with an orphan indication was 27%, and has increased in relative terms over the last 3 years of that period.

The most recent approval experience aside, the share of approvals sponsored by the set of population firms (top 50) matches closely the historical average for all approvals from 1987 to 2010 (22% for top 50 firms versus 23% of all approvals). The survey firms were nearly indistinguishable from the population non-survey firms by this metric (21% versus 23%).

31 Drugs for these indications, with some notable exceptions, tend to garner lower sales given limited patient populations. This contention is supported by recent data analysis conducted by IMS Health (Divino et al., 2014). They found that sales in the United States for orphan indications varied from only 4.8% to 8.3% of total pharmaceutical sales over 2007–2013. The analysts also projected that growth in orphan drug expenditures would slow over 2014–2018.

32 The result was calculated from information provided by the FDA on its website and included in a Tufts CSDD database of NMEs and therapeutically significant biologic approvals. The share of new drug approvals with orphan indications has increased very recently. The Orphan Drug Act was enacted in 1983, but it took several years for an appreciable number of such approvals to appear. From 1987 to 1999 the orphan drug share of all new drug approvals was 23%; the same share as for the 2000–2010 period. The orphan drug share was, however, unusually high for 2014 (41%), and above-average for 2011–2013 (approximately one-third of approvals).

33 An FDA analysis of Center for Drug Evaluation and Research (CDER) marketing applications for NMEs and new biologics for 2006 to 2010 found that approximately one-third of the applications were sponsored by small firms, and that 75% of the applications for first-in-disease therapies for orphan indications came from small firms (Lesko, 2011). Such firms may find a low R&D cost orphan disease oriented strategy attractive, given that typical sales and operating profit levels may still be sufficient to increase their market valuations.

29 The data included in the top 50 firm dataset were curated primarily from information contained in two commercial investigational drug pipeline databases that are available after payment of subscription fees. Additional information was obtained from freely available web sites. See Section 4 above for a description of data sources.

30 One of the participating firms was outside of the top 50.
The molecules approved in 2014 by orphan drug status.

Table 8

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Median</th>
<th>Range</th>
<th>% multiple indications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Orphan (n=17)</td>
<td>8.5</td>
<td>7.0</td>
<td>1–4</td>
<td>88%</td>
</tr>
<tr>
<td>Orphan cancer (n=9)</td>
<td>10.9</td>
<td>9.0</td>
<td>1–24</td>
<td>89%</td>
</tr>
<tr>
<td>Non-orphan (n=22)</td>
<td>2.7</td>
<td>2.0</td>
<td>1–7</td>
<td>73%</td>
</tr>
<tr>
<td>All approvals (n=39)</td>
<td>5.3</td>
<td>3.0</td>
<td>1–24</td>
<td>79%</td>
</tr>
</tbody>
</table>

* Therapeutic new molecular entities (NMEs) and new biologic entities (NBEDs) approved by the Center for Drug Evaluation and Research (CDER) of the United States Food and Drug Administration (FDA).

The cost survey sample contained two compounds that were approved originally for orphan indications. The average clinical period cost for these two compounds was nearly the same as the average for all sample approved compounds (94% of the overall average). One of the compounds, though, was relatively low cost, while the other was relatively high cost. This may reflect the experience of molecules approved for orphan indications generally, as total molecule cost depends not only on the approved indication, but critically, on the total number of indications (orphan and non-orphan) pursued.

To investigate this point further, we examined the development histories of all new orphan drugs and biologics approved in the United States in 2014. We studied the records for these compounds in two commercial pipeline database (IMS R&D Focus and Corellis), as well as the clinicaltrials.gov website. Table 8 demonstrates that, even with a conservative notion of what constitutes different indications, molecules approved for orphan indications were investigated in a substantial number of indications prior to original marketing approval. This was particularly true for compounds approved for treating orphan cancer indications, and, in general, the orphan drugs tended to be investigated in many more indications prior to approval than was the case for non-orphan compounds.

7.3. Taxes and R&D expenditures

As in our previous studies, the cost estimates presented here are pre-tax. Our objective was to measure the level of and trends in the private sector real resource costs of developing new drugs and biologics. As discussed in DiMasi et al. (2003), if one is calculating after-tax rates of return for R&D one would need to include the effect of taxes. Under current U.S. corporate income tax accounting practices, firms are able to deduct R&D expenses at the time they incur the costs. This is in contrast to many other investments, such as plants and equipment, which must be amortized and depreciated over a longer time period. This treatment reflects the difficulty of appropriately depreciating an intangible asset such as R&D. Later, when the company earns profits from the sales of approved pharmaceuticals it cannot depreciate the R&D investment for income tax purposes. The advantage for R&D investment over investment in plant and equipment is the timing of tax payments on net income. If one were calculating the rate of return on R&D investments one would need to take into account the tax implications. Making these adjustments is complicated by the fact that major firms operate in multiple tax jurisdictions.

In DiMasi et al. (2003) we also discussed several tax credits available in the United States to firms in the biopharmaceutical industry. In particular, we examined the Research & Experimentation tax credit for increasing qualified research expenditures, which we concluded had little impact on large multinational pharmaceutical firms. Since then, the Qualifying Therapeutic Discovery Project tax credit was created as part of the Patient Protection and Affordable Care Act of 2010 (http://grants.nih.gov/grants/funding/QTDP_PIJM); accessed 14.08.14). However, it is quite restrictive in that it applies to discovery projects for small firms with a limit of $5 million per taxpayer. Recently, the U.S. Congress Joint Committee on Taxation (2013) estimated tax expenditures for fiscal years 2012–2017 for the credit for increasing research activities, the Qualifying Therapeutic Discovery Project tax credit, and the advantage from expensing, as opposed to amortizing, research and experimental expenditures to be, in aggregate, in the range of $10 billion to $12 billion per year for fiscal years 2012–2017 across all U.S. corporations engaged in research activities. It is not clear how much of this is accounted for by the biopharmaceutical industry.

We also examined in DiMasi et al. (2003) the impact of tax credits for orphan drug research, and found them to be quite small in relation to total R&D expenditures for large pharmaceutical firms. The reporting requirements for orphan drug credits are such that many companies do not take the credit. The major financial incentive of the orphan drug program appears to be the intellectual property protection that is created from the granting of 7 years of marketing exclusivity. With respect to the magnitude of orphan drug tax credits utilized in the United States, the U.S. Congress Joint Committee on Taxation (2013) estimated that expected tax credits for orphan drug research are fairly small at between $700 million and $1 billion per year from fiscal years 2012–2017.

To put these tax credits and tax advantages in perspective, Battelle and R&D Magazine’s 2014 Global R&D Funding Forecast (http://www.battelle.org/docs/tpp/2014_global_rd_funding_forecast.pdf?sfvrsn=4; accessed 14.08.14) estimates that approximately $79 billion will be spent in the United States on R&D by the biopharmaceutical industry. Some other countries also have a number of tax credit incentives in place for R&D. However, it seems unlikely that, in aggregate, their value in relation to R&D expenditures for the biopharmaceutical industry is disproportionately higher than is the case for the United States. The Battelle and R&D Magazine’s prediction of global R&D spending by the biopharmaceutical industry is approximately $171 billion. In sum, in aggregate the value of R&D tax credits and the tax advantage of expensing versus amortizing R&D expenditures for the biopharmaceutical industry appear to be no more than one-sixth of total industry R&D expenditures (and perhaps significantly less than that).

7.4. Validation

We gathered publicly available data and performed a number of independent analyses on those data to corroborate our results. Details on methodology and data are provided in Appendix F of our online supplement. The validation efforts can be grouped into those

---

34 Analyzing orphan drug status for investigational compounds is problematic because the designation may be granted at any point during the development process. Thus, some compounds that might have been granted orphan drug status can be abandoned before that would occur.

35 Indications may be defined quite narrowly. We chose a broad definition that would limit the number of different indications pursued. Specifically, we considered all trials for the same disease and that applied to the same organ system as testing on the same indication. For example, oncology compounds may be tested as first-line treatment, second-line treatment, for refractory patients, as a monotherapy, in combination with other compounds, or for special patient populations. These cases were considered to be the same indication if they applied to the same organ (e.g., breast cancer or prostate cancer).

36 The impact may be greater for small firms if their R&D expenditures are growing more rapidly.

37 The report estimates that the industrial life sciences sector will spend $92.6 billion on R&D in the United States in 2014. However, the report also indicates that approximately 85% of all life sciences industrial expenditures are accounted for by the biopharmaceutical industry.
that utilize micro data on elements of the development process that are then used to develop growth rate estimates for portions of the process, and those that use publicly available aggregate financial time series data and compound approval statistics for biopharmaceutical firms as a check on our estimate of overall cost.

On a micro level, we examined survey data from the National Science Foundation (NSF), published estimates of trends in clinical trial complexity and clinical trial costs per subject, and published trade association times series data on R&D employment levels. Utilizing external data on costs per subject, along with clinical trial sizes and estimated clinical approval success rates from our analyses over time, we found a compound annual growth rate in real clinical trial costs between the study periods for our previous study and the current study of 9.9%, which is close to our clinical period cost growth rate of 9.2% for out-of-pocket costs shown in Table 5. We also examined measures of clinical trial complexity (number of procedures per trial) in the published literature (Getz et al., 2008; PAREXEL, 2005) and found a compound annual growth rate of 10.0% over our study period. Finally, we utilized trade association and 10-K information on R&D scientific and professional staff employment levels and NSF data on salary levels to estimate that labor costs increased at a rate of 8–9% per year across our study periods.

We examined PhRMA time series data on the R&D expenditures of its member firms. The reported growth rate for cost survey firms was 4.9%, compared to 4.2% for the PhRMA time series data for the portion of the survey period that could be compared.38 We also used the industry time series data, as we had in the previous study, in two ways to get a sense for the magnitude of overall costs per approved new molecule. In one approach, we estimated the portion of the reported time series expenditure levels that could be attributed to self-originated compound development. Next we determined the annual number of approvals of PhRMA-member firms that were self-originated. Finally, we used our study estimated time-expenditure profile to link aggregate R&D expenditures to approvals. For reasons expounded upon in the supplement, this will likely yield an upper bound estimate. Using this approach we found our out-of-pocket cost per approved molecule estimate to be 56% of the estimate derived from aggregate published industry data. The second approach focuses on the published industry self-originated R&D expenditure level for a single year, assumes that every self-originated member-firm approval (inclusive of failures) costs what we found to be our average out-of-pocket cost estimate, and uses our estimated time-expenditure profile to spread costs out over time to explain reported total R&D expenditures for the year considered. As with the previous method, the outcome would be problematic if using our average out-of-pocket cost estimate explained more than the reported aggregate R&D expenditure level. We found that this approach explained 57% of the reported expenditures.

Company total biopharmaceutical R&D expenditures reported for the cost survey are consistent with the audited financial statements of the firms in that the annual values are equal to or lower than company R&D expenses found in the financial statements.39 As another check on our overall results, we examined what survey company total biopharmaceutical R&D expenditures would be given our estimate of out-of-pocket cost per approved molecule and assuming that entry rates to survey company pipelines are in a steady state. That figure can then be compared to R&D expenditure levels reported for these firms for our cost survey (which, as noted, match audited financial statements). Full details of these calculations are in Appendix F of the supplement. Depending on assumptions, we found that we could account for between 51% and 94% of the reported total annual biopharmaceutical R&D expenditures in this way. Thus, all three approaches using aggregate R&D expenditure data suggest that our estimate of out-of-pocket cost per approved molecule is, if anything, conservative.

8. Conclusions

Studies of the cost of developing new drugs have long been of substantial interest to drug developers, drug regulators, policy makers, and scholars interested in the structure and productivity of the pharmaceutical industry and its contributions to social welfare. The interest has been strong and growing over the last few decades during which cost containment pressures for drugs approved for marketing have expanded and concerns have been raised about industry productivity in an environment in which industry structure has been evolving (Munos, 2009; Pammolli et al., 2011). The changing industrial landscape has featured consolidation among large firms, growing alliances among firms of all sizes, and the growth of a small firm sector.

We have conducted the fourth in a series of comprehensive compound-based analyses of the costs of new drug development. In the last study we reported average out-of-pocket and capitalized R&D costs of $403 million and $802 million in 2000 dollars ($524 million and $1044 million in 2013 dollars), respectively. For our updated analysis, we estimated total out-of-pocket and capitalized R&D cost per new drug to be $1395 million and $2558 million in 2013 dollars, respectively. To examine R&D costs over the entire product and development lifecycle, we also estimated R&D costs incurred after initial approval. This increased out-of-pocket cost per approved drug to $1861 million and capitalized cost to $2870 million. We validated our results in a variety of ways through analyses of independently derived published data on the pharmaceutical industry.

Our pre-approval out-of-pocket cost estimate is a 166% increase in real dollars over what we found in our previous study, and our capitalized cost estimate is 145% higher. Roughly speaking, the current study covers R&D costs that yielded approvals, for the most part, during the 2000s and early 2010s. Our previous study (DiMasi et al., 2003) generally involved R&D that resulted in 1990s approvals. The compound annual rates of growth in total real out-of-pocket and capitalized costs between the studies are 9.3% and 8.5%, respectively. These growth rates are both somewhat higher than those we found for the two previous studies (7.6% and 7.4%, respectively). Growth in out-of-pocket clinical period costs have moderated some from the 1990s, but the growth rate is still high at 9.2%. While the compound annual growth rate for out-of-pocket pre-human costs declined substantially for the previous study (from 7.8% to 2.3%), this study showed a substantially higher growth rate for pre-human costs in the new century (9.6%).

The success rate found for this study is nearly 10 percentage points lower than for the previous study. The overall change in the risk profile for new drug development by itself still accounted directly for a 47% increase in costs. It is difficult to know definitively why failure rates have increased, but a number of hypotheses worthy of testing come to mind. One possibility is that regulators have become more risk averse over time, especially in the wake of high profile safety failures for drugs that have reached the marketplace (most notably, VioxxTM, but there have been others as well). It may also be the case that the industry has generally focused more in areas where the science is difficult and failure risks are high as a result (Pammolli et al., 2011). Finally, the substantial growth in identified drug targets, many of which may be poorly validated, may have encouraged firms to pursue clinical development of more

38 As explained in the Supplement, the growth rate for the PhRMA time series may somewhat underestimate the true growth rate.
39 Biopharmaceutical R&D expenditures may be less than total company R&D expenditures if the firm engages in non-biopharmaceutical R&D.
compounds with an unclear likelihood of success than they otherwise would.

As can be seen from results cited in the supplement developed external to this study, as well as our own data, out-of-pocket clinical cost increases can be driven by a number of factors, including increasing clinical trial complexity (Getz et al. 2008), larger clinical trial sizes, inflation in the cost of inputs taken from the medical sector that are used for development, and possibly changes in protocol design to include efforts to gather health technology assessment information and, relatedly, testing on comparator drugs to accommodate payer demands for comparative effectiveness data. The expansion of the scope of the clinical trial enterprise during our study period is illustrated by the finding in Getz and Kaitin (2015) that for a typical phase III trial information had been gathered by sponsors on nearly 500,000 data points in 2002, but more than 900,000 data points in 2012.

Finally, it is difficult to assess whether and how regulatory burdens may have impacted changes in industry R&D costs over time. However, occasionally, an exogenous shift in the types and amount of information perceived as necessary for regulatory approval for particular classes of drugs can be instructive. For example, during our study period the FDA issued guidance (Food and Drug Administration, 2008) for the development of drugs to treat diabetes in late 2008 that highlighted a need to better assess and characterize cardiovascular risks for this class of compounds, after a number of cardiovascular concerns emerged regarding a previously approved drug (Avandia®). A number of development metrics positively related to R&D costs can be examined pre- and post-guidance. DiMasi (2015), for example, found that average U.S. clinical development times increased from 4.7 to 6.7 years for diabetes drugs approved in the United States from 2000–2008 to 2009–2014, respectively. In addition, Vierck and Boudes (2011) found that the number of randomized patients and patient-years in NDAs for diabetes drugs approved from 2005 to 2010 increased more than 2.5 and 4.0 times, respectively, before and after the guidelines were issued. Our sample data show that diabetes drugs were among the most costly (particularly for phase III [92% higher than the overall average]).

Our analysis of cost drivers indicates that the rate of increase observed in the current study was driven mainly by increases in the real out-of-pocket costs of development for individual drugs and by much higher failure rates for drugs that are tested in human subjects, but not particularly by changes in development times or the cost-of-capital. Continued analysis of the productivity of biopharmaceutical R&D should remain an important research objective.

Appendix. Supplementary data
Supplementary data associated with this article can be found in the online version, at http://dx.doi.org/10.1016/j.jhealeco.2016.01.012.

References


