Extraordinary claims require extraordinary evidence

Joseph A. DiMasi\textsuperscript{a,\ast}, Ronald W. Hansen\textsuperscript{b}, Henry G. Grabowski\textsuperscript{c}

\textsuperscript{a} Tufts Center for the Study of Drug Development, Tufts University, 192 South Street, Suite 550, Boston, MA 02111, USA
\textsuperscript{b} William E. Simon Graduate School of Business Administration, University of Rochester, NY, USA
\textsuperscript{c} Department of Economics, Duke University, NC, USA

\textit{JEL classification:} L65; O31

\textit{Keywords:} R&D cost; Pharmaceutical industry; Data validity

We thank Light and Warburton (2005) [hereafter LW] for providing a context for us to further discuss and clarify our studies on the cost of new drug development (DiMasi et al., 2003 [hereafter DHG]; DiMasi et al., 1991b [hereafter DHGL]; Hansen, 1979). We disagree, however, with their analyses, comparators, and conclusions. Although LW suggest, through the invocation of selected reported cost figures and other statistics, as well as questions about what may or may not have been included in our analyses, that new drug R&D costs are really substantially lower than are our estimates, their central criticism is that some of the underlying data are proprietary, and they claim, unvalidated. However, they do not address, or even acknowledge, the numerous validations of our results obtained from alternative data sources and analyses that were reported in DHGL (pp. 131–132, 137–140), in US Congress, Office of Technology Assessment (OTA) (1993, pp. 54–67), and in DHG (pp. 176–180).

We cross-checked our data against data from a variety of sources, including government healthcare statistics, other investigators’ analyses, data from pharmaceutical firms on drug development activities compiled by third-party vendors for use as business intelligence tools, internal checks against our own long-standing databases of drug development metrics, published industry-level trade association data, and audited company financial statements. The validations vary from checks on the comparability of key characteristics of the investigational compounds of survey and non-survey firms to data collected and reported by others...
that can be analyzed to check on rates of change in key cost components and on the absolute levels of R&D costs.

The OTA assessed the validity of the results from both the Hansen (1979) and DHGL studies. They ran numerous checks, including their own surveys of the growth in clinical trial sizes and changes in animal testing costs, industry R&D employment and National Science Foundation (NSF) survey data for scientists in life science industries, and an analysis of aggregate industry R&D expenditure and Food and Drug Administration (FDA) approval data. The OTA also considered the development timelines and success rates in both studies. Although they acknowledge that R&D cost estimates will always have some uncertainty associated with them, in the final analysis, the OTA was satisfied with accuracy of the underlying data. In particular, they noted, “To summarize, the estimates by DiMasi and colleagues of the cash outlays required to bring a new drug to market and the time profile of those costs provide a reasonably accurate picture of the mean R&D cash outlays for NCEs first tested in humans between 1970 and 1982.” (OTA [1993, p. 66]).

For the most recent study (DHG), we applied nearly all of the corroborative tests in DHGL and OTA (1993), as well as additional analyses of data from independent sources, to validate our results. The checks variously apply to growth rates relative to cost levels in our previous study and directly to absolute cost levels for the current study. As before, the tests support our results.

1. LW’s bullet points on data validity

Although they are not all completely independent of one another, LW offer six bullet points that purportedly raise doubts about the validity of the data examined in DHG. Their tax bullet point, however, actually does not have anything to do with data validity. We will address that concern first, followed by LW’s other bullet points in the order in which they were presented.

1.1. Taxes

LW and others (e.g., Public Citizen, 2001) argue that our results are not valid because we did not reduce our out-of-pocket costs for the deductions taken for R&D expenditures on US corporate income tax returns and for R&D tax credits. We explicitly labeled our estimates as pre-tax costs, so that neither data validity nor accuracy is an issue here. Our paper in fact contains an extensive discussion of tax issues (DHG, pp. 176–180), but it is worth mentioning a few salient points here.

The primary objective of our study was to estimate the private sector economic costs needed to get a new drug from discovery to market for one period, and to compare the results to those for earlier periods. Since tax structures change over time, tax-adjusted figures can misrepresent the extent to which resource costs have changed. We do note, however, that when explicitly considering the profitability of new drug development one should consider...
after-tax cash flows. Indeed, two of the authors have done so in rate of return studies (Grabowski and Vernon, 1990, 1994; Grabowski et al., 2002).

However, even if the intent is to measure the effective R&D cost to firms, as a stand-alone measure of R&D costs the approach advocated by LW and Public Citizen (2001) is highly misleading. In particular, it suggests that the deductions for R&D expenses allowed on corporate income tax statements are really tax breaks, whereby the public defrays a substantial amount of a company’s costs. This corporate welfare view of the R&D income tax deduction (which is implicit in both LW and Public Citizen, 2001) is erroneous. The corporate income tax must be understood as a tax on net income (i.e., corporate accounting profits). Deducting R&D and other costs from revenues is effectively just part of the method by which the targeted tax base (profits) is determined.

Alternatively, R&D tax credits (which apply to all industries) are tax subsidies designed to spur the growth of industrial R&D. However, as noted in our paper, it is difficult to include them with real precision in the types of estimates in which we were interested, and they appear, in any event, to have not been very financially significant for the type of firm that we analyzed. As reported in our paper, recent audited financial data for so-called big pharma firms suggests that realized R&D tax credits have been on the order of 2% of R&D expenditures. In addition, analysis of Congressional Research Service (CRS) data on orphan drug tax credits indicates that in aggregate these credits are much less empirically significant for big pharma firms than are even the R&D tax credits (DHG, p. 175).

1.2. Company cost accounting

LW suggest, without any evidence or even descriptive examples, that company cost allocation methods differ so much (“the degree of potential variation is large”) that this can make “any point estimate misleading.” It is unclear why this would bias a point estimate, as opposed to increase its variability. In any event, the survey respondents were required to meet our reporting needs regardless of their accounting systems. Our estimates depend mainly on when expenditures occurred over time during the development of a drug. It seems clear that they had the ability to distinguish whether costs were incurred prior to or subsequent to the initiation of clinical testing, or prior to or subsequent to marketing approval.

In assessing our previous study, the OTA addressed this issue in their report and dismissed it, noting “More careful analysis of project cost accounts and adjustment of estimates for different cost allocation rules would give a more consistent estimate across firms, but it is unlikely the resulting estimates of cash outlays would be very different, and probably not lower” (OTA, 1993, p. 66). The methodology for our current study was the same as for the earlier study. There is no reason to think that the conclusion in OTA’s report should be any different now.

1.3. Respondent veracity

With regard to the truthfulness of survey responses LW point to the OTA report, where the possibility that firms might overstate their costs for political reasons was noted. The
uses that the industry has made in the public arena of estimates of R&D cost and other drug development parameters have essentially consisted of assertions about the basic nature of drug development. Those familiar with the commercial drug development process uniformly recognize that, on average, drug development is in some sense costly, risky, and lengthy. There appears to be little reason for firms to fabricate to achieve particular results when the acknowledged reality supports such basic claims. Furthermore, the firms understood that we and other parties would subject the results to various validity checks. In any event, the numerous data checks from external and internal sources mentioned above serve to validate our results. Also, as noted above, after conducting their own checks, the OTA concluded that the data used for our previous study were valid.

1.4. Sample size and selection of firms

LW refer to our sample sizes as small. However, the sample sizes for the mean out-of-pocket costs for each of the clinical phases are large enough to meet the traditional statistical threshold (30 observations) for assuming large sample properties when analyzing sample means. In addition, the samples used to estimate development times, approval times, the clinical success rate, and clinical phase attrition rates each consisted of hundreds of observations across many firms.

As stated in DHG, the cost survey firms were selected from a Tufts Center for the Study of Drug Development (CSDD) ongoing database of investigational drugs. The information in this dataset includes development milestone dates, but not costs. The survey data for this database are checked against and supplemented by information in a number of commercially available databases of investigational drugs designed to provide pharmaceutical firms with business intelligence (Iddb3, NDA Pipeline, PharmaProjects, and R&D Focus). Data indicate that the CSDD dataset covers a high percentage of drugs taken into clinical testing in the United States (DiMasi et al., 1991a, 1994; DiMasi, 2001). Firms that had contributed the relevant milestone data for our study period were those that were selected for surveying.

Some firms declined to respond to our cost survey because of the burden involved in retrieving archival records and reporting in the manner we required. This task is more burdensome if their accounting systems had changed over time or if they were still preoccupied with transitional issues associated with a merger. As we reported in DHG, firms that did not respond offered one of these difficulties as the reason for declining. We see no reason, however, to suspect that the absence of the non-responding firms would bias the results either way.

The substantial mobility of industry scientists and managers and the growth in outsourcing parts of the development process during our study period argues against significant differences by firm in the costs of developing a drug of any given type. Size potentially could be a factor, but the survey firms were well distributed across size levels for traditional established pharmaceutical firms. As was the case for DHGL, the responding firms for DHG accounted for a substantial share of the R&D expenditures of the traditional pharmaceutical industry and the distribution of drugs by therapeutic class for these firms was nearly identical to that for the industry as a whole. For these and other reasons, the OTA noted the following regarding our previous study: “Thus, the sample of firms appears to pose no serious threat to the validity of the study” (OTA, 1993, p. 55).
1.5. Universe of new drugs

LW suggest that costs for the type of drugs that we sampled (self-originated new chemical and biopharmaceutical entities) are much higher than for “new drugs” as a whole. Even if true, this would not be a problem with the validity of our data since we clearly described what we sampled. Specifically, though, LW maintain that the costs associated with US new drug application (NDA) approvals for products that are not new molecular entities (NMEs) and the costs for licensed-in NMEs are substantially lower than are costs for the type of drug we sampled. LW also used the totals for FDA NDA approvals to suggest that self-originated new chemical and biopharmaceutical entities are a small percentage of all “new drugs”.

LW implicitly make the claim, echoed explicitly elsewhere, that in essence our estimates apply just to drugs “which require the most expensive types of research” (Public Citizen, 2001, p. 3). It is, therefore, worth examining the propriety of reporting an average cost figure that uses an NDA approval (as opposed to an NME approval) as the unit of observation. NME approvals are only a fraction of all NDA approvals. There are several reasons (both technical and conceptual) why spreading R&D expenditures over the FDA totals for NDAs is inappropriate.

Counting applications approved by the FDA is not at all the same as counting new drugs in any meaningful sense. The FDA data on NDA approvals covers a hodgepodge of widely differing types of regulatory actions. Aside from NMEs (products containing active ingredients that have never been approved for marketing), the NDA data apply to approvals for new salts or esters, new formulations or new indications for existing drugs, new combinations (where all active ingredients have been previously approved), a new manufacturer for an existing drug, and very old drugs that have been marketed without an approved NDA. All non-NME approvals are for drugs (i.e., active ingredients) that have already reached the marketplace, and some do not even represent new drug products.

One of the authors (DiMasi) is currently engaged in a project to examine the nature of and trends in the non-NME NDA approvals made by the FDA over the last several decades. At this writing data are still being collected, but we can report some of the findings for the last few years. Of the 2001–2003 non-NME approvals, 53% were for drugs that were first approved more than 14.5 years earlier, 40% were first approved more than 20 years earlier, and one-third were first approved more than 25 years earlier. Not surprisingly, then, our

---

2 LW also question the classification of drugs into self-originated and licensed. The classification is based on information in ongoing CSDD surveys that were independent of our cost survey. Since CSDD studies have traditionally reported survey results only in aggregated form, there does not appear to be any reason why firms would not be truthful about this element. In addition, CSDD checks its data and supplements it where necessary from information contained in commercial databases. These databases report on licensing agreements.

3 Most new indications are approved via a supplemental new drug application (SNDA), but some are approved via an NDA. Since 1994, the FDA has not included the new indication NDAs in their reported NDA totals. They are counted, along with the other new indication approvals, in the FDA’s efficacy supplement totals.

4 The significance of the 14.5-year threshold is that under U.S. law some of the patent life for a new drug that is lost during clinical testing and regulatory review can be restored. However, the maximum effective patent life (time since approval to loss of patent protection) that is allowed with patent term restoration is 14 years. An additional
analyses reveal that many of the non-NME NDA approvals were obtained by companies that had no link (i.e., through licensing, co-development, or acquisition) with respect to the drug in question to the firm that sponsored the original NME (43% of the 2001–2003 approvals). For the most part, the unrelated firms are generally small specialty pharmaceutical firms and generic drug manufacturers. In addition, some of the approvals are for drugs that are over-the-counter (OTC) or for Rx-to-OTC switches (8% of the 2001–2003 approvals). Thus, any analysis (e.g., Public Citizen, 2001) that attempts to link the prescription drug R&D expenditures of a subset of the firms (e.g., full members of the industry trade association, PhRMA) that obtain non-NME NDA approvals to the total number of non-NME NDA approvals is methodologically flawed.

Aside from technical problems associated with the indiscriminate use of the annual FDA totals on NDA approvals, there is a conceptual problem with averaging over all NDA approvals. The unit of observation for our estimates is a drug (active ingredient), not a drug product. Firms typically offer drugs in many different product presentations. Many NME approvals, in fact, cover a number of different strengths of a drug, each potentially sold at a different price and serving different consumer needs. Some NME approvals also cover several formulations.

Most of the non-NME NDA approvals are for new formulations of existing drugs and combinations of already-approved active ingredients, but many of these approvals simply reflect the firm filling out its product line to better serve consumer needs (e.g., oral solutions for children and others who have difficulty swallowing tablets or capsules). In any event, these approvals represent incremental additions and improvements to the product lines for existing drugs. These approvals concern new drug products, not new drugs.

While the incremental costs associated with various types of product line extensions are of some interest (mainly to developers who must decide if they are worth incurring\(^5\)), the extensions should not be treated alongside NME approvals as separate occurrences of the same unit of observation. The costs of incremental additions to product lines are usually significantly lower than they otherwise would be because many activities and investigations need not be repeated, and the knowledge generated by previous development of the drug will inform the firm of approaches that will likely not be fruitful. The cost of obtaining a non-NME NDA approval is, therefore, inextricably linked to the cost of obtaining the original NME approval. One cannot develop a new formulation without first having spent funds on discovering and developing the drug in its original formulation.

We believe that the best way to deal with R&D on line extensions is to use the active ingredient as the unit of observation and consider costs incurred for that unit over the entire lifecycle of the drug. This is also the way that drug companies view the process. One can usefully divide expenditures into costs incurred prior to and subsequent to the first approval of a drug. We did so in DHG and found an average capitalized cost per approved drug (not

---

\(^6\) The line extension will often cannibalize to some extent sales of existing product presentations of the drug, while also expanding the market for the drug to some degree.

---

six months can be obtained for testing in pediatric populations. Average effective patent lifetime for drugs first approved from the mid-1980s to the mid-1990s was 10–12 years (Grabowski and Vernon, 1996; Shulman et al., 1999). Thus, it is likely that the many of the non-NME NDA approvals are for drugs that were off-patent long before the non-NME approval.

---
drug product) over the entire drug lifecycle of US$ 897 million (with post-approval costs that are US$ 140 million out-of-pocket and US$ 95 million on a capitalized basis).\(^6\)

In discussing our full lifecycle cost estimate (US$ 897 million) in their concluding paragraph, LW claim that our post-approval cost estimate is constructed entirely from data on the costs of so-called “seeding trials.” In fact, none of the data used to obtain our estimate should include costs for seeding trials. Marketing costs are not incorporated in the data we used to derive the estimate. Marketing and sales departments pay for seeding trials. Such trials are not budgeted to R&D.\(^7\) What our post-approval cost data do include are expenditures that pharmaceutical firms routinely incur from pursuing approval by regulatory authorities of new indications, new dosage strengths, and new formulations for already-approved drugs. Such approvals must be based on valid scientific data.

Finally, based on a ratio that we employed in DHGL as part of a validation analysis that used aggregate published R&D expenditure and approvals data, LW maintain that R&D costs for licensed new drugs are much lower than for self-originated new drugs. They have grossly misconstrued the meaning of the ratio. The ratio accounted for R&D expenditures per approval for the self-originated and for the licensed-in or otherwise acquired drugs of the survey firms. Companies typically license-in drugs after discovery, preclinical development, and early clinical testing costs have already been incurred by other firms.\(^8\) Those earlier substantial expenditures were not included in the ratio. We do not have any reason to conclude that the full average costs for licensed or otherwise acquired drugs are significantly different than are the costs for self-originated drugs. Indeed, an analysis that uses published aggregate data on expenditures and approvals for self-originated and acquired drugs combined (and that corrects for errors in a similar analysis in Public Citizen, 2001), results in a range for the out-of-pocket cost per approved NME that was virtually identical to the range that we found in DHG using aggregate data for just self-originated drugs (DiMasi et al., 2004).

1.6. Government contracts to industry

LW claim that our company expenditure data includes contract funds paid by government agencies to the firms. This is not the case. The cost data that we obtained reflect only private resources. We also excluded government grants from the PhRMA industry R&D expenditure data that we utilized in our validation checks that examined published aggregate data. Even if government funds had not been excluded, the impact would have been trivial. For our validation analysis, we had examined the PhRMA annual R&D expenditure data from 1978 to 2000. During this period, the annual share of pharmaceutical R&D expenditures

---

\(^6\) Of course, if costs are to be compared to returns, then returns should be cumulated over all product presentations of the drug (as was done in the rate of return studies noted above).

\(^7\) Firms would put themselves in great jeopardy if they classified as R&D expenses what were really marketing expenditures. Since companies claim R&D tax credits, the IRS and other taxing authorities would take a very dim view of such practices. In addition, firms would potentially expose themselves to charges from the SEC and other similar foreign regulators that they misled investors, as well as to potential shareholder lawsuits.

\(^8\) The licensees “pay” for the value of this earlier R&D through upfront fees, milestone payments, and royalty fees.
accounted for by government grants varied from 0 to 0.6%, with a mean of 0.2% (and no trend).

2. Alternative cost figures in LW

To suggest that the overall R&D cost estimate in DHG is greatly overestimated, LW compare the full cost estimate in DHG directly to figures from three reports (Public Citizen, 2001; Love, 2003; OTA, 1993). The figures in the OTA report, however, cover the period analyzed for our previous study (DHGL). Thus, the comparison to our current study is inappropriate. In addition, as noted above, the OTA affirmed the validity of the data we used for the previous study.

Neither the Public Citizen (2001) nor the Love (2003) reports went through the usual anonymous peer review process. If they are to be used as credible alternatives, then they should be scrutinized as well. LW provides no such scrutiny. Before we indicate why the figures in these reports are either flawed or not comparable to our results, we note that two other recent reports put pharmaceutical R&D costs at levels that are similar to our results.

In a report on the future of pharmacogenomics, Boston Consulting Group (2001) provided an estimate of pharmaceutical R&D cost (US$ 880 million) based on their modeling and on interviews with industry executives that was modestly higher than our average cost. Bain & Company also recently reported estimates of the cost of getting a new drug to market (Gilbert et al., 2003). Based on their Bain Drug Economics Model, 2003, they estimated that the cost per approved compound was US$ 1.7 billion for 2000–2002 new drug approvals. This figure, however, included US$ 250 million in launch costs (which are not part of our R&D cost estimates). They also report figures for 1995–2000 approvals (a more relevant period for comparison to our study) that are much closer to our results.9 Like the Public Citizen (2001) and Love (2003) reports, these two reports were not vetted through the usual academic peer review process. However, they should also be noted here for completeness.

The Public Citizen (2001) report, which was issued on the advocacy group’s Web site in July 2001, purported to show that pharmaceutical industry R&D costs are much lower than what the industry had claimed. They took two basic approaches to challenging a figure that industry had been using (an extrapolation of our 1991 study estimate). The first approach simply consisted of ignoring the time costs of new drug development (described as “theoretical”, and so presumably not real) and then reducing the out-of-pocket cost from our previous study according to a statutory corporate income tax rate.10 Making the argument that opportunity costs for R&D investments exist and are relevant hardly seems necessary in

---

9 Their total, inclusive of launch costs, for 1995–2000 approvals is US$ 1.1 billion. Although they do not specify launch costs for this period, they note that launch costs increased nearly 50% for the later period. Using US$ 167 million as the launch cost value for 1995–2000, yields an R&D cost per approved drug of US$ 933 million. The average approval date for drugs in our sample was in 1997, but our sample included approvals back to 1990 and the data indicate that R&D costs have been increasing over time. In addition, although not stated, the Bain results likely are in year 2003 dollars.

10 Public Citizen later applied the same approach to the out-of-pocket cost in DHG and posted the result on its Web site. It is worth noting that acceptance of their no-opportunity cost/tax-adjusted figure is an implicit endorsement of the validity of our underlying data.
a journal such as this.\textsuperscript{11} In any event, our studies report both out-of-pocket and capitalized costs. We have already addressed the tax issue above in the context of one of LW’s bullet points.

In theory, Public Citizen’s second approach to estimating drug development costs could reflect on the validity of our data. However, their uses of US industry trade association (PhRMA) R&D expenditure and FDA approvals data are deeply flawed. Their ratios of expenditures to approvals are biased downward substantially. The numerators exclude many relevant expenditures and the denominators are inflated by many approvals that were not obtained by the firms that provided the data for the numerators.\textsuperscript{12} Thus, their estimates use incomplete and mismatched data in a way that dramatically understates R&D costs. Space limitations preclude presenting a comprehensive discussion of the flaws in Public Citizen (2001) here, but the interested reader can find our analysis on the Web (DiMasi et al., 2004).

LW also cite Love (2003) as evidence that R&D costs may be significantly less than our full capitalized cost estimate. There are numerous figures discussed in Love (2003), but the only cite therein to anything that is close to a full cost estimate is taken from a report of the Global Alliance on Tuberculosis Drug Development on the economics of tuberculosis drug development (Global Alliance for TB Drug Development, 2001 [hereafter TB Alliance]). We assume that numbers from this report are those to which LW refer.

The TB Alliance report provides a range of numbers that cannot be legitimately compared to our estimates of the average cost of new drug development. The report projects costs for what is essentially a very special case—the development of a hypothetical drug with unspecified chemical and pharmacological properties for use as a treatment for tuberculosis. The development program is modeled after the development of an antibiotic (rifapentine) that was approved under the orphan drug and accelerated approval programs at the FDA for a tuberculosis indication. While such a model may well be appropriate to use for a new tuberculosis drug, it certainly is not a representative, or average, drug development program for new drugs as a whole.

In addition, the TB Alliance report figures represent projections of costs for a single indication. Many new drugs are investigated clinically in a number of indications before the drug has been approved for anything.\textsuperscript{13} Our estimates are costs per drug, not costs per indication. This distinction can at least partially explain why the number of pre-approval clinical trial subjects posited in the report is only a quarter of what we found in DHG for the average number of subjects across all pre-approval clinical phases, or of an independent estimate of the average number of subjects in NDAs for NMEs approved during 1998–2001 (PAREXEL, 2002, pp. 108–111).

There are additional reasons why the overall cost figures in the TB Alliance report are not appropriate comparators to our total capitalized cost estimate (including the fact that costs

\textsuperscript{11} See, however, DiMasi (2002) for a brief discussion of the rationale for capitalizing pharmaceutical R&D costs.

\textsuperscript{12} In addition, Public Citizen uses the approved NDA as the unit of observation. As noted above, this is conceptually, as well as technically, inappropriate.

\textsuperscript{13} To use a drug tested for a tuberculosis indication as an example of the extent to which this can occur, the investigational drug SRL-172, which was studied as an adjuvant to standard tuberculosis therapy, and which failed in that indication in three large phase III trials (PJB Publications Ltd., 2002), has, according to commercial databases for investigational drugs (PharmaProjects, IDdb3), also been tested clinically in at least 14 other indications ranging from seasonal allergic rhinitis to a variety of cancers. The drug has not been approved for any indication.
in the report were discounted back to lead candidate optimization, as opposed to capitalized forward to marketing approval). As was the case for the Public Citizen report because of space limitations we refer the interested reader to our Web document (DiMasi et al., 2004).

3. Conclusions

LW question the validity of the data used to develop our pharmaceutical R&D cost estimates. However, they have not offered any evidence that the data are invalid, and they have ignored numerous validation checks using a wide variety of internal and external sources that corroborated our results. A number of their stated concerns about data validity also do not actually reflect on the validity of our underlying data. Finally, cost figures from other sources that they cite as alternatives to our results are either methodologically or conceptually flawed, or they are not appropriate comparators. It is always possible to get a “bad sample”, but we have no reason to believe that our estimates are not reasonably accurate for the period analyzed.

References


PJB Publications Ltd., 2002. SRL-172 fails again in tuberculosis. Scrip 2788 (October (9)), p. 27.

