Chapter Title: Is the Pharmaceutical Industry in a Productivity Crisis?

Chapter Author: Iain M. Cockburn

Chapter URL: http://www.nber.org/chapters/c0032

Chapter pages in book: (p. 1 - 32)
Is the Pharmaceutical Industry in a Productivity Crisis?

Iain M. Cockburn, Boston University and NBER

Executive Summary

Rising R&D expenditures and falling counts of new drug approvals since 1996 have lead many observers to conclude that there has been a sharp decline in research productivity in the pharmaceutical industry over the past decade. A close look at the underlying data, however, suggests that these trends are greatly exaggerated: properly measured, research output is unlikely to have fallen as much as these figures imply, while trends in R&D expenditure are seriously overstated by failing to account for inflation in R&D input costs. Some of the increase in R&D investment is a necessary, indeed welcome, response to new technological opportunities and can be expected to deliver a handsome return of innovative drugs in future years. The rising cost per new drug approved is nonetheless a serious cause for concern, particularly where this is driven by transactions costs and other inefficiencies in the market for basic research, and by late-stage abandonment of drug development projects on purely economic grounds. Policies that make "small" markets more attractive, build capacity in translational medicine, reduce the cost, time, and uncertainty of regulatory review, maximize access to basic research, and encourage greater cooperation and collaborative research within the industry that can all contribute to greater R&D efficiency.

I. Introduction: Crisis, What Crisis?

By many accounts, the pharmaceutical industry is experiencing a severe decline in research productivity. More and more money is being invested in R&D, but the rate at which new drugs are introduced is failing to keep pace. Recent years have seen a steady flow of reporting in trade journals and mass media referring to drug companies' "dry," "weak," or "strangled" pipelines, and as the FDA's books closed for calendar 2005 with only 20 new drug approvals, the New York Times concluded recently that the "research drought" has grown worse.
"The number of new drugs approved by the FDA has fallen by more than half since 1996..." while "R&D spending in the pharmaceutical industry more than doubled." Figure 1.1 replicates the New York Times' graphical display of these data.

Similar trends are apparent in worldwide data, with a recent survey by The Economist reporting estimates of global industry R&D spending rising from $30bn per year in 1994 to $54bn in 2004, with global drug launches falling from 40 per year to 26 per year over the same period. The obvious inference to be drawn from these figures is that the "bang for the buck" in biomedical research is in sharp decline. This is particularly puzzling in the light of the extraordinary advances in biomedical science in recent decades. Generous public funding of research in the U.S. and elsewhere has expanded fundamental biomedical knowledge at a remarkable rate. Landmark events like the sequencing of the human genome are representative of major advances in our understanding of basic biochemical processes and molecular and cellular biology. Yet so far the "payoff" in terms of new drugs has been disappointing.

Pharmaceutical R&D has paid off handsomely in the past, most visibly in areas like depression, cholesterol, and ulcers where new drugs have had a huge impact on the practice of medicine, costs of treatment, and health outcomes. More broadly, statistical studies show an historical correlation since the 1950s between the number of new drugs

![Figure 1.1](image_url)

The productivity crisis.
introduced and declines in mortality and other health indicators across a wide range of diseases and health problems. Nonetheless, progress has been disappointing in other areas. No new broad spectrum antibiotics have been marketed in almost 40 years, and chronic diseases and disorders such as atherosclerosis, diabetes, obesity, Alzheimer’s, Parkinson’s, and schizophrenia still lack effective and well-tolerated treatments.

The apparent disconnect between progress in basic science and development of new drugs has led regulators, academic researchers, investment analysts, and many other observers to the conclusion that the mechanism for translating science into drugs—profit-oriented research and development by pharmaceutical companies—has broken down. A report issued by the FDA in 2004, for example, expressed "growing concern that many of the new basic science discoveries made in recent years may not quickly yield more effective, more affordable, and safe medical products for patients," citing falling numbers of applications for approval of new drugs, and placing the blame squarely on an "increasingly challenging, inefficient, and costly" product development path."

If this productivity crisis is serious, it presents policy makers with some difficult questions. Taxpayers around the world support well over $25bn per year of biomedical research: do these apparently poor outcomes justify continued public investment at its current scale? What, if anything, can be done to turn the industry around?

Closer examination of the data in the light of underlying trends in the industry may give, if not answers, at least some further insight into these issues. This essay offers a glass-half-full, glass-half-empty interpretation of the productivity crisis. On the one hand, any decline in "true" research productivity is almost surely severely exaggerated by looking simply at ratios of new drugs approved to dollars spent on R&D. Recognizing the flaws in this measure leads to the conclusion that things are not as bad as the media reports suggest. Innovation in the industry, properly measured, is unlikely to have fallen as drastically as simple comparisons of counts of annual drug approvals to trends in R&D spending indicate. Quality-adjusted output, measured in ways that capture the full value of new drugs to consumers could even be rising. On the input side, real R&D spending has not risen as fast as the nominal totals, and some substantial portion of the increase in R&D is good news rather than bad news, reflecting a rational and welcome response by industry to a massive expansion of technological opportunities, and efforts to better address patients' needs.
On the other hand, falling rates of new drug approvals may reflect increasing focus on more challenging diseases, failure to invest in human and institutional capacity in "translational medicine," problems with adapting processes and standards for regulatory review to new research technologies, and reluctance of drug companies to bring forward products with low sales potential. And on the input side, some of the increase in R&D spending may reflect socially costly effects of the "dis-integration" and restructuring of the industry over the past few decades, as well as inefficiently low levels of collaboration and sharing of precompetitive data. Some of these causes of poor productivity performance suggest opportunities for policy interventions.

II. Measuring Productivity

Economists usually think about productivity as the ratio of the "output" of a process to some measure of the "inputs" utilized. Interpreting figure 1.1 in terms of outputs and inputs conveys a clear, and ominous, message about productivity. Since input (R&D expenditures) is rising much faster than output (the number of new drug approvals), their ratio is falling—with the clear implication that the productivity of biopharmaceutical R&D is in sharp decline.

For some economic activities, this type of calculation is easy to perform, and the results are straightforward to interpret. For a simple, repetitive labor-intensive task such as digging ditches, output per man-hour gives a meaningful measure of productivity. But serious difficulties emerge when the process has multiple, heterogeneous, and long-lived outputs and inputs, when some inputs or outputs are not directly measured or priced (e.g., knowledge spillovers), and when output is realized at a different point in time from when the inputs are utilized. These problems are particularly acute in biopharmaceutical R&D, where R&D expenditures are incurred over many years prior to product launch, advances draw extensively on un-priced spillovers from basic research (often conducted in the public sector), and where simple counts of regulatory approvals of particular products attributable to an R&D program may be a poor proxy for that program's true output.

Thus, tempting though it may be to look (explicitly or implicitly) at drug approvals per dollar of R&D spending as a measure of research productivity, this calculation can be seriously misleading. To make sense of the trends portrayed in figure 1.1, a closer look at both the numerator and the denominator of the productivity ratio is necessary.
III. Measuring Innovative Output in the Pharmaceutical Industry

Counts of Drug Approvals

The annual number of new drug approvals is a popular way to measure innovative output. Most analysts are careful to distinguish between regulatory approvals of drug products containing novel active ingredients (new molecular entities or "NMEs") and the much larger volume of approvals of products which are minor chemical modifications (new salts or esters) of existing drugs, new formulations or dosage strengths, new combinations of already approved drugs, or new indications. Figure 1.2 plots annual counts of NMEs approved by the FDA from 1965 to 2005, placing the downward trend in approvals since 1996 in historical perspective.\(^4\) From 1990 onwards, the series also includes new biotech drugs, often called "biological therapeutics," which historically have moved through a different approval process, and are frequently omitted from counts of drug approvals.\(^5\)

Although the approval of a new drug normally represents a significant advance in therapy, and therefore merits close attention, simply counting new drug approvals may present a significantly distorted picture of the outputs and impacts of biopharmaceutical research. Introductions of new products to the marketplace are a very restricted notion of innovative output, ignoring contributions to the pool of scientific knowledge that will continue to have economic value far into
the future. Even within the narrow product-oriented notion of output, by focusing attention exclusively on NMEs, which presume "breakthrough" innovation, this procedure gives zero weight to "incremental" product improvements—which have been shown in many other contexts to account for a very large fraction of total benefits from innovation. It also ignores the fact that not all new drugs are of the same quality, measured in terms of their impact on human health or by consumers' willingness to pay.

**All Drugs Are Not Equal**

Drugs vary significantly in their scientific significance, health impact, and economic value. This heterogeneity in "quality" of drugs means that simple counts of NMEs may seriously mismeasure R&D performance. Blockbusters with more than $1 billion in annual U.S. sales, for example, are given equal weight to newly approved drugs that achieve only $50 million in annual U.S. sales, and drugs which represent a major advance in the treatment of disease are given the same weight as the "me-too" products that appear in their wake. The obvious solution to this would be to weight each drug approval by a measure of quality, but systematically measuring differences in drug quality is surprisingly difficult. A number of productivity analysts of the pharmaceutical industry have taken a step towards addressing heterogeneity among drugs through indexing the volume of R&D output by weighting each of new drug approvals by its sales volumes. Comanor (1965), for example, calculated the output of the pharmaceutical industry as the sum of the first two years' sales of all new chemical entities. But since drugs are sold into imperfectly competitive markets, characterized by complex insurance contracts and attendant agency problems, government regulation, and negotiation of prices between manufacturers, third-party payors, and specialist intermediaries, it is not clear that prices and sales volumes are good measures of willingness to pay, and few analysts have attempted to compute the "correct" economic measure of innovative performance based on consumer and producer surplus.

Efforts to account for differences in quality have therefore tended to use multi-dimensional measures of quality. Vernon and Gusen (1974) decomposed the Comanor output measure into two parts: the number of newly approved chemical entities (a function of R&D), and the discounted sales per newly approved chemical entity over its first two years, hypothesized to reflect in part marketing promotional efforts.
Dranove and Meltzer (1994) defined drug quality in various ways, including scientific novelty as measured by whether the FDA granted the New Drug Application priority rather than standard review status; number of citations in medical textbooks, medical journals, and in subsequent patent applications; number of worldwide introductions; plus U.S. sales in first five years on the market. Dranove and Meltzer concluded that, based on various measures, higher quality drugs were being approved more rapidly by the FDA. Large-scale efforts by regulatory bodies to systematically rate the clinical effectiveness of different drugs (such as the NICE process in the UK) or to compute benefits in QALY or DALY units are another source of information that could be used to improve measures of innovative output.

Without a carefully conducted retrospective analysis of all of the hundreds of new drugs introduced in the past 30 years, it is hard to determine with certainty whether quality-weighted output has risen or fallen over time. Some economic indicators suggest that average "quality" is rising: new products continue to obtain premium prices in the face of competition from existing drugs, and R&D-based companies have seen steady growth in sales despite vigorous generic competition and increasing focus on cost control by purchasing institutions. Viewed in long-term perspective it is also clear that many of today’s drugs, developed using rational drug design methods and improved understanding of fundamental physiology and biochemistry, are significantly "better" than their predecessors in the sense of greater efficacy, fewer side-effects, and easier dosing. It is quite unlikely, therefore, that a properly constructed series on quality-weighted NMEs would trend downwards.

**Incremental Innovation and Product Improvements**

A further problem with focusing on counts of NMEs is that any benefits of incremental innovation are completely ignored. Figure 1.2 is notable for what is left out, i.e., regulatory approvals of new indications, formulations, and dosages of previously approved drugs. Drugs are approved on New Drug Application ("NDA") or Biologic License Application ("BLA") submissions to the FDA. As part of the submission, the sponsor typically provides clinical evidence in support of FDA approval for some particular medical condition, known as the primary "indication." This does not mean that clinical research stops. In many cases, companies carry out further research in so-called Phase IV trials,
and following submission of the initial NDA/BLA, develop evidence used to obtain subsequent FDA approvals for additional indications; this type of an application is called a supplemental NDA ("sNDA"). For example, the clinical trial that led Merck to voluntarily withdraw its acute pain and osteoarthritis agent, Vioxx, in September 2004 was a Phase IV study designed to obtain evidence in support of the use of Vioxx for preventing colorectal cancer.

Most analysts implicitly or explicitly exclude such "secondary" approvals when measuring output in terms of the number of new drug approvals. This flies in the face of considerable, albeit anecdotal, evidence that follow-on discoveries in medicine can generate very significant public health benefits, often for an indication unrelated to the initial major breakthrough. For example, Spivey, Lasagna, and Trimble (1987) have stated:

"Examples of this phenomenon include the protective effects of β-blockers against myocardial infarction and coronary death, the use of β-blockers to prevent migraine and reduce blood pressure, the antiarrhythmic actions of lidocaine, the use of amantidine to treat parkinsonism, the anti-epileptic efficacy of carbamazepine, the use of diazepam for status epilepticus, and the uricosuric effect of probenecid."

Research that supports the use of existing drugs in new indications can therefore generate substantial health benefits. One measure of these benefits is the utilization and sales volumes for new indications. Anecdotal evidence suggests that sales volumes for supplemental indications can in some cases be considerably larger than sales from the original primary approved indications. For example, while Zantac (ranitidine) was originally approved for treatment of a hypersecretory condition known as Zollinger-Ellison syndrome (a relatively rare condition) and for short-term treatment of active duodenal ulcer (a considerably more common condition, but limited to acute episodes), supplementary indication approvals obtained for Zantac included much larger populations and entailed considerably greater sales volumes, such as those for treatment for gastroesophageal reflux disease ("GERD," a severe but relatively common form of heartburn), and maintenance of healing of erosive esophagitis (a common condition requiring long-term treatment).

In addition to the use of a drug in new indications, innovation that takes the form of improved formulations, delivery methods, and dosing protocols may also generate substantial benefits associated with
improved patient compliance, greater efficacy as a result of improved pharmacokinetics, reduced side effects, or the ability to effectively treat new patient populations. Again, anecdotal evidence suggests that these innovations can generate significant increases in utilization and sales. The development of Valtrex (vancyclovir), a pro-drug version of acyclovir, for example, enabled utilization of the drug in suppression and prevention of genital herpes with once per day dosing, significantly expanding its use beyond its initial labeling.¹⁰

One economic indicator of the magnitude of these benefits is the extent to which supplemental indication approvals provide incentives for industrial R&D. The available evidence suggests that the prospect of additional sales beyond the initial indication provides commercial justification for extensive R&D expenditure.¹¹ For example, in their study of the costs of developing new drugs, DiMasi, Hansen, and Grabowski (2003) estimate that post-approval R&D is about 25.8 percent of total (pre- plus post-approval) out-of-pocket R&D costs ($140 of $543 million), whereas in capitalized costs it is about 10.6 percent of total costs ($95 of $897 million). CMR International estimates that 30 percent of industry R&D spending is devoted to "line extensions."¹²

Berndt, Cockburn, and Grepin (2005) looked at sales of drugs in three large and medically significant therapeutic classes (ACE inhibitors, SSRI/SNRI antidepressants, and anti-ulcer drugs) and decomposed sales of each drug according to whether the patient was given a diagnosis consistent with the drug's "primary" indication or was given a diagnosis consistent with "secondary" indications or off-label use. In two out of the three drug classes considered here, utilization in patients with diagnoses outside each drug's initially approved indication accounts for 70 to 80 percent of total use. While these classes may not be fully representative of the entire range of drugs, these results suggest that conventional measures of innovative output based on counting NMEs may seriously understate the productivity of research in this industry. While the number of new NDA/BLA approvals has declined or at best stayed roughly constant in the last decade, in these three therapeutic classes the number of sNDAs has been generally increasing over time, and these indicators of cumulative incremental innovation are associated with substantial medical and economic benefits.

Again, a broad-based, systematic adjustment to the "standard" output measures to address this flaw is a forbiddingly difficult task. Mason (2004) recommends correcting the traditional measures of innovative output by counting each new indication approval as equal to 0.5 of an
NME, and each major line extension as equal to 0.25 of an NME. Even this type of crude ad-hoc quality-adjustment would likely result in a substantial revision to perceptions of trends in output.

**Time Horizons, Inventory-clearing and Other Statistical Distractions**

A final problem with discussions of productivity trends based on output data such as figure 1.2 is the time horizon that is used. Media accounts have depicted a particularly dramatic decline in the productivity of biopharmaceutical R&D by focusing on the decline in approvals since their 1996 peak year. During the period 1996–2005, counts of NMEs have trended downwards, suggesting a sharp decline in research productivity.

This is, however, something of a statistical mirage: 1995–1996 were years in which exceptionally large numbers of NDAs were approved, and approval rates in subsequent years fall well within historical norms. Note also that while some of this “bumper crop” is simply the result of chance (new drug candidates do not enter the approval process on a deterministic schedule), it also appears to have been driven by the evolving regulatory environment and its impact on the FDA. The “spike” in approvals between 1995 and 1996 may well reflect the impact of the Prescription Drug User Fee Act (PDUFA) on FDA review times and approvals. The PDUFA legislation was passed in an attempt to reduce the time and cost of drug development, authorizing the FDA to collect fees from sponsors submitting an NDA, BLA, or a supplemental NDA, and enabling the FDA to hire additional review staff to facilitate more rapid review. Though the PDUFA legislation only legally obliged the FDA to “review and act on” NDA/BLA submissions, not necessarily approve them more rapidly. In essence, PDUFA mandated responses and action letters from the FDA, but not necessarily approvals. Nonetheless, review times appear to have fallen substantially in the early-to-mid 1990s, driving a substantial “inventory clearing” effect. Though approval times were already falling prior to PDUFA, a careful analysis by Berndt, Gottschalk, Philipson, and Strobeck (2005) of 662 New Molecular Entities submitted to the FDA between 1979 and 2002 shows that after controlling for other factors, PDUFA accelerated the annual percentage reduction of estimated FDA approval times from −1.7 percent pre-PDUFA to −9.3 percent during the five years following passage of PDUFA, and to −5.3 percent during the legislation’s second five year period.
Average approval times were about 20 months in 1992, but fell to less than 15 months by 2002, which has very substantial implications for the timing of annual numbers of approvals. Assuming that PDUFA had no impact on the number or timing of applications, Berndt et al.'s model can be used to estimate what approval rates would have been in the absence of PDUFA. The results are quite startling. Without PDUFA the peak in annual approvals would have been both lower (55 NMEs versus 62 in fiscal 1996) and later (fiscal 1998 versus 1996).15 Perhaps the most useful way to quantify the impact of PDUFA on drug approvals is to look at the model's predicted cumulative number of NMEs over time. Without PDUFA the cumulative number of NMEs approved between fiscal 1992 and 1997 would have been 187, rather than the actual 220—a reduction of 33 NMEs, or 15 percent. By the end of fiscal 2002, the cumulative number of NMEs approved since 1992 in the absence of PDUFA would have been 376, only 13 (or 3.3 percent) less than the 389 that actually occurred. Hence, in a world without PDUFA, although many patient lives would have been adversely affected by the delay in gaining access to new therapies, at least the apparent decline in the productivity of biopharmaceutical R&D would not have been nearly as dramatic.

Exceptional factors, plus the inherent "noisiness" in counts therefore make it very difficult to accurately assess short-term trends in innovative output from counts of new drug approvals. Statements like "lowest number of drugs approved in the past ten years" or "approvals hit new low" should therefore be viewed very skeptically. Perhaps the most important conclusion to be drawn from figure 1.2 is that annual numbers of new drugs approved have risen steadily over the past 30 years, with no statistically discernable departure from trend once exceptional factors like PDUFA are taken in account.

IV. Measuring Inputs to Drug Development

Turning to the input side of the productivity equation, figure 1.3 presents data on pharmaceutical R&D expenditures from 1964–2005. The series shown here is one which is commonly used to track pharmaceutical R&D: worldwide R&D expenditures reported by members of PhRMA, the trade association for US-based "Big Pharma" companies.16

The picture is dominated by the steady upwards growth in R&D expenditure, at an average rate of almost 12 percent per year. Though this growth rate clearly exceeds the long term trend growth rate of
new drug approvals, implying a slowdown in productivity, these data should also be treated with caution. The series is not adjusted for inflation, and since the prices of resources used in R&D have risen over time, increases in nominal R&D expenditures likely substantially overstate the real increase in resources applied to drug discovery and development. For example, the Biomedical R&D Price Index published by the U.S. National Institutes of Health rose by 55 percent between 1990 and 2004, considerably faster than the 33 percent increase in the Gross Domestic Product price deflator. As in all such efforts to account for price inflation, separating the effects of changes in prices from changes in the composition and quality of the Biomedical R&D Price Index components is difficult, and the reliability of any R&D deflator is difficult to assess. Nonetheless, since the numerator in this productivity ratio (number of new drug approvals) is not in monetary units, failing to deflate R&D expenditures in the denominator will automatically induce a downward bias in productivity trends. After using the NIH Biomedical R&D Price Index to express R&D in constant 2005 dollars, the growth rate of R&D spending is halved to six percent per year for the period 1964–2005. During the “crisis” period 1996–2005 depicted in figure 1.1, the difference between nominal and real growth rates is similar: 8.8 percent versus 5.4 percent.
Even after adjusting for input cost inflation, growth in R&D spending has nonetheless been substantial. But before jumping to the conclusion that productivity measured in NMEs per R&D dollar has fallen significantly, it is important to recognize that a contemporaneous comparison of R&D spending and new drug approvals is a deeply flawed measure of productivity. Drug development is a lengthy (and very risky) process. A substantial portion of total R&D spent on developing a drug precedes product approval by many years. The drug development process includes preclinical investigations (1-5 years), clinical studies (5-11 years), and regulatory review time (0.5 to 2 years). Thus new drug approvals in any given year to a great extent reflect R&D input expenditures incurred far in the past. This delayed impact of R&D during the various phases of development on future new drug approvals is not captured when R&D productivity is measured in terms of contemporaneous R&D expenditures and new drug approvals. Indeed, if historical relationships hold true, these lengthy lags in the development process suggest that the acceleration in R&D expenditure over the past decade is likely to be followed by a surge in new drug approvals in the next decade.

Perhaps the most worrying productivity statistics are those which are derived from careful project-by-project accounting of R&D costs and outcomes, taking into account the passage of time (i.e., the opportunity cost of capital) and the riskiness of development projects—the "dry holes" of failed drug candidates. The most recent in a series of studies over the years from the Tufts Center for the Study of Drug Development (Dimasi, Hansen, and Grabowski 2003) estimates the present value of R&D expenditures to bring a new drug to market to be $802 million per FDA-approved new drug. In year 2000 dollars, this $802 million amount is more than 70 percent larger than the $318 million figure in an earlier 1991 study, and almost six times larger than the $138 million figure calculated in a 1979 analysis. Recent industry estimates of this figure are now well in excess of $1 billion per successful new drug.

V. Why Are Drug Development Costs Going Up?

Rising costs per successful new drug and rising overall industry R&D expenditures are alarming. Worldwide total R&D spending by industry likely now exceeds $80 billion and growing pressures around the world, particularly in erstwhile "safe havens" like the U.S., to limit
drug expenditures call into question whether end-user demand can support substantial further growth. These trends are driven by a number of factors, some of which are indeed cause for concern, and point to a variety of policy responses. But careful consideration of the wider range of underlying causes of increased R&D spending indicates that things may not be as bad as many commentators suggest. So what is driving these trends?

**Mining Out**

One appealing hypothesis is "mining out"—the idea that the "easy" (i.e., cheap) scientific problems were solved in past decades, leaving the industry with the challenges posed by the biochemistry and disease pathology underlying complex, subtle, systemic diseases such as Alzheimer’s, which are much more difficult (i.e., expensive) to investigate. Many commentators have suggested that the pharmaceutical industry is facing sharply diminishing marginal returns to R&D. Drews (1998), for example, characterizes drug development during the 1970s and ‘80s as a matter of making minor chemical improvements to existing compounds directed at a static set of about 500 well-proven physiological "targets"—an activity that surely runs quickly into diminishing returns. While there is some truth to this view, it cannot be the whole story. Economists have long recognized that technological opportunities are not finite, and that industries experience "recharge" as well as "exhaustion" of opportunities and inventions. The extraordinary progress of basic biomedical sciences has substantially expanded technological opportunities: for example, by some estimates the number of "druggable targets" in the human body has risen from 500 to at least 3,000 over the past two decades.21

Indeed, the pipeline of compounds in early stages of development has never been fuller. One industry source identified almost 4,500 compounds in preclinical development in 2004, up from less than 2,900 in 1995; with nearly 900 in Phase I of the development process (preliminary clinical testing in humans) in 2004, up from just over 400 in 1995. Figure 1.4 shows these trends.

Not all of the data about the pipeline is good. At the other end of the drug development process, trends in the volume of new drugs submitted for approval are less clear. The FDA’s "Critical Path" White Paper reported a steady decline in submissions of NMEs for both traditional small molecule drugs and biological therapeutics between 1993 and 2003.
Is the Pharmaceutical Industry in a Productivity Crisis?

Preclinical compounds in development

Phase I compounds in development

Figure 1.4
The pharmaceutical pipeline.

But more recent data is more encouraging: in 2004 NME submissions hit a five-year high. It should also be recognized that the volume of submissions is driven partly by the movement of new drugs through companies’ development pipelines, largely driven by exogenous scientific factors such as trial protocols and success rates at each stage, and partly by companies’ decisions about how hard to “push” products and whether or when to submit them, which may be somewhat responsive to expectations about regulatory review standards or market opportunities.

Other indicators also point to vibrant activity in early stage research, with venture capital investments in life sciences reaching a five-year high in 2005. A significant fraction of the increase in R&D spending should therefore be understood as representing a rational (and welcome) effort to exploit these new opportunities.
Re-tooling and Industry Transformation

Along with increases in technological opportunities, the biopharmaceutical industry has also seen dramatic changes in the tools and methods used to exploit them. Technologies such as ultra high throughput screening, combinatorial chemistry, microfluidics, gene arrays, and bioinformatics represent multiple-order-of-magnitude improvements in the technology used to perform research, but have not been cheap to acquire. Paralleling the evolution of the R&D model from "random screening" of candidate molecules to "rational drug design" and "science-based" drug discovery, drug companies have had to acquire a wide range of costly specialized assets and human capital, and to invest in managerial and organizational infrastructure to deploy them. At a more abstract level, the scientific disciplines and knowledge used in the research process have changed—crude put, molecular biology has supplanted chemistry—and a variety of new disciplines, research communities, and bodies of knowledge are now important to drug discovery, such as genomics, proteomics, and metabolomics. All this has required substantial and sustained investments in acquiring new capabilities.23

Again, these expenditures represent a welcome and valuable investment. More broadly, this "re-tooling" process can be understood in terms of the normal process of industry transformation. "S-curves" have been observed in many industries and technologies, where marginal returns from exploiting a given technology or paradigm are initially low, become much larger as the technology takes off, and then decline as the technology matures. Eventually a new technology appears, typically developed by new entrants or industry "outsiders" which initially has poorer performance than the existing technology (which is dominated by successful incumbent firms). As the new technology enters its take-off phase, incumbents face a difficult and expensive transition to jump to the new "S-curve." The drug industry appears to be going through just such a period of transformation, with the previously successful chemistry-based drug development technology reaching maturity and experiencing falling marginal returns to R&D, and being supplanted by a new biology-based technology that is just beginning to payoff. This is sketched out below in figure 1.5.

These episodes of transition are typically characterized by economic turbulence and associated costs, and followed by periods of high marginal returns to R&D. Over the next decade, all else equal, we should
Is the Pharmaceutical Industry in a Productivity Crisis?

Figure 1.5
The evolution of drug development technology.

therefore expect R&D costs to stabilize, if not decline, as the new technology enters its "take-off" phase.

Failure Rates in the Development Process

As discussed above, the cost per approved new drug needs to take into account the large numbers of candidates that fail to meet criteria for progressing through the phases of clinical development, as well as the opportunity cost of capital. The mathematics of these calculations point to one of the major causes of increased R&D costs per approved drug: high failure rates, particularly in the later stages of development. On average, fully 75 percent of the fully capitalized cost of developing a new drug is the cost of failures. Notwithstanding scientific progress in basic research, these failure rates are persistently high and very troubling to industry insiders.

Drug development takes place in well-defined phases: Discovery, where candidate molecules are identified; Preclinical, where candidates are tested for toxicology in vitro and in animal models; Phase I, where the drug is tested for safety in small numbers of healthy human subjects, and some initial clinical data is collected; Phase II, where controlled trials are used to obtain evidence on efficacy and toxicity in patients affected by the disease; Phase III, where clinical trials are conducted on
large numbers of people to establish definitive data on likely efficacy, toxicity, and side effects of the drug in widespread use; followed by submission of the drug for regulatory review; and ultimately, regulatory approval that permits the drug to be marketed.

Kola and Landis (2004) examined causes of failure of development projects for the top ten pharmaceutical firms over the period 1991–2000, and found that only 11 percent of compounds tested in man made it through to approval for sale. Even in late stages of clinical development failure rates reported in this study are disturbingly high: 62 percent of drug candidates that make it through Phase I fail to pass Phase II, and 45 percent of those that do fail to pass Phase III. These late stage failures are extremely costly, both because of the expense of running large scale controlled trials, and because expenditures made much earlier in the development of the drug have accumulated substantial opportunity costs.

Some very interesting findings in this study relate to changes in the causes of failure in these data. Great progress was made between 1991 and 2000 in solving problems relating to pharmacokinetics and bioavailability (maintaining therapeutic but not toxic levels of the drug in the body). These accounted for more than 40 percent of drug failures in 1991, but less than 10 percent in 2000. Less success was achieved in solving failures due to toxicology and safety problems, which rose slightly to about 30 percent of failures in 2000, and lack of efficacy, which continues to account for about 25 to 30 percent of failures.

The scientific community in industry, academia, and government appears to be reaching some consensus as to why failure rates for these technical/scientific reasons are so high, and how they can be improved. High failure rates are thought to be attributable to a number of factors. These range from "straightforwardly fixable" problems such as inadequate training and workforce development in preclinical research and investigative medicine, too much weight placed on unreliable animal models, reluctance and regulatory obstacles to move drugs into humans more quickly, and poor communication and lack of interaction with regulators, through to much less tractable scientific challenges. Progress in "translational medicine" has been limited by the imperfect state of knowledge in systems biology: reductionist science has generated vast amounts of data and knowledge at the molecular and cellular level, but progress in understanding whole-organism processes and disease pathology has been much slower. There also seems to be growing recognition that lack of collaboration in precompetitive and
preclinical research, along with excessive secrecy, "data hoarding" and efforts to gain exclusive rights to basic research tools and data through the patent system may be increasingly counterproductive.25

Some of the solutions to these problems involve changes to clinical testing methodology, such as use of surrogate endpoints and "biomarkers" to provide early quantitative evidence of efficacy, deployment of new technologies such as advanced medical imaging to measure clinical outcomes, identification of patient subgroups who respond differentially, flexible protocols involving adaptive dosing or "enrichment" of the sample of patients based on early identification of positive responses, as well as greater use of modeling, simulation, and advanced information technology to collect more and better data and predict outcomes. Others involve developing greater capacity in translational medicine, in mechanisms for encouraging collaboration between institutions, and across the "profit divide" between industry, government, and academia.

But perhaps the most alarming finding from the Kola and Landis study is the reported increase in the fraction of failures due to essentially economic problems: prohibitively high manufacturing costs, and unspecified "commercial" reasons. The share of failures for these reasons rose from five percent in 1991 to 30 percent in 2000. This points to a very important role of economic and competitive pressures in driving up R&D costs.

Vicious Circles? The Blockbuster Syndrome

Some observers believe that rising R&D costs and falling productivity are the result of "addiction to blockbusters." Faced with pressure from financial markets to grow earnings and realize high rates of return, drug companies have found the extraordinary profitability of successful one-size-fits-all products sold into large markets irresistible.26 The attractiveness of these opportunities combined with billion-dollar costs of developing new products appears to have led many companies to set very high commercial hurdles for drug candidates. In order to meet these high initial sales targets, the Willie Sutton Theorem ("that's where the money is") dictates focusing development efforts on the needs of very large patient populations. But these are typically crowded, highly competitive markets where development costs are high (intensive clinical development programs demand more and larger clinical trials) and market conditions and sales forecasts are subject to a great deal
of uncertainty. In these circumstances, lowered sales forecasts are not unlikely, and if these then fail to meet the hurdle, can halt development quite late in the process. Higher late stage failure rates in turn have the perverse effect of raising ex-post average drug development costs—and thus the height of the bar for future candidates. The search for blockbusters may also prompt companies to "swing for the fences" with drug candidates that have novel mechanisms of action, whose development is both more expensive, requiring novel clinical protocols and more interaction with regulators, and more likely to fail.

Blockbusters seem likely to remain a compelling goal for drug developers. But unsatisfactory results from pursuing this strategy, new business models that emphasize "targeted development" and niche products, and the potential for "personalized medicine" based on genetic profiling are driving companies towards a portfolio of mixed blockbuster and niche products.

Dis-integration, Resource Allocation, and Transactions Costs

Cockburn (2004, 2006) speculates that other economic forces relating to the "vertical dis-integration" of the industry may also underlie rising R&D expenditures. A variety of legal and institutional changes during the 1980s prompted a surge of entry into the industry at the interface between for-profit industrial R&D and public sector research institutions. These small, entrepreneurial, research-focused companies ("the biotechs") have become an important source of new drugs. Relatively few of them have succeeded in bringing new drugs to market through internal development, but Danzon et al. (2005) report that over 1/3 of new drugs approved between 1963 and 1999 originated in alliances between industry participants. To an increasing extent, resource allocation in drug discovery is moving away from the internal capital markets of large, vertically integrated firms towards a "market for technology."

There may well be substantial beneficial effects on R&D productivity from this industry restructuring. Specialization of activities is normally associated with greater efficiency, and allows a superior market-based allocation and pricing of risk. Entrepreneurial firms may be able to offer more powerful and more carefully tailored incentives to employees. Entry into an industry typically prevents incumbents from "shelving" or delaying promising technologies and forces inefficient incumbents to upgrade or exit. Large firms often incur substantial costs associated with costly, rigid, and conservative internal bureaucracies that are
necessary to control and coordinate their activities. And opening up a "market for technology" in the form of licensing deals and alliances may well result in a more efficient allocation of resources through competition and price signals.\textsuperscript{27}

In the other hand, industry restructuring may be responsible for some inefficiencies in R&D, and some (socially) unnecessary spending. In a world with perfect information, competitive markets, and no transactions costs, there is no need for vertical integration. But stepping away from this benchmark, it has long been clear that large vertically integrated firms are an efficient response to a number of real world problems. These include the inability to diversify risk where capital markets are incomplete or imperfect, the inability to minimize transactions costs when complete contracts cannot be written, the inability to capture spillovers or other externalities, and a variety of familiar difficulties that arise from flaws in markets for information. In fact, there is a strong presumption that vertical integration is the first best solution to economic problems such as financing and management of multiple projects which are long-term, risky, complex, involve activities which are costly to monitor, require substantial project-specific unrecoverable investments, and have shared costs and vertically complementary outcomes—i.e., pharmaceutical R&D!

It is far from clear, therefore, whether small entrepreneurial firms in this industry have any long run productivity advantage over large, vertically integrated incumbents. It is worth noting that of the many thousands of well-financed entrants with strong patent portfolios and exciting science, only a few hundred have survived. These firms face significant problems due to their small size, lack of diversification, and dependence on outside investors. Lerner and Merges (1998) and Lerner, Shane, and Tsai (2003) show, for example, that the terms of contractual arrangements between biotech firms and downstream licensees are sensitive to their financial condition and capital market access. Small firms may also have significant agency problems. For example, Guedj and Scharfstein (2004) show that the management of "one horse" biotech firms can inefficiently pursue their only project past the point at which a more diversified organization would abandon it.

Competitive pressure may also be responsible for socially wasteful over-investment in R&D when companies face "first past the post" incentives in technology races, or induce defensive investment by incumbents who need to strengthen their bargaining position with respect to entrants. Resources may also get wasted on bargaining costs,
payments to intermediaries, extra organizational overhead dedicated to seeking out, structuring, and operating collaborative ventures, as well as on developing (and litigating) excessively large patent portfolios. One strategic response of large firms to upstream entry has been agglomeration, consolidating their control over access to downstream markets during the merger wave of the 1990s. While the companies involved frequently claimed that these mergers were prompted by the pursuit of R&D efficiencies, this contradicts estimates of economies of scale and scope in drug discovery reported by Henderson and Cockburn (1996), whose results suggest that the productivity benefits from increasing size and diversity were exhausted at much smaller scale than the research efforts of today's industry leaders.

A final, and perhaps more subtle issue is that the efficiency of the "market for technology" in allocating research resources is open to question. Prices in the market for technology licenses and alliances may be significantly distorted by informational asymmetries, thin markets, bargaining outcomes that reflect large disparities in the size, sophistication, and financial situation of the parties, and a variety of other transactions costs. Using market prices as signals for resource allocation works well from a social perspective when prices reflect the marginal opportunity cost of the resources employed. But when market failure drives a wedge between prices and marginal opportunity costs, markets send the wrong signals, and poor decisions result.

Evidence on all of these points is scarce. Whether the new, vertically disintegrated industry structure has higher or lower aggregate productivity than the previous configuration remains open to question. Danzon et al. (2005) are optimistic about the productivity benefits of collaborative research arrangements, but a meaningful counterfactual is difficult to construct. Cockburn (2004) points out that it will take decades before enough data accumulates to decide the issue.

VI. Productivity in the Long Run: Relationships between Open Science and Industrial R&D

Profit-oriented commercial research and publicly funded "Open Science" have always been closely linked in the life sciences. The simplistic "waterfall" model of innovation whereby product-oriented industrial research feeds on a steady flow of basic scientific knowledge and data generated by upstream institutions like universities, public labs, and foundations is clearly counterfactual. Scientific knowledge,
materials, and personnel have always moved in both directions across the "profit divide" and the notion that there is a sharp division of labor between "upstream" basic research with no immediate practical application and "downstream" applied research focused entirely on marketable products is demonstrably false. Industry conducts a significant amount of basic research—increasingly within specialist firms located at the "blurred boundary" between public sector institutions and profit-oriented organizations—that has made major contributions to fundamental biological and biochemical knowledge and Nobel prizes for industry scientists. At the same time, many publicly funded labs and researchers are engaged in activity that is indistinguishable in many important senses from industrial research: screening compounds, conducting clinical trials, building molecular libraries and so forth.

Nonetheless, it is clear that publicly-funded Open Science, with its curiosity-driven, investigator-initiated agenda and priority and publication-based incentives, is a distinctive and vital component of the biomedical innovation system. Over the long run, biopharmaceutical research productivity depends critically on the contributions of Open Science. Some of these contributions are easy to see, such as the generation of new knowledge, new models, new data, and trained personnel that are available to industry. Others are more subtle. For example, some of the unique institutions of Open Science such as peer review, publication, and replication of experiments provide important "managerial infrastructure" to commercial science, where pharmaceutical companies use their employees' participation in the wider scientific community to monitor and reward research activity. Open Science also plays an important role as a public "truth-telling mechanism" on complex and difficult questions relating to safety, efficacy, and utilization of drugs.

Perhaps the most significant contribution of Open Science to the productivity of pharmaceutical research is its pursuit of a research agenda that is largely independent of commercial pressures. Industry can, and does, fund "blue sky" research with no obvious application, as well as a certain amount of projects directed at economically unattractive markets such as tropical diseases or rare disorders. But by and large industry research necessarily (and quite appropriately) focuses on topics with more obvious, and more immediate, application, and overweights its R&D effort towards products intended for a relatively narrow range of medical needs. By contrast, the agenda of Open Science, though not
immune to the "demand pull" of market forces, is driven to a great extent by other factors, principally the curiosity of individual investigators and community consensus on the intrinsic scientific interest of research topics and questions. This independent research agenda can overlap the range of topics that industry would fund in the absence of public science, but is not confined to it. Over the long run this research activity is responsible for generating the ideas, data, and paradigms that are currently not economically viable, but which significantly expand the technological opportunities available for future exploitation by industry.

Weakening the institutions of Open Science may therefore prove to be very costly in future decades. Science is becoming increasingly "propertized" by the enthusiastic participation of universities and academic researchers in the patent system, and a shift in the locus of intellectual energy in life sciences towards "just off campus" entrepreneurial companies. This has obvious and potentially very serious consequences for the direction of academic research. Limited (or just more expensive) access to proprietary research tools and data may limit replication and experimentation, and lead researchers to avoid important areas or topics. Academic researchers may also move effort away from basic research toward commercially attractive topics. However relatively little evidence has been found to suggest that these problems are currently having any "first order" impact on the conduct of Open Science. Some surveys have shown a decline in data sharing in some academic disciplines (e.g., Blumenthal et al. 1997), and an intriguing study by Murray and Stern (2005) shows a small but significant impact of the issuance of a patent on subsequent citations to its "twin" academic publication. But other surveys have found little evidence of substantive obstacles to accessing materials or research tools in university research and quantitative studies of the patenting and publishing behavior of individual academics have found no evidence of a substantial substitution of effort away from "pure" research. And while the dystopian prospect of a patent-driven "anticommons" in biomedical research raised by Heller and Eisenberg (1998) cannot be ruled out, this type of problem remains thus far only a hypothetical cause of declining research productivity.

Unfortunately, as with the effects of industry restructuring, it is likely to take many decades before the full impact of these institutional changes in the conduct and culture of Open Science on the productivity of industrial research is felt.
VII. "Fixing" Research Productivity

Careful consideration of the factual basis for claims that the biopharmaceutical industry is facing a productivity "crisis" suggests that these are overblown. Declining counts of new drug approvals in recent years are worrisome, but look less dramatic when statistical anomalies are accounted for, and when it is recognized that these figures are very noisy measures of innovative performance that completely neglect other important outcomes from R&D performed in this industry. Similarly, the trend of increasing R&D expenditures is both overstated to some degree, and also a signal of a likely surge in approvals of new (and better) drugs in the next ten years.

There are, nonetheless, real grounds for concern, which also present opportunities for policy initiatives to positively influence trends in research productivity and the costs of drug development.

Lack of capacity in translational medicine can be addressed by refocusing public research support towards relevant disciplines, and by investing in appropriate education and training. Academic medical centers are critical institutions in this area, combining clinical investigation with basic research and training, and bringing "bench and beside" together. Historically, much of this research has been funded by cross-subsidization from payments for patient care. Azoulay and Tay (2003) have documented a significant negative impact of changes in health care reimbursements on academic medical centers, where the impact has fallen disproportionately on the research budgets of clinical investigators and physician scientists rather than laboratory researchers. Greater attention to these adverse consequences of efforts to control health care costs, and development of alternative direct mechanisms for supporting this type of research could play an important role in building and sustaining capacity in translational medicine.

To the extent that the vertical struggle for rents between the biotechs and "Big Pharma" is depressing research productivity by inducing unproductive defensive expenditures and distorting allocation of research effort across competing opportunities, steps to encourage more efficient vertical relationships and greater collaboration may also be helpful. Public support of (and participation in) research consortia, patient pools, and open databases may be helpful in this respect, along with close scrutiny of the terms of access to publicly funded basic research embodied in university technology licensing agreements.
Finally, in thinking about research productivity it may be worth reflecting on the role played by the pricing and profitability of pharmaceutical products in directing research expenditures. The evidence that drug companies terminate large numbers of drug development projects on commercial grounds suggests that many more drug candidates will be brought forward for regulatory review if “small” markets can be made more economically attractive, thus raising productivity. Conversely, policies such as price controls, government purchasing, or weakened patent protection that are intended to reduce drug spending may carry with them significant long term costs in terms of lower rates of innovation and reduced research productivity. Scherer (2001) and others have found a strong contemporaneous link between drug companies’ profitability and their R&D spending. Reducing the profitability of existing products, and lowering the anticipated returns to future products will likely cause drug companies to reduce R&D spending, and one important mechanism by which this will occur is decisions to terminate drug development projects. Absent any functioning mechanism to “rescue” such abandoned projects, these decisions will be socially very costly in terms of wasted R&D costs sunk during the early stages of developing these products as well as forgone opportunities to improve human health.

It may also be worth noting that another immediate effect of policies that attempt to shift the burden of financing biopharmaceutical R&D away from consumers is likely to be a financial collapse in the biotech sector, with magnified consequences for innovation in the industry as a whole. The biopharmaceutical industry now relies very heavily on tools and new products generated by these small and financially fragile specialist firms, which, unlike Big Pharma, have very limited ability to finance continued research out of internal cash flow. Concerns are already being expressed about the adverse impact of the drying up supply of research tools from this sector as venture capitalists changed their focus towards “product” companies in recent years. A substantial decline in investors’ willingness to keep injecting resources into this sector may therefore result in a socially costly loss of critical research capacity.

Recent increases in R&D spending in biopharmaceuticals will generate a future payoff in the form of innovative new medicines. But the size of this payoff—the “bang for the buck” ultimately realized—is contingent on a favorable policy environment. As patent policy, health care finance, medical education and other issues come to the forefront of the
policy agenda, their impact on research productivity in this important industry merits careful consideration.

Acknowledgements


Endnotes


4. NME counts for 1990-2005 are from the FDA website, and for 1964–1989 are from Graham (2005).

5. Like "small molecule" drugs, there are large numbers of approvals of biotech products which are for new formulations, indications etc. These counts are taken from Tufts Center for Study of Drug Development publications (Reichert 2004), where "novel biological therapeutics" (e.g., monoclonal antibodies and rDNA-derived proteins) are defined analogously to NMEs, excluding additional indications and formulations, as well as blood products and vaccines.

6. Other early studies of R&D productivity include Baily (1972) and Wiggins (1981), each of which use the number of new chemical entities as the measure of output.

7. An earlier study by Wardell and DiRaddo (1980) discusses using a compromise of commercial and technological success measures to quantify innovation, where technological success was determined by a consensus expert panel.

8. Occasionally, several distinct indications are simultaneously approved with the initial NDA/BLA.


10. See Corey et al. (2004).

11. Critics of the industry often argue that this R&D expenditure is unnecessary, directed principally at artificially extending the innovator’s "franchise" beyond the period of patent protection associated with the original approval.


14. For further details, see Carpenter, Chernew, Smith, and Fendrick (2003).

15. Annual counts of approvals can be quite different depending on whether calendar year or fiscal year data are used, driven by a strong "December effect" present in the timing of approvals. See Graham and Berndt (2006).

16. It is important to recognize that this series does not include R&D conducted by companies based in Europe or Japan, expenditure by non-PhRMA members (principally biotech companies) or public sector research.


18. As noted above, expenditures by PhRMA members are only a fraction of the worldwide R&D effort that generates new drugs. Since 1990, R&D by European-based companies has been equivalent to about 80–90 percent of the amount spent by U.S.-based companies, and R&D by Japanese companies has been 30–50 percent of the U.S.-based amount, though these figures are muddied by exchange rate movements and other reporting problems. The growth rate of the PhRMA series is probably a reasonable proxy for the growth of total worldwide R&D spending by pharmaceutical companies, provided spending by non-U.S. companies is in roughly constant proportion. But the growth rate of total commercial R&D is likely understated by the PhRMA series, since expenditures by biotech companies are omitted and these have increased significantly over time.


20. $39bn reported by U.S.-based PhRMA members, $12bn by non-PhRMA members based in the U.S. (Burrill & Co survey) plus at least $25bn by Europe-based biopharmaceutical companies (EPFIA), and at least $8bn in Japan, Australia, and countries with an emerging research capability.


24. Other studies have found similar attrition rates for small molecules, for example DiMasi (2001).

25. See the FDA "Critical Path" White Paper, and e.g. Korn and Stanski (2005).

26. Even more so where senior managers' compensation has a high-powered stock price-based component and investors focus disproportionately on blockbusters.


31. See Azoulay, Ding, and Stuart 2006; and Breschi, Lissoni, and Montobbio 2005.

References


Is the Pharmaceutical Industry in a Productivity Crisis?


