Regulatory Review Time and Pharmaceutical R&D

Anna Chorniy, James Bailey,
Abdulkadir Civan, and Michael Maloney

MERCATUS WORKING PAPER

3434 Washington Blvd., 4th Floor, Arlington, Virginia 22201
www.mercatus.org

Anna Chorniy, James Bailey, Abdulkadir Civan, and Michael Maloney. "Regulatory Review Time and Pharmaceutical R&D." Mercatus Working Paper, Mercatus Center at George Mason University, Arlington, VA, July 2019.

Abstract

In the United States, all newly developed drugs undergo a lengthy review process conducted by the Federal Food and Drug Administration (FDA). These regulatory delays are costly for drug manufacturers and patients. We collected data on review times of drugs approved between 1999 and 2005 and found that, in medical indication categories where it takes the FDA longer to approve drugs, fewer drugs are developed. On average, three additional months of the review process result in one fewer drug in development in that drug category, suggesting that the length of the regulatory delay matters for pharmaceutical firms' research and development decisions.

JEL codes: I18, O38

Keywords: regulation, pharmaceutical R&D, Federal Food and Drug Administration, drug development

Author Affiliation and Contact Information

Anna Chorniy
Center for Health and Wellbeing
Princeton University, New Jersey
achorniy@princeton.edu

James Bailey
Department of Economics
Providence College, Rhode Island
jbailey6@providence.edu

Abdulkadir Civan
Gediz University, Turkey
akcivan@hotmail.com

Michael Maloney
John E. Walker Department of Economics
Clemson University, South Carolina
maloney@clemson.edu

This paper can be accessed at https://www.mercatus.org/publications/regulation/Regulatory-Review-Time-Pharmaceutical-RD

Regulatory Review Time and Pharmaceutical R&D

Anna Chorniy, James Bailey, Abdulkadir Civan, and Michael Maloney

Introduction

We expect that pharmaceutical companies, like any economic agent, respond to incentives. In this paper, we look at the pharmaceutical industry's research and development (R&D) behavior in response to the US Federal Food and Drug Administration (FDA) policies on drug approval. We examine the correlation between the time it takes the FDA to review and approve drugs and the number of drugs that currently are in manufacturers' development pipelines, classified by disease category. The maintained hypothesis is that in ailment categories where it takes the FDA longer to approve drugs, fewer drugs will be developed. The magnitude of the extent to which pharmaceutical companies react to the regulatory delay in their choice of which conditions to target is an important policy question.

The FDA and its counterparts in other countries face a tradeoff between preventing entry of unsafe drugs to the market and facilitating the entry of safe and effective drugs. Owing to the complex nature of pharmaceuticals and the limited accumulated knowledge of human physiology, determining whether a certain drug is safe is a challenging task. Assessing the efficacy of the drugs is even more so. This is why the FDA requires pharmaceutical companies to conduct costly and lengthy studies before drugs are allowed on the market.

If the FDA applies stringent rules on safety and efficacy, many patients will be unnecessarily deprived of potentially successful treatment. On the other hand, if the rules are loose, unsafe and ineffective drugs are likely to enter the market and cause potentially fatal damage. Finding the optimum balance between those two effects is challenging.

The FDA is criticized from both sides of the field. Many argue that the FDA is too lax in regulating pharmaceutical companies and that it does not require them to undertake the necessary amount of clinical trials, potentially allowing too many unsafe drugs to the market. Others argue that the FDA requirements are too stringent and that the agency does not allow enough potentially beneficial drugs to the market owing to relatively minor side effects. Another critique is that the FDA requires excessive clinical trials and a prolonged period of approval processing of these drugs so that, in the meantime, many patients lose their lives or suffer unnecessarily (for example, Philipson et al. 2008).

In this paper, we concentrate on measuring the responsiveness of pharmaceutical companies to the time taken by the regulatory process. We use a dataset of all prescription drugs approved by the FDA between 1999 and 2005. For each new drug application (NDA), we collect information on the time it took the FDA to approve it. To test our hypothesis on the responsiveness of drug manufacturers to review time, we sort NDAs into categories by their medical indications and link them to categories from the drug development pipeline data. Additionally, we look at the relationship between the burden of diseases (health consequences) and the R&D efforts. The hypothesis is that mortality and morbidity approximate the size of the market (Lichtenberg 2001, Blume-Kohout and Sood 2013). That is, the more lives lost and the more suffering a condition causes, the more incentive there is to develop a drug for the condition, all else constant. Finally, to control for the financial incentives of drug manufacturers, we use drug prices and all-payer spending for existing drugs as a rough proxy for potential market revenues.

We find evidence that, in disease categories where the FDA has taken longer to approve drugs in the past, fewer drugs are currently in the development pipeline. Where review times

were three months longer, there is, on average, one fewer drug in development in that drug category. This suggests that pharmaceutical firms are responsive to this margin of regulation. Although it does not follow from our research that the review process should be shortened without further consideration, our conclusion should not be ignored in the policy debate.

Drug Approval Process and Prior Research

Under current regulatory structure, developing drugs and proving their safety and efficacy require substantial financial resources and time. Following a drug discovery, medications-to-be must go through preclinical trials. In this stage, the drug manufacturer develops potential remedies and tests them on animals. Preclinical trials last from one to six years (DiMasi, Hansen, and Grabowski 2003). If the results of those tests seem promising, the company submits an investigational NDA to the FDA. In case the application is approved, the company starts clinical trials on humans, which last approximately 10 years. The manufacturer then applies to the FDA for final market approval of the drug. If the drug is deemed both safe and effective by the FDA, it will be granted an approval. Otherwise, the FDA can request additional clinical trials or reject the application altogether. The FDA review time may be as short as two months, but it can also last five years or more.

Several changes in the FDA legislation have provided opportunities for researchers to study the effects of the FDA review process. The first significant change occurred in response to the thalidomide tragedy of the late 1950s. Thalidomide was taken by pregnant women to treat morning sickness, and it was later discovered to cause birth defects in many children whose mothers had taken it during pregnancy. This episode greatly increased concerns among policymakers and the public about the safety and efficacy of drugs. Sam Peltzman (1973)

provides an excellent quotation by a former drug company's medical director that summarizes this view on the drug industry:

> Industry spokesmen would have us believe that all research is on wonder drugs or better medicinal products. They stress that there are many failures for each successful drug. This is true. . . . The problem arises out of the fact that they market so many of their failures. . . . Most [industries] must depend on selling only their successes . . . [but] with a little luck, proper timing, and a good promotion program a bag of asafetida with a unique chemical side chain can be made to look like a wonder drug. The illusion may not last, it frequently lasts long enough. By the time the doctor learns what the company knew at the beginning it has two new products to take the place of the old one. (Administered Prices: Drugs 1961.)

In 1962, the Kefauver Harris Amendment to the Federal Food, Drug, and Cosmetic Act required premarket approval of all drugs by the FDA; the agency was made responsible for ensuring the safety and efficacy of the drugs in the market. The main rationale for the amendment was that pharmaceutical companies' practice of developing minor variants of existing drugs and selling those at higher prices was an act of patent protection that was believed to result in deadweight loss in the form of higher R&D and marketing costs. The amendment required pharmaceutical companies prove that new drugs are "effective" in addition to being "safe." This new requirement significantly increased premarketing costs to drug companies owing to higher levels of information requirements, which must be obtained from costly clinical trials. It also prolonged the FDA approval times. These effects were found to be welfare reducing. For example, Peltzman finds that the effectiveness standard leads to reduced innovation and that this cost to consumers greatly exceeds the benefit of preventing ineffective drugs from entering the market (Peltzman 1973). He concludes that the amendment reduced social welfare by slowing down innovation.

A more recent statute, in contrast, shortened the regulatory review process. The Prescription Drug User Fee Act (PDUFA) of 1992 required the FDA to speed up its procedures in exchange for user fees paid by pharmaceutical firms. The PDUFA was renewed in 1997, 2002,

and 2007, with minor changes. Philipson and his coauthors (2008) find that the 1992 and 1997 versions of the PDUFA reduced review times by 6 to 7 percent and 3 to 4 percent per year, respectively. A substantial portion of these cuts in the length of the review time was because of the availability of increased FDA resources from user fees (Olson 2004). These payments currently fund 40 percent to 50 percent of the agency's new drug reviews (Olson 2013). However, there have been some concerns about the agency's integrity given its new financing structure. Many believe that an FDA dependent on industry user fees would ease up on approval standards; or, even if there was no intentional easement on the standards, speeding up the process would naturally increase the number of erroneous decisions (Olson 2002).

Several studies look at the effect of the PDUFA on drugs safety. Michael Friedman and his coauthors (1999) analyze four drug withdrawals from the market after the PDUFA became effective and reject the hypothesis that reduced review processing time was the reason for the drugs' removal. Ernst Berndt and his coauthors (2005) also find no effect of the PDUFA on the rate of drug withdrawals. K. I. Kaitin (2005) reaches a similar conclusion.

Since drug withdrawals take place only in extreme cases, some researchers analyzed other proxies for the safety of the drugs as well. The FDA maintains a database of adverse drug reactions. Physicians and patients can report adverse drug reactions to the FDA, and the agency uses that information for withdrawal or label change decisions. Thus, data on adverse drug reactions may potentially represent drug safety. Henry Grabowski and Y. Richard Wang (2006) do not find any relationship between FDA review speed and adverse effects; however, in a series of papers, Mary Olson (2002, 2004, and 2008) concludes that faster reviews are associated with increased adverse drug reactions.

On the other hand, benefits of increased speed of the FDA review process are not well studied. Shortened review procedures have static and potentially dynamic gains. Static gains are straightforward in the sense that performing drugs enter the market sooner; patients have access to those beneficial drugs earlier, so consumer surplus rises. Moreover, since the pharmaceutical companies can start marketing earlier, their cash flow is positively affected. Taking the cost of capital in pharmaceutical markets as 11 percent (Grabowski et al. 2002), gaining market access even several months earlier can affect manufacturer profit substantially. Increases in producer surplus may encourage new R&D and result in additional novel drug introductions. This would increase the consumer surplus even further. Even relatively small increases in drug R&D can improve social welfare substantially (Lichtenberg 2006, Lichtenberg 2011, and Civan and Maloney 2009).

The first comprehensive study on the benefits of reduced review time was carried out by Philipson and his coauthors (2008). Using US drug sales data, they estimate the life-cycle sales projections to evaluate private and social surpluses associated with pharmaceutical drugs. They calculate a welfare effect of the PDUFA using earlier work by Berndt and his coauthors (2005) that showed that the PDUFA reduced the review period by 5 percent to 6 percent per year. Philipson and his coauthors (2008) find that the PDUFA improved consumer welfare by $7 billion to $20 billion, producer surplus by about $7 billion to $11 billion, and social welfare by $14 billion to $31 billion. Their estimate of the upper bound of damages owing to unsafe drugs slipping into the market after the PDUFA was enacted is substantially smaller than the estimated benefits. Note that Philipson and his coauthors (2008) analyze static gains only. Numerous papers have shown that increases in the potential profits boost R&D investments as well (Blume-Kohout and Sood 2013).

The research closest to ours was done by John Vernon and his coauthors (2009). They analyze the influence of FDA approval times on pharmaceutical companies' R&D investments, taking advantage of the speed-up in review times following the passage of the PDUFA in 1992. They use survey data on firm-level R&D expenditures of seven large US-based pharmaceutical companies from 1990 to 1999. They find that a 10 percent decrease in FDA approval times results in an increase in R&D spending of 0.9 percent to 1.7 percent. Vernon and his coauthors (2009) average the FDA approval times for the entire set of drugs in a given year. As a result, their analysis is based on a very small number of observations (nine).

We contribute to the earlier research by focusing on a cross-sectional rather than time-series variation. Previous work took advantage of plausibly exogenous policy changes that slowed down (Peltzman 1973) or sped up (Vernon et al. 2009) FDA review times. But these papers look at how a single change in policy affected the total number of drugs in the pipeline, making it difficult to pin down whether changes in subsequent drug development were truly owing to changes in review times rather than other changes that happened to occur around the same time. By using cross-sectional data on review times and the drug pipeline across hundreds of drug categories, we bring to bear new information that exploits the variation in the time the FDA takes to review drugs in different disease categories, controlling for an average amount of effort firms put into proving that a drug is safe and effective.

To do so, we collect novel information on the length of the drug application paperwork that includes details on safety and efficacy reviews as provided by the FDA. A concern undermining our empirical strategy is the possibility that there might be unobserved underlying reasons for companies to develop fewer drugs in a given category that also require the FDA to review these drugs for a longer time period. We believe the complexity and amount of NDA

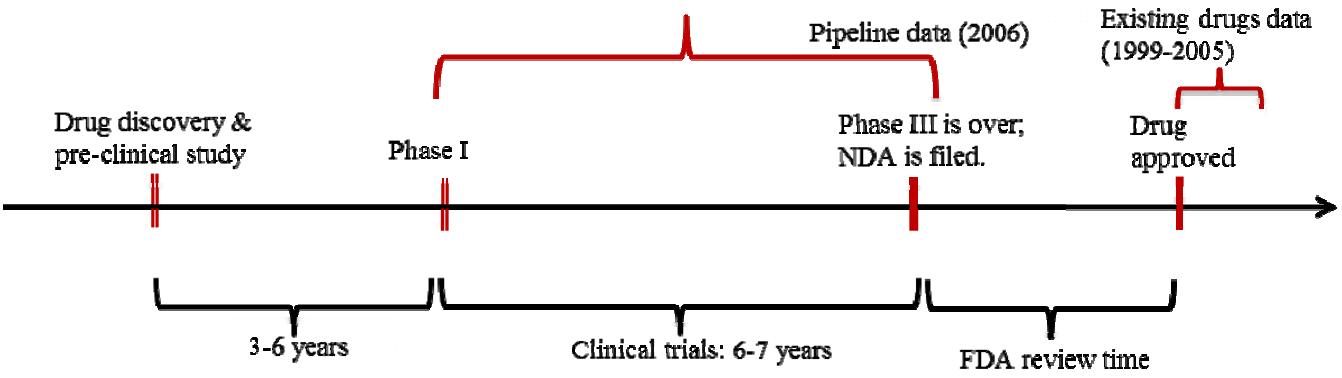
paperwork allows us to approximate the cost of drug development and to alleviate the endogeneity problem.

Methodology

This paper investigates the effects of the delay caused by the regulatory process on pharmaceutical R&D. We measure R&D by counting the drugs in the development pipeline. Drugs in the development pipeline are then organized by medical indications, or the conditions these drugs are supposed to treat. We match medical conditions from the pipeline data to the conditions indicated for the approved drugs. Regulatory review time, or delay from the drug manufacturers' standpoint, is the time NDAs were under review by the FDA for the drugs approved between 1999 and 2005.

Although regulatory delay occurs at various stages of the drug development process, we focus on the time between the day a drug manufacturer filed an NDA and the day the drug was approved by the FDA. This period follows Phase III of trials, when most drug tests are concluded and the company has amassed sufficient evidence that its new molecule is safe and has beneficial therapeutic effects. This period of time is entirely within the purview of the FDA. Figure 1 shows the process timeline.

Figure 1. Drug Development and Data Collection Timeline



We generate data on review times from the Drugs@FDA database maintained by the FDA. For each drug, the FDA supplies information on application and approval dates, as well as whether the drug had priority review or orphan drug status. This database also contains the entire set of communications between the NDA applicant (the drug manufacturer) and the FDA, including the chemical reviews and documents showing drug efficacy and safety. We collect information on the number of pages in each section of the amassed paperwork. Our dataset includes all 171 drugs approved by the FDA between the beginning of 1999 and the end of 2005, the time period we take as relevant for R&D decisions about drugs in the pipeline in 2006. Data on review times are available only for the drugs that were eventually approved.[1] The oldest drug in the FDA database is Pletal, used to treat intermittent claudication, approved on January 15, 1999. The newest drug in the database is Vaprisol, used to treat euvolemic hyponatremia, approved on December 29, 2005.

[1] The FDA provides nearly no public information on rejected drugs, in sharp contrast to their impressive transparency with approved drugs. A recent publication from internal FDA researchers indicates that the overall acceptance rate was 73 percent over the 2000–2012 period (Sacks et al. 2014). It suggests that the time it takes for the FDA to review drugs is as important as clinical trials: there is no certainty that the molecule will be approved, and the company might find it more profitable to divert the resources to a drug that could pass the process with more certainty and less time.

Data on drugs in the development pipeline come from the Adis R&D Insight database by Wolters Kluwer Health (2006). This database includes medications currently in clinical trials or at the FDA for review, and it allows search by indication and by status for drugs in the FDA application process. Drugs reported in Adis R&D Insight first appear in the data with the early laboratory reports and continue through to world market launch.[2] For the drugs in the pipeline measure, we include drugs in Phase I through Phase III of clinical trials and for which an NDA was filed. Our use of 2006 data allows some time for firms to incorporate earlier (1999–2005) review times in their decisions about which candidate drugs to push forward.

We collapse some of the categories to facilitate the match of the pipeline data to the current drugs data. For instance, the database shows five different indications for HIV-related health conditions: HIV infections, HIV infections treatment, HIV-1 infections, HIV-2 infections, and HIV-associated nephropathy. We sum all the drugs listed in these categories. As a result, we end up with a dataset of 721 indication categories and 4,621 drugs in the pipeline. The condition categories exhibit the vagaries of drug development. Some categories are sweepingly general: cancer (102 drugs in development), solid tumors (267). Others are fairly specific: breast cancer and prostate cancer (each 106). Some are even more specialized: acute hypoxia (1).

The FDA approves drugs for specific indications. Both our datasets, drugs in the pipeline and existing drugs, list these indications, often more than just one. These indications do not map directly into any single comprehensive list of medical conditions. They are descriptive categories lacking a formal coding system. To standardize the descriptive indications and match them

[2] The Adis R&D Insight database is described in an earlier paper (Civan and Maloney 2009). The database is compiled from information collected from many sources: direct contact with companies involved with research and development; information collected from drug and therapeutic literature published in medical and biomedical journals; attendance at international meetings and conferences; company annual reports; news services; and press releases. Highly regarded, the Adis R&D Insight database is one of the leading data sources for professionals and researchers in pharmaceutical R&D, universities, and healthcare institutions.

between the pipeline data and existing drugs data, we turn to one of the most comprehensive lists of medical conditions: ICD-10 codes. ICD-10 codes have multiple levels, where conditions are aggregated into broader categories when possible. Specifically, all medical conditions are classified into 22 "chapters." These are the broadest categories. For example, the "G" chapter has all "Diseases of the Nervous System," while chapter "J" has all "Diseases of the Respiratory System." These chapters, in turn, have multiple blocks that define more narrow conditions.

Naturally, if we cross-walk all the indications from the drugs in the pipeline to the narrowest ICD-10 blocks, they are likely not to have an exact match among existing drugs. However, matching based on the narrower categories would be more precise. We select four-digit codes as the most refined level at which to perform matching. For example, attention-deficit disorder with hyperactivity would be coded as F90.1; F90 (a less detailed code) would denote "Attention-deficit disorders" in general; and F90–F98 would stand for "Behavioral and emotional disorders with onset usually occurring in childhood and adolescence." An ever more aggregated ICD-10 chapter in this case is "Mental, Behavioral, and Neurodevelopmental Disorders" (F01–F99). Suppose the drug pipeline contains a drug with an indication for "Oppositional defiant disorder." The most detailed ICD-10 code for that is F91.3. Matching this code to existing ADHD drugs (F90.1) will not be possible. However, at a higher level of "Behavioral and emotional disorders with onset usually occurring in childhood and adolescence," we will have a match of a drug in the pipeline to existing ADHD drugs market characteristics. In other words, using broader ICD-10 categories produces more matches but uses a more lax definition of the "market."

We code all these levels and create an indicator that shows on what level of detail the pipeline category matches an existing drug category. We assign each category the average

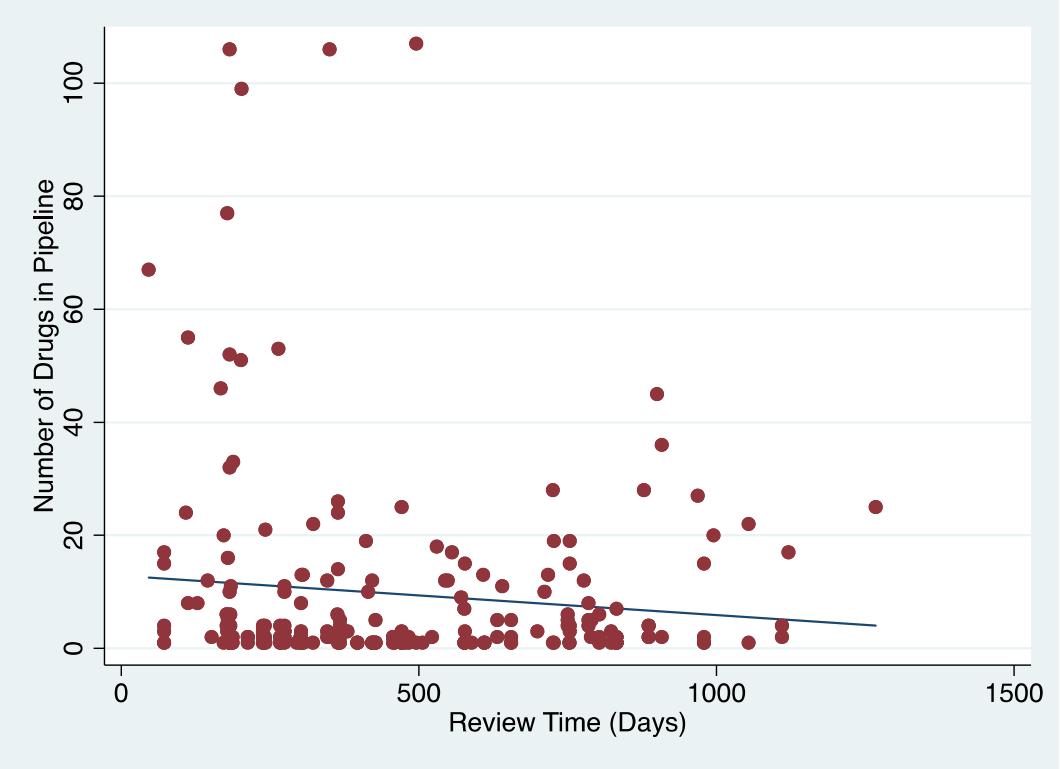
review time of the existing drugs in the category. The closer the match between the pipeline and existing drug category, the better, but using only perfect matches leaves us with fewer observations. Broader matches allow for the possibility that pharmaceutical companies consider review times for previous drugs with similar indications or with indications that are not the primary use of a drug, not only the exact same specific primary indication. For example, Abilify is counted only toward the schizophrenia category when using perfect matching, as this was the only indication it was originally approved for. When using broader matching, it may count toward related ICD-10 categories, such as schizoaffective disorder, or broader categories, such as psychotic disorders.

We choose the three-digit ICD-10 level for our main results and report our summary statistics (table 1) using a relatively narrow match. Following our main analysis, we run robustness checks using all other levels of detail (results are provided in appendix table 1). Closer matches yield larger but less precisely estimated coefficients for the effect of review time on the number of drugs in development, while broader matches yield smaller but more precisely estimated coefficients, so that the level of the match has little effect on statistical significance. Larger point-estimates for closer matches suggest that drug manufacturers, indeed, are guided more by the review times in the specific indication category in which the drug is being developed. Following earlier work, we use proxies to control for the demand conditions that may influence R&D decisions. They include the number of existing drugs in each category; the severity of the medical condition as measured by morbidity and mortality; and drug prices and all-payer expenditures collected from the Medical Expenditures Panel Survey (MEPS) prescription drug files between 1999 and 2010. We also add a novel control variable to account

for the difficulty and cost of developing a drug in a particular category by collecting the information on the length of the drug application paperwork.

Figure 2 shows the relationship between review times and number of drugs in the pipeline. It is negative, as expected; that is, indication categories with longer review times in the 1999–2005 approval data show fewer drugs in the pipeline in 2006. This correlation is quite noisy, however, since many other factors likely influence the number of drugs in the pipeline, factors we aim to control for.

Figure 2. Review Times 1999–2005 and the Drug Pipeline in 2006



Sources: Pipeline data are from Wolters Kluwer Health, "Adis R&D Insight," 2006, https://www.springer.com/gp/adis/products-services/adisinsight-databases/r-d-insight. Review time data are from Center for Drug Evaluation and Research (US), *Drugs@FDA* (Washington, DC: Food and Drug Administration, 2009), www.fda.gov/drugsatfda.

Data on drug prices are notoriously hard to obtain owing to the lack of transparency surrounding the issue. We use two very different sources in our attempts to get around this issue. First, we use Medco (now Express Scripts), which is the largest pharmacy benefit manager (PBM) in the United States and serves as PBM for many health insurance plans, including some universities. Insurance plans allow their enrollees to view the prices of drugs available under the plan. We collected retail prices for 144 of the 171 drugs in the FDA approval dataset. The process of assigning prices to drugs is somewhat complex given that frequency of use varies and many drugs come in multiple forms and strengths; those details are provided in the appendix. Second, we collect the average wholesale price from the Red Book, a product of Truven Health Analytics that contains the latest drug product pricing and packaging information on prescription and over-the-counter drug products. These prices, however, are not matched to the drug dosing and frequency-of-use information as is done for Medco prices (see the appendix). As an alternative to drug prices, we also gather data on total spending by drug from all payer sources (out-of-pocket, Medicaid, private insurance, etc.) from the MEPS in 1999–2010 to serve as proxy for the potential revenue that new drugs can fetch (Lichtenberg 2014).

Some pipeline indication categories have multiple existing drugs approved for treatment. The number of existing drugs in a pipeline category arguably signals the size of the market for drugs of that kind. At the same time, the more existing drugs there are in a given category, the stronger the competitive pressure on price and the lower the price a new drug may fetch. How these forces balance out is an empirical question. Even so, the number of existing drugs in a pipeline category is likely to be related to the number of drugs in development.

Finally, we measure the severity of health consequences of each disease category with the 2004 World Health Organization (WHO) data on morbidity and mortality. WHO constructs data

for most of the countries of the world and for 128 disease categories. The matching process between WHO data and pipeline indication categories is not straightforward. WHO categories include war, mayhem, and accidents, which are largely not relevant for our purposes.[3] Also, WHO categories do not perfectly match our pipeline categories even for illnesses linked to parts of the human anatomy. In general, WHO categories are broader. As a result, we find that many of the pipeline categories fit into the same WHO category. Also, some pipeline categories fit into multiple WHO categories. As a result, we match 85 WHO categories to 633 pipeline categories.

In earlier research, we showed that drug development by disease is positively related to disease incidence only in the United States. Here, we include both measures provided by the WHO: morbidity and mortality. The WHO morbidity measure, called the disability-adjusted life year (DALY), captures the years of life lost to illness in addition to death. The WHO DALY statistic evaluates the number of years of life spent in illness. For instance, if the average migraine sufferer has six attacks a year that last three days each, that is 18 days per year of healthy life lost, or 5 percent of a person-year. Multiplying this by the number of people suffering from the disease obtains the total number of person-years lost.

We analyze the determinants of the number of drugs in the development pipeline using a linear regression of the form

$$\text{NDrugsPipeline}_c = \alpha + \beta * \text{lnReviewTime}_c + \mathbf{Controls}_c * \gamma + e_c,$$

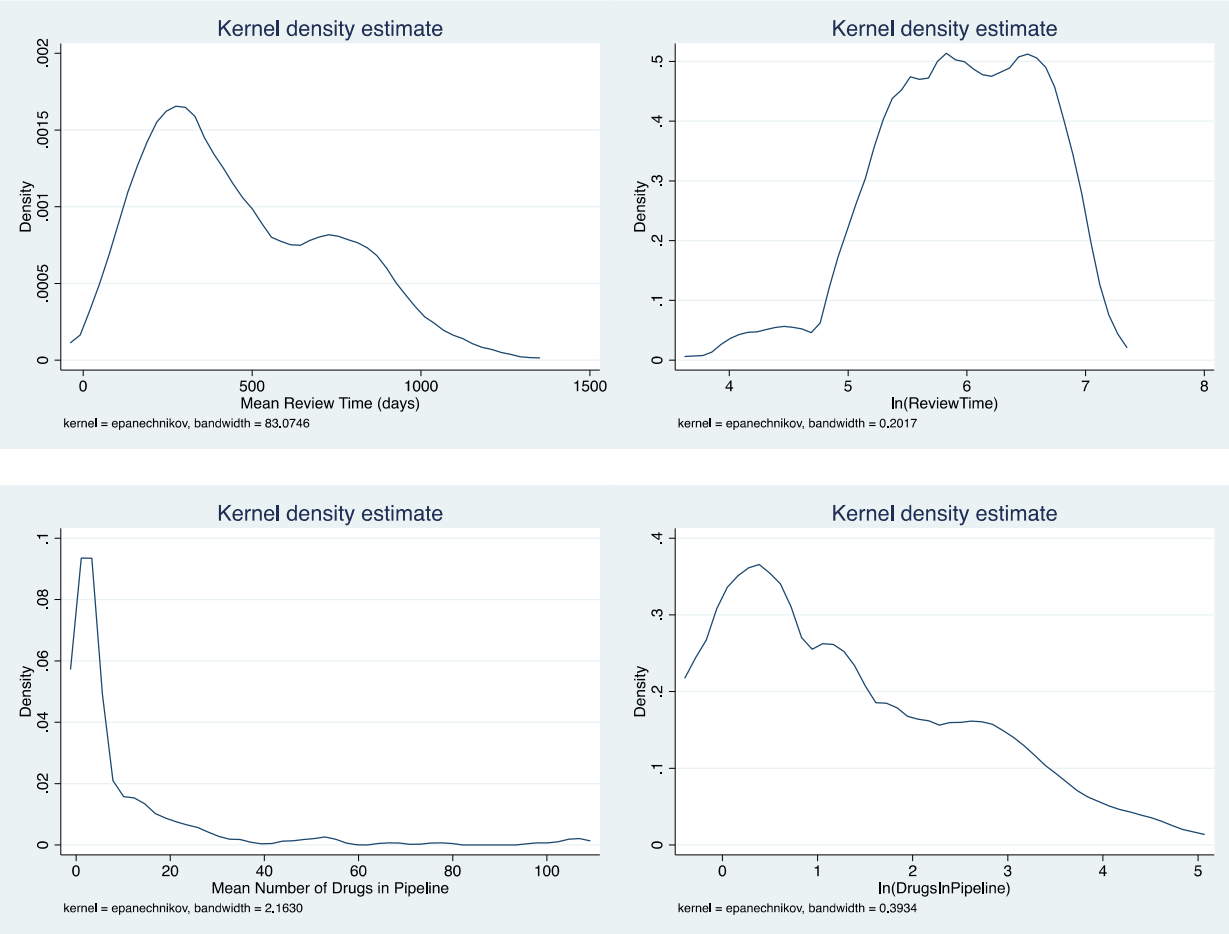
where c represents predetermined pipeline categories (defined according to medical criteria of the diseases and conditions that drugs treat). *Drugs in the pipeline* is the dependent variable; it is the number of drugs that are undergoing clinical trials or that have an NDA submitted for review as of 2006. Available controls include mortality, morbidity, page length of FDA review

[3] Things like accidents could be relevant, if the accident were related to a particular part of the human body. There are drugs to treat things like brain trauma and the like. However, the WHO accident categories are based on the cause of the accident, such as traffic or drowning.

paperwork, priority and orphan review status, and the number, price, and reported expenditures on the approved drugs.

Figure 3 shows the kernel density graphs of the number of drugs in the pipeline and review times, before and after taking natural logs. Both variables are right skewed, implying that logs may be appropriate. In our main specification, we use the natural log of review time

Figure 3. Distributions of Review Times and Number of Drugs in Pipeline



Note: The figures use 216 pipeline categories matched to the existing drugs at the preferred quality match level; that is, the number of drugs in the pipeline distribution is conditional on there being any existing drugs in a condition category.
Sources: Pipeline data are from Wolters Kluwer Health, "Adis R&D Insight," 2006, https://www.springer.com/gp/adis/products-services/adisinsight-databases/r-d-insight. Review time data are from Center for Drug Evaluation and Research (US), *Drugs@FDA* (Washington, DC: Food and Drug Administration, 2009), www.fda.gov/drugsatfda.

but not the number of drugs in the pipeline, as we believe this is the easiest specification to interpret (how percentage changes in review time are associated with the number of drugs in the pipeline); log-log and level-level specifications yield qualitatively similar results.

Data and Results

Table 1 shows summary statistics for the dependent variable and the set of controls. The number of drugs in pipeline disease categories varies significantly. On average, for the 216 included categories, there are about nine drugs being developed in a category.

Table 1. Summary Statistics

Variable	Number of observations	Mean	Standard deviation	Minimum	Maximum
Number of drugs in pipeline	216	9.6	17.8	1.0	107.0
Review time, days	216	466.2	270.5	46.0	1,268.0
Disability-adjusted life year	192	593.7	455.9	0.0	2,495.0
Mortality	192	52.2	60.9	0.0	327.8
Average Wholesale Price	214	2,070.0	4,695.0	15.0	25,055.0
Spending (MEPS)	216	131,784.0	247,470.0	396.7	1,484,689.0
Review paperwork, pages	213	604.2	230.6	148.0	1,271.0

Sources: Author calculations using data from Wolters Kluwer Health, "Adis R&D Insight," 2006, https://www.springer.com/gp/adis/products-services/adisinsight-databases/r-d-insight; Center for Drug Evaluation and Research (US), *Drugs@FDA* (Washington, DC: Food and Drug Administration, 2009), www.fda.gov/drugsatfda; World Health Organization, *The Global Burden of Disease: 2004 Update*, 2008, https://www.who.int/healthinfo/global_burden_disease/GBD_report_2004update_full.pdf; Medco, "Price a Medication," 2010, https://www.express-scripts.com/medco/consumer/homearticle.jsp?pageid=JP_Medication_Pricing&lh=; Truven Health Analytics, "RED BOOK Online," 2009, https://www.micromedexsolutions.com/micromedex2/4.34.0/WebHelp/RED_BOOK/Introduction_to_REDB_BOOK_Online.htm; and *Medical Expenditure Panel Survey (MEPS)* (Rockville, MD: Agency for Healthcare Research and Quality, 2018), https://www.ahrq.gov/data/meps.html.

In the FDA-supplied data, the average review time for drugs approved after 1999 is 466 days, or about 1.3 years, but there is wide variation in how quickly drugs are approved. In our

data, it takes from 46 days (Eloxatin) to 1,827 days (Prialt) for a drug to complete the review process. Likewise, the average review is 604 pages long, but review length ranges from 148 pages to 1,271. The average drug in the sample is about seven years old, as of December 2009.

We collect prices for drugs that were in the market in 2009. As discussed, price is measured as the cost of a daily dose to the average patient (Medco prices) and as a raw wholesale price (Red Book). The mean price is $189, but this varies from a few cents per day to over $5,000 daily. The upper end of the distribution is populated by several cancer and HIV drugs. The most expensive drug in the sample is Fuzeon, an HIV treatment, which is $5,269 per day. The monthly, annual, or lifetime cost of the medication will vary by how often and how long the drug is administered.

WHO data on morbidity and mortality are mapped into the pipeline disease categories. On average, there are 52,000 deaths and 593,000 DALYs lost to disease per category.

Table 2 shows a partial list of drug-pipeline categories that can be matched to our approved drugs. The table shows the top 25 of 171 pipeline categories based on the number of drugs in development. As mentioned, the categories vary widely in terms of the specificity of the disease. The diseases in the top 25 are generally well known.

Table 2. Drug Counts by Indication Categories

Drug indication	Number of drugs	
	Pipeline 2006	Approved 1999–2005
Type 2 diabetes mellitus	107	5
Breast cancer	106	3
Prostate cancer	106	2
Cancer (general)	102	1
Non-small cell lung cancer	99	2

(continued on next page)

Drug indication	Number of drugs	
	Pipeline 2006	Approved 1999–2005
Rheumatoid arthritis	77	1
Colorectal cancer	67	1
Alzheimer's disease	65	3
Asthma	56	1
Multiple myeloma	53	1
Acute myeloid leukemia	51	2
HIV infections	46	8
Renal cancer	46	1
Schizophrenia	45	2
Depression	41	1
Chronic obstructive pulmonary disease	36	6
HIV-1 infections	33	8
Chronic lymphocytic leukemia	32	1
Lymphoma	31	2
Obesity	28	1
Parkinson's disease	28	2
Type 1 diabetes mellitus	26	2
Hypertension	25	2
Migraine	25	3
Myelodysplastic syndromes	24	1

Note: The table shows the top 25 diseases by the number of drugs in the development pipeline.

Sources: Pipeline data are from Wolters Kluwer Health, "Adis R&D Insight," 2006, https://www.springer.com/gp/adis/products-services/adisinsight-databases/r-d-insight. Approved drug data are from Center for Drug Evaluation and Research (US), *Drugs@FDA* (Washington, DC: Food and Drug Administration, 2009), www.fda.gov/drugsatfda.

Table 3 reports the main results. We regress the number of drugs in the pipeline by indication category on a set of covariates. The coefficient on the FDA review time is negative and statistically significant in all specifications. The negative coefficient estimate supports our hypothesis that drug manufacturers are sensitive in their R&D decisions to the length of the review process. A doubling of the review length is associated with approximately six fewer drugs in the development pipeline in that disease category. This implies that a one-sixth increase

in review length is associated with approximately one fewer drug in development; with a mean review length of 466 days, this implies that each 78 extra days of review is associated with one fewer drug in development. Perhaps surprisingly, the control variables all tend to have no statistically significant association with the number of drugs in the pipeline.[4]

Table 3. Predictors of the Number of Drugs in the Pipeline

	(1)	(2)	(3)	(4)	(5)	(6)
Ln (review time)	−6.39***	−5.72**	−6.47***	−6.66***	−8.23**	−6.97**
	(2.37)	(2.31)	(2.44)	(2.36)	(3.64)	(3.25)
Morbidity	0.006		0.006	0.006	0.006	0.006
	(0.005)		(0.005)	(0.005)	(0.005)	(0.005)
Mortality		0.03				
		(0.03)				
Ln (Red Book price)			0.02		0.52	0.707
			(0.86)		(0.11)	(1.09)
Spending ($thousands)				0.009		
				(0.008)		
Priority status					−4.21	−3.34
					(5.32)	(5.07)
Orphan status					−1.10	−2.12
					(4.26)	(4.06)
Review pages						−.008
						(.006)
Intercept	44.6	42.9	44.9	17.5	54.8	50.7
	(14.2)	(14.0)	(16.3)	(22.6)	(25.9)	(24.2)
Number of observations	192	192	190	192	190	187

Notes: The dependent variable is the number of drugs in pipeline. Morbidity is the WHO DALY measure divided by 1000. Robust standard errors are in parentheses. ** represents $p < 0.05$; *** represents $p < 0.01$.

[4] Table A2 considers some alternative control variables and finds that in some specifications, higher morbidity and a higher number of approved drugs in a disease category lead to more drugs in the pipeline.

One key limitation of these results is the possibility that FDA review times are endogenous to new drug development. We take some comfort from the DiMasi, Milne, and Tabarrok (2014) study of how FDA review times vary across divisions, showing that much of the variation is possibly owing to exogenous differences in staffing and efficiency.[5] Because we are studying the effect of past review times on future drug development, we can at least rule out direct reverse causation of future drug development on past review times. However, it is possible that review times and drug development are jointly determined by a third variable, such as the scientific and technical complexity of a drug class. High complexity could lead to both longer reviews and fewer drugs in a category. Alternatively, if both drug companies and the FDA prioritize drugs that are expected to have a larger medical impact, this could lead to shorter reviews in categories with more drugs. We have attempted to account for these avenues of endogeneity by including covariates that measure medical importance (morbidity, mortality, and FDA priority status) and medical, pharmacologic, and clinical complexity (using the page length of the FDA review paperwork). Although our approach is limited in its ability to account for underlying endogeneity, we believe these factors constitute the most important ones that threaten identification.

[5] As their introduction states,

> The Neurology division took the most time (nearly 600 days), almost three times as long as the approval period for the Oncology and Anti-Viral divisions, both of which clocked in at under 200 days. These differences are suggestive of big gaps in productivity, but a number of other factors could be at work to explain the wide disparity in timing. Speedier approvals might depend on one division having fewer problems with its applications, for instance, or more resources than another. But even when those factors, along with safety considerations, were taken into account, the reason for the gaps still appeared to be varying levels of productivity; that is, faster divisions used their time and resources in a more efficient and effective way than slower divisions did.

Conclusion

We have investigated the effect of the regulatory delay on the propensity of companies to engage in pharmaceutical R&D for a given medical condition. We compare the number of drugs in the development pipeline to the time it took the FDA to approve existing drugs used to treat these conditions. Stratifying drugs by disease allows us to look at how existing drugs used to treat these diseases fare economically, as well as to gauge the potential size of the market. Controlling for other factors, we find a substantial and statistically negative relation between drug development and the time the FDA takes to process NDAs. Cross-sectional analysis of drug development by indication category enables us to make use of much greater variation in review times than previous time-series work could. This methodology contributes to the literature by introducing a novel measure of the cost and complexity of drug development by analyzing the paperwork submitted to the FDA as an NDA. It allows us, to some degree, to alleviate potential endogeneity problems.

These results suggest that firms are sensitive to regulatory delay. If review time is prolonged by about three months, there is, on average, one fewer drug in development in each drug category. This alone does not constitute a policy recommendation. However, it indicates that the time efficiency of FDA review is not something to be ignored in policy debates.

References

Administered Prices: Drugs. 1961. Hearing before the US Senate Judiciary Committee, 87th Cong., 1st sess., June 27.

Berndt, Ernst R., Adrian H. B. Gottschalk, Tomas J. Philipson, and Matthew W. Strobeck. 2005. "Industry Funding of the FDA: Effects of PDUFA on Approval Times and Withdrawal Rates." *Nature Reviews Drug Discovery* 4, no. 7: 545–54.

Blume-Kohout, Margaret, and Neeraj Sood. 2013. "Market Size and Innovation: Effects of Medicare Part D on Pharmaceutical Research and Development." *Journal of Public Economics* 97:327–36.

Center for Drug Evaluation and Research (US). 2009. *Drugs@FDA*. Washington, DC: Food and Drug Administration, www.fda.gov/drugsatfda.

Civan, Abdulkadir, and Michael T. Maloney. 2009. "The Effect of Price on Pharmaceutical R&D." *BE Journal of Economic Analysis & Policy* 9, no. 1.

DiMasi, Joseph A., Ronald W. Hansen, and Henry G. Grabowski. 2003. "The Price of Innovation: New Estimates of Drug Development Costs." *Journal of Health Economics* 22, no. 2: 151–85.

DiMasi, Joseph A., Christopher-Paul Milne, and Alex Tabarrok. 2014. "An FDA Report Card: Wide Variance in Performance Found among Agency's Drug Review Divisions." Manhattan Institute Working Paper, New York, NY, April.

Friedman, Michael A., Janet Woodcock, Murray M. Lumpkin, Jeffrey E. Shuren, Arthur E. Hass, and Larry J. Thompson. 1999. "The Safety of Newly Approved Medicines: Do Recent Market Removals Mean There Is a Problem?" *JAMA* 281, no. 18: 1728–34.

Grabowski, Henry, and Y. Richard Wang. 2006. "Determinants of New Drug Safety in the United States, 1993–2003." Unpublished manuscript, February.

Grabowski, Henry, John Vernon, and Joseph A. DiMasi. 2002. "Returns on Research and Development for 1990s: New Drug Introduction." *Pharmacoeconomics* 20, no. 3: 11–29.

Kaitin, K. I. 2005. "Drug Safety Withdrawals in the US Not Linked to Speed of FDA Approval." *Tufts Center for the Study of Drug Development Impact Report* 7:1–4.

Lichtenberg, Frank R. 2001. "The Allocation of Public Funds for Biomedical R&D." NBER Working Paper, National Bureau of Economic Research, Cambridge, MA. http://www.nber.org/chapters/c7642.pdf.

———. 2006. "The Effect of Using Newer Drugs on Admissions of Elderly Americans to Hospitals and Nursing Homes: State-Level Evidence from 1997 to 2003." *Pharmacoeconomics* 24, no. 3: 5–25.

———. 2011. “The Quality of Medical Care, Behavioral Risk Factors, and Longevity Growth.” *International Journal of Health Care Finance and Economics* 11, no. 1 (March): 1–34.

———. 2014. “The Impact of Pharmaceutical Innovation on Disability Days and the Use of Medical Services in the United States, 1997–2010.” *Journal of Human Capital* 8, no. 4: 432–80.

Medco. 2010. “Price a Medication.” https://www.express-scripts.com/medco/consumer/homearticle.jsp?pageid=JP_Medication_Pricing&lh=.

Medical Expenditure Panel Survey (MEPS). 2018. Rockville, MD: Agency for Healthcare Research and Quality. https://www.ahrq.gov/data/meps.html.

Nursing 2008 Drug Handbook. 2008. 28th ed. New York: Wolters Kluwer/Lippincott Williams & Wilkins.

Olson, Mary K. 2002. “Pharmaceutical Policy Change and the Safety of New Drugs.” *Journal of Law and Economics* 45, no. S2: 615–42.

——— . 2004. “Are Novel Drugs More Risky for Patients than Less Novel Drugs?” *Journal of Health Economics* 23, no. 6: 1135–58.

——— . 2008. “The Risk We Bear: The Effects of Review Speed and Industry User Fees on New Drug Safety.” *Journal of Health Economics* 27, no. 2: 175–200.

——— . 2013. “Eliminating the US Drug Lag: Implications for Drug Safety.” *Journal of Risk and Uncertainty* 47:1–30.

Peltzman, Sam. 1973. “An Evaluation of Consumer Protection Legislation: The 1962 Drug Amendments.” *Journal of Political Economy* 81, no. 5: 1049–91.

Philipson, Tomas, Ernst R. Berndt, Adrian H. B. Gottschalk, and Eric Sun. 2008. “Cost-Benefit Analysis of the FDA: The Case of the Prescription Drug User Fee Acts.” *Journal of Public Economics* 92, no. 5: 1306–25.

RxList. 2018. “Drugs A–Z.” www.rxlist.com.

Sacks, L. V., H. H. Shamsuddin, Y. I. Yasinskaya, K. Bouri, M. L. Lanthier, and R. E. Sherman. 2014. “Scientific and Regulatory Reasons for Delay and Denial of FDA Approval of Initial Applications for New Drugs, 2000–2012.” *JAMA* 311, no. 4: 378–84.

Truven Health Analytics. 2009. “RED BOOK Online.” https://www.micromedexsolutions.com/micromedex2/4.34.0/WebHelp/RED_BOOK/Introduction_to_REDB_BOOK_Online.htm.

Vernon, John A., Joseph H. Golec, Randall Lutter, and Clark Nardinelli. 2009. “An Exploratory Study of FDA New Drug Review Times, Prescription Drug User Fee Acts, and R&D Spending.” *Quarterly Review of Economics and Finance* 49, no. 4: 1260–74.

Wolters Kluwer Health. 2006. “Adis R&D Insight.” https://www.springer.com/gp/adis/products-services/adisinsight-databases/r-d-insight.

World Health Organization. 2008. *The Global Burden of Disease: 2004 Update*. https://www.who.int/healthinfo/global_burden_disease/GBD_report_2004update_full.pdf.

Appendix: Constructing Price

Drugs come in multiple forms and strengths, and our dataset contains 353 observations on price for 144 drugs. To average different prices for each drug and to compare prices across drugs, we normalize price based on the daily dose that is normally prescribed for the average patient.[6] That is, if the reported price is $59 for 30 pills of 50mg, we calculate the price per milligram, which is 4 cents. We then find the recommended dose per day, say 200mg, and multiply our price per milligram by the daily dosage, which yields a price per day for the medication of $7.87. In other words,

$$\text{Price per day} = \frac{\text{Price per package}}{\text{Total amount of active ingredient in package}} \times \text{Daily dosage} \qquad (1)$$

Our price measure represents the cost of a daily dose of the medicine. Other datasets provide information on drug prices for a 30-day supply, but this approach cannot be uniformly applied when the drug is not administered in such a fashion.[7,8]

[6] All the information on dosages comes from *Nursing 2008 Drug Handbook* (2008) and RxList (2018). The latter is owned and operated by WebMD, an online medical resource dedicated to offering detailed and current pharmaceutical information on brand and generic drugs. Both sources provide information on dosages according to the severity of a disease, its stage, and particular characteristics of the patient (such as age, weight, etc.). We average the maximum dosage per day and minimum dosage per day if there is no "recommended target dose." Otherwise, the target dose is used. For cases when calculation of a dose requires patient characteristics, we use those for an average person: average adult weight of 70kg and average body area of 1.82m^2. We also specify average weights for relevant children ages. For drugs that are taken once a month on a monthly cycle, we divide by 30 to get the daily price. However, for drugs that are taken just once, we use the full price as the daily price. Some other assumptions include the calculation of the amount of a solution or cream applied each time. For example, for eyedrops we assume that one drop is equal to 0.05ml.

[7] If price data were not available from Medco, we searched other online price sources, such as drugstore.com, and used those prices if available. Some drug prices are not available because the drug is distributed only through hospitals. In other cases, the drug has been discontinued or otherwise taken off the market.

While this method gives a consistent measure of price per day across strengths and forms and even across drugs, drugs vary according to the frequency of their administration. Because our interest is in obtaining a proxy for the revenue stream associated with a drug, we assign drugs to four different groups based on the frequency with which they would be taken over the life of an average patient. Frequency I labels drugs that are taken for a very short period of time, such as drugs treating infections that rarely return. Frequency II is for drugs that are taken for a longer period of time but are not expected to be repeated. This group includes cancer drugs, for example. Frequency III is for drugs that are taken intermittently but consistently for a long period of time, such as allergy drugs and drugs for the treatment of asthma. Finally, Frequency IV includes drugs that are taken frequently, maybe daily, over the patient's life. Diabetes drugs and some forms of heart medication fall into this group.

The primary purpose of these groups is to expand the daily price information to obtain a sense of the revenue that may be obtained from a particular drug. Importantly, the frequency variable is not a categorical variable in this case. It serves as a cardinal relationship between each frequency type. We use this feature in building the regression; that is, a combination of price per day and frequency is our proxy for the revenue stream from an existing drug.

[8] We compare our prices to those available from Medicare Part D data. Medicare reports data from Part D insurance plans on the "30-day" cost of each drug. We are able to match 161 of our 353 price observations to Medicare data. For these, there is a general similarity of prices. It appears that, by and large, Medicare prices assume that dosages are associated with strengths, which we do not, but this does not impact our analysis because we average over forms and strengths. It is also clear that Medicare prices are not "30-day" prices when the drug is not administered in this fashion, and when drugs come in different forms, the comparability of different Medicare prices for the same drug is questionable. While there are a few substantial differences between our prices and Medicare reported prices, we are confident that our methodology is fundamentally correct and that one must use care in employing the Medicare prices.

Table A1. Predictors of the Number of Drugs in the Pipeline with Narrower and Broader Drug Category Matching

Broadness of match	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
Ln (review time)	−11.5	−10.5	−8.23**	−5.41**	−5.25**	−5.19**	−4.01**
	(7.14)	(5.82)	(3.64)	(2.18)	(2.19)	(2.16)	(2.03)
Number of observations	59	89	190	437	457	471	629

Notes: The dependent variable is the number of drugs in the pipeline. Matching goes from narrowest on the left (only including definite matches) to broadest on the right. Broader matches yield more observations because they allow more drugs in the pipeline to be matched to an existing drug and so assigned a review time; insisting on a very narrow or precise match means that most pipeline drugs have no existing drug review time to match to and so are dropped. Results for control variables are omitted; controls included morbidity, price, and priority and orphan status. Morbidity is the WHO DALY measure divided by 1000. Robust standard errors are in parentheses. ** represents $p < 0.05$.

Table A2. Additional Predictors of the Number of Drugs in the Pipeline

	(1)	**(2)**	**(3)**	**(4)**	**(5)**
Ln (review time)	−4.36**	−5.44**	−8.91***	−8.36**	−8.85
	(2.06)	(2.37)	(3.85)	(4.12)	(5.35)
Morbidity		0.006	0.014**	0.013	0.013
		(0.004)	(0.007)	(0.007)	(0.007)
Number of approved drugs		4.23***		2.23	2.49
		(1.30)		(1.89)	(1.95)
Ln (Medco price)			0.06	0.02	−0.26
			(0.96)	(0.96)	(1.04)
Priority status					−4.93
					(9.00)
Orphan status					0.02
					(6.85)
Review pages					−0.02
					(.014)
Intercept	35.6	35.0	60.5	53.7	69.8
	(12.9)	(14.4)	(24.3)	(27.1)	(37.0)
Number of observations	216	192	96	96	93

Notes: The dependent variable is the number of drugs in the pipeline. Morbidity is the WHO DALY measure divided by 1000. Robust standard errors are in parentheses. ** represents $p < 0.05$; *** represents $p < 0.01$.