MANAGEMENT SCIENCE
Vol. 53, No. 9, September 2007, pp. 1452–1466
ISSN 0025-1909 | EISSN 1526-5501 | 07 | 5309 | 1452

informs

DOI 10.1287/mnsc.1070.0703

Valuing R&D Projects in a Portfolio: Evidence from the Pharmaceutical Industry

Karan Girotra
INSEAD, Boulevard de Constance, 77305 Fontainebleau Cedex, France, karan.girotra@insead.edu

Christian Terwiesch, Karl T. Ulrich
The Wharton School, University of Pennsylvania, Philadelphia, Pennsylvania 19104
{terwiesch@wharton.upenn.edu, ulrich@wharton.upenn.edu}

Understanding the value of a product development project is central to a firm's choice of project portfolio. The value of a project to a firm depends not only on its properties but also on the other projects being developed by the firm. This is due to interactions with the other projects that address the same consumer need and require the same development resources. In this study, we empirically investigate the structure and significance of these *portfolio-level project interactions*. Using a self-developed pharmaceutical industry data set, we conduct an event study around the failure of phase III clinical trials and their effect on the market valuation of the firm. The study exploits the natural experiment of a product development failure to give us a measure of the value of a drug development project to a firm. We then explain the variance in the value of projects based on interactions with other projects in the firm's portfolio. We find that the presence of other projects targeting the same market and a build-up of projects that require the same development resources reduce the value of a development project. In addition to providing evidence on the significance and structure of these portfolio-level project interactions, the empirical model estimated in this paper also provides a data-driven approach to valuing projects that may be relevant to licensing transactions.

Key words: product development; pharmaceuticals; development pipeline; portfolio properties; backup projects; portfolio management
History: Accepted by Christoph H. Loch, technological innovation, product development, and entrepreneurship; received January 25, 2005. This paper was with the authors 11 months for 3 revisions. Published online in *Articles in Advance* July 20, 2007.

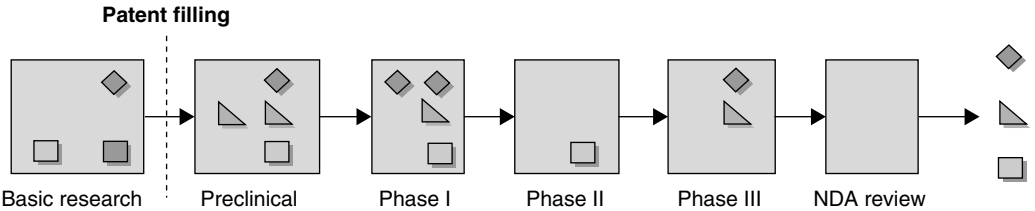
1. Introduction

Understanding the value of a product development project is essential to the scientific management of the product development process. The value of a project to a firm depends not only on the project's properties but also on those of the other projects the firm is developing. This is due to interactions with the other projects in the firm's same development portfolio. Understanding these *portfolio-level project interactions* is central to a firm's choice of project portfolio and development capacity (Loch et al. 2001a, Kavadias and Loch 2003). Decision support models for portfolio choice have provided analytical models of these interactions (cf. Ding and Eliashberg 2002, Loch and Kavadias 2002). In this study, we empirically investigate the structure and significance of these interaction effects.

We use the natural experiment of a product development failure to estimate the value of an individual project. We design an event study around the failure of a late-stage development project. This event study gives us a metric of the change in the firm's value (as measured by the stock markets) due to the failure of a development project with all other factors affecting the firm's value being held constant (MacKinlay 1997). This change in firm value is an empirical measure of the value of the failed project to the firm. We then explain the variance in the value of all failed projects in our sample based on the interactions of the project with other projects in the portfolio. Specifically, we investigate how the value of a project to a firm may depend on the presence of other projects in the firm's portfolio that address the same customer need or utilize the same development resources.

The specific context of our empirical examination is the pharmaceutical industry. New product development in the pharmaceutical industry is regulated and thus, proceeds along a series of well-defined steps illustrated in Figure 1. Drug development starts with an investigation of the chemical and biological properties of a compound in the lab (basic research), followed by animal trials (preclinical studies) and, finally, three stages of clinical trials or trials in human subjects (phase I, II, and III). Our study is designed around the failure of development projects currently undergoing phase III clinical trials, which is the final stage in the development process, where the efficacy

Figure 1 Drug Development Process



	Basic research	Preclinical	Phase I	Phase II	Phase III	NDA review
Key issue	Compound discovery and screening	Animal testing for safety and metabolism	Safety in small sample of healthy volunteers	Safety and efficacy in sample of patients	Safety and efficacy in large sample of patients	FDA review of filling
Average duration (years)		1.5	1	2	2	1
Average output rate (compounds/year)	12.5	6.25	3.75	1.25	1.13	1
Average cost of phase per unit of output (MM$/compound)	12	13	32	141	220	24

Notes. Symbols ◇ □ ◺ ● indicate different patient needs (clinical target indicators). The values are for a typical pharmaceutical company and are normalized based on an average annual output of one compound. Development projects that target the same market or indication are denoted with the same symbol. Estimates were obtained from the Parexel pharmaceutical R&D Statistical Source Book, 2002/2003.

and safety of the drug is investigated in a large sample of patients. Common causes for failures at this stage include adverse side effects of the drug and harmful interactions with other drugs. For a detailed description of the drug development process, the reader is referred to Pisano and Rossi (1994) and Girotra et al. (2004).

The pharmaceutical industry presents an ideal domain of enquiry for our study. It is a large industry where product development is central. Although the product development process closely resembles the classic *stage-gate* development process prevalent in most industries, the role of regulators in the later phases of the drug development process significantly simplifies the empirical design of our study. The different stages in the product development process are clearly and uniformly defined by the regulator, and all pharmaceutical firms must pass their development projects through the same development stages. This allows us to identify projects at the same stage of development across different firms in the industry. The results of each stage of the development process are public knowledge. This allows us to create our data set of product development failures from public sources. Finally, the late-stage product development portfolio of each firm is public knowledge. Thus, when the stock markets value failures, they have the information on the portfolio-level project interactions which we are investigating.

We find empirical evidence for two portfolio-level project interactions. First, we find that the impact of the failure, which is our measure for the value of the project to the firm, is smaller when the firm is developing other projects for the same market as the failed project. When the firm is developing multiple projects for the same market, the failure of one of them does not preclude the firm from earning sales in that market. Thus, the marginal value of any of the multiple projects is smaller than the value of a single project being developed for the market.

Second, we find evidence that the value of a compound or the impact of a failure is smaller if the firm has more projects than the anticipated number in its portfolio, which require the same resources as those used by the failed project (even if the other projects are not "backups" for the same market). A failure leads to the freeing up of resources shared by the failed and other projects. These freed-up resources can be redirected to other projects, which may then be brought to the market sooner than they would be if there was no failure. Thus, failures in portfolios with more than the anticipated number of projects that utilize the same resources as the failed project lead to the

acceleration of the other compounds in the portfolio and have a smaller impact.

This study enhances our understanding of portfolio-level project interactions. We build and validate a theory on the financial effect of these interactions. We find that these interactions significantly alter the value of a development project to a firm and are thus crucial to portfolio choices. Our empirical evidence also allows for a critical examination of the existing analytical literature on portfolio and capacity choices with respect to the modeling of project interactions. This can help us understand the reasons behind the limited application of this literature in practice (Loch et al. 2001a, Loch and Kavadias 2002, Shane and Ulrich 2004) and inspire new improved analytical models, which take into account the empirical regularities that we find. Finally, our results also provide a data-driven model that aids in valuing individual projects in the context of the product development portfolio of a firm. This is useful in valuing development projects available for in-licensing and comparing alternative development projects.

2. Prior Literature

Two streams of academic work are relevant to this study: (1) research on portfolio choices and (2) research on the financial impact of product development outcomes. An established body of literature in operations research attempts to provide optimal product portfolio decisions. Initially, optimization models were developed in a static and deterministic setting, with the decision modeled as one-shot choice under complete information, often with a mathematical programming formulation (see e.g., Lucas 1971). More recent work has emphasized the stochastic, dynamic, or process nature of the problem and has analyzed capacity and congestion effects (Loch and Terwiesch 1999) as well as strategies for search and information gathering (Loch et al. 2001b, Dahan and Mendelson 2001).

Portfolio-level project interactions are central in many of the contemporary models on portfolio selection. Loch and Kavadias (2002) present a dynamic model of portfolio selection within a budget constraint, taking into account multiple project interactions, including those arising out of shared markets in a general setting. In their study on the number of development approaches to pursue for a given market, Dahan and Mendelson (2001) model the interactions between projects of different quality that target the same market. Ding and Eliashberg (2002) investigate the number of development approaches to pursue for a given market in a staged development process. They build an analytical model of the interactions between projects targeting the same market. In their model, unlike Dahan and Mendelson (2001), all successful projects are assumed to have identical quality. In our paper, we empirically examine interactions similar to those investigated by Ding and Eliashberg (2002).

Adler et al. (1995) build a model of project interactions due to shared development resources. Using a development project as their unit of analysis they find that if development resources are stretched, the project completion times are longer. In contrast to Adler et al. (1995), we take an empirical approach and study the effect of shared development resources at the portfolio level. We examine the impact of one project on other projects in the portfolio. Further, we find the impact of these interactions on the financial value of the project as opposed to the development lead time.

Multiple studies have focused on the impact of product development events on financial value, notably Hendricks and Singhal (1997) on the impact of product development delays. They find significant negative stock returns associated with the announcement of product introduction delays. Industry competitiveness and the firm's degree of diversification influence the size of this impact. Chaney et al. (1991) and Chaney and Devinney (1992) study the stock market reaction to announcements of new products across a wide range of industries. Bayus et al. (2003) study the impact of new product introductions in the personal computer industry on profit rate, profit rate persistence, and asset growth. Robertson et al. (1995) and Chen et al. (2005a) study the impact of new product announcements on competing firms. Chen et al. (2005b) examine the effect of product introduction delays on industry rivals. Sharma and Lacey (2004) compare the impact of pharmaceutical successes and failures on firm value. This body of work quantifies the impact of these product development events. Further, they explain the variance in the financial impact of product development with product or industry properties, but not the portfolio.

We build on the rigorous methodologies developed in this literature to empirically value projects. However, in contrast to this literature, we relate the impact of the product development outcomes (our measure for the financial value of projects) to key properties of the product development portfolio—the presence of other compounds in the portfolio that target the same unmet market need, and the availability of research opportunities that can utilize resources freed up due to the failure.

3. Theory Development

Failure of a late-stage development project represents the loss of potential future sales for a firm, which should lead to a decrease in the value of the firm.

In the pharmaceutical industry, a drug undergoing phase III clinical trials has an average approval probability of about 80% (Parexel 2002/2003). On approval, an average drug has the potential to generate sales of hundreds of millions of dollars. When a phase III failure occurs, these potential sales are lost; the 20% probability of failure is updated to certainty. This gives us our baseline hypothesis:

Hypothesis 1. *The value of a pharmaceutical firm falls when a compound fails in phase III clinical trials.*

Not surprisingly, previous research has proposed and found evidence for similar hypotheses. Sharma and Lacey (2004) propose a similar hypothesis in their study of stock market reactions to news from the pharmaceutical industry.

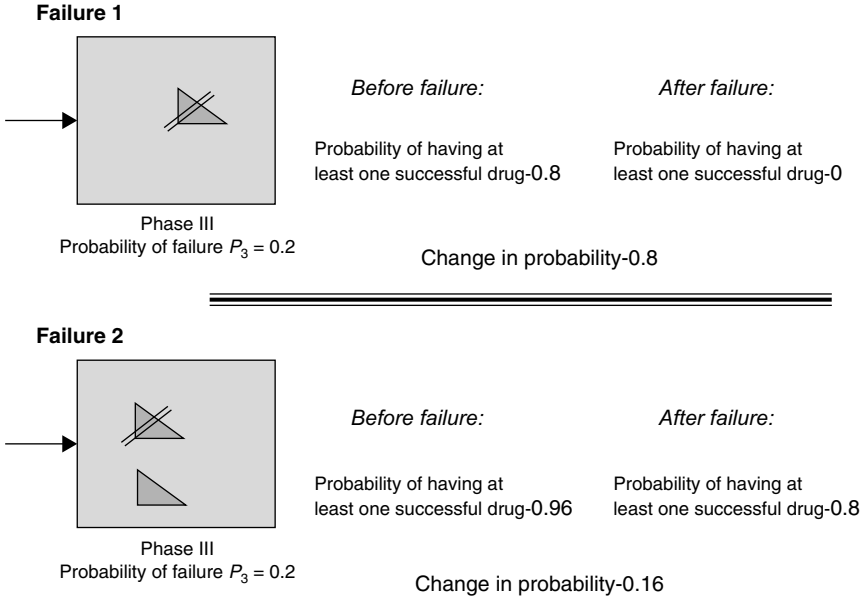
3.1. Effect of Projects Targeted at the Same Market

Drug development, like most other product development, is associated with long development lead times and low odds of success. The clinical trials phase of the drug development process alone takes an average of six years to complete, and only one out of six drugs that enter clinical trials makes it to the market. Fortunately, there are often multiple, unrelated technological approaches available to address the same medical need. For instance, there are multiple chemical compounds that pharmacologically inhibit the COX enzymes, which provide relief from the symptoms of inflammation and pain. These compounds differ in their side effects and thus, their success or failure in late-stage clinical trials are largely unrelated. In such settings, firms follow a parallel development strategy that increases the likelihood of developing a viable product for a given lucrative market within a reasonable time frame. The candidate compound farthest along in the development process is referred to as the lead compound, and the other compounds are referred to as backup compounds. Such parallel development strategies have been shown to be optimal in a variety of product development settings where the odds of success are low, the development lead times are long, and the correlation between the successes of different concepts is low (Loch et al. 2001b, Ding and Eliashberg 2002).

Although firms often develop multiple compounds to address a given market opportunity, firms rarely market more than one of them. Introducing several drugs for the same medical need earns the same sales as introducing only one successful compound (Ding and Eliashberg 2002, Girotra et al. 2004). Firms typically introduce the first one that passes clinical trials, and then cease development of all other drugs targeted toward this market. Consequently, the probability of having at least one successful drug for a given market is a crucial metric related to the financial value of a product development pipeline. By developing multiple compounds for the same market, firms increase this probability. Thus, the marginal value of a compound to a firm is proportional to how much the compound increases the probability of having a successful drug in its market for the firm. To illustrate this concept consider the two scenarios provided in Figure 2.

In Figure 2, the compounds of concern have a probability of failure of 20%. In failure 1, the firm is

Figure 2 Example Illustrating the Role of Backup Compounds



developing only one compound for the indication. Prior to the failure, the firm has an 80% probability of obtaining at least one successful drug. After the failure, this probability changes to 0%. Due to the failure, the probability of earning the sales from this indication falls by 80 percentage points. Equivalently, the change in the expected future cash flows from the pipeline due to the failure is proportional to 80 points. Alternatively, consider failure 2, where the firm is developing two compounds for the indication. Prior to the failure, the firm had a 96% $(1-(0.2)^2=0.96)$ chance of having at least one successful drug in the indication. After the failure, this changes to 80%. As a result of failure 2, the change in probability is 16 percentage points. The change in the expected future cash flows from the pipeline is thus proportional to 16 points. The presence of a backup compound in the case of failure 2 mitigated the impact of the failure. This should be reflected in the stock market reaction to the failures, or the valuation of these development projects. Next, we extend this logic to a general portfolio with compounds in each of the three phases of development.

Consider a pipeline with n_i candidate drugs undergoing phase i $(i=1,2,3)$ development for the market in question. Let p_i denote the probability of failure of a drug currently in phase i trials *during phase i or in any subsequent phase.*[1] From this pipeline the firm could earn the sales from the market under three mutually exclusive realizations of the clinical trials: (1) One of the compounds currently undergoing phase III trials succeeds[2] (probability $(1-p_3^{n_3})$); (2) all the compounds currently undergoing phase III trials fail, but one of the compounds currently undergoing phase II trials succeeds in phase II trials and in phase III (probability $p_3^{n_3}(1-p_2^{n_2})$); (3) all the compounds currently undergoing phase III and phase II trials fail, but one of the compounds currently undergoing phase I succeeds in phase I trials and all subsequent phases (probability $p_3^{n_3}p_2^{n_2}(1-p_1^{n_1})$).

Under each of the three scenarios, the firm earns the sales from the market; however, scenarios 2 and 3 have longer times to market. Thus, their sales should be discounted using appropriate discount factors, α_2 and α_1, respectively. The expected net present value of sales from this indication is thus given as

$$\mathrm{E}[NPV]=\mathrm{E}[M]*[(1-p_3^{n_3})+\alpha_2 p_3^{n_3}(1-p_2^{n_2})+\alpha_1 p_3^{n_3}p_2^{n_2}(1-p_1^{n_1})].$$

[1] If the probability of success in phase i is given as $s_i\in[0,1]$; $p_i=1-\prod_{k=i}^{3}s_i$. Thus, $p_1\geq p_2\geq p_3$ with the equality arising iff the probability of success in any stage is 1. We assume that all drugs at the same stage of trials and targeted for the same market have the same probability of success. Danzon et al. (2005) provides empirical support for this assumption.

[2] We assume clinical trials of different drugs are independent.

Here, $\mathrm{E}[M]$ is the expected value of sales from the target market conditional on having a successful drug. Equivalently,

$$\mathrm{E}[NPV]\propto(1-p_3^{n_3})+\alpha_2 p_3^{n_3}(1-p_2^{n_2})+\alpha_1 p_3^{n_3}p_2^{n_2}(1-p_1^{n_1}),$$

where the expected sales serve as the proportionality constant. We refer to the right-hand side of the above expression as the time-adjusted probability of indication success (or TAPIS),

$$TAPIS(p_1,p_2,p_3;n_1,n_2,n_3)=[(1-p_3^{n_3})+\alpha_2 p_3^{n_3}(1-p_2^{n_2})+\alpha_1 p_3^{n_3}p_2^{n_2}(1-p_1^{n_1})]. \quad (1)$$

TAPIS is proportional to the expected financial returns from the development projects. The marginal value of a phase III compound, or the impact of losing a phase III compound in the portfolio, is thus proportional to the consequent change in TAPIS $(n_3\rightarrow n_3-1)$:

$$\Delta TAPIS=TAPIS(n_3)-TAPIS(n_3-1)=(P_3^{n_3-1}-P_3^{n_3})[1-\alpha_2(1-P_2^{n_2})-\alpha_1 P_2^{n_2}(1-P_1^{n_1})]. \quad (2)$$

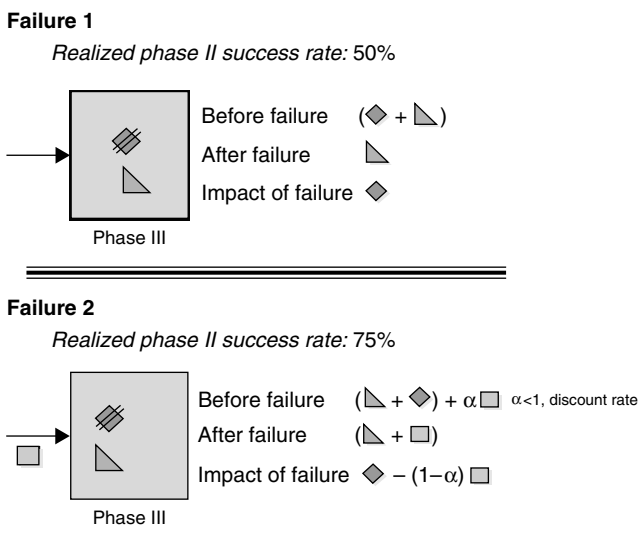
Equation (2) is the change in the time-adjusted probability of earning the sales from the target market. This is proportional to the reduction in the expected value of the future cash flows from the sales in the target market. Thus, failures for which the expression in (2) is large should have a higher financial impact.

Hypothesis 2. *The change in the value of a pharmaceutical firm when a phase III compound fails is negatively correlated with the change in the time-adjusted probability of success for an indication* ($\Delta TAPIS$ *in Equation* (2)).

3.2. Effect of Other Compounds in the Portfolio That Utilize the Same Development Resources

Phase III trials involve the examination of the safety and efficacy of the drug in a large sample of patients. The primary resources required at this stage are clinical trial sites and bio-statisticians. Irrespective of the disease or indication, all compounds draw from the same pool of resources. It is quite costly to scale the capacity of these resources up or down in the short term because these resources are mostly professionals hired with a multi-year expectation of employment, or fixed assets, which take time to set up. Consequently, firms must make long-run commitments to the capacity of these sticky phase III resources. Firms set up phase III capacity well in advance of observing the results of the most recent phase II trials or the current demand for phase III resources. On the basis of the long-run or expected probability of success in phase II and the phase II capacity, firms forecast the demand for phase III

Figure 3 Example Illustrating the Effect of Different Realized Phase II Success Rate



resources, and make long-term commitments to R&D capacity (Girotra et al. 2004).

For instance, consider a firm that from its past experience estimates that the average phase II success rate across all diseases is 50%. If the firm has the capacity to process four compounds per year in phase II clinical trials, it expects that two of the four compounds will succeed and proceed to phase III resources. Thus it establishes phase III capacity as 50% of phase II capacity, two compounds per year.

The *actual* demand for phase III resources at any point in time is a function of the *actual* realizations of recent phase II trials. Continuing with our example of the firm that has set phase III capacity to two compounds per year, consider two possible realizations of phase II trials (Figure 3): (1) The realized success rate is 50% (two of the four compounds in phase II trials succeed); (2) the realized success rate is 75% (three of the four compounds in phase II trials succeed). In case 2, the firm has more compounds than the available phase III resources, thus one compound has to wait in a "buffer" before phase III.

Now consider that out of the two compounds currently undergoing phase III trials in each of the above scenarios, one compound fails; e.g., the compound denoted by the rhombus in Figure 3. In both cases, the firm loses the potential future sales from the failed compound; however in case 2, there is a mitigating effect: One compound is waiting in the buffer (denoted by the square in Figure 3), which can now take advantage of the freed-up phase III resources. The value of this waiting compound actually increases as it can now enter phase III trials and be brought to the market earlier. There is no such mitigating effect of the failure in case 1. Thus, the impact of the failure in case 2 should be *smaller* than the impact of the failure in case 1.

Failures that come at a time when the realized phase II success rate in the firm's pipeline is higher than expected lead to an acceleration of the other compounds in the pipeline, and their impact should therefore be lower. Further, the benefit associated with the acceleration should depend on the unanticipated demand, or the number of compounds that are waiting in the "buffer," captured by the degree to which the realized success rate was higher than the expected phase II success rate.

Although the above example uses the notion of a buffer to illustrate this effect, literal presence of a buffer is not required for the acceleration. As long as a higher resource utilization of resources leads to longer processing times, the freeing up of resources due to failures will have the beneficial impact of accelerating compounds in the pipeline. The higher the utilization, the higher the benefit, because the extent of the acceleration and the number of projects that benefit from this acceleration are both higher at higher levels of utilization.

Hypothesis 3. *The decrease in firm value from a phase III failure is lower* (*higher*) *if the firm has experienced an above* (*below*) *average phase II success rate in the period prior to the failure*.

Note that Hypothesis 3 is based on the number of successful phase II compounds in the recent past across all target markets; whereas Hypothesis 2 is concerned with "backup" compounds in all phases but only in the target market of the failed compound.

4. Data Source

We use drug development data from the R&D Insight database developed by ADIS international. It is compiled by a team of scientific editors that track more than 17,000 drugs in active development from over 200 pharmaceutical companies. The primary sources of information are: direct contact with companies, information collected from medical and biomedical journals, attendance at international meetings and conferences, company annual reports, news services, press releases, licensed Lehman Brothers' Pharma-Pipelines data, and public information from the Food and Drug administration. Drug development is tracked from the earliest laboratory report and continues through world market launch. Every scientific or commercial development advancing the drug's progress to market is assessed, evaluated, and verified for authenticity before being reported in the database. The database is used by many leading pharmaceutical companies to monitor the competitive landscape. A sample entry for a failed drug from the ADIS database

Table 1 Firms in the Study

Firm/subsidiaries	No. of phase III failures (1994–2004)	Annual sales+ (MM$)	No. of employees+ ('000s)	R&D expenditure+ (MM$)	R&D/sales (%)
Pfizer Inc.	19	21,107.97	67.82	3,788.98	17.95
Aventis	18	18,159.86	82.81	2,248.14	12.38
Bristol Myers Squibb, Mead Johnson	13	17,146.50	48.88	1,994.80	11.63
Glaxo SmithKline	13	20,237.11	77.07	2,784.24	13.76
Wyeth, Wyeth Vaccines	9	13,546.55	56.91	1,628.51	12.02
Chiron Co., Chiron Vaccines	9	976.37	4.63	302.31	30.96
Genentech	8	1,544.39	4.01	439.66	28.47
Novartis	8	22,852.80	82.24	2,886.46	12.63
Amgen, Amgen Boulder	7	3,581.10	6.73	1,092.44	30.51
AstraZenca	6	12,998.73	44.24	1,960.31	15.08
Pharmacia, Monsanto	5	10,707.44	36.91	1,547.00	14.45
Abbot GMBH, Abbott Labs	5	13,511.70	56.60	1,262.00	9.34
Eli Lilly	5	9,355.40	33.97	1,706.10	18.24
Merck	3	29,708.38	60.31	2,147.03	7.23
Bayer	1	30,645.40	131.56	2,352.43	7.68
Novo Nordisk	1	2,985.00	15.02	462.82	15.50
Schering AG	1	4,854.75	25.12	875.22	18.03
Alcon	1	2,929.33	11.56	307.00	10.48
Median	6.5	13,255.215	46.56	1,667.305	14.10

+Data refer to values reported in the COMPUSTAT Industrial Annual database averaged over the period of the study (1994–2004).

is provided in §EC.4 of the online supplement (provided in the e-companion).[3]

We use the ADIS database to identify dates of drug failures, the associated indication, the ownership pattern of the drug, and the portfolio properties (other compounds in development (n_1, n_2, n_3), and the recent success rates) on the failure date. Finally, we look at industry-wide historical data on successes and failures from the ADIS database to estimate the success probability of compounds in each indication (p_1, p_2, p_3).

To verify the data on the firms' portfolio (n_1, n_2, n_3) properties, we imputed the pipeline for one firm in our data set (Merck & Co), and compared it with information obtained from within the firm. The two portfolios were identical. To verify the authenticity of the failure announcements, we checked a sample of failures with the lead pharmaceutical analyst at a financial firm. Failures were found to be accurate in both date and indication specification. Some of the more prominent failures in our sample received extensive coverage in the popular press. We compare the dates of these failures from the ADIS database with news reports (in the Factiva database) and find the data from ADIS to be accurate. A news report and the associated ADIS database entry is provided in §EC.5 of the e-companion.

We restrict our attention to phase III failures that originated at publicly traded firms with common stock listed on any U.S. market at the time of the failure. We get stock price data from the CRSP financial database.[4] We identify the ownership and holding pattern of the originator firm(s) at the time of the failure using the Corporate Affiliations data set maintained by the Lexis-Nexis group. Some descriptive statistics on the firms included in this study are provided in Table 1. The annual sales (averaged over the period of the study) for firms in our data set range from over US$20 billion for the big pharmaceutical firms such as Merck, Pfizer, and Bristol Myers Squibb to just under US$1 billion for biotechnology firms such as Chiron. The median firm in our data set has annual sales of US$13.26 billion, employs 46,560 employees, spends US$1.6 billion annually on R&D (14.10% of sales), and experiences 6.5 phase III failures during the period of our study.

During the time period of the study, 1994–2004, there were 132 phase III failures for publicly traded pharmaceutical firms in our database; they represent our sample. Less than 2% of the events are within a month of other related events, thus assuaging any concerns about clustering.

[3] An electronic companion to this paper is available as part of the online version that can be found at http://mansci.journal.informs.org.

[4] The CRSP Database provides access to NYSE, AMEX, and NASDAQ daily and monthly securities prices, as well as to other historical data related to over 20,000 companies. The data is produced and quarterly updated by the Center for Research in Security Prices (CRSP), a financial research center at the Graduate School of Business at the University of Chicago.

5. Methodology and Variables

5.1. Measuring the Impact of Drug Failures

To measure the implications of a late-stage failure, we use an event study methodology (MacKinlay 1997, Kothari and Warner 2007). Event studies have been applied to quantify the impact of a wide variety of firm-specific and economy-wide events. Notable applications from the finance and accounting literature involve measuring the impact of mergers and acquisitions, earnings announcements, issue of new debt or equity, and announcements of macroeconomic variables (trade deficits, unemployment data, interest rates). Notable applications from the strategy literature include studies on the impact of CEO succession, name changes, diversification, takeovers, and competitive entry. In the product development and supply chain management literature, they have been employed to estimate the impact of new product introductions (Chaney et al. 1991), delays in new product introductions (Hendricks and Singhal 1997), supply chain disruptions (Hendricks and Singhal 2005), ISO 9000 certification (Corbett et al. 2005), and of excess inventory (Singhal 2005).

Using the prices of a firm's tradable securities in financial markets, an event study measures the impact of a specific event on the value of a firm as measured by the price of its common stock. The logic behind this approach is the efficient-market hypothesis: Given rationality and information in the marketplace, the impact of an event should be reflected by the change in the stock price of the firm.

Event study methodologies provide a rigorous foundation to isolate the change in stock price due to the event from the change in stock price due to other factors known prior to the event. A model for the returns given the information prior to the event is first estimated using historical data over the *estimation period* for each event in the study. This estimated model is then used to predict the expected returns on the day of the event, conditional on no new information or events. To ensure robustness of our findings, we use three alternative return-generating models for predicting the expected returns, the comparison period model (CP), the market model (MM), and the Fama-French three-factor model (FF). Details of the three models are provided in §EC.2.1 of the e-companion. These models give us the expected returns on the day of the failure, taking into account the impact of market and firm specific factors, but in absence of the failure.

The component of the return that *cannot* be explained by the return-generating models, or the difference between the actual return and the expected return, is attributed to the event—in our case the failure of the phase III clinical trials. This component is commonly referred to as the abnormal return. If no economically relevant information is available, we expect this abnormal return to be zero.

Often, the impact of the event is not limited to the day of the occurrence, but extends a few days before and after the event. This is called the *event window*. We use multiple event windows, including those suggested by looking at trading volumes using the techniques proposed by Tkac (1999) (detailed in §EC.2.2 of the e-companion). We then aggregate the abnormal return for each day in the respective event windows to obtain our main dependent variable, the cumulative abnormal return or CAR_i. This variable captures the financial impact of losing a compound while controlling for other factors that influence firm value. This is an empirical measure of the value of each compound to the associated firm, our dependent variable.

For Hypothesis 1, we test the null hypothesis that $CAR_i = 0$. We report the cross-sectional standard deviation test (the "standard approach" from MacKinlay 1997), the standardized Patell-Z test statistic (Patell 1976),[5] a test that controls for cross-sectional dependence between individual security returns ("crude dependence adjustment test" from Brown and Warner 1980, p. 233),[6] a nonparametric generalized sign-z statistic (Sprent 1989),[7] and the nonparametric Wilcoxon signed rank test. We test this hypothesis for several typical event windows as well as for the event window implied by excessive trading volumes.

To test Hypotheses 2 and 3, we run a linear regression with CAR_i as the dependent variable, the two explanatory variables, $\Delta TAPIS$ and the *phase II buffer*, in addition to the control variables. We describe the construction of the two explanatory variables and the control variables in the next three subsections.

5.2. Explanatory Variable: $\Delta TAPIS$

To test Hypothesis 2, we need to construct our explanatory variable, $\Delta TAPIS$ (Equation (2)). $\Delta TAPIS$ is a function of the number of compounds at each stage of development (n_1, n_2, n_3) and the probabilities of success of each compound (p_1, p_2, p_3). n_1, n_2, and n_3 are obtained from the ADIS database as the number of distinct compounds in each of the three stages of trials for the same market as the failed

[5] Unlike the cross-sectional standard deviation test, in computing the Patell-z statistic, each abnormal return is standardized using the estimated variance of the abnormal return obtained from the estimation period model.

[6] This test uses a single variance estimate for the entire portfolio thereby avoiding the potential problem of cross-sectional correlation of security returns.

[7] The nonparametric sign-z test tests the null hypothesis that the number of positive and negative return is the same (Sprent 1989).

compound. To estimate the probabilities, we use data on all clinical trials in the ADIS database. A vast majority of these trials are run by firms that are not publicly traded and do not otherwise appear in our sample. Danzon et al. (2005) find that the indication explains the largest fraction of the variance in success probabilities between different drugs. Thus, we estimate p_i at the level of an indication and assume that all drugs for an indication at the same stage of development have the same odds of failure. For example, to determine p_i for an Alzheimer's drug, we look at the performance of all Alzheimer's drugs irrespective of originating firm. The estimated-indication phase-specific probabilities and the detailed procedure are provided in §EC.3 of the e-companion. Finally, we use an annual discount rate of 12%, to compute α_1 and α_2.[8] A minority of compounds (25 out of 116) in our sample fail for more than one indication at the same time (often due to safety concerns), thus they have more than one $\Delta TAPIS$ value associated with them. For these compounds, we compute an aggregated $\Delta TAPIS$ as the sum of the multiple $\Delta TAPIS$ values and present our results using the same. We also tested our results using the average and the maximum of the multiple $\Delta TAPIS$ values and found similar results.

5.3. Explanatory Variable: *Phase II Buffer*

Hypothesis 3 posits that the impact of the phase III failure is proportional to the unanticipated demand for phase III resources at the time of the failure. The unanticipated demand is captured by the difference between the recent and the expected phase II success rate for the firm in question (*phase II buffer*). To compute the phase II success rates, we divide the number of phase II successes (across all indications) by the total number of phase II trials (sum of the number of successes and failures across all indications) over the relevant time period for the firm in question. For the expected success rate, we look at the number of successes and failures for the firm over all periods of the data set, 1994 to 2004. To compute the recent success rate, we look at the number of successes and failures in the 300-day period preceding the day of the failure announcement.[9] The difference of the two is used as the explanatory variable in our regressions (Equation (3)).

$$\begin{aligned} &\textit{Phase II buffer} \\ &\quad = \textit{Recent success rate} - \textit{Long-run success rate} \\ &\quad = \frac{\#\ \textit{of Successes}_{t\in[-300,\,0]}}{\#\ \textit{of Successes}_{t\in[-300,\,0]} + \#\ \textit{of Failures}_{t\in[-300,\,0]}} \\ &\qquad - \frac{\#\ \textit{of Successes}_{\forall t}}{\#\ \textit{of Successes}_{\forall t} + \#\ \textit{of Failures}_{\forall t}} \end{aligned} \tag{3}$$

To test the robustness of our results, we also run our regression models with just the recent success rate, the absolute number of recent successes, and the log of the recent success rate minus the log of the long-run success rate. Our results are robust to all these formulations.

5.4. Control Variables

We control for the properties of the compound in question and the firm in question. At the compound level, we include three variables: First, the number of active trials at the time of the failure in the same indication as the focal compound across all firms present in the ADIS database (*NActiveTrials*). Previous research (Nicholson et al. 2005) finds that this variable is highly correlated with the revenue potential of the compound. Second, the number of licensees for the compound in question (*Nlicensees*). Depending on the structure of the licensing agreement, this variable is associated with the firm's financial stake in the compound. Third, the number of originating firms associated with the compound (*NOriginators*). This variable is also related to the firm's financial stake.

We also include two firm-specific control variables, sales in the quarter of the failure (*Sales*) and R&D expenses incurred by the firm in the quarter of the failure (*R&D Expenses*). These capture the firm-specific heterogeneity, namely the size of the firm and the R&D organization associated with the failure. We estimate the model in Equation (4).

$$\begin{aligned} CAR_i = {} & a_0 + a_1(NActiveTrials)_i + a_2(NOriginators)_i \\ & + a_3(NLicensees)_i + a_4 Sales_i \\ & + a_5 R\&D\ Expenses_i + a_6 \Delta TAPIS_i \\ & + a_7(Phase\ II\ buffer)_i + \varepsilon_i \end{aligned} \tag{4}$$

Hypothesis 2 implies negative and significant estimates for the coefficient a_6. Hypothesis 3 implies positive and significant estimates for the coefficient a_7.

Descriptive statistics of the explanatory variables are provided in Table 2. The median failed compound has one originator firm, is targeted at one indication, has one backup compound, and has no licensees. Table 3 shows the Pearson correlation coefficients among the variables.

[8] A compound in phase I (II) of development takes on average four (two) years longer to reach the market than a compound in phase III of development, thus it is worth significantly less to the firm. $\alpha_1 = (1 + x/100)^{-4}$ and $\alpha_2 = (1 + x/100)^{-2}$, where x% is the cost of capital. The results are presented using $x = 12\%$, implying $\alpha_1 = 0.636$ and $\alpha_2 = 0.797$. The results remain unchanged for a wide range of values that we tested ($x = 5\%$ to 35%).

[9] We consider alternate specifications that look at periods of 200 days and 400 days before the failure. The results are near identical.

Table 2 Independent Variables

Independent variables	Mean	Median	Maximum	Minimum	Standard deviation
NActiveTrials	24.99	15	169	0	27.35
NLicensees	0.76	0	8	0	1.32
NOriginators	1.20	1	4	1	0.56
Sales (MM$)	3,767.27	3,627.01	13,982.00	117.58	2,494.14
R&D expenses (MM$)	655.98	483.00	3,266.08	67.45	563.77
Δ*TAPIS*	0.7006	0.6417	2.2661	0.0294	0.5013
Phase II buffer	0.0413	0.0791	0.4138	−0.8095	0.2169

6. Results and Discussion

6.1. Identification of Event Window: Abnormal Trading Volume

We first estimate a model for the benchmark trading volume for each firm and find the days associated with unexpectedly high trading. The abnormal trading volume data are illustrated graphically in Figure 4. Abnormal trading volume peaks at 816 million shares over the expected volume two days before the announcement of the failure. Abnormal trading volumes in the time period $(-2, 4)$ are found to be statistically different from zero. This implies that most of the information about the event failure is incorporated in the value of the firm during this period. Thus, the abnormal returns in the period $(-2, 4)$ should capture the effects of the event on the firm's valuation. To ensure robustness of our results with respect to the choice of the event window, we also conduct all further analyses for alternative event windows ($(-3, 3)$ and $(-4, 4)$), which are often used in the event study literature.

6.2. Impact of Phase III Failures: The Average Cumulative Abnormal Return

We estimate three benchmark expected return models—the comparison period model (CP), the market model (MM), and the Fama-French three-factor model (FF). These models are estimated individually for each security event in our database using data from the event window $(-255, -10)$. For three data points, fewer than 10 trading days of market data are available in this estimation period. These points are excluded from our sample. The estimated benchmark return models are used to compute the daily and cumulative abnormal returns as described in §5.1.

Table 4 reports the results for the mean and median cumulative abnormal returns. The mean cumulative abnormal returns are negative and significant for a wide variety of model specifications, event windows, and test specifications. These results provide evidence for Hypothesis 1 predicting a negative effect of a drug failure on firm value. Over the time period of the event window, $(-2, 4)$, a phase III drug failure leads to an average loss of 1.46% in the value of the firm (according to the Fama-French model, estimates range from −1.07% to −1.61% using different models and event windows). In dollar terms, these losses correspond to a decrease in the firm value by US$405 million.

Chaney et al. (1991) report in their study of announcements of product *successes* a cumulative abnormal return of 0.21% using the market model for the pharmaceutical firms in their sample. For an average phase III compound, with an 80% chance of success, the increase in probability on completion of a successful trial is 20% as opposed to an 80% reduction in probability on failure. Thus, we expect the financial impact from the study by Chaney et al. (1991) to be four times smaller. Our results from the market model are consistent with this plausibility check.

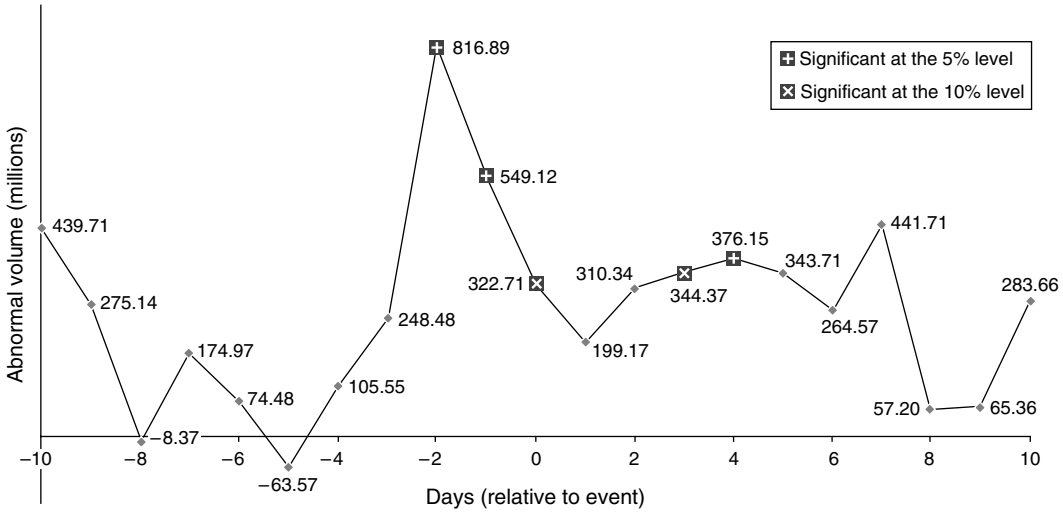
Sharma and Lacey (2004) construct in their comparative study of drug development failures and successes a data set of 41 FDA rejections of new drug approval (NDA) applications (the last stage in Figure 1, subsequent to phase III success). An NDA rejection indicates managerial failure in implementing a firm's internal controls and systems to ascertain drug safety. This is a more serious and rarer event than the medical failure of a drug undergoing phase III clinical trials. Thus, the expected impact of these events should be larger than that of our event. They find abnormal returns as high as 21% associated with the announcement. This corresponds to a financial impact of over US$7 billion, which is much higher than the net present value of the sales of even the

Table 3 Pearson Correlation Between Independent Variables

	NActiveTrials	*NLicensees*	*NOriginators*	*Sales* (net) (MM$)	*R&D expense* (MM$)	*Phase II buffer*	Δ*TAPIS*
NActiveTrials	1						
NLicensees	−0.18*	1					
NOriginators	−0.14*	0.07	1				
Sales (net) (MM$)	0.29***	−0.28***	−0.29***	1			
R&D expense (MM$)	0.2*	−0.22**	−0.27**	0.61	1		
Δ*TAPIS*	0.33***	−0.07	−0.08	0.22**	−0.08	1	
Phase II buffer	0.12	0.14*	−0.08	0.06	0.02	0.04	1

*Significant at the $p < 10\%$ level, **significant at the $p < 1\%$ level, ***significant at the $p < 0.1\%$ level.

Figure 4 Abnormal Trading Volume



biggest blockbuster drugs. This finding suggests that investors may be losing confidence in the firm's management on account of this kind of rare failure and may be penalizing it for much more than just the lost compound.

The cross-industry study of Hendricks and Singhal (1997) reports that on announcement of product development delays firm values drop by an average of 5.25%, or US$119 million. Chen et al. (2005b) report that on announcement of delays firm values drop by 11.4%.

6.3. Effects of Backup Projects and Recent Success Rate

Table 5 provides results from the OLS estimation of the model in Equation (4). Results are provided for the three return-generating models described in §5.1 and three event windows. We obtain similar results using a WLS regression, with the precision of the estimated abnormal returns as the weights. The R^2 for our models ranges from 12% to 19%, which is comparable to other studies of this type (e.g., Chaney et al. 1991). The regressions using the dependent variable created from the market model and the Fama-French model are all significant at the $p < 0.01$ or higher level. The regressions using the comparison period model are significant at the $p < 0.05$ level. Diagnostic tests reveal no problems with heteroskedasticity or multicollinearity. We also examine our estimation procedures for influential observations (Belsley et al. 1980) and find our results to be robust to outliers.

Table 4 Cumulative Abnormal Returns

		Mean abnormal return					Median abnormal return		Change in market capitalization	
Model	Window	Mean cumulative abnormal return (%)	Cross sectional std. dev test	Patell-Z statistic$^+$	Crude dependence adjustment test$^\&$	Generalized sign-$Z^\wedge$	Median cumulative abnormal return (%)	Wilcoxon signed rank test	Mean change in market capitalization (MM$)	Median change in market capitalization (MM$)
Comparison period model	(−2, 4)	−1.14	−2.18**	−2.20**	−2.15**	−1.64**	−1.12	−1.53*	−143.29	−32.63
	(−3, 3)	−1.19	−2.17**	−2.69***	−2.25**	−0.77	−1.08	−1.59*	−12.91	−16.52
	(−4, 4)	−1.23	−2.16**	−2.31***	−2.06**	−1.46*	−1.52	−1.90**	−184.17	−40.12
Market model	(−2, 4)	−1.07	−2.04**	−2.15**	−2.15**	−0.19	−0.38	−1.59*	−170.59	−4.90
	(−3, 3)	−1.20	−2.27**	−2.76***	−2.41***	−0.53	−0.48	−2.09**	−95.41	−5.07
	(−4, 4)	−1.38	−2.44***	−2.58***	−2.44***	−1.93**	−0.85	−2.81***	−210.28	−36.12
Fama-French three-factor model	(−2, 4)	−1.46	−3.17****		−2.90***	−1.79**	−0.90	−3.54****	−404.99	−21.77
	(−3, 3)	−1.48	−3.14****		−2.95***	−1.44*	−1.13	−3.89****	−383.86	−51.77
	(−4, 4)	−1.61	−3.18****		−2.83***	−1.79**	−1.11	−3.90****	−451.02	−36.94

Note. Significance levels from a one-tail t-test: *10% significance, **5% significance, ***1% significance, ****0.1% significance. $N = 132$.

$^+$Unlike the cross-sectional standard deviation test, in computing the Patell-z statistic, each abnormal return is standardized using the estimated variance of the abnormal return (Patell 1976), standardized tests are not available for the Fama-French Model.

$^\&$This test uses a single variance estimate for the entire portfolio thereby, avoiding the potential problem of cross-sectional correlation of security returns.

$^\wedge$The nonparametric sign-z test tests the null hypothesis that the number of positive and negative return is the same (Sprent 1989).

Table 5 Regression Model: Parameter Estimates

	Comparison period model			Market model			Fama-French model		
Variables	(−2, 4)	(−3, 3)	(−4, 4)	(−2, 4)	(−3, 3)	(−4, 4)	(−2, 4)	(−3, 3)	(−4, 4)
Intercept	0.0307	0.0186	0.0346		0.0206	0.031		0.0128	0.0176
	(1.37)	(0.76)	(1.45)	(1.31)	(0.9)	(1.34)	(0.48)	(0.64)	(0.87)
NActiveTrials		−0.7618	−0.4597	−0.9947**	−1.09**	−1.02*	−0.5481	−0.8145*	−0.8076*
1,000s	(−1.25)	(−1.35)	(−0.83)	(−1.99)	(−2.04)	(−1.89)	(−1.26)	(−1.75)	(−1.73)
NOriginators		−0.355	−2.12*	−0.606	−0.0004	−1.275	−0.68	−0.257	−0.804
100s	(−1.27)	(−0.32)	(−1.96)		(0)	(−1.21)		(−0.28)	(−0.88)
NLicensees		−0.219	−0.526	−0.908**	−0.606	−0.984**	−0.589*	−0.273	−0.656*
100s	(−1.17)	(−0.49)	(−1.2)	(−2.25)	(−1.41)	(−2.26)	(−1.72)	(−0.74)	(−1.78)
Sales		7.22**	5.88*	5.6*	6.44**	5.68*	6.42**	6.88**	5.78**
1,000,000s	(2.11)	(2.24)	(1.86)	(1.96)	(2.12)	(1.84)	(2.59)	(2.58)	(2.16)
R&D expenses		−4.012***	−2.877**	−2.33*	−3.011**	−2.174*	−2.476**	−3.273***	−2.39**
100,000s	(−2.46)	(−2.93)	(−2.14)	(−1.93)	(−2.34)	(−1.67)	(−2.35)	(−2.89)	(−2.11)
Δ*TAPIS*		−2.476**	−2.514**	−3.029***	−2.956***	−2.831**	−1.863**	−2.23**	−2.178**
100s	(−2.61)	(−2.05)	(−2.12)	(−2.84)	(−2.6)	(−2.46)	(−2.01)	(−2.23)	(−2.18)
Phase II buffer		4.762*	5.684**	6.471***	4.964*	5.707**	6.038***	4.226*	5.384**
100s	(2.39)	(1.71)	(2.07)	(2.63)	(1.89)	(2.15)	(2.81)	(1.83)	(2.33)
Adj *R*-square (%)		6.36	8.70	13.88	9.12	12.43	12.35	8.67	10.69
R-square (%)		12.06	14.26	19.17	14.70	17.81	17.69	14.23	16.13
N#		116	116	116	116	116	116	116	116
F-value		2.12	2.57	3.63	2.63	3.31	3.32	2.56	2.97
Pr > *F* (%)		4.79	1.75	0.15	1.50	0.31	0.31	1.77	0.70

Note. Significance levels from a two-tail *t*-test: *10% significance, **5% significance, ***1% significance.
#Of the 132 failures for publicly traded firms, pipeline data is available only for 116 failures.

We find support for Hypothesis 2. The coefficient for the variable Δ*TAPIS* is found to be significant. A 1% difference in the change in TAPIS leads to 1.86 basis points difference in the abnormal return associated with the failure (according to the Fama-French (−2, 4) model; estimates from other models range from 1.86 to 3.02 basis points). In the illustrative example of Figure 2, our regressions suggest that failure 1 would hurt the firm by an extra 1.19%, or US$334.1 million, compared to failure 2. Further, for one standard-deviation difference in the value of Δ*TAPIS* (calculated from our sample), the difference in financial impact corresponds to 0.92%, or US$258 million. Support for Hypothesis 2 demonstrates that the presence of other compounds for the same market leads to lower financial impact of a failure or lower valuation of a compound.

We also find support for Hypothesis 3. The coefficient for the variable *phase II buffer* is significant. Failures that follow a period where the phase II success rate was higher than the long-run average success rate lead to a less negative impact on the firm value. Conversely, failures that follow periods of below average phase II success hurt firm value more. A 1 percentage point difference in success rate (phase II buffer) leads to a 6.04 basis point difference in the abnormal return associated with the failure. (According to the Fama-French (−2, 4) model, estimates from other models range from 4.2 basis points to 6.5 basis points.) In the illustrative example of Figure 3, our regressions suggest that failure 2 would hurt the firm by an additional 1.52%, or US$419.5 million as compared to failure 1. Further, for one standard deviation difference in the phase II buffer, the difference in financial impact corresponds to 1.27%, or US$351.7 millions.

Support for Hypothesis 3 demonstrates that presence of additional projects that utilize the same development resources as the failed project mitigates the impact of a failure. Put differently, the value of a compound for a portfolio in which there are many other projects that utilize the same resources is smaller than for a portfolio in which it is the sole claimant to these resources.

Support for this hypothesis also empirically validates the anecdotal phenomena that a failure (an in-licensing opportunity) at a time when the product development pipeline is "congested" or has more compounds than expected hurts (helps) the firm less vis-à-vis a failure (an in-licensing opportunity) at a time when the development pipeline is lean or has fewer compounds than expected (cf. Landers and Lublin 2003 on the impact of failures and a lean pipeline on Merck Pharmaceuticals).

In developing the Δ*TAPIS* variable in §3.1, we ignored the effects of compounds for the same market that are present in the pipelines of competing firms. If the competition has late stage compounds in development, we would expect the mitigating effect of the

firm's early stage backup compounds to be smaller. To capture this effect, we construct an alternate version of $\Delta TAPIS$ that quantifies the change in the odds of being the first firm to launch a compound (as opposed to the odds of launching a compound, in the original metric). In our sample, this metric is highly correlated with the original metric suggesting that there is not significant variation in the competitive situation for the failed compounds. Not surprisingly, we find that this metric also has a statistically significant effect on the impact of a failure.[10]

In testing Hypothesis 3, we argued that the difference in the phase II success rate measures the excess of work or shortfall of phase III resources. Alternately, this variable may also measure a perception of the firm's capabilities: A higher than average phase II success rate may indicate high or improving capabilities, and vice versa. To test this alternate interpretation, we construct a variable measuring the difference in the recent phase III success rate versus the long-run phase III success rate, and run our regression with this variable instead of the phase II buffer variable used in the study. This variable is arguably a more direct measure of firm capabilities, but does not measure the work build-up for phase III. We find no statistically significant impact of this variable on the impact of the failure. This discredits the alternate interpretation of the variable used to test the hypothesis.

Taken together, support for Hypotheses 2 and 3 suggests that portfolio-level project interactions significantly alter the value of a project. Ignoring these interactions would lead to estimation errors in a project's value by an order of millions of dollars when the average values of a project is approximately US$500 million. This could lead to highly suboptimal portfolio and capacity choices. Thus, a product development manager interested in maximizing shareholder returns would benefit from using decision support systems that acknowledge and model these interactions based on our empirical observations.

Although we developed a detailed nonlinear model for Hypothesis 2, relating the number of backup compounds to the financial value of a compound, we did not do so for Hypothesis 3. The financial impact of the congestion effects central to Hypothesis 3 may not be linear in the proxy for utilization.[11] An appropriate queuing model that analytically captures these effects remains the subject of future work.

7. Conclusion

The results of our empirical investigation suggest that a late-stage failure of a project is associated with a significant decline in firm value; for an average failure in our data set, this corresponds to a decline in value of US$405 million. We find support for our hypothesis predicting that decline in firm value is mitigated by the presence of backup projects. Put differently, the value of a project within a portfolio that contains multiple projects targeting the same market is smaller than within a portfolio in which it is the sole claimant to the market. We also find support that this impact is mitigated if the firm has an above average phase II success rate prior to the failure leading to a more than expected number of compounds that will utilize the same development resources as the failed project. Put differently, the value of a project within a portfolio containing more projects that utilize the same resources as the failed project is smaller than a project within a portfolio where it is the sole claimant to the resources.

In addition to validating our intuition about portfolio-level projects interactions, support for our hypotheses validates our theoretical metrics for the financial impact of these interactions. Finally, the magnitude of our results suggests that these portfolio-level project interactions significantly alter the value of a project.

Our method offers a data-driven approach for valuation and comparison of in-licensing opportunities available to a pharmaceutical firm. Using the natural experiment of failures, we have built a predictive model of the impact of different compounds on the firm's valuation, taking into account the portfolio-level project interactions or the fit of the compound in the firm's portfolio. This model can be used to predict the increase in a firm's value if a particular compound is added to its portfolio. This should be the maximum fair price that the firm should pay for this compound.

Although the coefficients estimated in this paper are applicable only for late-stage failures in the pharmaceutical industry, the insights, framework, and empirical methodology can be employed more generally for product development portfolios in environments with high uncertainty, which gets resolved in consecutive phases of testing. The development of alternate approaches akin to backup compounds is common in many product development settings. The "winner-takes-all" payoff structure is typical for industries where alternate approaches are investigated to address one user need. The notion of a shared, fixed development capacity and the associated economics are also typical of many R&D environments. Product development environments such as consumer packaged goods with test markets, multiphase defense development contracts, etc. are all amenable to the methods and insights developed in this paper.

Of course, a generalization of our results must be approached with caution. Our methodology rests

[10] We thank an anonymous referee for suggesting this.

[11] We thank the anonymous referee for highlighting this.

on the assumption that markets accurately estimate the factors that influence profits from drug development. This is a reasonable assumption for the pharmaceutical industry, which has high investments by sophisticated institutional investors, extensive regulatory and scientific scrutiny, high levels of disclosure, and exogenously defined and publicly measured metrics of product performance. However, this assumption may not apply equally well to all industries. Although the actual impact predicted from the failure may be less accurate when the assumption does not hold, the insights into the relationship between firm value and properties of the product development portfolio should remain applicable as long as there is no systemic irrationality correlated with our product development variables.

This study empirically identifies a direction for developing decision support models for portfolio and capacity choice in risky development environments. Decision support models that realistically represent the portfolio-level project interactions we identify in this paper can be useful for product development managers and could help address the hitherto limited relevance of academic research for product development choices in industrial practice.

8. Electronic Companion

An electronic companion to this paper is available as part of the online version that can be found at http://mansci.journal.informs.org.

Acknowledgments

The authors are grateful to Christoph Loch, Taylor Randall, seminar participants at The Wharton School, Stanford Graduate School of Business, Harvard Business School, the Sloan School of Management, and INSEAD. They thank the three anonymous reviewers and the associate editor for many helpful comments, Sean Nicholson for helping them procure data for this study, Lifei Cheng and Brian Kelly from Merck Research for numerous thought-provoking discussions, and Jim Birchenough and Tom Miner from Lehman Brothers for their comments on this research.

References

Adler, P. S., A. Mandelbaum, V. Nguyen, E. Schwerer. 1995. From project to process management: An empirically based framework for analyzing product development time. *Management Sci.* **41** 458–484.

Bayus, B. L., G. Erickson, R. Jacobson. 2003. The financial rewards of new product introductions in the personal computer industry. *Management Sci.* **49** 197–210.

Belsley, D. E., E. Kuh, R. Welsch. 1980. *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*. John Wiley & Sons, New York.

Brown, S. J., J. B. Warner. 1980. Measuring stock price performance. *J. Financial Econom.* **8** 205–258.

Chaney, P. K., T. M. Devinney. 1992. The impact of new product introductions on market value of firms. *J. Bus. Finance Accounting* **19** 677–695.

Chaney, P. K., T. M. Devinney, R. S. Winer. 1991. The impact of new product introductions on the market value of firms. *J. Bus.* **64** 573–610.

Chen, S. S., K. W. Ho, K. H. Ik. 2005a. The wealth effect of new product introductions on industry rivals. *J. Bus.* **78** 969–996.

Chen, S. S., T. Y. Chung, K. W. Ho, C. F. Lee. 2005b. Intra-industry effects of delayed new product introductions. Working paper, National Taiwan University, College of Management, Taipei, Taiwan.

Corbett, C. J., M. J. Montes-Sancho, D. A. Kirsch. 2005. The financial impact of ISO 9000 certification in the United States: An empirical analysis. *Management Sci.* **51** 1046–1059.

Dahan, E., H. Mendelson. 2001. An extreme value model of concept testing. *Management Sci.* **47** 102–116.

Danzon, P. M., S. Nicholson, N. S. Pereira. 2005. Productivity in pharmaceutical-biotechnology R&D: The role of experience and alliances. *J. Health Econom.* **24** 317–339.

Ding, M., J. Eliashberg. 2002. Structuring the new product development pipeline. *Management Sci.* **48** 343–363.

Girotra, K., C. Terwiesch, K. T. Ulrich. 2004. New drug development at Merck & Co., Wharton School Teaching Case, Philadelphia, PA.

Hendricks, K. B., V. R. Singhal. 1997. Delays in new product introductions and the market value of the firm: The consequences of being late to the market. *Management Sci.* **43** 422–436.

Hendricks, K. B., V. R. Singhal. 2005. Association between supply chain glitches and operating performance. *Management Sci.* **51** 695–711.

Kavadias, S., C. H. Loch. 2003. *Project Selection Under Uncertainty: Dynamically Allocating Resources to Maximize Value*. Kluwer Academic Press, Boston, MA.

Kothari, S. P., J. B. Warner. 2007. Econometrics of event studies. B. Espen Eckbo, ed. *Handbook of Corporate Finance.* Elsevier/North-Holland, Amsterdam, The Netherlands.

Landers, P., J. S. Lublin. 2003. Under a microscope: Merck's big bet on research by its scientists comes up short. *Wall Street J.* (November 28) A1.

Loch, C. H., S. Kavadias. 2002. Dynamic portfolio selection of NPD programs using marginal return. *Management Sci.* **48** 1227–1241.

Loch, C. H., C. Terwiesch. 1999. Accelerating the process of engineering change orders: Capacity and congestion effects. *J. Product Innovation Management* **16** 145–159.

Loch, C. H., C. Terwiesch, S. Thomke. 2001b. Parallel and sequential testing of design alternatives. *Management Sci.* **47** 663–678.

Loch, C. H., M. T. Pich, C. Terwiesch, M. Urbschat. 2001a. Selecting R&D projects at BMW: A case study of adopting mathematical programming models. *IEEE Trans. Engrg. Management* **48** 70–80.

Lucas, R. 1971. Optimal management of a research and development project. *Management Sci.* **17** 679–697.

MacKinlay, C. A. 1997. Event studies in economics and finance. *J. Econom. Literature* **35** 13–39.

Nicholson, S., P. M. Danzon, J. Mccullough. 2005. Biotech-pharmaceutical alliances as a signal of asset and firm quality. *J. Bus.* **78** 1433–1464.

Parexel. 2002/2003. *PAREXEL's Bio/Pharmaceutical R&D Statistical Sourcebook*. Parexel, Waltham, PA.

Patell, J. M. 1976. Corporate forecasts of earning per share and stock price behavior: Empirical tests. *J. Accounting Res.* **14** 246–274.

Pisano, G. P., S. Rossi. 1994. Eli Lilly and Co.: The flexible facility decision–1993. HBS Case 9-694-074, Harvard Business School, Boston, MA.

Robertson, T. S., J. Eliashberg, T. Rymon. 1995. New product announcement signals and incumbent reactions. *J. Marketing* **59** 1–15.

Shane, S. A., K. T. Ulrich. 2004. Technological innovation, product development, and entrepreneurship in Management Science. *Management Sci.* **50** 133–144.

Sharma, A., N. Lacey. 2004. Linking product development outcomes to market valuation of the firm: The case of the US pharmaceutical industry. *J. Product Innovation Management* **21** 297–308.

Singhal, V. R. 2005. Excess inventory and long-term stock price performance. Working paper, Georgia Institute of Technology, Atlanta, GA.

Sprent, P. 1989. *Applied Nonparametric Statistical Methods*. Chapman and Hall, London, UK.

Tkac, P. A. 1999. A trading volume benchmark: Theory and evidence. *J. Financial Quant. Anal.* **34** 89–114.

MANAGEMENT SCIENCE
DOI 10.1287/mnsc.1070.0703ec
pp. ec1–ec8

e-companion

ONLY AVAILABLE IN ELECTRONIC FORM

Electronic Companion—"Valuing R&D Projects in a Portfolio: Evidence from the Pharmaceutical Industry" by Karan Girotra, Christian Terwiesch, and Karl T. Ulrich, *Management Science*, DOI 10.1287/mnsc.1070.0703.

Online Supplement

EC.1. FAQS for Hypothesis 2

Q. What do the discount factors mean?

A. These discount factors capture the impact of a delay in product introduction to the market. This could be due to the time value of money or changed market conditions- lower margins, increased competition, loss of market share, etc. Hendricks and Singhal (1997) rigorously study the impact of such delays.

Q. Is the proportionality constant the same for different drugs?

A. Different drugs may have different market sizes and thus the proportionality constant may not be the same across all failures in our sample. Accurate estimates on the market size of a drug are hard to come by, since they are strongly influenced by the drug label, which is only issued by the FDA *after* the successful completion of all clinical trials. Also, the competitive position of the drug is not easy to predict before the results of all trials are known. Therefore, we do not include the market size in our TAPIS metric, and our developed hypothesis will test the role of interaction between compounds targeting the same market, ceteris paribus. In other words, we test if *on an average*, the change in probability is related to the financial impact of the failure. This process reduces the explanatory power of our estimation and makes it less likely that we will find support for our hypothesis. If we had accurate market size data, we could include that in the TAPIS metric and proceed with the same empirical design and possibly have better results.

Q. Does this hypothesis tell me how many backup drugs should I develop?

Q. Are backup drugs always useful even if they cost a lot?

Q. Why don't you include the additional development costs because of backup drugs in your hypothesis?

A. Note that in formulating the financial impact of a failure, we did not include the development costs already incurred by the firm for development of multiple compounds. A firm that has more backup compounds has already invested in the development of these multiple compounds till the current stage of development and *has already incurred* higher development costs; these *sunk* development costs are not influenced by the failure and thus do not appear in our hypothesis. Our hypothesis captures the value of a compound or the impact of a failure, *given* the choices made by the firm with respect to the number of parallel development efforts. It *does not* advise on the optimal number of development efforts. Ding and Eliashberg (2002) provide guidance on the number of parallel development efforts to pursue.

There is an effect of failure on development costs *subsequent* to the failure, related to the failure-induced higher probability of phase I, phase II compounds now being brought into phase III trials, and thus higher development costs. We believe this is a minor second order effect and we don't capture this effect in our hypothesis.

Q. What if the success/failure of different drugs in a firm's portfolio is not independent?

A. In developing the hypotheses, we keep in mind that for empirically testing the developed theory, we will need to compute these probabilities using available data. With the data available to us, we

pp. ec1–ec8; suppl. to *Management Sci.*, DOI 10.1287/mnsc.1070.0703, © 2007 INFORMS

can not estimate the correlation structure between different drugs and thus, we assume that the trials are independent. If detailed data would be available such that the correlation structure between the different compounds can be estimated (for instance on the basis of the mechanism of action of the drug); the expressions in our analysis can be suitably modified to capture the correlations. The rest of the empirical analysis would proceed in the same way.

Even with the availability of data, developing a scientific approach to estimating the correlation structure between different compounds is a challenging problem that is not well understood by product development researchers or the pharmaceutical industry. Addressing this is the focus of our ongoing research.

EC.2. The Event Study Methodology

EC.2.1. Return Generating Models

The three return-generating models for expected returns that we use are–

The Comparison Period Mean Model (*CP*). The comparison period mean model uses the average return on the firm's stock over the estimation period as the expected return on the stock. The abnormal return is computed as the difference between the return on the days of the event and the average return on the firm's stock over the estimation period.

The Market Model (*MM*) (MacKinlay 1997)*:* The Market model posits a linear relationship between the return on a firm's stock and the return on the market portfolio. According to the market model, for any firm i, the expected return at time t, R_i^t, is given by Equation (EC1); where R_M^t is the period-t return on the market index, Ω^{t-1} is the information set available to the market about the firm at time $t-1$.

$$\mathrm{E}[R_i^t \mid \Omega^{t-1}] = \alpha_i + \beta_i R_M^t \tag{EC1}$$

We use the CRSP equally weighted index[EC1] as a proxy for the market portfolio. α_i and β_i are firm-specific parameters that are estimated using estimation-period historical data for *each* event in the sample.

The Fama-French Three Factor Model (*FF*) (Fama and French 1996): The model posits the return generating process described in Equation (EC2). β_i is the sensitivity of the stock returns to market returns; s_i is the sensitivity of the stock returns to the difference between small and large capitalization stock returns and h_i is the sensitivity of the stock returns to the difference between the returns on value and growth stocks.[EC2]

$$\mathrm{E}[R_i^t \mid \Omega^{t-1}] = \alpha_i + \beta_i R_M^t + s_i SMB^t + h_i HML^t \tag{EC2}$$

We use the broad based CRSP equally weighted index as a proxy for the market portfolio. α_i, β_i, s_i, and h_i are firm-specific parameters that are estimated using historical data for each event in the sample.

The component of the return that *can not* be explained by the above described market models, is attributed to the event, in our case the failure of the phase III clinical trials. This component is commonly referred to as the abnormal return, AR_i^t described in Equation (EC3).

$$AR_i^t = R_i^t - \mathrm{E}[R_i^t \mid \Omega^{t-1}] \tag{EC3}$$

If no new information is available between the period $t-1$ and t, we would expect that $\mathrm{E}[AR^t] = 0$.

[EC1] The CRSP equally weighted index is an index provided by the Center for Research in Security Prices (CRSP), a financial research center at the Graduate School of Business at The University of Chicago. The index gives all securities an equal weight.

[EC2] SMB^t is defined as the average return on three small market capitalization portfolios minus the average return on three large market capitalization portfolios; HML^t is defined as the average return on two high book-to-market equity portfolios minus the average return on two low book-to-market equity portfolios. (See Fama and French 1993 for detailed explanations of SMB^t and HML^t.) The factors- SMB^t and HML^t are obtained from Prof. Kenneth R. French's website (http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/index.html).

EC.2.2. Finding the Event Window Using Trading Volumes

Often the impact of an event is not limited only to the day of the announcement. Information about the event may leak before the announcement (a speculative article in the press, rumors at the financial market, etc.) or markets may take a few sessions to fully price the impact of the event. Moreover, the exact time of the day when the drug failure announcement is made may not be known. The announcement may have been made after trading hours on the day of the announcement and in such a scenario, the stock price is not affected until the next trading session. Most event studies include the abnormal return for a few days before and after the event is used when constructing a measure to incorporate the full impact of the event. This measure is often referred to as the cumulative abnormal return or CAR_i (Equation (EC4)), where the period (t_1, t_2) is called the event window.

$$CAR_i = \sum_{t=T+t_1}^{t=T+t_2} AR_i^t \tag{EC4}$$

To identify the appropriate event window, Tkac (1999) suggests using the number of shares traded or the trading volume. As the market incorporates the new information, market participants adjust their positions leading to higher trading volume than under normal circumstances. Tkac (1999) also provides a theoretical model for the normal or the expected trading volume (Equation (EC5)).

$$\mathrm{E}[V_i^t] = \gamma_i + \delta_i MV^t \tag{EC5}$$

In Equation (EC5) the expected trading volume for firm i at time t, $\mathrm{E}[V_i^t]$, is explained using the trading volume in the market MV^t. The factors γ_i and δ_i are estimated using historical data for *each* event in the sample. The abnormal volume, AV_i^t, is then given as the excess of the actual volume over the expected volume (EC6).

$$AV_i^t = V_i^t - \mathrm{E}[V_i^t] \tag{EC6}$$

We compute the abnormal trading volume averaged across the events and define the event window to include days before and after the event that have a statistically significant positive average abnormal trading volume.[EC3] The event window identified allows us to compute the Cumulative Abnormal Return and the loss in market capitalization[EC4] associated with each drug failure in our sample.

EC.3. Probability Estimation

To estimate the probabilities (p_1, p_2, p_3), we use data on all clinical trials in the ADIS database. A vast majority of these trials are run by firms that are not publicly traded and don't otherwise appear in our sample. Danzon et al. (2005) finds that the indication explains the largest fraction of the variance in the odds of success and failure between different drugs. Thus, we estimate p_i at the level of an indication and assume that all drugs for an indication at the same stage of development have the same odds of failure. For example, to determine p_i for an Alzheimer's drug, we look at the performance of all Alzheimer's drugs irrespective of originating firm. p_i is the probability of failure of a drug currently undergoing phase i trials, *during phase i trials or in any subsequent stage*(*s*) for the disease in question. If the probability of success in phase i is given as $s_i \in [0, 1]$; $p_i = 1 - \prod_{k=i}^{3} s_i$. To compute p_1, p_2, and p_3, we use estimates on s_1, s_2, and s_3. Each s_i is estimated by looking at the average success rate in the indication of concern.

$$s_i = \frac{\#\ of\ Successful\ Phase\ i\ Trials}{Total\ \#\ of\ Completed\ Phase\ i\ Trials} \tag{EC7}$$

In Table EC.1, we provide the values of (p_1, p_2, p_3), computed from the estimated (s_1, s_2, s_3). The numbers in parens denote the number of compounds or data-points used to estimate (s_1, s_2, s_3). For a small fraction of indication-phase pairs (36 out of 282), the number of total completed trials is less than 4. For these indication-phase pairs the estimates on s_i may not be reliable, in such cases, we substitute the indication-specific success rate with an average success rate across all indications. Such indication-phase pairs are denoted by a "*" in Table EC.1. The number in parens for such cases is the total pooled number of phase i compounds used for that estimation.

[EC3] This is analogous to the market model for stock returns, and is often referred to as a volume event study.

[EC4] Change in Market Capitalization = Shares Outstanding * Price * Cumulative Abnormal Return.

Table EC.1

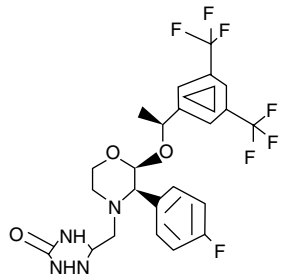
Sno.	Indication/disease	p_1	p_2	p_3
1	*Acute myeloid leukaemia*	0.826 (20)	0.677 (13)	0.4 (5)
2	*Alzheimer's disease*	0.926 (79)	0.76 (47)	0.636 (22)
3	*Amyotrophic lateral sclerosis*	0.973 (11)	0.833 (6)	0.8 (5)
4	*Anxiety disorders*	0.992 (55)	0.935 (28)	0.833 (12)
5	*Arrhythmias*	0.987 (28)	0.943 (30)	0.714 (7)
6	*Arterial thrombosis*	0.785 (4)	0.732 (6)	0.197* (3,147)
7	*Asthma*	0.809 (152)	0.709 (109)	0.226 (53)
8	*Atherosclerosis*	0.831 (28)	0.732 (9)	0.197* (3,147)
9	*Atopic dermatitis*	0.718 (25)	0.657 (10)	0.143 (14)
10	*Attention-deficit hyperactivity disorder*	0.617 (14)	0.464 (12)	0.286 (7)
11	*Bacterial infections*	0.616 (60)	0.385 (42)	0.109 (46)
12	*Basal cell cancer*	0.967 (6)	0.8 (6)	0.8 (5)
13	*Benign prostatic hyperplasia*	0.58 (27)	0.542 (16)	0.083 (12)
14	*Bladder cancer*	0.563 (18)	0.444 (12)	0.167 (6)
15	*Breast cancer*	0.599 (120)	0.52 (59)	0.114 (35)
16	*Cancer*	0.862 (181)	0.777 (74)	0.176 (34)
17	*Cardiovascular disorders*	0.738 (9)	0.55 (5)	0.25 (4)
18	*Catabolism*	0.552* (5,900)	0.4* (3,740)	0.197* (3,147)
19	*Cerebrovascular disorders*	0.678 (6)	0.599 (6)	0.197* (3,147)
20	*Chemoprotection*	0.79 (19)	0.625 (10)	0.375 (8)
21	*Chronic obstructive pulmonary disease*	0.752 (41)	0.589 (23)	0.273 (11)
22	*Colorectal cancer*	0.705 (104)	0.621 (41)	0.182 (11)
23	*Congestive heart failure*	0.73 (34)	0.535 (26)	0.364 (11)
24	*Coronary artery restenosis*	0.858 (16)	0.74 (12)	0.375 (8)
25	*Depression*	0.911 (78)	0.747 (60)	0.525 (40)
26	*Diabetes mellitus*	0.826 (26)	0.657 (14)	0.4 (5)
27	*Diabetic complications*	0.785 (10)	0.732 (9)	0.197* (3,147)
28	*Diabetic neuropathies*	0.948 (27)	0.825 (19)	0.667 (6)
29	*Diabetic retinopathy*	0.356 (4)	0.197 (5)	0.197* (3,147)
30	*Diagnostic imaging*	0.516 (32)	0.269 (23)	0.269 (26)
31	*Diarrhoea*	0.742 (9)	0.502* (3,740)	0.333 (6)
32	*Emesis*	0.59 (16)	0.427 (20)	0.182 (11)
33	*Epilepsy*	0.836 (30)	0.619 (17)	0.353 (17)
34	*Erectile dysfunction*	0.52 (22)	0.4 (16)	0.2 (15)
35	*Fracture treatment*	0.518 (4)	0.4* (3,740)	0.197* (3,147)
36	*Gastric ulcer*	0.738* (5,900)	0.625 (4)	0.25 (12)
37	*Glioblastoma*	0.678 (15)	0.599 (6)	0.197* (3,147)
38	*Glioma*	0.926 (27)	0.8 (16)	0.6 (5)
39	*Graft-versus-host disease*	0.896 (14)	0.714 (14)	0.636 (11)
40	*Gram-positive infections*	0.555* (5,900)	0.402* (3,740)	0.2 (5)
41	*HIV-1 infections*	0.833 (28)	0.688 (16)	0.375 (8)
42	*Head and neck cancer*	0.842 (47)	0.697 (22)	0.444 (9)
43	*Head injuries*	0.993 (9)	0.9 (10)	0.875 (8)
44	*Heart failure*	0.814 (33)	0.665 (39)	0.346 (26)
45	*Hepatitis B*	0.291 (27)	0.237 (24)	0.036 (28)
46	*Herpes simplex virus infections*	0.841 (12)	0.688 (16)	0.444 (9)
47	*Herpes zoster*	0.586* (5,900)	0.333 (5)	0.333 (6)
48	*Huntington's disease*	0.4* (5,900)	0.197 (4)	0.197* (3,147)
49	*Hypercholesterolaemia*	0.552 (22)	0.497 (13)	0.067 (15)
50	*Hyperlipidaemia*	0.91 (37)	0.739 (17)	0.444 (9)
51	*Hypertension*	0.783 (88)	0.638 (76)	0.19 (63)
52	*Irritable bowel syndrome*	0.904 (28)	0.748 (17)	0.571 (7)
53	*Ischaemic heart disorders*	0.988 (28)	0.924 (23)	0.75 (12)
54	*Leukaemia*	0.62 (19)	0.492 (7)	0.111 (9)
55	*Malaria*	0.455 (11)	0.333 (6)	0 (5)
56	*Malignant hypertension*	0.552* (5,900)	0.4* (3,740)	0.197* (3,147)
57	*Migraine*	0.737 (42)	0.558 (21)	0.286 (14)
58	*Multiple myeloma*	0.651 (31)	0.484 (16)	0.25 (8)
59	*Multiple sclerosis*	0.877 (37)	0.706 (17)	0.5 (10)
60	*Muscle wasting*	0.552* (5,900)	0.4* (3,740)	0.197* (3,147)
61	*Musculoskeletal pain*	0.354* (5,900)	0.167 (4)	0.167 (6)

Table EC.1 (Continued)

Sno.	Indication/disease	p_1	p_2	p_3
62	*Myocardial infarction*	0.766 (37)	0.622 (32)	0.364 (22)
63	*Neuropathic pain*	0.731 (23)	0.625 (10)	0.25 (4)
64	*Neutropenia*	0.707 (5)	0.531 (8)	0.375 (8)
65	*Non-small cell lung cancer*	0.877 (108)	0.715 (46)	0.563 (16)
66	*Osteoarthritis*	0.506 (21)	0.404 (26)	0.032 (31)
67	*Ovarian cancer*	0.816 (77)	0.707 (42)	0.353 (17)
68	*Pain*	0.647 (74)	0.523 (46)	0.156 (32)
69	*Pancreatic cancer*	0.905 (58)	0.772 (31)	0.583 (12)
70	*Parkinson's disease*	0.721 (47)	0.543 (23)	0.3 (20)
71	*Peptic ulcer*	0.929 (21)	0.794 (27)	0.444 (18)
72	*Peripheral nerve disorders*	0.678 (4)	0.599 (4)	0.197* (3,147)
73	*Picornavirus infections*	0.552* (5,900)	0.4* (3,740)	0.197* (3,147)
74	*Postmenopausal osteoporosis*	0.51 (47)	0.313 (39)	0.162 (37)
75	*Postoperative pain*	0.557 (18)	0.416 (14)	0.091 (11)
76	*Psychotic disorders*	0.868 (19)	0.667 (12)	0.5 (8)
77	*Renal cancer*	0.925 (72)	0.814 (23)	0.571 (7)
78	*Rheumatic disorders*	0.76 (12)	0.621 (16)	0.241 (29)
79	*Rotavirus infections*	0.5* (5,900)	0.331 (6)	0.197* (3,147)
80	*Schizophrenia*	0.711 (36)	0.628 (25)	0.154 (13)
81	*Septic shock*	0.992 (19)	0.913 (13)	0.875 (8)
82	*Small cell lung cancer*	0.753 (25)	0.63 (9)	0.333 (9)
83	*Somatotropin deficiency*	0.325 (6)	0.1 (6)	0.1 (10)
84	*Spinal cord injuries*	0.678 (6)	0.599 (4)	0.197* (3,147)
85	*Stem cell mobilisation*	0.518 (4)	0.4* (3,740)	0.197* (3,147)
86	*Stroke*	0.939 (54)	0.772 (38)	0.667 (18)
87	*Systemic lupus erythematosus*	0.839 (6)	0.799 (4)	0.197* (3,147)
88	*Thrombocytopenia*	0.964 (12)	0.825 (10)	0.75 (4)
89	*Thrombosis*	0.862 (65)	0.697 (34)	0.357 (14)
90	*Transplant rejection*	0.784 (20)	0.619 (7)	0.333 (9)
91	*Tuberculosis*	0.669 (7)	0.518 (5)	0.197* (3147)
92	*Type-2 diabetes mellitus*	0.612 (76)	0.515 (48)	0.105 (19)
93	*Ulcerative colitis*	0.905 (30)	0.762 (18)	0.571 (7)
94	*Unstable angina pectoris*	0.818 (15)	0.658 (18)	0.385 (13)

EC.4. Sample Data Entry from the ADIS Data Set

Aprepitant (EMEND, L 754030, MK 0869, MK 869)



Chemical Name: 5-(2(*R*)-(1(*R*)-(3,5-Bis(trifluoromethyl)phenyl)ethoxy)-3(*S*)-(4-fluorophenyl) morpholin-4-ylmethyl)-3,4-dihydro-2H-1,2,4-triazol-3-one
Molecular Formula: C23 H21 N4 F7 O3

CAS Number: 170729-80-3; 170902-81-5 ((1S)-isomer); 172822-28-5 ([2S(1S),3S]-isomer); 172822-29-6 ([2S(1R),3R]-isomer)

WHO ATC Code: A04A-D (Other antiemetics);
N06A-X (Other antidepressants)

EphMRA ATC Code: A4A9 (Other Antiemetics and Antinauseants);
N6A (Anti-Depressants and Mood Stabilisers)

Originator Companies: Merck & Co
Last Update: 2003-12-15
Accession Number: 9602

Drug Development History
15-12-2003—Clinical data from a media release have been added to the pharmacokinetics section (9029028)
11-12-2003—Sales forecasts reviewed by Lehman Brothers
14-11-2003—Discontinued—Phase-III for Depression in USA (PO) → Failure in Phase III Clinical
07-08-2003—Studies have been added to the adverse events and Advances in the Treatment of Nausea and Migraine therapeutic trials sections (897818,897819)
23-07-2003—Sales forecasts reviewed by Lehman Brothers
05-06-2003—Data presented at the 39th Annual Meeting of the American Society of Clinical Oncology (ASCO-2003) have been added to the Nausea and migraine therapeutic trials section (9021600)
22-05-2003—Launched for Emesis in USA (PO)
08-05-2003—Data has been added to the drug interactions section (941574)
16-04-2003—Data from a media release have been added to the adverse events and Nausea and migraine therapeutic trials sections (9019583)
10-04-2003—Sales forecasts reviewed by Lehman Brothers
01-04-2003—Registered for Emesis in USA (PO)
31-03-2003—Merck & Co. Inc. has received an approvable letter from the FDA for aprepitant in acute and delayed nausea and vomiting associated with highly emetogenic chemotherapy (www.merck.com)
13-12-2002—Preregistration for Emesis in USA (PO)
12-12-2002—Sales forecasts reviewed by Lehman Brothers
12-09-2002—Sales forecasts reviewed by Lehman Brothers
31-05-2002—Phase-III clinical trials in Depression in USA (PO)
30-05-2002—Sales forecasts reviewed by Lehman Brothers
19-03-2002—Phase-III clinical trials in Emesis in USA (PO)
04-12-2001—Sales forecasts reviewed by Lehman Brothers
24-07-2001—Discontinued-II for Schizophrenia in USA (PO)
14-05-2001—Predicted peak sales have been added
25-08-2000—A study has been added to the Advances in the Treatment of Nausea and Migraine therapeutic trials section (835407)
09-06-1999—Data have been added to the adverse events and the Affective disorders therapeutic trials sections (747155)
17-05-1999—Phase-II clinical trials for Schizophrenia in USA (PO)
17-05-1999—Phase-II clinical trials for Anxiety disorders in USA (PO)
17-02-1999—A study has been added to the Advances in the treatment of nausea and migraine therapeutic trials section (736323)
29-01-1999—Phase-II clinical trials for Emesis in USA (PO)
29-01-1999—Suspended-II for Depression in USA (PO)
22-09-1998—A study has been added to the adverse events and the Affective disorders therapeutic trials events sections (632240)
09-01-1998—Phase-II clinical trials for Depression in USA (PO)
09-01-1998—New profile

EC.5. Associated News Report from Factiva

Merck Ends Research on Once-Promising Antidepressant

By Peter Landers
1123 words
13 November 2003
The Wall Street Journal
B1
English

MERCK & CO. abandoned a depression drug it had been working on for more than a decade.

The company said the drug, known generically as aprepitant, failed to show advantages over a placebo in a large human trial. The decision leaves Merck, once renowned for its research prowess, with one of the thinnest pipelines among major drug makers. Without new products, Merck's future is in danger because its anticholesterol blockbuster Zocor will lose U.S. patent protection in 2006.

The failure of aprepitant is also a major setback for depression patients who don't respond to the class of drugs called selective serotonin reuptake inhibitors, such as Eli Lilly & Co.'s Prozac, Pfizer Inc.'s Zoloft and GlaxoSmithKline PLC's Paxil, or who can't handle their side effects. Doctors say some 40% of people who take SSRIs experience sexual problems, fueling the search for an alternative. "It's really too bad. I definitely was hoping for some new options," says Carol Peyser, a psychiatrist in Redwood City, Calif., who frequently prescribes antidepressants.

The failed drug trial may raise new questions about Merck Chief Executive Raymond Gilmartin, who angered investors last month when he announced that the company wouldn't meet its profit forecast for this year. Mr. Gilmartin is under growing pressure to prove he can turn around Merck, which expects net income to fall this year to $2.90 to $2.95 per share, down from $3.14 per share in 2002.

Antidepressants are a $17 billion market, and Merck was hoping to grab a big chunk of that. Aprepitant has its origins in research that Merck undertook in the early 1990s on a mysterious brain chemical called substance P. (The "P" stands for "powder," the form in which substance P was first isolated in 1931.) In the 1980s, scientists believed that substance P was involved in transmitting pain messages among the brain's neurons, and drugs blocking substance P were tried out as painkillers.

Those studies bore little fruit, but scientists at Merck came to believe that substance P actually worked as a mood depressant—one that had a useful function in normal people but might cause disease if it got out of control. Merck chemists, meanwhile, cooked up aprepitant, a pill that blocks substance P from landing at receptors in the brain's neurons.

The project's shining moment came in 1998, when Merck published a study in the journal Science showing that aprepitant was significantly superior to a placebo in treating depression in a trial of 213 patients. Equally important, the study showed fewer cases of sexual dysfunction in those taking aprepitant compared with Paxil.

But as Merck continued testing aprepitant, the project took a negative turn. The results of a trial that had been intended to ascertain the proper dose of aprepitant shocked the company: None of the proposed doses was superior to a placebo. The study was never published, although the company discussed it at scientific meetings in 2001.

Many companies would have stopped their research at that point. But Mr. Gilmartin said in an interview this past June that Edward Scolnick, Merck's research chief at the time of the dosage study, urged him to keep funding the depression program. One idea was to try out a separate substance P blocker called "compound A" in severely depressed patients. That trial showed good results and convinced Merck that substance P was indeed a factor in depression.

Separately, Merck scientists developed an elaborate brain-scanning system using radioactive tracers to test whether aprepitant was sufficiently hitting the substance P receptors in the brain. The scientists feared the failed trial might have been caused by the drug's difficulty in reaching the brain, but the scans showed that the drug was indeed hitting its target.

Perhaps the most important reason for Merck to keep betting on its depression program was its shortage of other good drugs in development. Currently, its pipeline has only three other drugs in large-scale testing: a tweak to its existing painkiller, Vioxx; a vaccine against cervical cancer; and a diabetes drug bought from a Japanese company. As recently as Oct. 22, Mr. Gilmartin named aprepitant as one of the drugs in Merck's pipeline that suggested the company has a bright future.

Monday-morning quarterbacks may question Merck's decision to fund the big trial in the face of the earlier failed trial and the lack of solid scientific evidence that substance P plays a role in depression. Mr. Gilmartin, a former executive for a medical-devices company, had no experience in pharmaceuticals prior to taking the Merck job in 1994 and often deferred to Dr. Scolnick, a Harvard Medical School graduate who had supervised blockbusters such as Vioxx.

"In a pharmaceutical company it is often difficult to stop advanced projects, especially if you have a strong R&D director," says Max Noble, managing director of Britannia Pharmaceuticals Ltd., a small company based near London.

pp. ec1–ec8; suppl. to *Management Sci.*, DOI 10.1287/mnsc.1070.0703, © 2007 INFORMS

Peter Kim, who replaced Dr. Scolnick as head of research at the beginning of this year, defended the decision to pursue aprepitant. "There are significant challenges in scientific research, and unfortunately, sometimes disappointments," Dr. Kim said in a statement. He said Merck would continue studying other psychiatric drugs.

The failure of aprepitant illustrates how difficult depression trials can be. Doctors have trouble distinguishing patients who are clinically depressed from those who are simply going through a rough patch. In addition, even depressed patients may get better temporarily for no particular reason. As a result, the placebo arm of depression trials can show a response rate of 50% or more.

Pfizer is still working on substance P blockers in depression. Some companies are pursuing ways to block corticotrophin releasing factor, or CRF, which may be overactive in depressed patients. AstraZeneca PLC is trying an approach to improving serotonin flow. "I am hopeful that over the next decade we will have therapies," says Gil Block, who heads the company's depression research.

One small consolation prize for Merck: In its exhaustive examination of aprepitant's properties, it discovered that the drug has an antinausea effect. This year, the company received approval to sell aprepitant under the brand name Emend as a treatment for nausea induced by chemotherapy. U.S. sales of Emend were $4 million in the third quarter, or 0.1% of Merck's total U.S. sales.

References

See references list in the main paper.

Fama, E. F., K. R. French. 1993. Common risk factors in the returns on stocks and bonds. *J. Financial Econom.* **33** 3–56.

Fama, E. F., K. R. French. 1996. Multifactor explanations of asset pricing anomalies. *J. Finance* **51** 55–84.