VALUING R&D PROJECTS IN A PORTFOLIO: EVIDENCE FROM THE PHARMACEUTICAL INDUSTRY

Karan Girotra, Christian Terwiesch and Karl T. Ulrich

The Wharton School, 500 Huntsman Hall, University of Pennsylvania, Philadelphia, PA 19104 USA
{karang, terwiesch, ulrich}@wharton.upenn.edu

January, 2005; Revised December, 2005, July, 2006

Abstract
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 the project's properties but also on the *other* projects being developed by the firm. This is due to interactions with other projects in the portfolio that address the same consumer need and with other projects that require the same development resources. In this study, we empirically investigate the structure and significance of these *portfolio-level project interactions*. Using a pharmaceutical industry dataset that we develop, we conduct an event study around the failure of phase III clinical trials, which 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 the focus of licensing transactions.

Key words
product development, pharmaceuticals, development pipeline, portfolio properties, backup projects, portfolio management

Acknowledgements
We are grateful to Taylor Randall, seminar participants at the Wharton School, Stanford Graduate School of Business, Harvard Business School, the Sloan School of Management, INSEAD; the three anonymous reviewers and the associate editor, for many helpful comments. We would like to thank Sean Nicholson for helping us 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.

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 the other projects that the firm is developing. This is due to interactions with other projects in the firm's 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, 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 which address the same customer need and on other projects which 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. (In Figure 1 and all subsequent illustrations, each small symbol indicates a unique development project. Projects that target the same market are illustrated in the same shape.) Drug development starts with an investigation of the chemical and biological properties of a compound in the lab (Basic Research), followed by animal trials (Pre-clinical Studies) and, finally three stages of clinical trials or trials in human subjects– Phase I, II and III clinical trials. Our study is designed around the failure of development projects currently undergoing phase III clinical trials, the final stage in the

development process, where the efficacy 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. While the product development process closely resembles the classic *phase-gate* development process prevalent in most industries, the role of regulators in the later phases of this 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 for each firm is public knowledge. Thus, the stock markets have information on portfolio-level project interactions that we are investigating, when they value failures.

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 any one project does not preclude the firm from earning sales in that market. Thus, the marginal value of any one of the multiple projects is smaller than the value of a lone 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 than the anticipated number of projects in its portfolio that require the same resources as those used by the failed project. 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 waiting for these resources, which may then be brought to the market sooner than they would have 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), and 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 alternate development projects.

2 Prior Literature

Two streams of academic work are relevant to this study: research on portfolio choices and 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. Dahan and Mendelson (2001) in their study on the number of development approaches to pursue for a given market, 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. In this investigation, they build an analytical model of the interactions between projects targeting the same market. Unlike Dahan and Mendelson (2001), in this model all successful projects are assumed to have identical quality. In contrast with this decision theoretic literature on portfolio selection, we take an empirical approach to investigate the portfolio-level project interactions. We empirically investigate interactions similar to the ones modeled by Ding and Eliashberg (2002).

Adler et al. (1995) in their analysis of project development time, build a model of project interactions due to shared development resources. Using a development project as their unit of analysis they find– if development resources are stretched, the project completion times are longer. In contrast with Adler et al. (1995), we 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 take an empirical approach and find the impact of these interactions on the financial value of the project as opposed to the development lead-time.

There have been multiple studies that focus on the impact of product development events on financial value, notably Hendricks and Singhal (1997) on the impact of product development delays. Hendricks and Singhal (1997) find significant negative stock returns associated with the announcement of product introduction delays. They find that industry competitiveness and the firm's degree of diversification are predictors 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. (2005b) study the impact of new product announcements on competing firms. Chen et al. (2005a) 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 events using the properties of the product (which influence the sales potential) or industry, but *not the portfolio*.

We build on the rigorous methodologies developed in this literature to empirically value projects. However, in contrast to this body of 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 and so, should lead to a decrease in the value of the firm. In the pharmaceutical industry, a typical 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 realized with 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. Hendricks and Singhal (1997) also hypothesize that a firm's market value falls when there is an announcement of a product development delay.

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 6 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 customer need. For instance, in the pharmaceutical industry, there are multiple chemical compounds that pharmacologically inhibit the COX enzymes and provide relief from the symptoms of inflammation and pain. These compounds often have different side-effects and thus, their success or failure in late-stage clinical trials is largely unrelated. In such settings, firms follow a parallel development strategy that increases their 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)).

While firms often develop multiple compounds from the same class to address a given market opportunity, firms rarely introduce more than one drug for the market. Even if the firm has more than one successful drug for the market, it would earn the same sales from the market as it would if it had only one successful compound (Ding and Eliashberg (2002), Girotra et al. (2004)). Firms typically introduce any of the candidate drugs that pass clinical trials, and then cease development of all other drugs targeted towards 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 only developing one compound for the disease. Prior to the failure, the firm has an 80% probability of having at least one successful drug in the indication of question from the drugs in phase III development. 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. Alternately, consider failure 2, in this case the firm is developing two compounds for the disease market. 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*.[1,2] From this pipeline, the firm could earn the sales from the market at any one of three different points in time corresponding to three mutually exclusive realizations of the clinical trials– (1) One of the compounds currently undergoing phase III trials succeeds. (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 all subsequent phases (Phase III). (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.

Assuming that clinical trials of different concepts are independent, the probability of scenarios 1, 2 and 3 are given as $\left(1-p_3^{\,n_3}\right)$, $p_3^{n_3}\left(1-p_2^{\,n_2}\right)$ and $p_3^{n_3}p_2^{n_2}\left(1-p_1^{\,n_1}\right)$ respectively. Under each of the three scenarios, the firm earns the sales from the market, however under scenario 2 it takes the firm longer to bring the

[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.

[2] 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. (2004) provides empirical support for this assumption.

product to the market, and even longer under scenario 3. Thus, the sales under scenario 2 and 3 should be discounted using appropriate discount factors– α_2 and α_1 respectively. The expected net present value of sales from this indication are thus given as

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

where $E[M]$ is the expected value of sales from the target market conditional on having a successful drug. Equivalently,

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

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),

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

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 $\left(n_3 \to n_3 - 1\right)$ and that leads to a change in the expression in (1) given by (2).

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

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. This leads us to our next hypothesis—

> *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 (ΔTAPIS in equation(2)).*

3.2 Effect of other compounds in the portfolio that utilize the same development resources

Clinical trials for compounds can be viewed as a three-stage process, with each phase of clinical trials as a stage. Compounds that are successful in phase II trials constitute the demand for phase III resources. Phase III trials involve the examination of the safety and efficacy of the drug in a large sample of patients and the primary resources required at this stage are clinical trial sites and bio-statisticians. Irrespective of the disease or indication associated with the compound, all compounds draw substantially from the same pool of resources. It is quite costly to scale the capacity of these phase III 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 resources, and make long-term commitments to R&D capacity to meet the same (Girotra et al. (2004)).

For instance, consider a firm that from its past experience believes that the average phase II probability of success or 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 the phase II capacity– 2 compounds/yr.

While the firm establishes capacity of phase III resources on the basis of the expected phase II probability of success, 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 phase III capacity of 2 compounds/yr., consider two possible realizations of phase II trials (Figure 3)– (1) where the realized success rate is 50% (2 of the 4 compounds in phase II trials succeed); and (2) where the realized success rate is 75% (3 of the 4 compounds in phase II trials succeed). In case 2, the firm has more compounds than the available phase III resources, thus a compound may have 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; say 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. In case 2, there is a compound 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 earlier than before and thus, can 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 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 may 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.

While the above example uses the notion of a buffer to illustrate this effect, buffering of demand is not necessary for observance of this phenomenon. As long as higher utilization of resources leads to longer processing times, the freeing up of resources due to failures will have a beneficial impact with respect to accelerating compounds in the pipeline. Further, the benefit associated with this acceleration is higher at higher levels of utilization, since the extent of the acceleration and the number of projects that benefit from this acceleration are both higher at higher levels of utilization. Formally, our hypothesis states—

> *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 successful 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. The database 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 used by the editors are: direct contact with companies involved with research and development, 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 to 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 is provided in Section 4 of the online supplement.

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

To verify data on the firm's portfolio (other compounds in development (n_1, n_2, n_3), and the recent success rates), we imputed the pipeline for one firm in our dataset (Merck & Co), and compared it with the private data for this firm. The pipeline data imputed from the ADIS database and in the firm's records 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 receive 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 data-base entry is provided in Section 5 of the online supplement.

We restrict our attention to all phase III failures of compounds catalogued in our data source that originated at publicly traded firms with common stock listed on any US market. We get stock price data from the CRSP financial database.[3] We identify the ownership and holding pattern of the originator firm(s) at the time of the failure using the Corporate Affiliations dataset 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 dataset range from over US$20 billion for the big pharmaceutical firms like Merck and Co., Pfizer Inc. and Bristol Myers Squibb to just under US$1 billion for biotechnology firms such as Chiron Co. The median firm in our dataset has annual sales of US$12.9 billion, employs 44,240 employees, spends US$1.6 billion annually on R&D (14.44% of sales), and experiences 7 phase III failures during the period of our study.

During the time-period of the study, 1994-2004, there were 169 phase III failures for publicly traded pharmaceutical firms in our database. We run our event study around these events. The year 1999 contributes the largest number of events in our study (29 phase III failures). The median year in our study has 15.5 phase III failures. Less than 2% of the events are found to be within a month of other related events, thus assuaging any concerns about clustering of events.

5 Methodology & 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 (2004)). 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

[3] 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 updated quarterly, by the Center for Research in Security Prices (CRSP), a financial research center at the Graduate School of Business at The University of Chicago.

new debt or equity, announcements of macro-economic variables (trade deficits, unemployment data, interest rates, etc.). 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)), the consequences of delays in new product introductions (Hendricks and Singhal (1997)), the effect of supply chain disruptions (Hendricks and Singhal (2003)), the impact of ISO 9000 certification (Corbett et al. (2004)) and the financial consequences of excess inventory (Singhal (2005)).

Using the prices for 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 stock price of the firm.

Event study methodologies provide a rigorous foundation to isolate the component of change in stock price due to the event and 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 Section 2.1 of the online supplement. 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 *can not* 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 event, but to a few days before and after the event, referred to as the event window. We use multiple event windows, including those advised by looking at trading volumes using the techniques proposed by Tkac (1999) (detailed in Section 2.2 of the online supplement). We then aggregate the abnormal return for each day in these 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 which influence firm value. This is an empirical measure of the value of each compound to the associated firm. We test our three hypotheses on this 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)),[4] a test that controls for cross-sectional dependence between individual security returns (Crude Dependence Adjustment Test from Brown and Warner (1980), pg. 233),[5] a non-parametric generalized sign-z statistic (Sprent (1989))[6] and the non-parametric Wilcoxon Signed Rank test. We test this hypothesis for several typical event windows *as well as* 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–ΔTAPIS 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 sub-sections.

5.2 Explanatory Variable: ΔTAPIS

To test hypothesis 2, we need to construct our explanatory variable, $\Delta TAPIS$ (given by the expression in (2)). $\Delta TAPIS$ is a function of the number of compounds at each stage of development (n_1, n_2, n_3) and the

[4] 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.

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

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

probabilities of success of each compound (p_1, p_2, p_3). (n_1, n_2, n_3) are obtained from the ADIS database as the number of distinct compounds in each of the three stages of trials being developed for the same market as the failed 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 don't otherwise appear in our sample. Danzon et al. (2004) 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. The estimated indication-phase specific probabilities and the detailed procedure are provided in Section 3 of the online supplement. Finally, we use an annual discount rate of 12%, to compute α_1 and α_2.[7] A minority of compounds (25 out of 169) 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 find similar results.

5.3 Explanatory Variable: Phase II Buffer

Hypothesis 3 posits that the impact of the phase III failure is proportional to the extent of the unanticipated demand for phase III resources at the time of the failure captured by the difference between the recent success rate and the average phase II success rate for the firm in question (Phase II buffer). To compute the phase II success rates, we divide 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 average, long-run or expected success rate, we look

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

at the number of successes and failures for the firm over the full time-period of our data-sets– 1994 to 2004. To compute the recent success rate we look at the number of successes and failures in 300-day period preceding the day of the failure announcement.[8] The difference of the two is used as the explanatory variable in our regressions (Equation (3)).

$$\begin{aligned} \textit{Phase II Buffer} &= \textit{Recent success rate} - \textit{Long run success rate} \\ &= \frac{\#of\ Successes_{\ t\in[-300,0]}}{\#of\ Successes_{\ t\in[-300,0]} + \#of\ Failures_{\ t\in[-300,0]}} - \frac{\#of\ Successes_{\forall t}}{\#of\ Successes_{\forall t} + \#of\ Failures_{\forall t}} \end{aligned} \tag{3}$$

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 compound in question across all firms present in the ADIS database (NActiveTrials). Previous research (Nicholson et al. (2003)) 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 financial stake of the firm in the compound. Third, the number of originating firms associated with the compound (NOriginators). This variable is also related to the financial stake of the firm in the compound in question.

We also include two firm-specific control variables- the sales in the quarter of the failure (Sales) and the R&D expenses incurred by the firm in the quarter of the failure (R&D Expenditure) to account for firm-specific heterogeneity. These variables capture the size of the firm and the R&D organization associated with the failure.

We estimate the model in equation (4).

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

$$CAR_i = a_0 + a_1(NActiveTrials)_i + a_2(NOriginators)_i + a_3(NLicensees)_i + a_4 Sales_i + a_5 R\,\&\,D\,Expenditure_i + a_6 \Delta TAPIS_i + a_7\left(Phase\,II\,Buffer\right)_i + \varepsilon_i \quad (4)$$

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 for the explanatory variables constructed for this study 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 between the variables used in this study.

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 volume. The abnormal trading volume data are illustrated graphically in Figure 4. Average trading volume peaks to 816 million shares over the expected trading volume two days before the announcement of the failure. Trading volumes in the time period (-2, 4) are found to be statistically significant and 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 with abnormal returns for alternative event windows ((-3, 3) and (-4, 4)) 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. These estimates are used to compute the daily and cumulative abnormal returns as described in Section 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) in their study of announcements of product *successes* report 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 in line with the results in this paper.

Sharma and Lacey (2004) in their comparative study of drug development failures and successes, construct a dataset of 41 outright FDA rejections of New Drug Approval (NDA) applications (the last stage in Figure 1). An NDA rejection indicates a serious mismanagement in the firm's internal controls and systems to ascertain drug safety. Typical reasons for the rare failure at this stage may be misrepresentation of results, improper design of clinical trials, not following standard procedures during trials, withholding data during trials, etc. This is a much more serious and rarer failure than the medical failure of a drug undergoing phase III clinical trials. Thus, the expected impact of these events should be much 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 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. (2005a)

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 Section 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 25%, 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 multi-collinearity. We also examine our estimation procedures for influential observations using guidelines suggested by Belsley et al. (1980) and find our results to be robust to outliers.

We find support for Hypothesis 2. The coefficient for the variable ΔTAPIS is found to be significant. A one percentage point difference in the change in TAPIS leads to 1.63 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 1.23 to 3.6 basis points). In the illustrative example of Figure 2, our regressions suggest that failure 1 would hurt the firm by an extra 1.05% or US$294.2 million compared to failure 2. Further, for one standard-deviation difference in the value of ΔTAPIS (calculated from our sample), the difference in financial impact corresponds to 0.81% or US$227 million.

Support for hypothesis 2 demonstrates the following– First, the presence of other compounds for the same market leads to lower financial impact of a failure or lower valuation of a compound. Second, in developing ΔTAPIS, we assume that all successful compounds are equally good, have the same probability of success at the same stage of development, and that the successes of different compounds in clinical trials are independent (like Ding and Eliashberg (2002)). Our empirical evidence shows that ΔTAPIS is correlated with the financial valuation of compounds by stock markets. Thus, our assumptions

are representative of the collective wisdom of the stock markets.

We find support for Hypothesis 3. The coefficient for the variable 'Phase II buffer' is found to be 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 one percentage point difference in success rate (Phase II buffer), leads to a 6.58 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 6.8 basis points). In the illustrative example of Figure 3, our regressions suggest that failure 2 would hurt the firm by an additional 1.65% or US$457 million as compared to failure 1. Further, for one standard deviation difference in the Phase II buffer, the difference in financial impact corresponds to 1.38% or US$382.2 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. Further, in developing hypothesis 3 we assumed that development resources can not be scaled up or down in the short-term. Empirical evidence suggests that this is a representative assumption.

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 (28th November, 2003) in the Wall Street Journal on the impact of failures and a lean pipeline on Merck Pharmaceuticals).

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 an even more significant 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 errors in estimation of the value of a project by the order of 100's of millions of dollars when the average values of a project is approximately US$500 million. This could lead to highly sub-optimal portfolio and capacity choices. Thus, a product development manger interested in maximizing the returns for his shareholders would be benefited by using decision support systems that acknowledge and model these interactions based on our empirical observations.

Finally, while for hypothesis 2, we developed a detailed non-linear model relating the number of backup compounds to the financial value of a compound; for hypothesis 3, we did not do so. The financial impact of the congestion effects central to hypothesis 3 may not be linear in the proxy for utilization.[9] An appropriate queuing theoretic model that analytically captures these effects remains the subject of our 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 for a portfolio in which there are multiple projects targeting the same market is smaller than that for a portfolio in which it

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

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 more than expected number of compounds that will utilize the same development resources as the failed project. Put differently, the value of a project to a portfolio in which there are more projects that utilize the same resources as the failed project is smaller than to a portfolio in which it is the sole claimant to the resources.

In addition to validating our intuition on the portfolio-level projects interactions, support for our hypothesis validates the assumption around the structure of portfolio-level project interactions that we used to develop our theoretical metrics on 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 results also provide 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 were added to its portfolio. This should be the maximum fair price that the firm should pay for this compound.

While the coefficients estimated in this paper are applicable only for late stage failures in the pharmaceutical industry, the insights, framework and empirical methodology developed can be employed for analyzing product development portfolios in R&D environments with high uncertainty that is 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 shared fixed development capacity and the associated economics that we investigate are also typical of many R&D environments. Product development environments such as the development of consumer packaged goods with test markets, multi-phase defense development contracts, etc. are all amenable to the methods and insights developed in this paper.

Our study suffers from some potential limitations. The methodology employed rests on the assumption that markets accurately estimate the factors that influence profits from drug development. This is a reasonable assumption for pharmaceutical industry which has high investments by sophisticated institutional investors, extensive regulatory and scientific scrutiny, high levels of disclosure, exogenously defined and publicly measured metrics of product performance. However, this assumption may not apply equally well to all industries. While 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 are equally applicable, as long as there is no systemic irrationality correlated with our product development variables.

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

REFERENCES

Adler, P. S., A. Mandelbaum, V. Nguyen, and E. Schwerer (1995): "From Project to Process Management: An Empirically-Based Framework for Analyzing Product Development Time," *Management Science*, 41, 458-484.

Bayus, B. L., G. Erickson, and R. Jacobson (2003): "The financial rewards of new product introductions in the personal computer industry," *Management Science*, 49, 197-210.

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

Brown, S. J., and J. B. Warner (1980): "Measuring Stock Price Performance," *Journal of Financial Economics*, 8, 205-258.

Chaney, P. K., and T. M. Devinney (1992): "The Impact of New Product Introductions on Market Value

of Firms," *Journal of Business, Finance and Accounting*, 19, 677-695.

Chaney, P. K., T. M. Devinney, and R. S. Winer (1991): "The Impact of New Product Introductions on the Market Value of Firms," *Journal of Business*, 64, 573-610.

Chen, S. S., T. Y. Chung, K. W. Ho, and C. F. Lee (2005a): "Intra-industry Effects of Delayed New Product Introductions," *Working Paper*.

Chen, S. S., K. W. Ho, and K. H. Ik (2005b): "The wealth effect of new product introductions on industry rivals," *Journal of Business*, 78, 969-996.

Corbett, C. J., M. J. Montes-Sancho, and D. A. Kirsch (2004): "The Financial Impact of ISO 9000 Certification in the US: An Empirical Analysis," *Working Paper*.

Dahan, E., and H. Mendelson (2001): "An Extreme Value Model of Concept Testing," *Management Science*, 47, 102-116.

Danzon, P. M., S. Nicholson, and N. S. Pereira (2004): "Productivity in Pharmaceutical-Biotechnology R&D: The Role of Experience and Alliances," *Working Paper, The Wharton School*.

Ding, M., and J. Eliashberg (2002): "Structuring the New Product Development Pipeline," *Management Science*, 48, 343-363.

Girotra, K., C. Terwiesch, and K. T. Ulrich (2004): "New Drug Development at Merck & Co.," *Wharton School Teaching Case*.

Hendricks, K. B., and 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 Science*, 43, 422-436.

— (2003): "An Empirical Analysis of the effect of supply chain disruptions on operating performance," *Working Paper*, 2003.

Kavadias, S., and C. H. Loch (2003): *Project Selection under Uncertainty: Dynamically Allocating resources to Maximize Value*. Boston, MA: Kluwer Academic Press.

Kothari, S. P., and J. B. Warner (2004): "Econometrics of Event Studies," Elsevier/North-Holland.

Landers, P., and J. S. Lublin (28th November, 2003): "Under a Microscope: Merck's Big Bet on Research By Its Scientists Comes Up Short --- As 4 of 11 Drugs in Pipeline Fail, Some Investors Clamor For

Merger or New Strategy --- Board Stands by Mr. Gilmartin ".

Loch, C. H., and S. Kavadias (2002): "Dynamic Portfolio Selection of NPD Programs using Marginal Return," *Management Science*, 48, 1227-1241.

Loch, C. H., M. T. Pich, C. Terwiesch, and M. Urbschat (2001a): "Selecting R&D Projects at BMW: A Case Study of Adopting Mathematical Programming Models," *IEEE Transactions on Engineering Management*, 48, 70-80.

Loch, C. H., and C. Terwiesch (1999): "Accelerating the Process of Engineering Change Orders: Capacity and Congestion Effects," *Journal of Product Innovation Management*, 16, 145-159.

Loch, C. H., C. Terwiesch, and S. Thomke (2001b): "Parallel and Sequential Testing of Design Alternatives," *Management Science*, 47, 663-678.

Lucas, R. (1971): "Optimal Management of a Research and Development Project," *Management Science*, 17, 679-697.

Mackinlay, C. A. (1997): "Event Studies in Economics and Finance," *Journal of Economic Literature*, 35, 13-39.

Nicholson, S., P. M. Dranzon, and J. Mccullough (2003): "Biotech-Pharmaceutical Alliances as a Signal of Asset and Firm Quality," *Working Paper, The Wharton School*.

Parexel (2002/2003): *PAREXEL's Bio/Pharmaceutical R&D Statistical Sourcebook*. Parexel.

Patell, J. M. (1976): "Corporate Forecasts of Earning per Share and Stock Price Behavior: Empirical Tests," *Journal of Accounting Research*, 14, 246-274.

Pisano, G. P., and S. Rossi (1994): "Eli Lilly and Co.: The Flexible Facility Decision– 1993," *Harvard Business School Case # 9-694-074*.

Robertson, T. S., J. Eliashberg, and T. Rymon (1995): "New Product Announcement Signals and Incumbent Reactions," *Journal of Marketing*, 59, 1-15.

Shane, S. A., and K. T. Ulrich (2004): "Technological Innovation, Product Development, and Entrepreneurship in Management Science," *Management Science*, 50, 133-144.

Sharma, A., and N. Lacey (2004): "Linking product development outcomes to market valuation of the

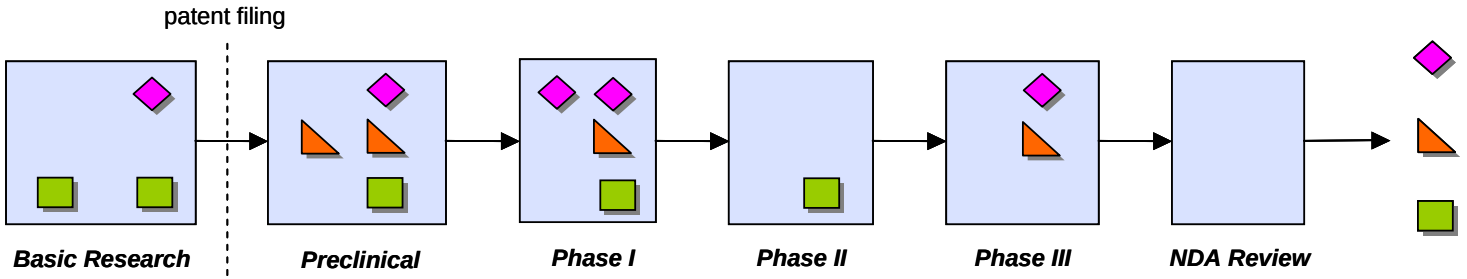
firm: The case of the US pharmaceutical industry," *Journal of Product Innovation Management*, 21, 297-308.

Singhal, V. R. (2005): "Excess Inventory and Long-term stock performance," *Working Paper, Georgia Institute of Technology*.

Sprent, P. (1989): *Applied Nonparametric Statistical Methods*. London: Chapman and Hall.

Tkac, P. A. (1999): "A Trading Volume Benchmark: Theory and Evidence," *Journal of Financial and Quantitative Analysis*, 34, 89-114.

TABLES AND FIGURES



	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 filing
average duration (years)		1.5	1	2	2	1
average output rate (comp/yr)	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

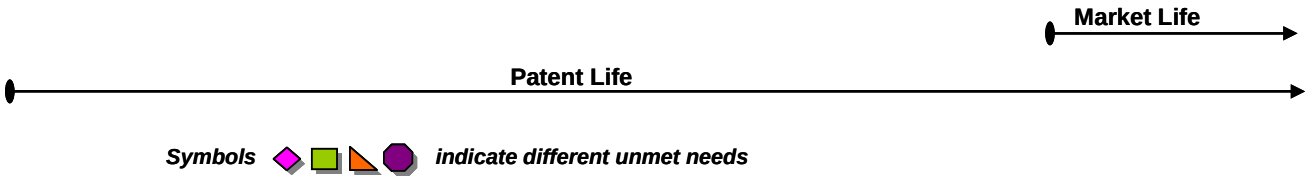


Figure 1: The drug development process Values are for a typical pharmaceutical company and are normalized based on an average annual output of one compound. Development projects that target the same market or indication are denoted with the same symbol. Estimates were obtained from the Parexel Pharmaceutical R&D Statistical Source Book, 2002/2003.

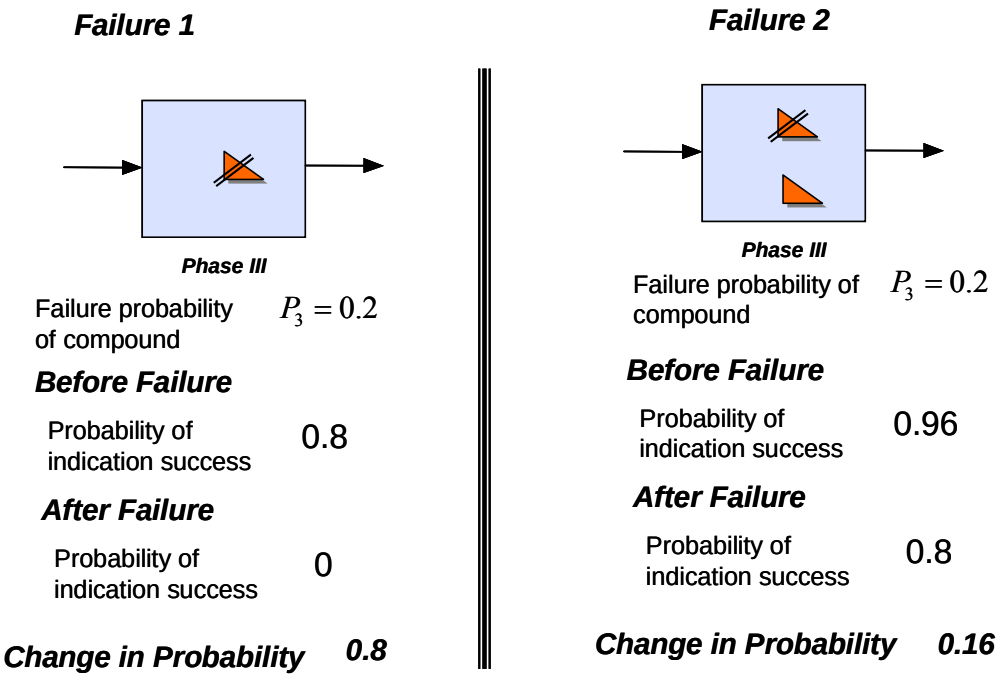


Figure 2: Example illustrating the role of backup compounds.

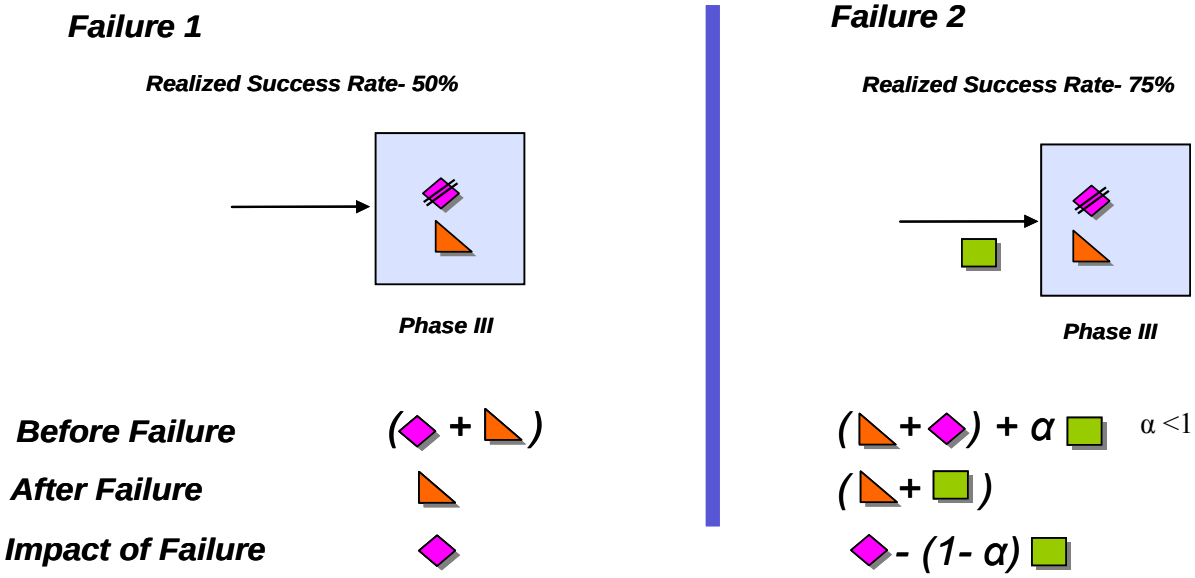


Figure 3: Example illustrating the effect of different realized Phase II Success Rate

Firms in the Study

Firm/ Subsidiaries	No. of phase III failures (94-04)	Annual Sales+ (MM$)	No. of Employees+ ('000s)	R&D Expenditure+ (MM$)	R&D/Sales (%)
Pfizer Inc.	20	21,107.97	67.82	3,788.98	17.95
Aventis	19	18,159.86	82.81	2,248.14	12.38
Glaxo SmithKline	17	20,237.11	77.07	2,784.24	13.76
Bristol Myers Squibb, Mead Johnson	15	17,146.50	48.88	1,994.80	11.63
Novartis	12	22,852.80	82.24	2,886.46	12.63
Genentech	11	1,544.39	4.01	439.66	28.47
Chiron Co., Chiron Vaccines	10	976.37	4.63	302.31	30.96
Wyeth, Wyeth Vaccines	10	13,546.55	56.91	1,628.51	12.02
Sanofi Synthelbo	9	7,330.08	31.31	1,184.92	16.17
Amgen, Amgen Boulder	7	3,581.10	6.73	1,092.44	30.51
AstraZenca	7	12,998.73	44.24	1,960.31	15.08
Pharmacia, Monsanto	7	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
Schering AG	5	4,854.75	25.12	875.22	18.03
Bayer	4	30,645.40	131.56	2,352.43	7.68
Merck	3	29,708.38	60.31	2,147.03	7.23
Novo Nordisk	2	2,985.00	15.02	462.82	15.50
Alcon	1	2,929.33	11.56	307.00	10.48
Median	7	12,998.73	44.24	1628.51	14.44

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

Table 1: Firms in the Study

Independent Variables

Independent Variable	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$)	3767.27	3627.01	13982.00	117.58	2494.14
R&D Expenses (MM$)	655.98	483.00	3266.08	67.45	563.77
ΔTAPIS	0.7006	0.6417	2.2661	0.0294	0.5013
Phase II Buffer	0.0413	0.0791	0.4138	-0.8095	0.2169

Table 2: Independent Variables

Pearson Correlation Coefficients

	NActiveTrials	NLicensees	NOriginators	Sales (Net) (MM$)	R&D Expense (MM$)	Phase II Buffer	ΔTAPIS
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		
ΔTAPIS	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<.0.1% level, **-significant at the p<1% level, *-significant at the p<10% level***

Table 3: Correlation between Independent Variables

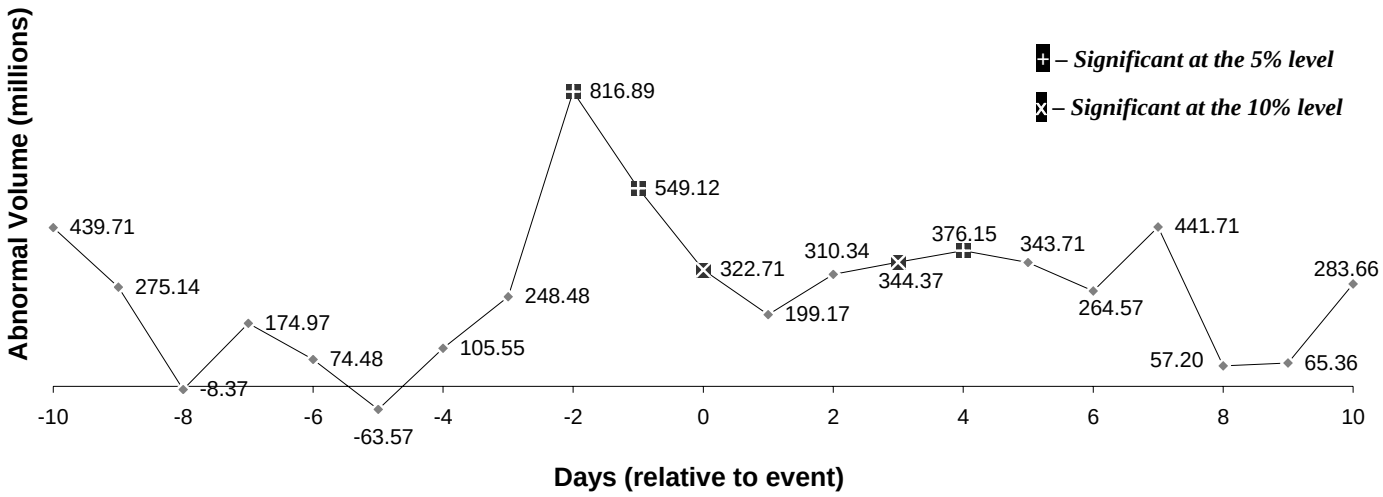


Figure 4: Abnormal Trading Volume

Cumulative Abnormal Return

Model	Window	Mean Abnormal Return					Median Abnormal Return		Change in Market Capitalization	
		Mean Cumulative Abnormal Return	Cross Sectional Std. Dev Test	Patell-Z Statistic+	Crude Dependence Adjustment Test&	Generalized Sign-Z^	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

Significance Levels from a one-tail t-test:*-0.1% significance ***-1% significance, **-5% significance, *-10% significance.***
+- 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 ^- The non-parametric sign-z test, tests the null hypothesis that the number of positive and negative return is the same (Sprent (1989).

Table 4: Cumulative Abnormal Returns

Regression Model: Parameter Estimates

	Comparison Period Model			Market Model			Fama French Model		
Variable	**(-2,4)**	**(-3,3)**	**(-4,4)**	**(-2,4)**	**(-3,3)**	**(-4,4)**	**(-2,4)**	**(-3,3)**	**(-4,4)**
Intercept	0.02	0.0107	0.025	0.0413**	0.027	0.0408*	0.0275	0.0258	0.0376*
	(0.97)	(0.49)	(1.11)	(1.96)	(1.29)	(1.86)	(1.36)	(1.29)	(1.84)
NActiveTrials	-0.405*	-0.5573**	-0.3953	-1.04**	-1.04**	-0.9308*	-0.9582**	-1.27***	-1.3***
1,000s	(-1.78)	(-2.3)	(-1.59)	(-2.13)	(-2.14)	(-1.83)	(-2.13)	(-2.82)	(-2.85)
NOriginators	-1.091	-0.335	-1.766	-1.53*	-0.547	-2.277***	-1.91**	-0.989	-2.124***
100s	(-1.09)	(-0.31)	(-1.61)	(-1.95)	(-0.7)	(-2.78)	(-2.45)	(-1.27)	(-2.68)
NLicensees	-0.639	-0.422	-0.707	-0.927**	-0.616	-0.992**	-0.717**	-0.416	-0.866**
100s	(-1.56)	(-0.97)	(-1.58)	(-2.24)	(-1.5)	(-2.3)	(-2.02)	(-1.18)	(-2.41)
Sales	6.48**	7.11**	5.73*	6.28**	6.58**	6.49**	4.93*	5.41**	4.07
1,000,000s	(2.29)	(2.36)	(1.85)	(2.23)	(2.35)	(2.21)	(1.81)	(2)	(1.48)
R&D	-2.223*	-2.872**	-1.927	-2.475**	-2.986**	-2.204*	-2.38**	-3.187***	-2.237**
100,000s	(-1.88)	(-2.28)	(-1.49)	(-2.02)	(-2.45)	(-1.72)	(-2.25)	(-3.03)	(-2.09)
ΔTAPIS	-1.902*	-1.234	-1.491	-3.585***	-3.172***	-3.188***	-1.633*	-1.968**	-1.787**
100s	(-1.68)	(-1.02)	(-1.2)	(-3.34)	(-2.98)	(-2.85)	(-1.71)	(-2.06)	(-1.97)
Phase II Buffer	6.826***	5.949**	6.195**	5.789**	4.225*	4.666*	6.577***	4.241**	5.276**
100s	(2.69)	(2.21)	(2.23)	(2.45)	(1.8)	(1.89)	(2.99)	(1.96)	(2.37)
Adj R-Sq	9.22%	7.32%	6.71%	17.06%	10.37%	17.90%	19.18%	13.65%	20.32%
R-Square	14.43%	12.64%	12.07%	21.64%	15.31%	22.42%	24.01%	18.82%	25.08%
N#	123	123	123	128	128	128	118	118	118
F-Value	2.77	2.38	2.25	4.73	3.1	4.95	4.97	3.64	5.26
Pr>F	1.07%	2.63%	3.47%	0.01%	0.49%	0.00%	0.00%	0.14%	0.00%

Significance Levels from a two-tail t-test: ***-1% significance, **-5% significance, *-10% significance.

 # Abnormal returns can not be estimated for some events– where the security did not trade long enough to estimate the expected return models.

Table 5: Parameter Estimates