UNIVERSITY OF GOTHENBURG
SCHOOL OF BUSINESS, ECONOMICS AND LAW

Value Relevance of R&D Advances in the US Biotechnology Industry

Master Thesis

June 2020

Authors: Gerardo Ortiz & Ida Tobiasson

Supervisor: Dawei Fang

Graduate School

School of Business, Economics and Law

University of Gothenburg

Abstract

We investigate the value relevance of non-financial information in the US biotechnology industry by studying the stock market reactions to announcements of drug development outcomes for a sample period of 10 years (from January 2010 to December 2019). We find that the market reacts strongly to these announcements and that the magnitude of the market reaction differs between announcements made in different phases of the drug development process as well as between positive and negative announcements. Consistent with our theoretical predictions, we find that the market reacts more strongly to late-stage announcements than to early-stage announcements and more strongly to negative announcements than to positive announcements.

Keywords: Biotechnology, Non-financial information, R&D announcements, Stock market reactions, Value relevance

Acknowledgements Our deepest gratitude goes to our supervisor, Assistant Professor Dawei Fang at the University of Gothenburg, whose comments have been invaluable to us. We would also like to thank Hans Jeppsson for providing material and answering questions about the research topic, Bassam Ribbfors for helping us with programming issues and our opponents Simon Möller and Angelica Yngvesson for their extensive review which has improved the papers a lot.

Authors Contact Details

Gerardo Ortiz E-mail: gjo0510@gmail.com

Ida Tobiasson E-mail: ida.tobiasson@gmail.com

Contents

1 Introduction

In this paper we investigate the value relevance of non-financial information in the biotechnology industry. Our study focuses on biotechnology firms involved in the research and development of new drugs. Disclosures of accounting information are considered value relevant (Botosan, 1997). However, relying solely on accounting information to valuate firms is only appropriate when future performance resembles past performance, which is clearly not the case for biotech firms in the process of developing a new drug. The drug development process is long and characterized by heavy investments in R&D.[1] It often costs more than 1 billion USD to develop a drug (Ernst & Young, 2017). In fact, even though the US biotechnology industry is becoming a mature industry with a growing number of profitable companies (Nasdaq, 2017), most biotech firms are still in the early life-cycle stage where they typically spend more than 100% of revenues on R&D. Consequently, as US accounting standards require companies to expense their significant value enhancing investments in internally developed intangible assets, many biotech firms report large losses year after year. These industry characteristics give reasons to believe that non-financial information, such as progress announcements, is more value relevant than traditional accounting information for biotechnology firms.

The drug development process consists of several research phases. For a drug to be approved by the Food and Drug Administration (FDA), it has to pass the last clinical trial phase, Phase III. However, passing Phase III requires the drug to pass all earlier phases, including the pre-clinical phase, Phase I, and Phase II, and the drug can fail at any of these phases.

Overall, there appears to be a consensus regarding the value relevance of non-financial information in the biotech industry (Callen et al., 2010; Dedman et al., 2008; Ely et al., 2003; McConomy and Xu, 2004; Xu et al., 2007).[2] Therefore, in this study, we are interested not only in whether the market reacts to disclosure of drug development information but

[1] The largest R&D costs occurs during the clinical trials (Ernst & Young, 2017).

[2] notwithstanding Hand (2005)

more in how the market reacts. Does the market react more strongly to disclosure of drug development information in earlier phases or in later phases? Does the market react more strongly to positive announcements or to negative announcements?

To answer these questions, we first build a simple model. Our model predicts that stock price reacts more strongly to late-stage negative announcements than to early-stage negative announcements and, if passing an early stage is more likely than passing a late stage (conditional on passing the early stage), stock price reacts more strongly to late-stage positive announcements than to early-stage positive announcements. Moreover, under the condition that passing a phase is more likely than failing at a phase, stock price reacts more strongly to negative announcements than to positive announcements.[3] We then bring these predictions to US data using a 10 years sample period (from January 2010 to December 2019). Consistent with our theoretical predictions, we find that market price does react more strongly to late-stage announcements than to early-stage announcements and more strongly to negative announcements than to positive announcements. Moreover, the behavior of trading volumes is similar, which also react more strongly to late-stage than to early-stage announcements and more strongly to negative than to positive announcements.

Finally, we also look at firm- and project-specific variables to explain the cross-sectional variation in abnormal returns. In line with our predictions, we find that the larger the investment (number of patients) in the clinical trial, the larger the market reaction, and the larger and more diversified the firm is, the smaller the market reaction.

This study provides the following contributions to the accounting and finance literature. First, it documents how the market reacts in aggregate to US biotechnology firms' disclosure of non-financial information. Second, it shows to which extent firm- and project- specific

[3]In our data, the average success probability of Phase I and Phase III clinical trials are 0.65 and 0.59. For Phase II trials the average success probability is only about 0.33. For Phase I/II and Phase II/III it is 0.56 and 0.48, respectively.

information is relevant to investors. This knowledge is important for investors evaluating biotechnology firms as all else equal, a firm with fewer drugs in a certain development phase has a considerably different risk profile than a firm with more drugs in that same phase. Regardless of the importance of these factors to investors, it is very difficult to find reliable measures of these differences in the literature (Jeppsson, 2010).

Next, to our knowledge this study is the first to include historical success rates per therapy area for biotechnology firms' R&D advances. Many studies conducted in the past have used DiMasi (2001) success rates which are based on pharmaceutical companies, not biotechnology firms (e.g., Ely et al., 2003; Jeppsson, 2010; Xu et al., 2007). Using biotechnology-specific success rates give a more reliable prediction of the success probability for different types of drugs in specific phases of the R&D process.[4]

Finally, as far as we know this study is the first to evaluate the market effects of announcement of combined clinical trial phases (Phase I/II and Phase II/III) in the US biotechnology industry. These combined phases are supposed to allow research questions to be answered more quickly or with a smaller sample of patients. Although, there is a lot of literature discussing the effects of individual-phase announcements (Phase I, Phase II and Phase III), we have not been able to find any literature studying if the same effects can be found for combined phases. This study documents the difference in market reactions to both individual and combined-phase announcements.

[4]DiMasi (2001) classified the pharmaceutical therapy areas' success rates in ten categories updated as of December 31, 1999. In this study we use Biotechnology Innovation Organization's (2016) biotech success rates. These rates were estimated using information of biotechnology companies clinical and regulatory phase transitions from 2006 to 2015 for fifteen different therapy areas.

2 Literature Review

Our study is related to the literature on the value relevance of non-financial information in the following ways. First, we study the difference in price reaction to early- and late-stage R&D announcements and we find empirical support for our hypothesis, that the market will react more strongly the closer a drug is to receiving final approval. Our result is in line with most of the existing empirical results. More specifically, Dedman et al. (2008), Jeppsson (2010) and McConomy and Xu (2004) find that the strongest market reaction comes from later stage announcements (i.e. Phase III status updates). One exception is Ely et al (2003), who find that price reactions are largest and significant only for Phase II announcements. The reason why Ely et al. do not find a significant market reaction for the later stage announcements, could be that they do not have the correct event date or because of an over-sampling of drugs in Phase II (Joos, 2003). According to Joos (2003) information on Phase III and FDA submission typically becomes available through alternative information sources, such as medical journals, conference abstracts and analyst meetings before companies make the public announcement. However, because disclosures can influence the companies' stock price, securities laws prohibit companies from providing this information elsewhere before first making it publicly available (i.e., there is a rule against "selective disclosure" (Fisher, 2002)). Failing to follow Securities and Exchange Commission (SEC) regulation may lead to lawsuits. The oversampling of Phase II drugs Joos argues is due to there possibly being a reporting bias in the PhRMA surveys, caused by under-reporting of Phase I clinical trials, for competitive reasons.

Second, we study the difference in price reaction to positive and negative R&D announcements and find empirical support for our hypothesis, that the market will react more strongly to negative announcements. Our result is consistent with the findings of asymmetric market reactions to positive versus negative R&D announcements documented by empirical studies using European data (Jeppsson, 2010) and using less recent US data (Xu, 2009). The positive (negative) return on day zero for Phase I, Phase II and Phase III respectively, are 2.53% (-20.18%), 2.68% (-15.28%) and 4.83% (-39.04%). Also, Callen et al. (2010) document a

strong market reaction to negative news (mean abnormal return -23%) which they conclude shows the economic importance of each drug under development. Jeppsson (2010), on the other hand, argues that the large reaction to negative news suggests that negative news are largely unanticipated by investors.

Finally, our documentation of the value relevance of disclosure of drug development information is consistent with the broader results on value relevance of non-financial information. Amir and Lev (1996) investigate the value relevance of financial information, such as earnings, book values and cash flows of US cellular companies, and find that these variables alone are unrelated to security valuation. Further, they find that non-financial information is highly value relevant and that combined with non-financial information some financial variables, like earnings, do contribute to explaining market value. When McConomy and Xu (2004) investigate if the same is true for biotechnology firms they find that the market reacts more strongly to non-financial disclosures than to earning announcements. Their findings explain why for these companies, analysts and investors tend to focus on the revenue-generating potential of products in development (Yang, 2008) and hardly ever mention earnings in their industry analyses (Amir and Lev, 1996). Hirschey, Richardson and Scholz (2001) find similar results when studying a broader sample of high-tech companies, including among others, biotech, computer and communications firms. They more specifically find that patent quality data is most helpful when investors assess ongoing value creation by R&D investments in high-tech industries. In fact, R&D announcement information, which is what we study, is just one type of non-financial information and patent data which is another type has been used extensively in prior studies of the biotechnology industry. For example, patents have been used as a measure of R&D success (e.g., Joos, 2002), as a proxy for the firm's knowledge base and future earnings potential (e.g., Callen et al., 2010; Yang, 2007), and as a determinant of biotech firms' disclosure strategy (e.g., Guo et al., 2004). Callen et al. (2010) find positive (negative) abnormal returns around the announcement of patent grants and patent infringement lawsuits in favor of (against) biotech firms, hence providing evidence of patent data being another value relevant non-financial measure in the biotechnology industry.

3 Background

There are several research phases during the drug development process. In the pre-clinical phase (which we do not investigate in this study), basic questions about a drugs safety are answered through testing in animal subjects. During the clinical trials, the drug's safety and possible adverse effects are evaluated through tests on human subjects. More specifically, Phase I trials evaluate if the drug is safe to take, what the ideal dosage is and how it should be administered. While Phase I is performed on healthy volunteers, Phase II is an evaluation in patients with the target disorder. Phase II trials test the efficacy and further study side effects of taking the drug. Phase III, which is the last clinical trial phase before FDA approval, is meant to demonstrate that the drug is at least as safe and effective as existing treatment options (US Food and Drug Administration, 2018). Finally, Phase I/II (mostly used in cancer drug development) tests for safety and side effects and Phase II/III (mostly used in cancer drug development) tests for efficacy, side effects and adverse reactions in a shorter period, using smaller samples than traditional separate trial phases (National Cancer Institute, 2020).

Manipulation of R&D progress updates is prohibited and the FDA acts as a gatekeeper who decides whether to grant approvals and monitors companies' press releases to make sure that they do not contain misleading information. Decisions not to disclose material information when there is a duty to disclose or to disclose inaccurate information may lead to private lawsuits, SEC enforcement actions, and even criminal prosecutions. In the context of biotechnology firms the duty to disclose is triggered once one of two events occurs: 1) trading in the biotechnology firm's stock by insiders, or 2) a statement is made by the firm (including officers speaking for the firm) that misleads investors unless the firm also discloses material information it has been withholding (Fisher, 2002).[5]

[5]For example, any comment by a biotechnology firm that even implies effective performance may lead to a litigation later if the firm fails to simultaneously disclose negative test results (Fisher, 2002).

4 Theory and Research Hypotheses

Consider a simple model of the drug development process, consisting of two phases, an early phase and a late phase. For simplicity, we assume only two phases. As we mentioned in the introduction, in practice, there are more than two phases. Extending our analysis to the case of more than two phases is straightforward, but it will not change the qualitative results obtained from our simple two-phase model. The drug, if successfully developed, will generate a profit for the firm, which is normalized to 1. The probability of the drug passing the early phase is $p \in (0, 1)$ and the probability of the drug passing the late phase (conditional on the drug passing the early phase) is $q \in (0, 1)$. For analytical convenience, assume that the firm incurs no cost of developing the drug.

In what follows, we compute the expected firm value with the drug in different scenarios. Denote the value of the firm by π. The unconditional expected firm value (i.e., the expected firm value before the early phase) is given by

$$\mathbb{E}[\pi] = pq + v \tag{1}$$

where $v > 0$ represents the value from all other projects and we assume that v is a fixed constant throughout the drug development process. The expected firm value conditional on the drug successfully entering the late phase is given by

$$\mathbb{E}[\pi | \textit{entering the late phase}] = q + v. \tag{2}$$

The expected firm value conditional on failing at either stage is v. The expected firm value conditional on passing the late phase (the last phase in our model) is $1 + v$. The above results can be used to compute market reactions to different announcements. We measure market reaction to a certain announcement by the difference in expected firm value before and after the announcement. Let $m_{k,a}$ be the market reaction to Phase $k \in \{early, late\}$ announcement $a \in \{S, F\}$, defined as a percentage change of expected firm value before

and after this announcement, where S denotes success and F denotes failure. For example, $m_{early,S}$ denotes the market reaction to an announcement of success in the early phase, and $m_{late,F}$ denotes the market reaction to announcement of failure in the late stage. Following our definition, we obtain that

$$m_{early,S} = \frac{\mathbb{E}[\pi|entering\ the\ late\ phase] - \mathbb{E}[\pi]}{\mathbb{E}[\pi]} = \frac{q - pq}{pq + v} \tag{3}$$

$$m_{early,F} = \frac{\mathbb{E}[\pi|failing\ the\ early\ phase] - \mathbb{E}[\pi]}{\mathbb{E}[\pi]} = \frac{-pq}{pq + v} \tag{4}$$

$$m_{late,S} = \frac{\mathbb{E}[\pi|passing\ the\ late\ phase] - \mathbb{E}[\pi|entering\ the\ late\ phase]}{\mathbb{E}[\pi|entering\ the\ late\ phase]} = \frac{1 - q}{q + v} \tag{5}$$

$$m_{late,F} = \frac{\mathbb{E}[\pi|failing\ the\ early\ phase] - \mathbb{E}[\pi|entering\ the\ late\ phase]}{\mathbb{E}[\pi|entering\ the\ late\ phase]} = \frac{-q}{q + v} \tag{6}$$

Observations. First, it is obvious from equations (4) and (6) that $m_{early,F}$ and $m_{late,F}$ are negative. Second, because $p < 1$ and $q < 1$ it is obvious from equations (3) and (5) that $m_{early,S}$ and $m_{late,S}$ are positive. Thus, successfully passing any phase will lead to an increase in firm value and failing any phase will lead to a decrease in firm value. Based on these observations we formulate our first hypothesis.

Hypothesis 1: The abnormal share price reaction will be positive for disclosures of positive R&D announcements and negative for disclosures of negative R&D announcements.

Next we notice that

$$m_{late,F} - m_{early,F} = \frac{-q}{q + v} - \frac{-pq}{pq + v} = \frac{pq(q + v) - q(pq + v)}{(pq + v)(q + v)} = \frac{pqv - qv}{(pq + v)(q + v)}. \tag{7}$$

Because $p < 1,\ q > 0$ and $v > 0$ we have $pqv - qv < 0$. It is thus clear from the above equation that $m_{late,F}$ is more negative than $m_{early,F}$. This result implies that the market reacts more strongly to disclosure of late-stage negative announcements than to disclosure of early-stage negative announcements. Second, whether $m_{late,S}$ is more positive than $m_{early,S}$ (i.e., whether the market reacts more strongly to late-stage positive announcements than to early-stage positive announcements) is ambiguous. However, note that

$$m_{late,S} - m_{early,S} = \frac{1-q}{q+v} - \frac{q-pq}{pq+v} = \frac{v(1-2q+pq)+pq-q^2}{(q+v)(pq+v)}. \tag{8}$$

If $p > q$, then $1 - 2q + pq > 1 - 2q + q^2 = (1-q)^2 \geqslant 0$ and $pq - q^2 > 0$, in which case the last expression in the above equation is positive. Thus, late-stage positive announcements have a stronger effect on market price compared to early-stage positive announcements if passing the early stage is easier than passing the late stage (i.e., if $p > q$). The average success probabilities in our data suggest that passing an early stage is indeed easier than passing later stages (i.e., the probability of passing Phase I is higher than the probability of passing Phase II and Phase III). The above analysis implies the following hypothesis:

Hypothesis 2: The abnormal share price reaction will be larger for disclosures of late-stage R&D announcements than disclosures of early-stage R&D announcements.

Further, we investigate whether the market reacts more strongly to negative announcements or to positive announcements by computing the difference between $|m_{early,S}|$ and $|m_{early,F}|$ and between $|m_{late,S}|$ and $|m_{late,F}|$ (in absolute value terms).

$$\mid m_{early,F} \mid - \mid m_{early,S} \mid = \frac{pq}{pq+v} - \frac{q-pq}{pq+v} = \frac{q(2p-1)}{pq+v} \tag{9}$$

$$\mid m_{late,F} \mid - \mid m_{late,S} \mid = \frac{q}{q+v} - \frac{1-q}{q+v} = \frac{2q-1}{q+v} \tag{10}$$

Equations (9) and (10) are both positive if $p > 1/2$ and $q > 1/2$. Thus, if passing a phase is more likely than failing at a phase, negative announcements will have a more profound effect than positive announcements. These conditions are satisfied in our data for Phase I and Phase III. In our data, the average success probability of Phase I and Phase III clinical trials are above 0.65 and 0.59, respectively. We formulate the third hypothesis as:

Hypothesis 3: The abnormal share price reaction will be larger for disclosures of negative R&D announcements than disclosures of positive R&D announcements.

For Phase II trials the average success probability is only about 0.33. For Phase I/II and Phase II/III it is 0.56 and 0.48, respectively. Because the combined phases did not have any historical success probabilities reported in the BIO (2016) study (see Appendix 2 for details) we calculated the probabilities used in this study as the average between Phase I and Phase II and between Phase II and Phase III. Hence, the actual success probability for Phase II/III could well be above 0.5. We therefore test both hypothesis 2 and 3 using Phase I (Phase I/II) and Phase III (Phase II/III) clinical trial announcements.

Finally, we run the same tests for trading volumes, as for returns. The studies of price-volume relation can be traced all the way back to Osborne (1959); however, it was not until Ying (1966) that price-volume correlation was first documented in the same data set (Karpoff, 1987). Using previous and current research as support, Karpoff (1987) established two empirical relations. First, volume is positively related to the absolute price change. Second, in equity markets, volume is also positively related to the price change per se. These findings are further supported by more recent studies. For example, Matilla-García, Marín and Dore (2013) found empirical results of a causal relationship between stock returns and trading volumes. Based on these findings we formulate our fourth hypothesis.

Hypothesis 4: The abnormal volume reaction will be larger for disclosures of late-stage R&D announcements than disclosures of early-stage R&D announcements and larger for negative R&D announcements than disclosures of positive R&D announcements.

5 Data and Methodology

This section first describes the sample selection process and second the research methodology used in this paper.

5.1 Sample and data

Our sample consists of companies included in the NASDAQ Biotechnology Index (NBI) after the annual re-ranking in December 2019. All firms included in the sample are involved in the research and development of new drugs. To ensure a homogeneous sample we excluded 31 of the 214 initially identified firms, since they are not classified as Biotechnology according to the Industry Classification Benchmark. We collected, stock prices, trading volumes as well as all other financial and accounting data from the Bloomberg Terminal.

In total we have a sample of 954 R&D announcements, which are hand-collected, from the corporate websites of 132 biotech firms. The other 51 firms in the NBI were dropped at this stage due to lack of R&D announcements during the sample period 2010 to 2019 or because they lacked enough stock price data for the estimation window. We categorize an R&D announcement both according to the associated phase and the type of announcement. There are three possible types of announcements, initiation, positive, and negative. Initiation refers to an announcement of entering a particular phase. Positive announcements refer to announcements that include phrases such as "positive results" and "primary endpoint was met". Negative announcements refer to announcements that include phrases like "negative results" and "primary endpoint was not met" or words like "discontinue" and "terminate" (see Appendix 1 for more details). Table 1 reports the number of announcements in our sample for each announcement category. From Table 1 we observe that there are 609 positive and negative announcements of which 90% are positive. This finding is consistent with Dedman et al. (2008), Ely et al. (2003) and Jeppsson (2010). Dedman et al. report that one of their most important findings is that firms fail to release negative news, after their sample revealed only 5% negative R&D announcements. Ely et al. and Jeppsson had samples with 15% and 25% negative R&D announcements, respectively.

Table 1. Clinical trial phases and R&D announcement classification

Announcements category	*Stage*	*Number of announcements*
Initiation		73
Results (positive)	Phase I	127
Results (negative)		3
Initiation		114
Results (positive)	Phase II	201
Results (negative)		24
Initiation		108
Results (positive)	Phase III	172
Results (negative)		31
Initiation		40
Results (positive)	Phase I/II	44
Results (negative)		0
Initiation		10
Results (positive)	Phase II/III	7
Results (negative)		0
Total		954

Note: This table reports the different types of announcements and the different phases of the R&D process that we consider in our analysis.

5.2 Event study methodology

To analyze the effect of R&D announcements on stock price and trading volume we follow the short-term event study methodology as suggested by MacKinlay (1997). MacKinlay describes multi-factor models for event studies as suitable only for samples where firms have a common characteristic, like for example being members of the same industry. However, prior studies of the biotechnology industry which performed an event study (e.g., Callen et al. 2010; Dedman et al. 2008; Jeppsson 2010) have used the market model, which is a single factor model, to estimate predicted returns. Hence, we decide to follow standard practice (i.e., use a single-index model) and choose to use the NBI as a proxy for the market portfolio. Although all pharmaceutical companies in the NBI are excluded from the sample we still choose to use the index as a proxy for the market portfolio since pharmaceuticals make up less than 9.5 percent of the index market value.

We use the following formula to calculate the mean cumulative abnormal return:

$$\overline{CAR}(t_1, t_2) = \Sigma_{t=t_1}^{t_2} \overline{AR}_t \tag{11}$$

where

$$\overline{AR}_t = \frac{1}{N} \Sigma_{i=1}^{N} AR_{it} \tag{12}$$

and where

$$AR_{it} = R_{it} - (\hat{\alpha} - \hat{\beta} R_{mt}) \tag{13}$$

AR_{it} is the abnormal return for each firm calculated as the actual return minus the predicted return. Where the predicted return is estimated using estimates from the market model. The estimation period used is 130 days (from day -150 to day -21). We do the same procedure to get abnormal volumes, using the following formulas:

$$\overline{CAV}(t_1, t_2) = \Sigma_{t=t_1}^{t_2} \overline{AV}_t \tag{14}$$

where

$$\overline{AV}_t = \frac{1}{N} \Sigma_{i=1}^{N} AV_{it} \tag{15}$$

and where

$$AV_{it} = v_{it} - (\hat{\alpha} - \hat{\beta} v_{mt}) \tag{16}$$

Because the number of shares traded increases as the firm size increases, we first scale volume to the total number of shares outstanding to get actual volume (Ajinkya and Jain, 1989).

For the above calculations, all overlapping events (i.e., if there was any announcement the day before, the same day or the day after) are excluded in order to avoid confounding events and for announcements disclosed when the market is closed, we assign the next trading day as the event day. To check robustness, the cumulative abnormal returns (CAR) for a three-, five- and thirteen-day event window are estimated (Jeppsson, 2010).

We test hypotheses 2, 3 and 4 by performing pair-wise analysis of the differences in mean. The first test examines the stock market's reaction (with respect to price and volume) to early-stage (Phase I and Phase I/II) and late-stage (Phase III and Phase II/III) R&D announcements. The second test analyzes the stock market's reaction (with respect to price and volume) to positive and negative R&D announcements. The Games-Howell pair-wise comparison test is run in the statistical software SPSS.

5.3 Cross-sectional regression

In addition to testing the hypotheses we are also interested in explaining the cross-sectional variations in abnormal return. As suggested by Joos (2003), more recent papers (e.g., Callen et al., 2010; Guo et al., 2005; Jeppsson, 2010; Xu et al., 2007) have collected a rich set of micro level data in the hope of it providing more insight into how R&D contributes to the value of a biotechnology firm. While there is no clear indication of which variables should be included there are some variables that are recurring. For example, different measures of success probability, collaboration, drug portfolio diversification and patents are used extensively in the existing literature. Although, patent data has been found to be value-relevant non-financial information (e.g., Callen et al., 2010) there are problems with simply counting the number of patents since patent applications usually are filed in the pre-clinical stage when the drug is far from FDA approval (Xu et al., 2007). Given that only 5 in 5000 compounds make it to human clinical trials and, at most, 1 of those 5 is approved, patents reflect R&D output with great uncertainty (Dedman et al., 2008). For comparability with previous literature we in this study use a regression model with the same three project-specific variables (complexity, risk sharing and investment) and two out of three firm-specific

variables (diversification and region) used by Jeppsson (2010).[6] We run all regressions in the statistical software Stata.

$$AR_i = \alpha_0 + \alpha_1 Complexity_i + \alpha_2 Risk\_sharing_i + \alpha_3 Investment_i + \alpha_4 Diversification_i + \sum_{j=1}^{6} \beta_j Region_i + e_i \quad (17)$$

In what follows, we will explain each of the variables in the above regression equation, starting with the dependent variable, Abnormal return (AR).

Abnormal return (AR)

We run three regression models with the following dependent variables: (i) AR on day zero for all positive R&D announcements (excluding Phase I/II and Phase II/III), (ii) AR on day zero for positive Phase I announcements, and (iii) AR on day zero for positive Phase II announcements.

The following variables are all the independent variables in the above regression model.

Complexity

The historical success rate per therapy area is used as a proxy for the complexity of the research project. The therapy areas and their historical success rates are based on the findings of the Biotechnology Innovation Organization (2016). The grounds for including complexity in the regression is that positive news from clinical trials with low success rates are expected to have a larger stock market reaction. We predict that *Complexity* will have a negative sign. The more likely the project is to succeed (i.e., the larger the value of *Complexity*), the

[6] Jeppsson (2010) also used the independent variable market-to-book value of equity (MTB) which is measured as the market value of equity divided by the book value of equity. The reason we do not include this variable in our regression model is that we do not believe investors in the biotechnology industry consider MTB value relevant. This assumption is supported by Jeppsson finding very low coefficients for MTB in all models. For example, the coefficient in model (i) is only 0.002 (t-statistic 0.41).

smaller the market reaction to positive news. See Appendix 2 for detailed information on the historical success rates.

Risk_sharing

We include the dummy variable *Risk_sharing* which shows if a project is developed in collaboration with a partner. Because it is costly to develop a drug many biotechnology companies seek to collaborate with an experienced partner for R&D and/or marketing purposes. Given that collaboration agreements typically are very costly, these will only be undertaken when there is a strong belief in the project (Callen et al. 2010). We predict that *Risk_sharing* will have a negative sign. If a project is developed in collaboration with a partner the firm's risk is lowered, but so is the profit to the firm that now will be shared with the partner. Therefore, the market reaction to positive news should be smaller if the project is developed with a partner.

Investment

This variable is the logarithmic value of the number of patients. The number of patients enrolled in clinical trials typically increases with each phase. The number, however, differs a lot between firms and depends on how much the firm invests in the project. Therefore, this variable signals the firm's belief in the project. We predict that *Investment* will have a positive sign. The more patients that have been studied in a phase the more reliable the results and thereby the more likely it is that the next phase will have the same positive outcome. Therefore, the market reaction should be larger when the positive news are from a larger study.

Diversification

Unlike Xu et al. (2007) we do not collect data on the number of drugs and diseases targeted to measure the companies drug portfolio diversification. Instead we use a less sophisticated but more easily attainable proxy, namely the logarithmic market value of equity from day -24 to day -5, relative to the R&D announcement. One could argue that this variable measures size and not diversification, but as a company with fewer or no approved products will

have a smaller product portfolio and therefore generally are smaller in size (Dedman et al. 2008) we believe it to be an appropriate proxy. We predict that $Diversification$ will have a negative sign. The market should react less strongly to a firm making an announcement of the progress of one project if the firm has a large project portfolio (i.e., is well diversified).

Region

Following Jeppsson (2010), region dummies are included in order to control for the institutional characteristics between countries and biotech clusters in the US. California is used as the benchmark region relative to New England, New York, Other US regions, Europe, Asia and Canada. See Appendix 3 for the total number of announcements made by companies in each of these regions.

Panel A of Table 2 reports the summary statistics of the independent variables (Phase I/II and Phase II/III excluded). The average success rate in our sample is 0.50. Only 18.3% of the clinical trials that reported results during the sample period were conducted in collaboration with a partner. The average number of patients enrolled was 219. However, as can be seen, the number of patients varies a lot. The largest clinical trial enrolled 3285 patients and the smallest only 3 patients. The average market capitalization during the days leading up to the announcement is 6090 million USD. Here too the size differs substantially. The smallest firm's market capitalization before the R&D announcement was just above 27 million USD and the largest had a market capitalization of 129 billion USD. Investment and diversification are both just the logarithmic values of the number of patients and market capitalization respectively, and hence are not interpreted here. Panel B contains a correlation matrix which shows that the pair-wise correlation between the independent variables are not high enough to cause any problem with multicollinearity.

Table 2. Descriptive statistics

Panel A. Summary statistics

	Number of observations	Mean	Min	Median	Max	St. dev
Complexity	557	0.501	0.237	0.566	0.848	0.170
Risk_sharing	557	0.183	0	0	1	0.387
Number of patients	557	218.8	3	91	3285	364.8
Investment	557	4.557	1.099	4.511	8.097	1.305
Market capitalization	557	6090	27.20	1170	129000	16400
Diversification	557	21.04	17.12	20.88	25.58	1.594

Note: This table shows desriptive statistics of the independent variables. *Complexity*, *Risk_sharing* and *Investment* are project-specific variables included in the regression, and *Diversification* is a firm-specific variable. *Complexity* is the historical success rate per therapy area. *Risk_sharing* is a dummy which equals 1 if the project is developed with a partner, and zero otherwise. *Number of patients* is the number of patients in the clinical trial and *Investment* is the logarithmic value of the number of patients. *Market capitalization* is the average market value of equity (in million USD), from day -24 to day -5 relative to the R&D announcement. *Diversification* is the logarithmic value of *Market capitalization*. For comparability we only include observations for drugs in phase I, II and III (phase I/II and II/III are excluded).

Panel B. Correlation matrix

	Complexity	Risk_sharing	Investment	Diversification
Complexity	1			
Risk_sharing	0.0760*	1		
Investment	0.0504	0.1435**	1	
Diversification	-0.1007**	0.3048***	0.2700***	1

Note: This table shows the pair-wise correlations and the 1%, 5% and 10% significance levels, denoted by ***, ** and *.
All variables are detailed in Table 2, panel A.

6 Results

6.1 Stock market reaction to R&D announcements

Table 3 shows the mean abnormal return, mean cumulative abnormal return, mean abnormal volume, and, the mean cumulative abnormal volume for R&D announcements. From Table 3 we can see that the price reaction results (both on day zero and cumulative) for Phase I, Phase II and Phase III give strong support for our first hypothesis (i.e., that the stock market reacts positively to all positive R&D announcements on day zero, and negatively to all negative R&D announcements). For combined phases the price reaction to positive announcements is consistent with Hypothesis 1, however, the results are statistically not significant.[7] Note that whenever we use the word consistent we refer to the sign, not significance. In this case the mean price reaction to positive Phase I/II announcements is 0.82 which, like Hypothesis 1 predicts, is positive, but the result is statistically not significant and hence it offers no support for the hypothesis.

Moreover, Table 3 shows that on day zero all volume reactions to positive and negative announcements are significantly positive except for negative Phase I. One possible explanation for negative Phase I announcements not being statistically significant is the small number of negative Phase I announcements in our sample. Further, we observe that the mean cumulative abnormal volume over the three-day event window is lower than the mean abnormal volume on day zero (by 5% or more), for all positive and negative R&D announcements.

The market reaction to clinical trial initiation is ambiguous. For initiations of Phase I and Phase III we have positive and significant price reactions (0.74%, t-statistic 2.46 and 0.79%, t-statistic 2.74, respectively), while Phase II, Phase I/II and Phase II/III seem to have negative, although not significant, price reactions.[8] Wrong event dates and the fact that phase

[7] There are no negative combined-phase announcements in our sample.

[8] Jeppsson (2010) found significant reactions for initiations of Phase I and Phase II clinical trials, but not for Phase III.

initiations are not always good news, because firms might start the clinical trial for a more limited set of indications than initially thought, are two possible explanations for this ambiguity (Joos, 2003). From Table 3 we can also see that the volume reactions to initiations are positive and significant for all but Phase I (positive but not significant) and Phase II/III (negative and not significant).

Table 3. Mean abnormal return, mean cumulative abnormal return, mean abnormal volume, and, mean cumulative abnormal volume for R&D announcements

Stage of R&D process	Event	Abnormal return (%)			Abnormal volume (%)		
		n	$\overline{AR_0}$ *(t value)*	$\overline{CAR_{-1+1}}$ *(t value)*	n	$\overline{AV_0}$ *(t value)*	$\overline{CAV_{-1+1}}$ *(t value)*
Phase I	Initiation	73	0.74**	0.55	73	0.22	-4.54
			(2.46)	*(0.93)*		*(0.04)*	*(-0.62)*
	Results (Positive)	127	2.53***	3.29***	127	47.52***	34.41***
			(2.95)	*(3.08)*		*(6.11)*	*(4.68)*
	Results (Negative)	3	-20.18***	-23.84**	3	65.19	59.14
			(-2.88)	*(-2.48)*		*(1.37)*	*(0.86)*
Phase II	Initiation	114	-0.94	-0.28	114	18.47***	13.89*
			(-0.86)	*(-0.20)*		*(2.92)*	*(1.88)*
	Results (Positive)	201	2.68***	2.97***	201	69.70***	48.45***
			(5.17)	*(3.4)*		*(10.3)*	*(6.77)*
	Results (Negative)	24	-15.28***	-18.55***	24	136.28***	100.86***
			(-3.73)	*(-4.51)*		*(5.94)*	*(5.76)*
Phase III	Initiation	108	0.79***	0.69	108	10.85**	10.26
			(2.74)	*(1.49)*		*(2.39)*	*(1.64)*
	Results (Positive)	172	4.83***	6.04***	172	88.32***	64.70***
			(4.27)	*(4.16)*		*(9.19)*	*(8.62)*
	Results (Negative)	31	-39.04***	-41.46***	31	200.50***	119.28***
			(-4.64)	*(-5.02)*		*(7.18)*	*(6.62)*
Phase I/II	Initiation	40	-0.03	1.17	40	20.75***	11.57
			(-0.05)	*(1.22)*		*(2.95)*	*(1.25)*
	Results (Positive)	44	0.82	0.87	44	38.97***	33.11***
			(0.72)	*(0.54)*		*(3.57)*	*(3.03)*
	Results (Negative)	0	N/A	N/A	0	N/A	N/A
			N/A	N/A		N/A	N/A
Phase II/III	Initiation	10	-1.1	-2.52	10	-4.5	1.1
			(-1.28)	*(-1.41)*		*(-0.65)*	*(0.06)*
	Results (Positive)	7	5.29*	4.91	7	160.78***	45.44
			(2.16)	*(1.31)*		*(3.69)*	*(1.52)*
	Results (Negative)	0	N/A	N/A	0	N/A	N/A
			N/A	N/A		N/A	N/A

Note: This table reports mean abnormal return and mean abnormal volume for day zero and for the event window (-1 to +1 day relative to the announcement day). Mean cumulative abnormal return is estimated as: $\overline{CAR}(t_1,t_2) = \sum_{t=t_1}^{t_2} \overline{AR_t}$. $\overline{AR_t} = \frac{1}{N}\sum_{i=1}^{N} AR_{it}$. AR_{it} is calculated using estimates from the market model: $AR_{it} = R_{it} - (\hat{\alpha}_i + \hat{\beta}_i R_{m_t})$. Mean cumulative abnormal volume is estimated as: $\overline{CAV}(t_1,t_2) = \sum_{t=t_1}^{t_2} \overline{AV_t}$. $AV_t = \frac{1}{N}\sum_{i=1}^{N} AV_{it}$. AV_{it} is calculated using estimates from the market model: $AV_{it} = v_{it} - (\hat{\alpha}_i + \hat{\beta}_i v_{mt})$ *the values are based on robust standard errors. ***, **, and * indicate statistical significance at the 1%, 5% and 10% levels, respectively.*

Table 3 reveals two interesting patterns. Price and volume react (1) more strongly to late-stage announcements than to early-stage announcements, and (2) more strongly to negative than to positive announcements. For example, the day zero mean abnormal return is 2.53% (-20.18%) for positive (negative) Phase I results, and 4.83% (-39.04%) for Phase III results. Similarly, the day zero mean abnormal volume is 47.52% (65.19%) for positive (negative) Phase I results, and 88.32% (200.50%) for Phase III results. The first pattern can also be observed for the combined phases, where the mean abnormal return (volume) on day zero is 0.82% (38.97%) for positive Phase I/II and 5.29% (160.78%) for positive Phase II/III announcements. These results are consistent with Hypotheses 2, 3 and 4 and are tested formally in the pair-wise analysis of the differences in mean.

We find strong market reactions to negative R&D announcements, which suggests that negative announcements are largely unanticipated by investors. The largest price reaction to a negative Phase III announcement, in our sample, led to an 86% decrease in market value on day zero for the firm (not tabulated). From Table 3 it can further be seen that the mean cumulative abnormal return over the three-day event window (days -1 to + 1 relative to the announcement day) is not substantially larger than the mean abnormal return on day zero. The difference in market reaction to positive announcements range between 0.5% and 1% across all phases, and between 2% and 4% for negative announcements.[9] If negative announcements are largely unanticipated by investors, then the small differences that we can see between mean abnormal return on day zero and mean cumulative abnormal return over the three-day event window should be caused by trading on the following day, rather than before the event day. To check this the mean cumulative abnormal return prior to the event was calculated (day -10 to day -1). The data (not tabulated) shows that the market's reaction to negative announcements across all phases are statistically not significant and hence investors do not seem to anticipate negative results.

[9]For example, the three-day mean cumulative abnormal return to negative Phase I results is -23.84% (t-statistic -2.48), while the day zero mean abnormal return is -20.18% (t-statistic -2.88).

The robustness tests (five- and thirteen-day event windows) in Table 4 show that the price reactions for Phase I and Phase II are persistent over longer event windows, while for Phase III only the price reactions to negative announcements are persistent. Table 4 also shows that Phase II and Phase III volume reactions are persistent over longer event windows, while Phase I volume reactions show no persistence. The combined phases, Phase I/II and Phase II/III show no persistence over longer time periods for neither price nor volume reactions.

Table 4. Mean cumulative abnormal return, and, mean cumulative abnormal volume for R&D announcements

Stage of R&D process	Event	Abnormal return (%)			Abnormal volume (%)		
		n	$\overline{CAR_{-2+2}}$ *(t value)*	$\overline{CAR_{-2+10}}$ *(t value)*	n	$\overline{CAV_{-2+2}}$ *(t value)*	$\overline{CAV_{-2+10}}$ *(t value)*
Phase I	Initiation	73	0.43	-1.55	73	-1.72	-5.21
			(0.36)	*(-1.00)*		*(-0.27)*	*(-0.77)*
	Results (Positive)	127	4.08***	3.34**	127	20.50***	1.51
			(3.51)	*(2.07)*		*(3.02)*	*(0.23)*
	Results (Negative)	3	-21.37**	-24.47***	3	17.85	-27.58
			(-2.06)	*(-2.79)*		*(0.49)*	*(-1.20)*
Phase II	Initiation	114	-0.7	-2.68	114	-4.53	-0.33
			(-0.46)	*(-1.61)*		*(-0.62)*	*(-0.04)*
	Results (Positive)	201	3.78***	3.21**	201	29.68***	29.68***
			(3.79)	*(2.35)*		*(5.02)*	*(5.02)*
	Results (Negative)	24	-78.76***	-18.22***	24	58.81***	46.30***
			(-4.25)	*(-4.05)*		*(4.99)*	*(3.02)*
Phase III	Initiation	108	0.7	-0.58	108	0.23	5.93
			(1.31)	*(-0.51)*		*(0.04)*	*(0.85)*
	Results (Positive)	172	4.71***	2.33	172	44.78***	9.69*
			(3.21)	*(1.44)*		*(7.18)*	*(1.87)*
	Results (Negative)	31	-44.25***	-42.39***	31	87.03***	33.79**
			(-5.04)	*(-4.83)*		*(5.87)*	*(2.61)*
Phase I/II	Initiation	40	2.37**	2.82*	40	-14.66	-35.03***
			(2.22)	*(1.69)*		*(-1.21)*	*(-2.99)*
	Results (Positive)	44	0.73	2.31	44	24.65**	8.74
			(0.4)	*(0.71)*		*(2.55)*	*(0.68)*
	Results (Negative)	0	N/A	N/A	0	N/A	N/A
			N/A	N/A		N/A	N/A
Phase II/III	Initiation	10	-1.32	2.13	10	-31.81**	-34.36
			(-0.60)	*(0.99)*		*(-2.25)*	*(-1.49)*
	Results (Positive)	7	2.81	-1.74	7	25.06	13.53
			(0.97)	*(-0.27)*		*(1.54)*	*(0.53)*
	Results (Negative)	0	N/A	N/A	0	N/A	N/A
			N/A	N/A		N/A	N/A

Note: This table reports mean abnormal return and mean abnormal volume for two different windows (-2 to +2 day, and -2 to +10 day relative to the announcement day, respectively). Mean cumulative abnormal return is estimated as: $\overline{CAR}(t_1,t_2) = \sum_{t=t_1}^{t_2} \overline{AR_t}$. $\overline{AR_t} = \frac{1}{N}\sum_{i=1}^{N} AR_{it}$. AR_{it} is calculated using estimates from the market model $AR_{it} = R_{it} - (\hat{\alpha}_i + \hat{\beta}_i R_{m_t})$. Mean cumulative abnormal volume is estimated as: $\overline{CAV}(t_1,t_2) = \sum_{t=t_1}^{t_2} \overline{AV_t}$. $AV_t = \frac{1}{N}\sum_{i=1}^{N} AV_{it}$. AV_{it} is calculated using estimates from the market model $AV_{it} = v_{it} - (\hat{\alpha}_i + \hat{\beta}_i v_{mt})$, the *t* values are based on robust standard errors. ***, **, and * indicate statistical significance at the 1%, 5% and 10% levels, respectively.

6.2 Pair-wise comparison

Now we turn to look at the difference in share price reaction to early- versus late-stage and positive versus negative R&D announcements (i.e., Hypotheses 2 and 3). In general, the results in Panel A of Table 5 support both hypotheses. Panel A of Table 5 shows that the differences in mean abnormal return is 2.3% between positive early-stage and positive late-stage R&D announcements, and 18.86% for negative R&D announcements. Although we only find the market reaction between negative R&D announcements to be statistically significant, these results support Hypothesis 2 and are consistent for both positive and negative R&D announcements. Further, Panel A of Table 5 shows that the differences in the mean abnormal returns between positive and negative R&D announcements in the early stage are 17.65% and 34.21% in the late stage, both significant at the 1% confidence level. These results offer support for Hypothesis 3, for both early- and late-stage R&D announcements.

Panel B of Table 5 presents the pair-wise analysis of the differences in the mean abnormal volume for the early-stage and late-stage R&D announcements and for positive and negative R&D announcements (i.e., Hypothesis 4). Results show that there is a significant difference between early- and late-stage mean abnormal volumes both for positive and negative R&D announcements. Results further show that there is a significant difference between positive and negative late-stage R&D announcements (112.18%). Thus, the volume tests offer support for Hypothesis 4.

Table 5. Pair-wise comparison

Panel A. Abnormal return (%)				*Panel B. Abnormal volume (%)*			
	Positive R&D news	**Negative R&D news**	**Difference**		**Positive R&D news**	**Negative R&D news**	**Difference**
Early-stage (Phase I)	2.53*** (2.95)	20.18*** (2.88)	17.65***	**Early-stage (Phase I)**	47.52*** (6.11)	65.19 (1.37)	17.52
Late-stage (Phase III)	4.83*** (4.27)	39.04*** (4.64)	34.21***	**Late-stage (Phase III)**	88.32*** (9.19)	200.50*** (7.18)	112.18***
Difference	2.30	18.86**		**Difference**	40.80***	135.31**	

Note: This table reports pair-wise analysis in differences in mean abnormal return (Panel A) and mean abnormal volume (Panel B) for early-stage and late-stage R&D announcements, and, for positive and negative R&D announcements. Negative R&D news are reported in absolute values. Reported t-values are the results of the Games-Howell pair-wise comparison test with unequal variances. ***, **, and *, denote one tail 1%, 5% and 10% significance, respectively.

The pair-wise comparison of the mean abnormal returns and volumes for the combined phases are presented in Table 6. We do not compare for negative announcements because there are no negative announcements for combined phases in our sample. From Panel A of Table 6 we see that abnormal returns are higher for late-stage R&D announcements which is consistent with Hypothesis 2. The result, however, is not statistically significant. In Panel B of Table 6 we see that abnormal volumes are higher for late-stage R&D announcements which is consistent with Hypothesis 4. This result is significant on the 1% level, and thus, the volume test offers support for Hypothesis 4.

Table 6. Pair-wise comparison for the combined phases

Panel A. Abnormal return (%)				*Panel B. Abnormal volume (%)*			
	Positive R&D news	**Negative R&D news**	**Difference**		**Positive R&D news**	**Negative R&D news**	**Difference**
Early-stage (Phase I/II)	0.82 (0.72)	N/A N/A		**Early-stage (Phase I/II)**	38.97*** (3.57)	N/A N/A	
Late-stage (Phase II/III)	5.29* (2.16)	N/A N/A		**Late-stage (Phase II/III)**	160.78*** (3.69)	N/A N/A	
Difference	4.47			**Difference**	121.81***		

Note: This table reports pair-wise analysis in differences in mean abnormal return (Panel A) and mean abnormal volume (Panel B) for early-stage and late-stage R&D announcements, and, for positive and negative R&D announcements. Negative R&D news are reported in absolute values. Reported t-values are the results of the Games-Howell pair-wise comparison test with unequal variances. ***, **, and *, denote one tail 1%, 5% and 10% significance, respectively.

6.3 Cross-section regression results

We examine the cross-sectional variation in abnormal return on day zero, using the regression model shown in Table 7. The dependent variables are (i) AR on day zero for all positive R&D announcements (excluding Phase I/II and Phase II/III), (ii) AR on day zero for positive Phase I announcements, and (iii) AR on day zero for positive Phase II announcements.[10] Most of the regression results are in line with the predictions made in Section 5.3 for each of

[10]Table 8 in Appendix 4 reports the regression results for negative R&D announcements, which are in line with the results found for positive announcements.

the independent variables. Model (i) reports a positive and significant effect of $Investment$ (t-statistic 2.91) and a negative and significant effect of $Diversification$ (t-statistic -3.55).

When testing the different phases separately we find that no independent variable seems to have a significant effect on the abnormal return on day zero for Phase I announcements. However, model (iii) suggests a negative and significant effect of both $Diversification$ and $Risk\_sharing$ and a positive and significant effect of $Investment$, for Phase II announcements. The implications of these results are (1) the larger the investment (number of patients) in the clinical trial, the larger the market reaction, (2) the larger and more diversified the firm is, the smaller the market reaction, and (3) when there is a collaboration agreement (risk sharing), the market reaction is smaller.

Table 7. Cross-sectional regression results

$\text{AR}_i = \alpha_0 + \alpha_1 Complexity_i + \alpha_2 Risk\_sharing_i + \alpha_3 Investment_i + \alpha_4 Diversification_i + \sum_{j=1}^{6} \beta_j Region_i + e_i$

	Predicted sign	(i) All positive R&D announcements (n=499)	(ii) Positive phase I announcements (n=126)	(iii) Positive phase II announcements (n=201)	Positive phase I/II announcements (n=43)	Positive phase II/III announcements (n=7)
Intercept		0.196***	0.270	0.123	-0.394	N/A
		(3.44)	(1.33)	(1.75)	(-1.42)	
Complexity	-	0.013	-0.133	0.010	0.176	N/A
		(0.54)	(-1.05)	(0.15)	(0.88)	
Risk_sharing	-	-0.004	-0.017	-0.024*	0.009	N/A
		(-0.37)	(-1.17)	(-0.75)	(0.19)	
Investment	+	0.012***	-0.002	0.011*	-0.042*	N/A
		(2.91)	(-0.16)	(0.83)	(-1.97)	
Diversification	-	-0.012***	-0.008	-0.007**	0.021*	N/A
		(-3.55)	(-1.12)	(-2.06)	(1.78)	
Regional dummies	+/-	Yes	Yes	Yes	Yes	N/A

Note: This table reports the linear regression outputs. The dependent variable in the different models is (i) AR on day zero for all positive R&D announcements[1], (ii) AR on day zero for positive Phase I announcements, (iii) AR on day zero for positive Phase II announcements. *Complexity*, *Risk_sharing* and *Investment* are project-specific variables included in the regression, and *Diversification* is a firm-specific variable. *Complexity* is the historical success rate per therapy area. *Risk_sharing* is a dummy which equals 1 if the project is developed with a partner, and zero otherwise. *Investment* is the logarithmic value of the number of patients and *Diversification* is the logarithmic market value of equity from day -24 to day -5, relative to the R&D announcement. All regressions have robust standard errors.

[1]For comparability we only include observations for drugs in phase I, II and III (phase I/II and II/III are excluded) in model (i).

7 Discussion

In this section we compare the results found in this study to those found by Jeppsson (2010). The reason for comparing our results to those of Jeppsson is that we have the same research questions, but study different markets. While Jeppsson used a sample of European companies we study mainly US biotechnology firms.[11]

In general, our results are in line with Jeppsson's. First, when studying the mean abnormal returns, we similarly to Jeppsson found support for Hypothesis 1 (i.e., the market reacts positively to positive R&D announcements and negatively to negative R&D announcements). Second, the results of the pair-wise analysis of differences in the mean abnormal returns, Panel A of Table 5, are similar to what Jeppsson found (i.e., support both Hypotheses 2 and 3).[12] Also, the volume test results in Panel B of Table 5 are consistent with Jeppsson's findings. However, unlike Jeppsson, we found statistically significant differences in mean abnormal volume both between early- and late-stage R&D announcements, and between positive and negative R&D announcements. These results are in line with the positive relationship between stock price changes and trading volumes documented in the existing literature (e.g., Karpoff, 1987). Thus, our study contributes empirical evidence of this price-volume relationship in the biotechnology industry.

Looking at the cross-sectional regression results we find that the sign of the *Complexity* coefficient for Phase I is in line with the prediction (made in section 5.3). However, *Complexity* is not found to be significant in any of the models, which is inconsistent with what Jeppsson (2010) found. One possible explanation is that the success rates we use are not only based on projects that were announced to the market, but also on projects that might not have had any announcement reporting the outcome. Hence, our sample may not be representative of the whole population (as we sample announcements and not projects). So, even though the coefficient for *Complexity* is statistically not significant we cannot discard it from the model.

[11] See Appendix 3 for details on the regions.

[12] As mentioned in the literature review other studies have also found support for both Hypothesis 2 (e.g., Dedman et al., 2008; McConomy and Xu, 2004) and Hypothesis 3 (e.g., Xu, 2009).

In summary, the results in this study are largely in line with the results found by Jeppsson. Thus, the tested hypotheses are supported by data for biotechnology firms in both the US and European markets and over time.

8 Conclusion

In this paper we studied the value relevance of non-financial information in the US biotechnology industry. We did this by comparing the size of the market reaction to R&D announcements made in different phases of the drug development process and between positive and negative R&D announcements. We found that for individual-phase trials 1) stock price reacts positively to all positive R&D announcements on day zero, and negatively to all negative R&D announcements 2) the stock market (with respect to both price and volume) reacts more strongly to late-stage R&D announcements than to early-stage R&D announcements, and 3) the stock market (with respect to both price and volume) reacts more strongly to negative than to positive R&D announcements.

Due to the large market reactions to negative news, we expect firm managers to be more reluctant to publish negative R&D announcements, which is consistent with existing findings. For example, Dedman et al. (2008) found that firms fail to release negative news and Fisher (2002) claims that the tone in R&D announcements in many cases is overly optimistic.[13] Disclosure behaviour is out of the scope of this study. However, the results suggest a need for enhanced monitoring of disclosures of non-financial information in the biotech industry.[14]

Further, since we find this large reaction to negative announcements it might indicate that we have an over-representation of positive announcements in our sample. This limitation is common in studies sampling biotechnology firms' R&D announcements (e.g., Dedman et al.,

[13]We also found that some firms had a tendency to disclose negative news together with other more positive news, maybe in order to limit their price impact.

[14]See Jeppsson and Hamberg (2010) for more information about disclosure behaviour in the biotechnology industry.

2008; Ely et al., 2003; Jeppsson, 2010). Another limitation of this study is the small sample size for combined-phase announcements. As described in this paper, combined phases were created to accelerate the drug approval process of more complex or riskier types of therapy areas (e.g., oncology).[15] The limited sample size could explain why our results do not provide any strong evidence of the impact of combined-phase announcements on stock prices. Further investigation of the market's reaction to combined-phase (i.e., Phase I/II and Phase II/III) R&D announcements is necessary in order for any reliable conclusion to be drawn about their value relevance.

[15]We find no bias towards oncology in our combined-phase announcement sample.

Appendix

Appendix 1: Examples of R&D announcements

Example of positive R&D news announcement

ACADIA Announces Presentation of Data from Its Pivotal Phase III Parkinson's Disease Psychosis Study with Pimavanserin at the American Academy of Neurology Annual Meeting

SAN DIEGO--(BUSINESS WIRE)--Mar. 20, 2013-- ACADIA Pharmaceuticals Inc. (NASDAQ: ACAD), a biopharmaceutical company focused on innovative treatments that address unmet medical needs in neurological and related central nervous system disorders, announced that Jeffrey Cummings, M.D., Sc.D., Director of Cleveland Clinic Lou Ruvo Center for Brain Health, presented detailed results today from ACADIA's pivotal Phase III -020 Study with pimavanserin in patients with Parkinson's disease psychosis at the Emerging Science session of the 65thAmerican Academy of Neurology (AAN) Annual Meeting. The analysis of the full data set from the Phase III -020 Study showed robust and consistent efficacy of pimavanserin across a wide array of study measures and confirmed the positive top-line results previously reported.

Example of negative R&D news announcement

ACADIA Pharmaceuticals Announces Top-line Results from Phase 3 ENHANCE Trial of Pimavanserin as Adjunctive Treatment for Patients with Schizophrenia

- Pimavanserin did not achieve statistical significance on the primary endpoint, but showed a consistent trend in improvement of psychotic symptoms (p=0.0940)
- Significant improvements observed on secondary endpoint of PANSS negative symptoms scale sub-score (unadjusted p=0.0474)
- Conference call and webcast to be held today at 5:00 p.m. Eastern Time

SAN DIEGO--(BUSINESS WIRE)--Jul. 22, 2019-- ACADIA Pharmaceuticals Inc. (Nasdaq: ACAD), today announced top-line results from its Phase 3 ENHANCE study, which evaluated pimavanserin as an adjunctive treatment in adult schizophrenia patients with persistent inadequate response to their current antipsychotic therapy. A total of 396 patients with moderate-to-severe psychotic symptoms were randomized to receive either pimavanserin or placebo added to their current antipsychotic treatment. There is currently no FDA-approved adjunctive treatment for schizophrenia patients with inadequate response to existing therapies.

Appendix 2: Phase-specific success rates by therapy area

Therapy Area	Phase I to II	Phase II to III	Phase III to NDA/BLA
Hematology	0.733	0.566	0.75
Infectious Disease	0.695	0.427	0.727
Opthalmology	0.848	0.446	0.583
Other	0.667	0.397	0.696
Metabolic	0.611	0.452	0.714
Gastroenterology	0.756	0.357	0.606
Allergy	0.676	0.325	0.714
Endocrine	0.589	0.401	0.65
Respiratory	0.653	0.291	0.711
Urology	0.571	0.327	0.714
Autoimmune	0.657	0.317	0.622
Neurology	0.591	0.297	0.574
Cardiovascular	0.589	0.241	0.555
Psychiatry	0.539	0.237	0.557
Oncology	0.628	0.246	0.401

Notes: Phase success rates taken from Biotechnology Innovation Organization (2016)

Appendix 3: Regions

Region clusters	N
California	330
New York	105
New England	225
Other	179
Asia	16
Europe	94
Canada	5
Total	954

Note: Asia, Europe and Canada include companies from outside the United States with HQ in US, listed in Nasdaq and that are part of the NBI.

Appendix 4: Cross-sectional regression results

Table 8. Cross-sectional regression results

$$\mathrm{AR}_i = \alpha_0 + \alpha_1 Complexity_i + \alpha_2 Risk\_sharing_i + \alpha_3 Investment_i + \alpha_4 Diversification_i + \sum_{j=1}^{6} \beta_j Region_i + e_i$$

	Predicted sign	(iv) All negative R&D announcements (n=58)	(v) Negative phase I announcements (n=3)	(vi) Negative phase II announcements (n=23)
Intercept		-1.179***	N/A	-0.827
		(-3.19)		(-1.13)
Complexity	-	-0.404	N/A	-0.168
		(-1.27)		(-0.23)
Risk_sharing	+	-0.038	N/A	0.079
		(-0.36)		(0.53)
Investment	-	-0.142**	N/A	-0.082
		(-2.21)		(-1.02)
Diversification	+	0.090***	N/A	0.049
		(3.40)		(1.39)
Regional dummies	+/-	Yes	N/A	Yes

Note: This table reports the linear regression outputs. The dependent variable in the different models is (iv) AR on day zero for all negative R&D announcements[1], (v) AR on day zero for negative Phase I announcements, (vi) AR on day zero for negative Phase II announcements. *Complexity*, *Risk_sharing* and *Investment* are project-specific variables included in the regression, and *Diversification* is a firm-specific variable. *Complexity* is the historical success rate per therapy area. *Risk_sharing* is a dummy which equals 1 if the project is developed with a partner, and zero otherwise. *Investment* is the logarithmic value of the number of patients and *Diversification* is the logarithmic market value of equity from day -24 to day -5, relative to the R&D announcement. All regressions have robust standard errors.

[1]There are no negative R&D announcments for the combined phases (i.e., Phase I/II and Phase II/III).

References

Ajinkya, B. B., & Jain, P. C. (1989). The behavior of daily stock market trading volume. Journal of Accounting and Economics, 11, 331-359.

Amir, E., & Lev, B. (1996). Value-relevance of nonfinancial information: The wireless communications industry. Journal of Accounting and Economics, 22, 3–30.

Biotechnology Innovation Organization. (2016). Clinical Development Success Rates 2015-2016.

Botosan, C.A., (1997). Disclosure level and the cost of capital. The Accounting Review, 72, 323-349.

Callen, J. L., Gavious, I., & Segal, D. (2010). The complementary relationship between financial and nonfinancial informaion in the biotechnology industry and the degree of investor sophistication. Journal of Contemporary Accounting & Economics, 6, 61-76.

Dedman, E., Lin, S. W-J., Prakash, A. J., & Chang, C-H. (2008). Voluntary disclosure and its impact on share prices: Evidence from the UK biotechnology sector. Journal of Accounting and Public Policy, 27, 195-216.

DiMasi, J.A. (2001). New drug development in the United States from 1963 to 1999. Clinical Pharmacology and Therapeutics, 69(5), 286–296.

DiMasi, J. (2001). Risks in new drug development: Approval success rates for investigational drugs. Clinical Pharmacology & Therapeutics, 69, 297-307.

Ely, K., Simko, P. J., & Thomas, L. G. (2003). The usefulness of biotechnology firms' drug development status in the evaluation of research and development costs. Journal of Accounting, Auditing and Finance, 18, 163–196.

Ernst & Young. (2017). Biotechnology Report 2017: Beyond Borders - Staying the course.

Fisher, W. O. (2002). Key disclosure issues for life sciences companies: FDA product approval, clinical test results, and government inspections. Michigan Telecommunications and Technology Law Review, 8, 115-193.

Guo, R., Lev, B., & Zhou, N. (2004). Competitive costs of disclosure by biotech IPOs. Journal of Accounting

Research, 42, 319–355.

Guo, R., Lev, B., & Zhou, N. (2005). The valuation of biotech IPOs. Journal of Accounting Auditing and Finance 20(4), 423–459.

Hand, J.R.M. (2005). The value relevance of financial statements in the venture capital market. Accounting Review, 80, 613–648.

Hirschey, M., Richardson, V. J., & Scholz, S. (2001). The Relevance of Nonfinancial Information: The Case of Patent Data. Review of Qualitative Finance and Accounting, 17, 223-235.

Jeppsson, H. (2010). The value relevance of research progress in the European biotechnology industry. Working paper.

Jeppsson, H., & Hamberg, M. (2010). Market Timing and Equity Financing Decisions. Working paper.

Joos, P. (2002). Explaining Cross-sectional Differences in Market-to-Book Ratios in the Pharmaceutical Industry. Working Paper.

Joos, P. (2003). Discussion – The usefulness of biotechnology firms: Drug development status in the evaluation of research and development costs. Journal of Accounting, Auditing and Finance, 18, 197-205.

Karpoff, J. M. (1987). The Relation between Price Changes and Trading Volume: A Survey. Journal of financial and quantitative analysis, 22(1), 109-124.

MacKinlay, C. A. (1997). Event studies in economics and finance. Journal of Economic Literature, 35, 13-39.
Matilla-García, M., Ruiz Marín, M., & Dore, M. I. (2013). A permutation entropy based test for causality: The volume-stock price relation. Physica A, 398, 280-288.

McConomy, B., & Xu, B. (2004). Value creation in the biotechnology industry. CMA Management, 29-31.

Nasdaq. (2017). Nasdaq Leading the Way in Biotech.

National Cancer Institute. (2020). NCI Dictionary of cancer terms. Available at:
https://www.cancer.gov/publications/dictionaries/cancer-terms/def/phase-i-ii-clinical-trial. [2020-05-27].

Osborne, M. F. M. (1959). Brownian Motion in the Stock Market. Operations Research, 7, 145-173.

U.S. Food and Drug Administration. (2018). The drug development process. Step 3: Clinical Research. Available at: https://www.fda.gov/patients/drug-development-process/step-3-clinical-research. [2020-05-27].

Xu, B. (2009). Investment success and the value of investment opportunities: evidence from the biotech industry. Applied Financial Economics, 19, 527-537.

Xu, B., Magnan, M., & Andre, P. (2007). The stock market valuation of RD information in biotech firms. Contemporary Accounting Research, 24, 1291–1318.

Yang, Y. (2008). The value-relevance of nonfincial information: the biotechnology industry. Advances in Accounting, 23, 287-314.

Ying, C.C. (1966). Stock Market Prices and Volumes of Sales. Econometrica, 34, 676-686.