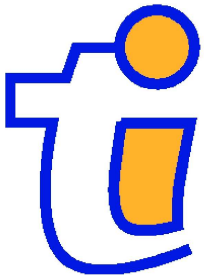


TI 2008-022/3
Tinbergen Institute Discussion Paper

Pharmaceutical Research Strategies

Sandra Phlippen[1]
An Vermeersch[2]

[1] *Dept. of Applied Economics, Erasmus University Rotterdam, and Tinbergen Institute;*
[2] *McKinsey and Company.*

Tinbergen Institute
The Tinbergen Institute is the institute for economic research of the Erasmus Universiteit Rotterdam, Universiteit van Amsterdam, and Vrije Universiteit Amsterdam.

Tinbergen Institute Amsterdam
Roetersstraat 31
1018 WB Amsterdam
The Netherlands
Tel.: +31(0)20 551 3500
Fax: +31(0)20 551 3555

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900
Fax: +31(0)10 408 9031

Most TI discussion papers can be downloaded at http://www.tinbergen.nl.

Complementarities and the R&D boundaries of the firm: a project level study on pharmaceutical R&D strategies.

Sandra Phlippen[1], An Vermeersch[2]

[1]Phlippen@few.eur.nl
[2]An_Vermeersch@mckinsey.com

1. Introduction

A firm's ability to innovate is increasingly the result of both internal R&D efforts and external knowledge sourcing (Gambardella, 1992; Freeman, 1991). External knowledge sourcing can be performed through informal personal interactions, formal collaborations, spin-out (and later spin-in) companies and consultancy or through job mobility (Abramovsky et al, 2007). Especially in the biopharmaceutical industry the complementarities between internal R&D and external sourcing through formal collaborations play an important role in large pharmaceutical innovation strategies (Arora & Gambardella, 1990; Henderson & Cockburn, 1996; Pisano, 1990).

Until recently, externally sourcing of R&D has been considered a substitute for in-house R&D activities, i.e. R&D has been perceived as either a make-, or a buy decision of the firm. Exemplifying in this respect is the seminal work of Pisano (1990) on the R&D boundaries of the firm. Although the author acknowledges complementarity between in-house R&D activities to be important, potential complementarities beyond the boundaries of the firm are ignored (Pisano, 1990).

Our study fills an important gap in the literature. While there exists a rich theoretical literature on complementarities between R&D activities, limited data availability has so far constrained empirical testing of some key features of these theories. Using unique data on more than 1300 early stage research projects of large pharmaceutical firms, we empirically study two important issues from the literature.

First, we investigate conditions under which in-house and external research are complementary to one another. Complementarity between two activities is defined[1] as "Adding an activity while the other activity is already performed has a higher incremental effect on performance than adding the activity in isolation" (Cassiman & Veugelers, 2006 pp. 70).[2] While the above definition of complementarity is elegantly simple, it does not provide any direction about when to expect complementarity and how these activities come about to be complementary. We therefore turn to Cohen and Levinthal's notion of absorptive capacity, which is similar to the notion of complementarity. The authors state that firms with a sufficient stock of relevant in-house R&D are better able to achieve complementarities from combining internal with external R&D. Internal know-how is used to effectively screen and absorb external knowledge and to exploit these findings internally (Cohen and Levinthal, 1989). So far, existing empirical work has not provided any answer as to what defines which knowledge is 'relevant' internal knowledge and how much

[1] Complementarity between two activities A_1 and A_2 arises only if $\prod(1,1) - \prod(0,1) \geq \prod(1,-) - \prod(0,0)$, whereby $\prod(A_1, A_2)$ represents performance, and each activity A either takes place (1) or does not (0).

[2] Milgrom and Roberts (1990) have first coined the term complementarity to describe synergies among organizational practices within the firm. In the strategy literature the concept of complementarity is better known as 'strategic fit' (Porter, 1980). Strategic fit is defined as: 'the degree to which the needs, demands, goals, objectives, and/or structures of one component are consistent with the needs, demands, goals, objectives, and/or structures of another component' (Nadler and Tushman, 1980: 36)

knowledge is a 'sufficient' stock of knowledge. Our results indicate which types of pharmaceutical R&D knowledge are relevant to achieve complementarities and, more importantly, we indicate a critical mass of prior R&D that is necessary for complementarities to occur.

Second, we investigate in detail how these complementarities occur. According to the theory of absorptive capacity (Cohen & levinthal, 1989; van den Bosch et al, 1999) there is a two-way knowledge flow between internal know-how and external knowledge that underlies the relationship between complementarities and performance. Knowledge flows or spillovers between activities cause learning effects and subsequently increase marginal returns on firm performance[3]. On the one hand, internal know-how is claimed to increase the marginal return to external sourcing through an increased ability to effectively screen and contribute to external projects knowledge (Lane & Lubatkin, 1998). On the other hand, external knowledge that has been absorbed needs to return into the organizations internal knowledge base in order to truly increase internal innovations. We will refer to the former as knowledge outflow and to the latter as knowledge inflow. We test both knowledge flows and their effect on performance directly. First, we measure how the performance of external R&D projects changes when the firm has generated sufficient in-house R&D in a similar knowledge domain. Second, we measure how internal R&D project performance changes when 'sufficient' external R&D is undertaken in a similar knowledge domain. Our findings are in line with the theory on absorptive capacity. This means that sufficient internal knowledge does not improve the performance of external R&D projects. External R&D projects that are selected do perform exceptionally well, but the main improvement occurs amongst the in-house R&D projects once a few (probable well screened) external projects are added to the 'group' of internal R&D projects. In other words, having a sufficient stock of internal R&D in a relevant knowledge domain enables the firm to screen (attract) the right external R&D projects from which it can learn and subsequently increase its marginal return on internal R&D. Finally, our findings indicate that this knowledge flow (or spillover) only occurs when relatively few external projects are added to the group of in-house R&D. We strongly suspect that the knowledge that is ultimately responsible for realized absorptive capacity does not flow or spill over between projects or even researchers, but only travels because the same people who are working on externally sourced projects are applying the obtained knowledge internally. This would naturally limit the number of external projects in relation to in-house capacity.

[3] Additionally it might be the case that complementarity raises competition (or tournament effects) between activities which leverages efficiencies and reduces organizational slack. Another potential driver of the relationship between complementarities and performance might be that investment in one activity improves selection capabilities for other activities[3] (Veugelers, 1997).

2. Literature

A number of elementary studies have identified complementarities and some studies have, what might be even more important, identified circumstances that drive complementarity. With our study, we hope to further our current understanding of why and how complementarities occur among research activities within-, and beyond the boundaries of the firm.

The earliest work on complementarities between different research activities goes back to Coase's (1937) work, where he argues that as a firm accumulates R&D experience internally the costs of internalizing new R&D decreases. More specifically, Nelson & Winter (1974) state that the 'ease' of internalizing new R&D depends on whether prior accumulated R&D is similar to the newly acquired R&D, and that the costs of these activities are reduced because of learning curve effects (Pisano, 1990). Also Milgrom and Roberts (1990) have focused their work on complementarities within the boundaries of the firm, although their definition of complementarity has been applied in some of the more fundamental work on complementarities between internal R&D and externally sourced R&D knowledge as well. The seminal work on complementarities between internal R&D and external sourcing of knowledge is the paper by Cohen and Levinthal (1989) on the two faces of R&D. The authors state that firms invest in R&D primarily to generate internal innovations. More interestingly, they discover that a side-effect of R&D investment is that it enables firms to appropriate external, publicly available spillovers more easily than firms which invest in R&D to a lesser extend. This side-effect is termed 'Absorptive capacity'. It is defined as a firm's relevant stock of prior knowledge that enables it to identify, assimilate and exploit knowledge from the environment. While the emphasis is on knowledge influx (absorbing external knowledge in), empirical work that followed from this study has often looked at the effect of internal R&D on successful external knowledge sourcing, which implies knowledge outflow (Lane & Lubatkin, 1998; Arora & Gambardella, 1990). The literature review on absorptive capacity by Zahra and George (2002) has brought structure to a growing conceptual ambiguity around absorptive capacity. The authors identify three stages within the construct of absorptive capacity which are essentially already incorporated in the original definition of Cohen & Levinthal (1989), namely: identification of external knowledge through external sourcing activities. Second, the assimilation or conversion of this knowledge back into the firm, and third the exploitation of the absorbed knowledge to new or improved products and processes. Based on these three phases, most existing work can be grouped as focusing on potential absorptive capacity (phase one and two) or on realized absorptive capacity (phase three) (Zahra & George, 2002). An interesting study that does cover all three phases of absorptive capacity (identification, incorporation and exploitation) is the study by Cassiman & Veugelers (2006). The authors find that firms which 'make and buy' R&D have a higher marginal return on innovation than firms who

only 'make' or only 'buy', especially if they are more heavily involved in basic R&D. More precisely, a 10% increase in reliance on basic R&D increases the likelihood of combining internal and external sourcing by 2.7 % (Cassiman & Veugelers, 2006 pp. 77).

The studies described above have convincingly argued that complementarities exist among R&D activities within and between firm boundaries. An important condition for complementarities to occur, as Cohen & Levinthal already mentioned, is that internal and external knowledge are relevant to one another. With two exceptions, existing contributions have largely ignored this condition of knowledge relevance. One exception is the work of Arora and Gambardella (1994) who test whether firms use their external linkages as complements. While complementarity itself is not precisely defined in this study, it is assumed that different types of external linkages are complements if they do not have overlapping (knowledge) purposes. Moreover, the authors argue that external linkages such as research agreements with universities and acquisitions of biotech firms are complementary strategies because they serve different purposes but are still correlated. While indeed, it is generally acknowledged that activities which are completely overlapping in terms of purpose or knowledge domain are considered as substitutes, it is not very clear whether non-overlapping features defines them as complements (Besanko, 2007). A more precise investigation into the relevance of knowledge between R&D activities is provided in the work of Lane and Lubatkin (1998). These authors argue that a firm's absorptive capacity is often seen as a firm-specific characteristic that determines its innovativeness vis-à-vis others to a large extent. However, Lane and Lubatkin (1998) state that a firm's ability to absorb external knowledge depends on the 'type' of external knowledge (and the partner carrying this knowledge) the firm is absorbing. In other words, absorptive capacity is a relational characteristic rather that an actor characteristic of the firm, since it differs with each external partner. A firm's absorptive capacity in this sense depends on the knowledge (cognitive) similarity between internal knowledge (experience), and external knowledge, and thereby becomes a relational characteristic of the firm. This 'new' notion of absorptive capacity furthermore implies that innovativeness, which is increased by absorptive capacity, differs for each activity where a firm is tapping into a new external source of knowledge. An important implication of this finding is that measuring absorptive capacity requires project-level information.

Our study differs and enriches these existing studies in a number of ways. To start with, all above mentioned studies are performed at the level of the firm, while our study enables a direct measure of performance at project level. Having project level information has important advantages. Not the least advantage is that it allows us to circumvent the danger of firm heterogeneity driving endogenous decisions of which projects are selected[4]. Furthermore, while being specific for the pharmaceutical industry in which our research is situated, we identify the type of knowledge where learning curve effects and subsequent complementarities occur. Most importantly however is that while previous studies have convincingly shown the existence of complementarity as a binary choice, our study allows us to treat complementarity as a continuous variable. More specifically, we identify a size threshold over which complementarities occur, which concretizes Cohen & Levinthal's (1989) notion of the 'sufficient' stock of knowledge required for achieving complementarities. Given the existence of complementarity, our data allow us to empirically disentangle whether these complementarities represent potential absorptive capacity (arising from improved external sourcing) or whether the firm has indeed managed to reintegrate and exploit external knowledge, which represents realized absorptive capacity.

3. Empirical setting: the pharmaceutical industry

A number of developments in the pharmaceutical industry over the last decade have made the quest for a strategy that achieves complementarities and subsequent innovation capacity an important and largely unanswered question in this industry. Within big Pharma there is an increased pressure on R&D due to decreased R&D productivity and approaching patent expirations. Along the pressure for big pharma to maximize shareholder value, strategies have shifted in the last two decades from being research driven to being market driven (Drews, 2003). In a research driven environment, emphasis of decision making was within R&D departments, where managers were geared at innovations originating from deep knowledge of disease pathways and pharmacology. While the industry was consolidating in the beginning of the nineties, the gravity of decision making has shifted toward marketing and finance departments (Drews, 2003). This strategic shift implied a more quantity based approach toward research whereby drug discovery was increasingly considered a statistical event (Booth & Zemmel, 2004). Today, it appears that this quantitative approach toward research is not yet paying off in terms of productivity, and analysts are revaluating the early days' in vivo empiricism based on disease knowledge (Erickson, 2003).

Furthermore, a series of findings suggest that alternatives to the traditional in-house R&D model of big pharma might be more successful. The first finding concerns the higher success probabilities of (new) biotechnologies. The proportion of newly admitted compounds using biotechnologies has increased with 20 -25% compared to the more traditional chemical

[4] For additional information on how to deal with the problem of unobserved firm heterogeneity we refer to Cassiman & Veugelers, 2006.

based compounds[5]. Biotechnologies are mainly exploited by small-, and medium sized biotechnology firms, while chemical based compounds mainly originate from big pharma. Second, newly developed compounds that are discovered in-house are being outperformed by compounds that are produced externally or through external collaborations[6].

Moreover, as biotechnologies increase the scope of research, big pharma increasingly realize that it is impossible to cover the whole spectrum of technologies themselves. While research for new drugs was traditionally conducted in-house, large pharmaceutical firms have by now build research portfolios where internal R&D efforts are combined with external R&D collaborations. Some analysts even go as far as to state that Pharma companies should be virtual in research, which means that they in-license all compounds from preclinical testing onwards.

Against the background of these developments it has become questionable whether big pharma still has to play a role in research. Wouldn't it be better if big pharma narrows down its core competences to downstream drug development and employ a virtual research model? Or are there still advantages to be obtained from in-house research? If the latter is true, which portfolio would generate complementarities between internal and externally sourced research projects?

4. Empirical strategy

Our empirical strategy is summarized in figure 1 below.

[insert figure 1]

4.1 Data selection

We selected the 20 highest R&D spending pharmaceutical firms in 2005 from Evaluate Pharma, a frequently used database in the pharmaceutical industry for forecasting and analysis services. Of the selected firms, we used Pharma Projects database to extract the whole research portfolio per firm at a certain time. PharmaProjects is a privately held project monitoring firm which continuously searches for information on both internal and externally sourced projects of large pharmaceutical firms through a number of search channels. Primarily, PharmaProjects visits events and conferences where pharmaceutical firms meet to exchange information about the projects running through their pipelines. This information is then verified and updated with press releases, website information and annual reports. Telephone surveys are regularly conducted to verify the accuracy of their database. As a deliberate strategy, no use is made of patent information, since "often..", our informant claims, "..the firm files patents on anything that lies around in the lab to create a smoke screen and hide their actual R&D strategy" (information based on telephone interview with Pharmaprojects data manager).

We focus our analysis on research projects that are active[7] at the earliest stage of research before entering clinical testing. In doing so, we follow earlier work by Pisano (1990) who convincingly argues that from clinical testing onwards, external contracting is often done by technological licensing instead of R&D contracting. Technological licensing and R&D contracting are two fundamentally different contracts. R&D contracts are typically long-term agreements where knowledge exchange and learning effects take place, while technological licensing agreements are one-time exchanges to obtain rights to an already developed technology (Pisano, 1990 pp. 163). As in our study we are interested in complementarities arising from learning through knowledge spillovers between and within organizations, we restrict our analysis to these early stage research projects. Another reason for this restriction is that our dependent variable (probability to enter clinical I) is more reliable when restricted to early stage research. If we were to measure the probability of a project to reach the market for example, we would be unable to distinguish complementarities from many other factors affecting whether a projects survives the 12 year (or longer) ride through the pharmaceutical production process.

4.2 Buiding a research portfolio

To determine each firm's research portfolio of 2002, we summed all research project that were announced as annual newly entering projects in early research from 2000 until 2002. We thereby follow a study by Phelps (2003) who shows that the average duration of R&D projects is three years (see also: Phlippen & van der Knaap, 2007). As a result, all newly announced projects in 2000, 2001 and 2002 are assumed to be part of a firm's 2002 research portfolio. Our choice of constructing the 2002 research portfolio and not a more recent portfolio is related to our dependent variable 'probability to reach clinical testing'. A project in early research can take (on average) up to four years to reach clinical testing which makes it necessary to track each project until the end of 2006 to know whether it has been successful in reaching clinical testing on humans. Our initial sample consisted of 1328 early stage research projects (before clinical testing I). Leaving out projects of which no success probability was known reduced our sample to 977 projects. Furthermore, we excluded all projects where no disease area information was given, which reduced our final work set to 762 projects.

[5] Tufts Center for the Study of Drug Development (95 – 99) New Biopharmaceuticals in the US

[6] 'Improving the pharma research pipeline' McKinsey Quarterly, August 2004.

[7] Projects only entered the database if they had a solid chemical structure and a therapeutic goal had been identified.

4.3 Variables

4.3.1 Depedent variable: success probability

Each project is defined as either successful, failure or unknown in reaching the first stage of clinical testing (SUCCESS). If we didn't find the project in our database for 4 years or longer, we decided to label it as a failure. Projects were labeled unknown if there was no information between 1 and 3 years. We used Binary logistic regression analysis to model the success probabilities of a research project.

4.3.2 Control variables

In every regression we controlled for a number of variables that are strongly associated with the probability of a project reaching clinical testing successfully (see: ***empirical setting***). The first is the variable indicating whether a project is being developed in-house or through external collaboration (EXTERNAL), as previous work has shown external collaborations to perform better on average than internal projects over the whole production process[8]. A second control variable indicates whether a project involves chemicals or biotechnologies (BIOTECH), as biotechnology based compounds are found to outcompete compounds based on chemical substance. We further controlled for the effect of a project aiming at a reformulation of an existing drug or aiming at a new drug (FORMULATION), since the former are assumed to be more likely to reach clinical testing. Each project is focused at a certain disease area. To prevent the disease area (e.g. anticancer or inflammation) itself to be driving the success probabilities of our projects we included dummies for all disease areas in most of our regressions (ANTICAN, INFLAM etc). At the level of the firm, we control for firm heterogeneity simply by adding dummies for all firms in most of our regressions (ROCHE, GSK, etc).

4.3.3 Explanatory variables

There are two main sets of explanatory variables. The first set is aimed at identifying the *conditions* for complementarity to arise, and the second set is aimed at understanding *how* complementarities occur between internal projects and external projects, i.e. whether they result from potential (knowledge outflow) or from realized absorptive capacity (knowledge inflow).

Conditions for complementarity

As we have argued in our introduction, Cohen & Levinthal (1989) and the work thereafter has emphasized the importance of a firm's 'sufficient stock of relevant prior knowledge' in order to absorb external knowledge effectively. We test what constitutes a 'sufficient' stock and what knowledge is 'relevant' to obtain complementarities.

Where to look for complementarities?

Which type of knowledge similarity generates complementarities? While it is typically assumed that projects aiming at similar therapeutical areas might compete and/or learn from each other and create complementarities by doing so, we test 3 other potential types of knowledge areas where complementarities might occur. First we add disease area as a potential knowledge area that might be an alternative for therapeutic area. TAsize (therapeutic area) is a more broadly defined group than DAsize (disease area). For example the therapeutical area named 'cognition enhancer' consists among others of the disease area 'Alzheimer'. Another knowledge area where complementarities might occur is the target that a drug in a project is aiming at (TARGETsize). A drug target can be the protein which the drug binds to, inhibits or activates (eg receptor subunits or enzymes). Finally we tested whether the pharmacological activity, which describes the beneficial or adverse effects of a drug on living matter (i.e. it describes *how* the drug works) might generate complementarities (PHARMACOsize). Based on each of these four knowledge areas we grouped all projects within each firm and tested whether more projects in each of these knowledge areas (within a firm) increases the average performance of projects.

When is a stock of knowledge sufficiently large?

The aim is to determine the effect of the number of 'similar' projects in a firm's research portfolio (SIZE) on the probability of a project reaching clinical testing. The knowledge area to which 'similarity' applies is to be defined as a first step in our analysis. This variable (SIZE) is categorized as either 'no or just one similar project' (SMALL), 2-9 similar projects (MEDIUM), or containing 10 or more similar projects (LARGE)[9]. Complementarities arise when more similar projects leads to higher average success probabilities. We thus expect that projects in category LARGE perform better than projects in category MEDIUM. Projects categorized as SMALL contain (nearly) isolated projects, which by definition do not measure complementarity. However, being a relatively large group of projects, we used this category as our reference category to benchmark our two other categories against.

[8] 'Improving the pharma research pipeline' McKinsey Quarterly, August 2004

[9] We explored different size categories and the currently used categories were most able to discriminate amongst success probabilities.

How do complementarities arise between internal and external projects?

The theory of absorptive capacity provides guidance on how complementarities are expected to arise between internal research and externally sourced research projects. While a firm's absorptive capacity implies knowledge flowing outside in, this theory argues that internal knowledge is first used to screen and absorb external knowledge, and as a second step the firm reintegrates the absorbed knowledge internally to improve internal R&D. We measure both parts of this process separately by identifying:
1. How the performance of external projects changes when a 'sufficient' amount of 'similar' internal projects are running (EXTERN_SUCCESS).
2. How the performance of internal projects changes when a 'sufficient' amount of 'similar' external projects are running (INTERN_SUCCESS).
The question about 'sufficient' amounts of internal and externally sourced projects essentially asks how a firm should design its research portfolio with regard to the ratio of internal and external R&D investment. Surprisingly we found no study that deals with this question explicitly and hence we explored the 'optimal' external/ internal ratio among 'similar' projects ourselves.

5. Results

5.1 Effect of individual project characteristics

Our analysis starts with the assessment of individual project characteristics that have in previous studies been identified as having a significant impact on the success probabilities of R&D. The two main characteristics are whether a project is conducted in-house or through external sourcing, and second whether a project builds on chemicals or on biotechnologies[10] [11].

[insert figure 2]

Externally sourced projects are significantly more likely to reach clinical testing (our indicator of success) than internal projects. While previous studies have already shown this differences between internal and externally sourced projects at the development stages in R&D (clinical I to III) (DiMasi, 2001), our findings indicate that the advantage of external collaboration already occurs during early research. We further found that projects involving biotechnologies, such as recombinant DNA technologies or monoclonal antibodies are also significantly more likely to be successful compared to projects based on chemical compounds. Figure 2 illustrates these findings by showing the differences in average success rates related to external collaborations and biotechnologies.

5.2 Conditions for complementarities

Identifying 'relevant' knoweldge

Our data allowed us to test four knowledge areas where complementarities might occur. More specifically, we grouped projects around the same disease area, the same target, the same pharmacological activity and around the same therapeutic area (within a company). For each grouping we then analyzed the effect on success probabilities, as is shown in table 1. After controlling for individual project characteristics we found that only disease area grouping has a positive significant impact on success probabilities, and that neither grouping by target, by pharmacological activity or by therapeutic area has a significant impact on success probabilities[12]. From here on, our analysis focuses on projects grouped around disease areas, as it appears to be the only relevant knowledge area for achieving complementarities.

[insert table 1]

Identifying a sufficiently large stock of knowledge

Previous work on complementarities has provided clear evidence that adding an R&D project to an already existing stock of R&D has a higher incremental effect on innovation than adding an R&D project in isolation (Milgrom & Roberts, 1990). Going one step further, we analyze whether the size of the existing stock of knowledge matters for achieving these complementarities. Intuitively, one could imagine that a critical mass of relevant knowledge must be achieved in order to truly benefit from complementarities. In order to test whether this is indeed the case we categorized projects as either belonging to a large disease area group, (i.e. containing 10 projects or more per firm), to a medium sized disease area group (containing 2 to 9 projects within the same disease area per firm), or as focusing on a (nearly) isolated disease area. In the latter category success probabilities are not caused by complementarities. Based on the actual distribution of projects over disease areas by our set of

[10] See "empirical setting: the pharmaceutical industry" for an overview.
[11] We also tested the effect of a project being a new formulation of an existing drug or a 'real' new drug.
[12] Interestingly, grouping based on therapeutic area appears to have a significant negative effect on success probabilities. It might be the case that on the higher aggregation level that therapeutical area represent, competition outweighs learning effects.

firms in figure 3 we chose the boundary between large and medium sized groups.

[insert figure 3]

Figure 3 reveals that the largest R&D spending pharmaceutical firms invest 46 percent of their total research projects in disease areas with on average 10 projects or more (within the 2 largest disease areas). If one can speak of a critical mass of knowledge that drives complementarities, we expect the boundaries to be around 10 or more projects[13].
Table 2 shows the binomial Logit regression results on the effect of the group size to which a project belongs on the probability that the project reaches clinical testing.

[insert table 2]

In order to exclude any effect of either firm heterogeneity or effects that are specific for any disease area, we added firm dummies and disease area dummies next to the usual project characteristics as control variables. To keep our results readable, we used a stepwise selection method for our variables with entry testing based on the significance of the score statistic, and removal testing based on the probability of a likelihood-ratio statistic based on the maximum partial likelihood estimates. Our main variable of interest is LARGE, which is offset against our base variable MEDIUM. As table 2 reveals, the variable LARGE is significant at the 10 percent level which indicates that projects active in a disease area where at least 10 other projects are running are more likely to reach clinical testing compared to projects in a disease area where only 2 to 9 similar projects are running. This finding confirms that projects in which a firm has built a critical mass of disease knowledge have a higher marginal return to performance than projects in which a firm has not built such a critical mass. More generally, complementarities between research projects arise once a firm has built a critical mass of knowledge in a similar knowledge domain.
However, one could argue that there is a critical danger of endogeneity driving our results i.e., projects do not become more successful because of complementarity effects, but the management of a firm puts its 'golden eggs' in one basket, namely in its main disease areas. This would imply that the expected success of certain projects drives management to focus their attention on these projects and create a large number of similar projects around these 'golden eggs'. To test whether this is the case, we analyze the effect of each firm's two main disease areas[14] on the success probabilities of projects in these DA's. Table 3 shows the results of this test, whereby we replicated the test on size effects (table 2) while interchanging the categories LARGE, MEDIUM, SMALL with the categories DA2largest and OTHER (reference category)[15].

[Insert table 3]

DA2largest measures whether a project is part of a firms two main disease areas or not. As becomes clear in table 3, projects that belong to one of the firm's two most important disease areas does not increase their probability of reaching clinical testing. Moreover, there seems to be a negative effect from being part of a firm's 2 main disease areas. This effect might be caused by the fact that the firms in our sample differ with respect to their R&D investment strategy. While some firms choose (or are able) to built a critical mass in a few disease areas, other firms rather spread their R&D projects over different disease areas. These strategy differences become clear when looking at figure 3. The number of projects that form a firm's 2 largest disease areas range from 2 to 37 projects. Interestingly, the strategy of spreading projects over different disease areas, referred to as risk diversification strategies, is appearing to be paying off as well. Although this study investigates complementarity effects among projects, the highly significant positive effect of nearly isolated projects (SMALL) in table 2 raises our suspicion that aiming for complementarities by building a critical mass of disease knowledge is not the only rewarding strategy.

5.3 How do complementarities arise between internal and external projects?

So far, we have found that complementarities arise from grouping a relatively large amount of research projects around a disease area. As a next step, we focus on these large disease area groups to find out how the ratio of internal R&D and external R&D affects complementarities among internal projects and externally sourced projects separately[16]. The ratio between internal and external R&D has been surprisingly little discussed within the literature on absorptive capacity. We argue however that it is a fundamental question since the capacity to absorb larger amounts of external information must depend greatly on a larger internal capacity to absorb this information. To some extend the work of Lane and Lubatkin (1998) recognizes the importance of this ratio by focusing on relative absorptive capacity. However, they do so at the level of the dyad (i.e. a relation

[13] We also explored shifting the boundaries to 9 or more projects and to 12 or more projects.
[14] Again, we chose to analyze a firms two main DA's since these represent 46 % of all research projects, while the third largest disease areas and beyond are strongly decreasing their contribution to the firms research portfolio (see figure 3).
[15] To prevent our main variable of interest DA2largest to be excluded from the results based on restricted entry testing used in table 2, we unconditionally let the variable DA2largest and the BIOTECH, FORMULATION and EXTERNAL variables enter our equation.
[16] By reducing our sample to projects in large disease areas we are unable to control for firm-, and disease fixed effects. This is due to the fact that only few firms are able to create a large number of projects in a few disease areas (e.g. anticancer and infection).

between internal R&D and one externally sourced project).
Once the ratio of internal versus external projects that generates complementarities is determined, we can test whether external projects or internal projects are most responsible for the increased success probabilities. In this part of the analysis we use a subsample of our data, namely only projects belonging to a firm's large disease area (10 or more similar projects).

A Firm's internal / external R&D ratio for large disease areas

Figure 4 below plots the number of projects that firms run in large disease areas at different internal / external ratios.

[insert figure4]

Based on the above plot we divided projects as either belonging to a disease area with few external projects and many internal projects (ratio externals 20/80 or less) or vice versa (ratio externals 20/80 or more)[17]. Dividing our projects into these to categories allows us to test if projects perform differently in each category. If they do, we can test whether internal projects benefit from a specific internal/external ratio or whether external projects benefit from this ratio. Projects in the 'poor performing' category serve as our benchmark. The results of these tests are displayed in table 4.

[insert table 4]

In table 4, the variable FEWEXTERNALS shows that projects in a disease area with an external ratio of 20/80 or less perform significantly better than projects in a disease area with a higher externals ratio. While this variable does not distinguish among internal and external projects, the addition of our interaction term (FEWEXTERNALS by EXTERNALS) controls for the (slightly negative) effect that external R&D projects have in the disease areas with a low external ratio. To put it differently, the higher performance of projects in DAs with low externals ratio is mainly attributable to the improved performance of internal projects. To clarify this point we add a fifth table where we only consider the subsample of internal projects in large disease areas.

[insert table 5]

Here, in table 5 the variable (INTERNAL_SUCCESS) indicates the performance of internal projects in large disease areas with a low external ratio in comparison to other internal projects in large disease areas. The results confirm the findings displayed in table 4, that internal projects perform better if they are part of a large disease area with few externally sourced projects involved.

[17] Obviously this choice is somewhat arbitrary. We explored different ratios and found this division to be discriminating our success probabilities.

6. Conclusions

For more than a decade firms in research driven environments such as the pharmaceutical industry, experience an increased pressure on R&D due to decreased R&D productivity and approaching patent expirations (Drews, 2001). In response to this, firms are exploring alternative ways to organize their R&D portfolio. While traditionally early stage R&D has been conducted mainly inside the firm's own R&D laboratories, the last decade has brought forward a huge increase in R&D collaborations with market based firms at all stages of the drug development process. This raises the question whether the make-*or*-buy decision of the firm should be replaced by a make-*and*-buy decision of the firm. The answer depends crucially on the extend to which internal R&D (make) and externally sourced R&D (buy) can be complementary, i.e. whether performing internal R&D in combination with external R&D generates higher marginal performance than only internal or only external R&D. In theory, complementarity between internal and external R&D exists if a firm has built a *sufficient* stock of *relevant* internal knoweldge to *effectively* absorp external knoweldge (Cohen & Levinthal, 1989). Our study is the first empirical work that has investigated the conditions for complementarity to arise, and the process through which it occurrs. More specifically, we examined what 'type' of knowledge is *relevant* for complementarity, how much of this knowledge is *sufficient*, and lastly we examined *how* this knoweldge flows between internal and externally sourced R&D. As it turns out, large pharmaceutical firms can achieve higher marginal returns on their R&D projects, if they group a relatively large number of projects (more than 10) around a specific disease area. By focussing on a specific disease area firms can develop deep in-house expertise, attract the best talent, and be a preferred partner for deal-opportunities outside the firm. Moreover, we conclude that if no more than 20 percent of these projects are externally sourced, complementarity effects are highest. This is caused by the fact that internal projects perform better when the number of externally sourced projects is relatively low. The reason for this might be that knowledge can only be transferred from external projects to internal projects if the same expert scientists are involved in both internal and external partners work. This would naturally limit the number of externally sourced projects. This interpretation is in line with the notion that knowledge required for pharmaceutical drug discovery is highly tacit and embedded in the scientists involved, which makes transfer of this knowledge between people leave alone projects difficult.

7. References

Abramovsky L., R. Harrison and H. Simpson (2007). University Research and the Location of Business R&D, *The Economic Journal*, 117 (519) 114-141.

Arora, A. and A. Gambardella, (1990). Complementarity and external linkages: the strategy of large firms in biotechnology. *Journal of Industrial Economics* 38. 361-379.

Arora, A. and A. Gambardella (1994). Evaluating technological information and utilizing it: Scientific knowledge, technological capability and external linkages in biotechnology.
Journal of Economic Behavior and Organization 24(1) 91-114.

Booth, B., and R. Zemmel (2004). Prospects for productivity. *Nature Reviews Drug Discovery* 3 451-457.

Cassiman, B., and R. Veugelers. (2006) In search of complementarity in the innovation strategy: Internal R&D and external knowledge acquisition. *Management Science* 52 (1) 68-82.

Coase, R. (1937). The Nature of the Firm. *Economica* 4(16) 386-405.

Cohen, W., and D. Levinthal, (1989). Innovation and learning: the two faces of R&D. *The Economic Journal* 99 569-596.

DiMasi J. (2001). Risks in New Drug Development: Approval success Rates for Investigational Drugs. *Clinical Pharmacology & Therapeutics*. May 297-307.

Drews, J. (2003). Strategic trends in the drug industry. *Drug Discovery Today*. 8 (9) 411-420.

Erickson, D. (2003) Wanted: drug hunters. *In Vivo* 45.

Freeman, C. (1991). Networks of innovators: a synthesis of research issues. *Research Policy* 20(5) 499-514.

Gambardella, A. (1992). Competitive advantages from in-house scientific research: The US pharmaceutical industry in the 1980s. *Research Policy* 21 391–407.

Henderson R, Cockburn I. (1996). Scale, scope, and spillovers: The determinants of research productivity in drug discovery. *RAND Journal of Economics.* 27 32-59.

Lane, P. J. and M. Lubatkin (1998). Relative absorptive capacity and interorganizational learning. *Strategic Management Journal* 19 461–477.

Milgrom, P. and J. Roberts (1990). The Economics of Modern Manufacturing: Technology, Strategy and Organization. *American Economic Review* 80(3) 511-528.

Nadler, D., and Michael Tushman. (1980). A Model for Diagnosing Organizational Behavior. *Organizational Dynamics* 9(2) 35-51.

Nelson, R.R., and S.G. Winter (1974). Neoclassical vs Evolutionary Theories of Economic Growth: Critique and Prospectus. *Economic Journal*, 84 886-905.

Phelps, C. (2003) Technological exploration: A longitudinal study of the role of recombinatory search and social capital in alliance networks. *Dissertation New York University*, Graduate School of Business Administration.

Phlippen, S. & G.A. van der Knaap (2007). When Clusters Become Networks. *Tinbergen Institute Discussion Paper* No. 2007-100/3

Pisano PG. (1990). The R&D Boundaries of the Firm: An Empirical Analysis. *Administrative Science Quarterly* 35. 153-176.

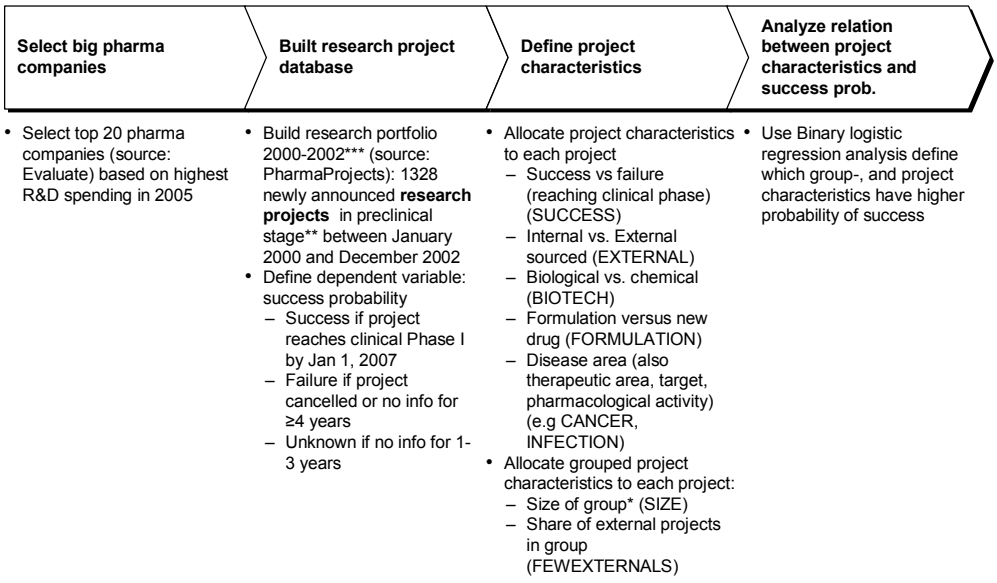
Porter, M.E. (1980) "Competitive Strategy", The Free Press, New York, 1980

Van den Bosch, F. H. J, H.W. Volverda & M de Boer. (1999) Coevolution of Firm Absorptive Capacity and Knowledge Environment: Organizational Forms and Combinative Capabilities. *Organization Science*. 10(5) 551 - 568.

Veugelers, R. (1997) Internal R&D Expenditures and External Technology Sourcing. *Research Policy*, Vol. 26(3). 303-315.

Zahra, S. A. and G. George (2002). Absorptive capacity: A review, reconcepualisation, and extension. Academy of Management Review 27(2) 185-203.

Figure 1. - Empirical strategy



* Number of projects that belong to same disease area (DASIZE), target (TARGETSIZE), therapeutic area (TASIZE) or pharmacological (PHARMACOSIZE) activity

** All stages of preclinical investigation including discovery, research, lead optimization.

*** Following Phelps (2003)5 we assume a project duration of three years. As a result, all projects that started between 2000 and 2002 make up the 2002 research portfolio.

Figure 2. – Individual project characteristics

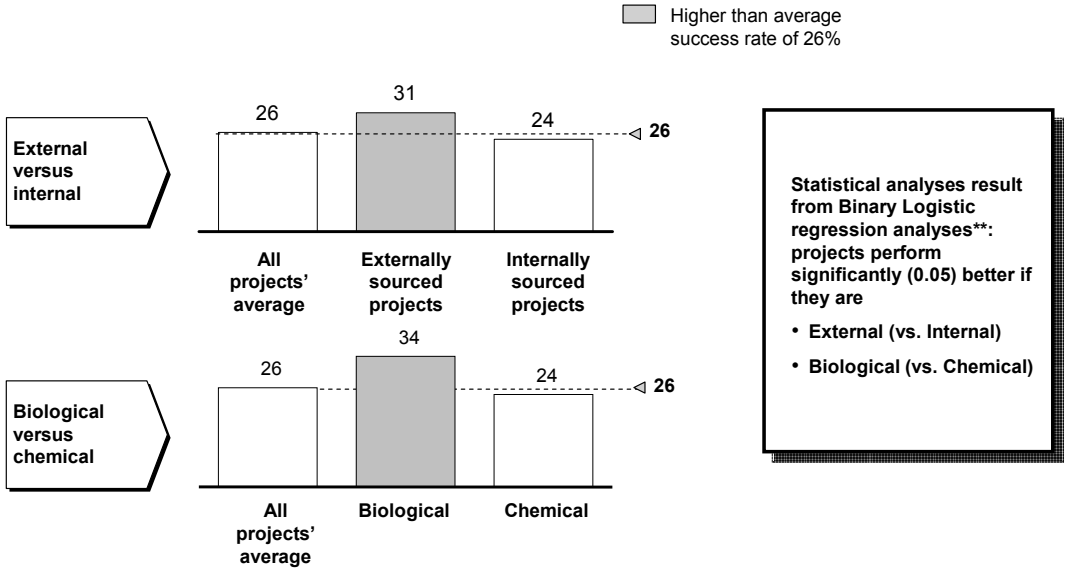


Table 1. – Identifying relevant knowledge areas: effect of grouping by therapeutic area, pharmacological activity, disease area, and target on a projects success probability.

		B	S.E.	Sig.	Exp(B)
Step 1(a)	BIOTECH	.724***	.234	.002	2.064
	FORMULATION	1.129***	.378	.003	3.092
	EXTERNAL	.360*	.195	.065	1.434
	Tasize	-.049	.016	.003	.952
	PharmacoclusterSize	-.015	.011	.168	.985
	Dasize	.034**	.017	.044	1.035
	TARGETsize	.089	.071	.209	1.093
	Daexternals	-.082	.050	.102	.921
	Constant	-1.008	.164	.000	.365

a Variable(s) entered on step 1: BIOTECH, FORMULATION, EXTERNAL, TAsize, PharmacoclusterSize, DAsize, TARGETsize, DAexternals.
Notes: coefficients significant at 1%***, 5%**, 10%*

Figure 3. – Number of projects per disease area.

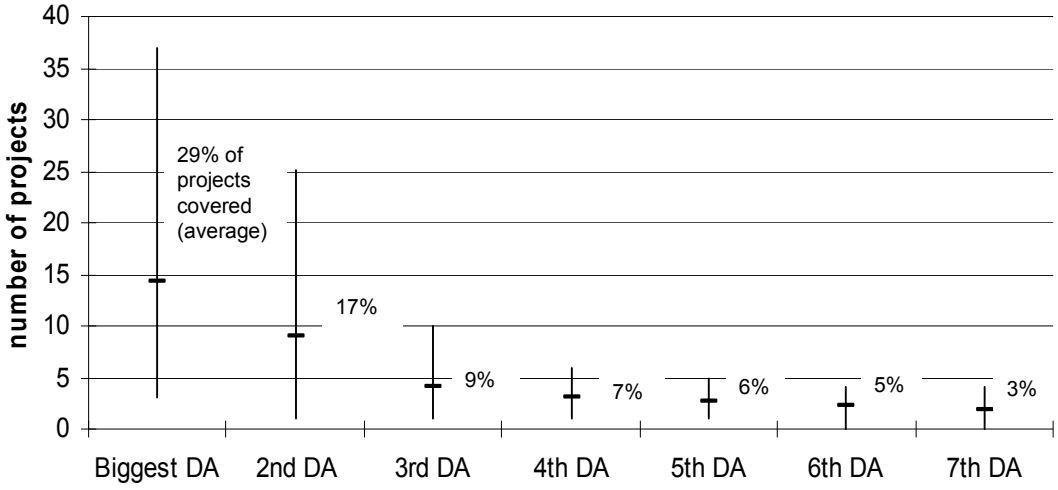


Table 2. – Identifying a sufficient stock of knowledge: the effect of the group size to which a project belongs on its success probability.

		B	S.E.	Sig.	Exp(B)
Step 1(a)	BIOTECH	.524	.198	.008	1.689
	FORMULATION	1.227	.363	.001	3.409
	EXTERNAL	.375	.164	.023	1.454
	LARGE	.367	.211	.082	1.444
	SMALL	1.020	.181	.000	2.773
	INFECTION	.461	.230	.045	1.585
	Constant	-1.815	.145	.000	.163
Step 2(b)	BIOTECH	.518***	.199	.009	1.679
	FORMULATION	1.229***	.363	.001	3.417
	EXTERNAL	.375**	.165	.023	1.456
	LARGE	.371*	.211	.079	1.448
	SMALL	1.035***	.182	.000	2.814
	Astrazeneca	.638**	.318	.045	1.893
	INFECTION	.495**	.230	.032	1.640
	Constant	-1.862	.148	.000	.155

a Variable(s) entered on step 1: INFECTION.
b Variable(s) entered on step 2: Astrazeneca.
Notes: coefficients significant at 1%***, 5%**, 10%*
Notes2: N = 977.

Table 3. – Effect of projects belonging to a firm's 2 largest disease areas on success probabiltities.

		B	S.E.	Sig.	Exp(B)
Step 1(a)	BIOTECH	.510	.221	.021	1.665
	FORMULATION	1.226	.374	.001	3.407
	EXTERNAL	.243	.183	.183	1.276
	DA2largest	-.311	.170	.068	.733
	Wyeth	.703	.336	.036	2.020
	Constant	-1.046	.124	.000	.351
Step 2(b)	BIOTECH	.510**	.221	.021	1.665
	FORMULATION	1.245***	.375	.001	3.472
	EXTERNAL	.237	.183	.196	1.268
	DA2largest	-.333**	.171	.052	.717
	GSK	.548**	.278	.049	1.730
	Wyeth	.759**	.338	.025	2.136
	Constant	-1.089	.126	.000	.336

a Variable(s) entered on step 1: Wyeth.
b Variable(s) entered on step 2: GSK.
Notes: coefficients significant at 1%***, 5%**, 10%*
Notes2: N = 977.

Figure 4. – The number of projects in large disease areas (within the firm) at different external/internal ratios

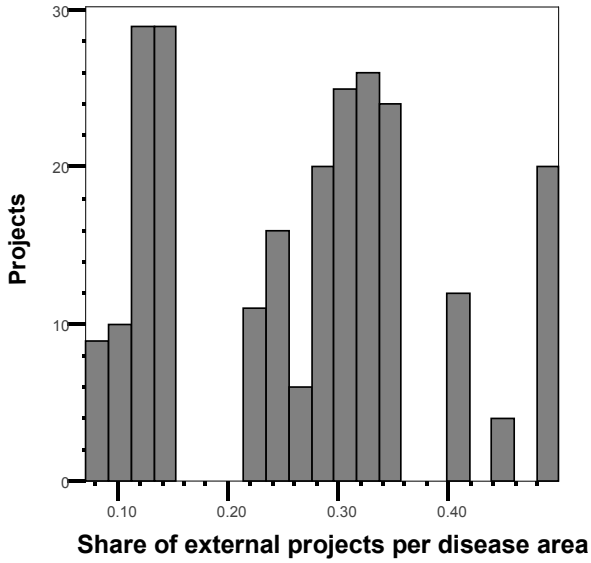


Table 5. - Effect of (internal) projects belonging to a large disease area with less than 20% external R&D projects on success probability. N = 136

		B	S.E.	Sig.	Exp(B)
Step 1(a)	BIOTECH	1.416***	.508	.005	4.120
	FORMULATION	-20.317	40192.970	1.000	.000
	EXTERNAL	1.022*	.561	.069	2.778
	FEWEXTERNALS	1.040**	.493	.035	2.829
	EXTERNAL by FEWEXTERNALS	-1.071	1.062	.313	.343
	Constant	-1.907	.391	.000	.148

a Variable(s) entered on step 1: BIOTECH, FORMULATION, EXTERNAL, FEWEXTERNALS, EXTERNAL * FEWEXTERNALS .
Notes: coefficients significant at 1%***, 5%**, 10%*

Table 6. - Internal project performance in large clusters with few externals compared to other internal projects.

		B	S.E.	Sig.	Exp(B)
Step 1(a)	BIOTECH	.923**	.418	.027	2.516
	FORMULATION	.380	1.182	.748	1.463
	INTERNAL_SUCCESS	.581*	.353	.100	1.788
	Constant	-1.479	.254	.000	.228

a Variable(s) entered on step 1: BIOTECH, FORMULATION, DA_large_extcat2_internalII.
Notes: coefficients significant at 1%***, 5%**, 10%*
Notes2: N = 182