DISA

Dipartimento di Informatica
e Studi Aziendali

Patent Disclosure and R&D Competition in Pharmaceuticals

Laura Magazzini, Fabio Pammolli,
Massimo Riccaboni, Maria Alessandra Rossi

UNIVERSITÀ DEGLI STUDI
DI TRENTO

2009/2

DISA Working Paper

DISA

Dipartimento di Informatica
e Studi Aziendali

2009/2

Patent Disclosure and R&D Competition in Pharmaceuticals

Laura Magazzini, Fabio Pammolli,
Massimo Riccaboni, Maria Alessandra Rossi

UNIVERSITÀ DEGLI STUDI
DI TRENTO

DISA WORKING PAPER

DISA Working Papers
The series of DISA Working Papers is published by the Department of Computer and Management Sciences (Dipartimento di Informatica e Studi Aziendali DISA) of the University of Trento, Italy.

Guidelines for authors
Papers may be written in English or Italian but authors should provide title, abstract, and keywords in both languages. Manuscripts should be submitted (in pdf format) by the corresponding author to the appropriate Associate Editor, who will ask a member of DISA for a short written review within two weeks. The revised version of the manuscript, together with the author's response to the reviewer, should again be sent to the Associate Editor for his consideration. Finally the Associate Editor sends all the material (original and final version, review and response, plus his own recommendation) to the Editor, who authorizes the publication and assigns it a serial number.

The Managing Editor and the Technical Officer ensure that all published papers are uploaded in the international RepEc public-action database. On the other hand, it is up to the corresponding author to make direct contact with the Departmental Secretary regarding the offprint order and the research fund which it should refer to.

Ricardo Alberto MARQUES PEREIRA
Dipartimento di Informatica e Studi Aziendali
Università degli Studi di Trento
Via Inama 5, TN 38100 Trento ITALIA
Tel +39-0461-882147 Fax +39-0461-882124
E-mail: ricalb.marper@unitn.it

Patent Disclosure and R&D Competition in Pharmaceuticals

Laura Magazzini[*], Fabio Pammolli[†],
Massimo Riccaboni[‡], Maria Alessandra Rossi[§¶]

April 18, 2009

Abstract

The prominent role played by patents within the pharmaceutical domain is unquestionable. In this paper we take an unusual perspective and focus on a relatively neglected implication of patents: the effect of patent-induced information disclosure (of both successes and failures) on the dynamics of R&D and market competition. The study builds upon the combination of two large datasets, linking the information about patents to firm level data on R&D projects and their outcome. Two case studies in the fields of anti-inflammatory compounds and cancer research complement our analysis. We show the important role played by patent disclosure in shaping firms technological trajectories through the possibility of reciprocal monitoring in a context of parallel research efforts, and suggest the importance of enhancing the diffusion of information concerning failures, not only to avoid wasteful duplication of innovative efforts, but also as a tool for the identification of promising research trajectories.

Keywords: Patent disclosure, Innovation, R&D competition.

JEL Classification: D23; D83; O34.

[*]Department of Economics, University of Verona
[†]University of Florence and IMT, Lucca
[‡]DISA, University of Trento
[§]CLEIS, University of Siena

[¶]This paper is the result of the "R&D competition" research project carried out jointly with Adrian Towse and Martina Garau of the Office of Health Economics, London, UK. A preliminary draft of the paper has been presented to the DRUID Summer Conference 2006 (Copenhagen), and to the 11th ISS Conference (Sophia-Antipolis).

1 Introduction

The pharmaceutical industry is a textbook example of a science based sector characterized by high R&D cost, uncertain and spillovers for which patent protection assures appropriability thus providing incentives for innovation. Indeed, Mansfield (1986) found that, absent patent protection, 60 percent of pharmaceutical inventions would not have been developed and 65 percent would not have been commercialized. Since then, the market segment for patented pharmaceutical products has gained a greater relevance than in many other industries. Moreover, patents play a crucial signalling role for venture capital and big pharmaceutical companies' investment decisions and dealmaking in markets for technologies. Against this background, the relationship between intellectual property rights, innovation and public health is at the heart of a blossoming international debate (WHO 2006). A recent empirical test of the "tragedy of the anticommons" (Heller and Eisenberg 1998) shows that, although modestly, upstream patenting in biotechnology might hinder the diffusion of scientific knowledge (Murray and Stern 2007).

This paper aims at contributing to the pharmaceutical patent debate by focusing on a relatively neglected function of patents, which was already characterized by Kenneth Arrow in his seminal contribution (Arrow 1962), that is information disclosure. In particular, we focus on the effect of patent-induced disclosure of information on the dynamics of both R&D and market competition. Most of the literature on patent disclosure has highlighted the trade-offs involved by the fact that patents tend to reveal "positive" information on a firm's technological advancements that might be prone to imitation by competing firms.

This paper, by contrast, identifies the crucial role patents play in providing negative information[1]. The value of negative information is particularly important in light of the fact that it wouldn't be otherwise available to competing firms due to the strong publication bias against negative results of clinical trials (Zarin and Tse 2008, Johnson and Dickersin 2007, Chan et al. 2004). Indeed, while information on biological and therapeutic properties of a pharmaceutical product can be easily obtained when it reaches the market, in most cases firms do not publicly disclose the reasons behind their failures, making the associated patent a fundamental source of information available.

In order to address this issue, the paper relies on the combination of two large datasets: the first includes all pharmaceutical and biotechnology

[1] See also David et al. (1992), Orsenigo et al. (2001).

patents granted by the USPTO since 1965, while the second comprises firm data at the level of specific R&D projects. Once integrated, the two dataset allows a comparison of the patterns of citations received by, on one side, patents associated to successful projects, i.e. those that led to a marketable product and, on the other side, patents associated to failed projects, i.e. projects that were discontinued in clinical trials. Citations, in turn, are taken as a measure of knowledge utilization and spillovers. One of the most interesting findings of the analysis is that, although discontinued patents do receive a lower overall number of citations, a large share of them continues to receive citations by other companies even after the project has turned out to be a failure, which suggests that the information these patents provide is relevant to competing firms.

To further corroborate our results, we present two case studies, based upon throughput analysis of patent documents and of related scientific literature and discussions with industry and scientific experts. The first case concerns anti-inflammatory drug development (p38 MAPK inhibitors), whereas the second one study the development of a family of anti-cancer drugs (DNA topoisomerase inhibitors). Both of them show patent disclosure's role in shaping firms' technological trajectories through the possibility of reciprocal monitoring they open up in a context characterized by parallel research efforts.

The case studies have been selected on the basis of the following criteria:

1. *Relevance.* The first case study on p38 MAPK inhibitors describe the efforts of almost all big pharmaceutical companies and biotech firms to find a safer and more powerful alternative to COX-2 inhibitors in the anti-inflammatory market that was worth more than 15 billion dollars after Vioxx withdrawal in 2004. The second case shows the development of inhibitors of the topoisomerases which have been useful in treating cell proliferative conditions, in particular, human cancers. Topoisomerases and has been clearly identified as a validated molecular target for a variety of widely prescribed anticancer drugs and collectively, the topoisomerase inhibitors comprise 6 percent of the total world market for cancer drugs in chemotherapy.

2. *Patent value.* Early patents are highly cited even if the associated compounds failed to reach the market. Fierce competition and lack of immediate feedbacks has "forced"companies to enter in the arena on the unique basis of the most recent scientific advances (patent disclosure) and projections without waiting for the final outcome of the

research on the first innovative compounds. In both cases there is no publicly available information of the reasons of failure of R&D projects.

3. *Parallel, ongoing R&D efforts.* In the first case, out of 68 candidate compounds half have been discontinued, half are still active and no drug is available on the market yet. In the second case, despite there are two drugs on the market they have heavy side effects and out of 90 R&D projects focused on the development of topoisomerase I inhibitors more than 60 percent are still active.

The results of this paper thus contribute to characterize pharmaceutical innovation as a domain characterized by "races" for reaching the market in which competitors pursue parallel research trajectories learning from both each other's successes and failures. However, building on a success does not lead to an increase in the probability of success, whereas we are not able to provide a clear-cut answer for projects building on a failure. Even though not precisely estimated, patents building on a previous failure experience a lower probability of success, but the reverse is true if the citation is made after the outcome of the project becomes known.

The paper is organized as follows. Section 2 provides a brief overview of the role of the patent system in information disclosure by identifying some elements of the economics debates surrounding the effectiveness of patent-induced information disclosures. Moreover, it highlights a few reasons why, *prima facie*, patent disclosure might be more important in the pharmaceutical than in other sectors. Section 3 describes the data, methods and summarizes empirical results. Section 4 presents the two case studies in the anti-inflammatory and anti-cancer research fields providing evidence of the role of disclosure in shaping the competitive environment. The final session presents a general discussion and draws some implications for public policy.

2 Patents and Information Disclosure

The patent system is meant to perform a disclosure function that is surprisingly neglected in recent theoretical analyses but enjoys a relatively high popularity with courts. In exchange of exclusive rights over inventions, patent-holders are required to disclose their protected inventions to the public so as to allow an effective diffusion of technological knowledge. This exchange is often referred to as a bargain between inventors and the State and it is in fact an inherent feature of the dual nature of patents (Arrow 1962). According to this view, patent protection increases the avail-

ability of scientific and technological knowledge that would otherwise be kept secret, inducing both direct benefits in the form of increased knowledge diffusion and indirect benefits in the form of a reduction of wasteful duplication of innovative efforts.

The general principle of patent disclosure is defined differently in different legal contexts, although its essence tends to be the same in any patent system[2]. The common principle of "enablement" requires inventors to disclose enough information to enable anyone skilled in the art to practice and reproduce the invention. This has become a worldwide minimum standard of adequacy of disclosure (Reichman 1995)[3].

The exact content of the disclosure requirement is difficult to spell out. Multiple doctrines have developed in different jurisdictions in order to clarify the implications of the disclosure requirement, not leading, however, to the emergence of an agreed-upon standard for disclosure. For instance, while in the US the legal standard involves a "best mode" requirement, i.e. patent applicants have to provide the information available at the time the application is filed on the best way to carry out the invention, no such requirement is explicitly provided for by EU law.

The patent disclosure function is surrounded by some degree of controversy also from an economic standpoint. On the critical side, patent disclosure function is held to be very limited because of patent applicants' incentives to withhold as much information as possible. Also, the concrete availability of the option to the keep the invention secret has been questioned (Plant 1934) and the hypothesis that only inventions that cannot be kept secret are in fact patented has been historically advanced (Machlup and Penrose 1950, Bessen 2005), pointing to the fact that patents, after all, do not facilitate the circulation of information that wouldn't otherwise be available. Finally, some theoretical models of patenting behavior have also highlighted the possibility that the disclosure obligation might induce inventors not to patent (Horstmann et al. 1985, Scotchmer and Green 1990).

On the positive side, some insights on the social value of patent disclosures have recently been offered by authors emphasizing the effects of dis-

[2]Art. 29(1) of the TRIPs Agreement states that "[m]embers shall require that an applicant for a patent shall disclose the invention in a manner sufficiently clear and complete for the invention to be carried out by a person skilled in the art...". Provisions similar to the one set forth in the TRIPs Agreement appear, for instance, in art. 83 of the European Patent Convention and in 35 U.S.C. §112, para. 1.

[3]This suggests the existence of some correlation between the scope of disclosure and the scope of claims, although it seems that a consensus is emerging on both sides of the Atlantic on the fact that this correspondence should not be considered excessively strict (Janis 2000).

closure on rent-seeking behavior and as a means to convey information on new uses of a given technology (Landes and Posner 2003). Moreover, some attention has been devoted to the role patents play in *indirectly* promoting information disclosure, with a special twist on strategic aspects (Anton and Yao 2003, Baker et al. 2005, Lichtman et al. 2000, Parchomovsky 2000). Indeed, strategic considerations may induce prospective patentees participating in a race to disclose valuable research results before their competitors apply for a patent for a variety of reasons, including the aim to foreclose the latter possibility and the aim to delay the end of the race by narrowing the extent of the "inventive step" that current inventions enjoy over existing prior art.

Differently from prior literature, our focus in this paper is on the *direct* effect exerted by patents in inducing information revelation through the fulfillment of the so-called "enablement requirement", e.g. the textbook disclosure function of the patent system.

The effectiveness of patents' disclosure function is, ultimately, an empirical question (Cohen et al. 2000, Levin et al. 1987).

Taking a different perspective, studies have also explored whether information disclosed through patents and available in patent databases are of any use to firms in different sectors. In principle, patent-induced information might be useful to firms in various ways: as means to monitor technological advances in their own sector; as a way to identify new applications of existing technologies in fields unrelated to the one in which they were developed; and as a way of gathering relevant legal information, such as information on the likelihood that one's own patent infringes someone else's patent or viceversa. In practice, patent databases are rarely consulted for reasons other than legal purposes in most industries (Oppenheim 1998, Tang et al. 2001). This holds particularly for small and medium enterprises (SMEs), due to the high costs involved in expert consultation of patent databases.

Against this background, the pharmaceutical industry is characterized by: a strong link of innovative activity to its scientific underpinnings, a high degree of cumulativeness at the sector level and a large presence of R&D spillovers, a remarkably high degree of R&D intensity, and high uncertainty both on the R&D and market sides[4].

[4]A rather uncontroversial account of the success rate of innovative efforts in the pharmaceutical industry reveals that less than 1 percent of the new compounds object of preclinical investigation reach the human testing stage and that only around 20 percent of compounds going through the clinical trials ultimately gains FDA approval (DiMasi 1995). As for market uncertainty, a hint of its relevance comes from observing the distribution of returns to R&D for new drug introductions, which is highly skewed, with few blockbuster

While the limited empirical evidence available to date supports at best a marginal role of patent-related information in most technological sectors, there are at least three reasons to think that the relevance of patent-induced information disclosure in pharmaceuticals is much higher than elsewhere. The first is that patents do play a much greater role overall in the pharmaceutical domain as compared to almost any other technological domain, as extensively documented by numerous empirical accounts of appropriability conditions in a range of sectors (Levin et al. 1987, Cohen et al. 2000).

The second reason is related to the importance of patents for the division of innovative labor between public research institutions, biotech companies and pharmaceutical corporations and market for technologies in general (Arora et al. 2001, Orsenigo et al. 2001). Innovation in this domain involves a range of actors, characterized by different motivations, incentives and ethos, especially with regard to the disclosure and diffusion of scientific information, with the consequence that patent disclosure becomes a particularly crucial means of bridging the gaps across the different innovative *milieu*. A link could thus be traced between the structure of innovative activity and the relevance of patent disclosures.[5]

The third reason is even more compelling and is more strictly related to the nature of R&D competition in pharmaceuticals that is the object of this paper. A number of important evolutionary trends have fundamentally reshaped the pharmaceutical industry in the past thirty years, strengthening the interactions between basic science and product development, with advances in physiology, pharmacology, enzymology, cell biology and later molecular biology strongly affecting the patterns of technological development (Gambardella 1995, Henderson et al. 1999).

The connectedness of drug development to its scientific underpinnings has at the same time increased the range of scientific opportunities available to players in the industry and increased the likelihood that firms pursue "parallel" trajectories of development, simultaneously working on compounds belonging to the same therapeutic class. Different firms pursue

drugs accounting for most of the returns (Grabowski et al. 2002). Moreover, while the degree of cumulativeness is overall high in the sector, cumulativeness at the firm level is low. Indeed, once a firm introduces a new product of major therapeutic value, its rivals explore the new line of research trying to develop similar or related drugs, and, in this race, the discovering firm seems not to have an advantage in discovering chemically related drugs. Sutton (1998) analyzed the top 50 selling drugs in 1960, 1973, and 1986 and provided evidence of high instability in the leadership within therapeutic classes.

[5]The evidence presented in this paper looks at this link but highlights the reverse direction of causation, the one running from patent disclosures to the structure of innovative activity.

alternative approaches to drug development that represent a particular instantiation of the inflow of scientific opportunities opened up by basic research results and, at the same time, contribute to validate and extend such results, in the context of a process that cannot aptly be characterized as linear (Orsenigo et al. 2001). This magnifies the impact on research productivity of knowledge spillovers across firms, as confirmed empirically by Henderson and Cockburn (1996), leading firms to actively seek the exploitation of such spillovers.

Patents play a key role in the context of parallel development described above, where firms' monitoring of competitors' achievements can rely on few means other than patent-induced information disclosures. This is true in two respects. On one side, firms learn from each other's successes, i.e. patent-disclosed information offers guidance on the biological and therapeutic properties of any product that reaches the market. On the other side, information conveyed by patents might turn out to be useful also when it concerns projects that are discontinued. Indeed, the innovation process in pharmaceuticals can be though of as a trial-and-error process, where firms learn from failures as they do from successes. This question – that of the role played by research failures in the pharmaceutical innovation race – has been so far neglected in the literature and constitutes the focus of this paper. In fact, this sort of information might be considered particularly important in light of the fact that in most cases firms do not publicly disclose the reasons behind their failures, making the associated patent the unique source of information available.

Disclosure function is guaranteed through scientific publications in peer reviewed journals (open science), patents (commercial science) or by patent-paper pairs (Murray and Stern 2007). However a vast literature has documented a strong bias in open science toward the publication of positive results. In the following we document the value of failures in pharmaceutical R&D. Since innovation in pharmaceuticals builds upon failures as well as on successes, we conclude that patents have a role in disclosing information about unsuccessful drug candidates which is not easily replicable by open science. Thus we suggest to extend compulsory disclosure for patented drugs to clinical trial results.

3 Data, Methods, and Results

For the sake of this project we have brought together two large datasets and linked them through an elaborate matching process: the first comprises

all pharmaceutical and biotechnology patents granted by the USPTO since 1965, including their patent citations[6]; the second comprises firm data at the level of specific R&D projects drawn from the Pharmaceutical Industry Database (PHID) maintained at the University of Siena, comprising the development history of all pharmaceutical R&D projects worldwide undertaken during the last 25 years[7].

Our analysis relies upon about 2,000 drug candidate-patent pairs, classified according to the final outcome of the R&D project (marketed or discontinued). In our sample 58 percent of patents are associated to discontinued compounds, whereas 42 percent are associated to a patent that has been successful in reaching the market[8].

Following the NBER tradition[9], we employ patent citations as an indicator of knowledge utilization and spillovers. The key assumption is that a citation made to a previous patent denotes a knowledge transfer from the cited patent to the citing one: more frequently cited patents contribute to a larger share of subsequent innovations[10]. Since we focus on R&D competition in the pharmaceutical industry, only citations from patents in the pharmaceutical domain have been taken into account. We distinguish self-citations from citations made by other companies, as they provide different indications about the nature of technology, appropriability conditions and knowledge spillovers. On the one side, self-citations are indicative of research trajectories strongly rooted within the firm/institution boundaries. On the other side, citations by firms/institutions other than the original innovator have been fruitfully employed in tracking knowledge spillovers (Hall et al. 2001, Jaffe et al. 2000).

First, we compare the raw number of citations received by marketed and discontinued projects.

Figure 1 plots the average citation functions, i.e. the likelihood that a

[6]US patents in selected IPC and US classes are included in the database.

[7]For a subset of projects, the database lists the patents protecting the compound under study. Old molecules and/or natural products, which do not have any associated patent have been omitted. For projects listing a patent granted by a patent office other than the USPTO, we considered the US patent in the same family as the one listed in the database. In case no US patent is identified, the record is not considered in the analysis.

[8]We refer to marketed/discontinued patents as the patents associated to marketed/discontinued R&D projects.

[9]See Jaffe and Trajtenberg (2002) and the literature referenced therein.

[10]We are aware that patent citation count is only a noisy proxy of the relevance of the knowledge disclosed, since citations might be included for strategic purposes or added by firm's lawyers or by patent examiners (Alcacer and Gittelman 2004). However, survey evidence shows that, even if noisy, patent citations are indicative of knowledge spillovers and communication among inventors (Jaffe et al. 2000).

patent will receive a citation as a function of the time elapsed from grant date, computed on the basis of observed data[11].

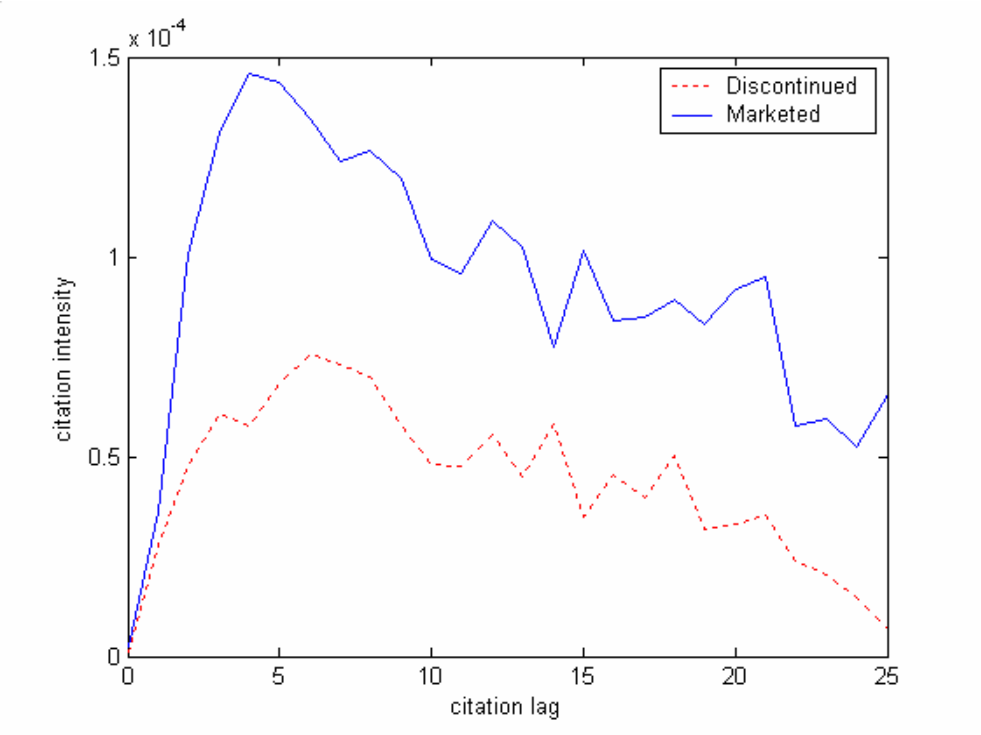

Figure 1: Observed average citation lag distribution

Consistently with previous literature showing that the number of citations received by a patent is positively associated to its value[12], discontinued patents receive, on average, a number of citations that is lower than the number of citation received by patents associated to marketed projects.

The distinction in the chemical journals between leadlike and druglike compounds is useful in interpreting the results. The failed project might be indicative of a dead end, and therefore the citation is the result of a negative outcome. But it could also be the case that the failed patent protects a compound, which could be either toxic or ineffective, but that presents good (even though sub-optimal) target binding affinity properties

[11] See Jaffe and Trajtenberg (1996) for details.

[12] Trajtenberg (1990); Lanjouw and Schankerman (1999); Harhoff et al. (1999); Jaffe et al. (2000); Trajtenberg et al. (1997); Jaffe and Trajtenberg (2002).

and is the antecedent of a new set of molecules that build around the failed one. Even though the compound provides no returns to the innovating firm, a positive social value is associated with the identification of the new mechanism of action or of the new compound, or on the negative side, of its ineffectiveness or its toxic effects, therefore pointing to research trajectories that shouldn't be pursued to avoid waste of resources.

Next, in order to analyze the patterns of technological competition, the role of information provided by science and R&D outcomes, we separately consider the pattern of citation (as a proxy for the diffusion and utilization of knowledge) before and after the outcome of the R&D project associated to the patent becomes known[13], i.e. the R&D project is either discontinued or ends with the launch of a product on the market.

By comparing the share of self-citations and the share of citations received by other companies, interesting conclusions can be drawn about the nature of the R&D competition in pharmaceuticals.

The first column of Table 1 reports the number (and shares) of discontinued and marketed patents respectively. Our sample comprises 1,243 discontinued patents and 897 marketed patents.

Next, we report the share of citations received, respectively, by discontinued and marketed patents after the project is terminated (either marketed or discontinued) with respect to the total number of citations received. The Table reports the figures both for citations by other firms and self-citations[14].

	N (%)	% citation after the outcome		avg. number of early citations	
Outcome		self-cits.	cits. by other	self-cits.	cits. by other
Discontinued pt.	1,243 (58%)	19.9	45.7	0.88	1.68
Marketed pt.	897 (42%)	43.1	63.3	0.87	1.95

Table 1: Share of citations received after the outcome of the projects, and early citations (within five years from patent application)

In the case of discontinued patents, about 80 percent of self-citations

[13]We compare the date of termination of the project (either marketed or discontinued) with the application year of the citing patents. Need it here to mention that average time to market or to discontinuation is not substantially different, being equal to 7.8 years for discontinued R&D projects and to 8.3 years for marketed R&D projects.

[14]100 represents the total number of observed citations (respectively, self-citations and citations by other companies/institutions) for each patent. For example, 19.9 percent of the self-citations made to discontinued patents take place after the outcome of the project becomes known. Put it differently, 80.1 percent of the self-citations made to discontinued patents take place before the project is terminated.

is made before the compound is known to be a failure, while the share decreases to 60 percent for marketed patents, i.e. a large part of patents associated to discontinued R&D projects are abandoned by the innovating firm after the properties of the associated compound are understood, but they still represent the ground for subsequent innovation by other companies: 45.7 percent of citations received by discontinued patents and made by a company other than the original innovator take place after the time the project is discontinued. The figure is higher for marketed patents, but it is not surprising that successful compounds, leading to products that are commercialised on the market, continue to induce research (i.e. to receive citations) after the commercialization of the compound. Available evidence shows that after the introduction of a new product of major therapeutic value, the rivals of the innovating firm explore new lines of research trying to develop similar or related drugs, therefore leading to citations of the patent protecting the original compound (Sutton 1998). On the other side, it is of interest to understand the rationale behind the citations to discontinued compounds, still taking place after the compound under study is known to be toxic or ineffective for the originally targeted disease.

In most cases failures are not the subject of publications, therefore very limited information is available about the reasons behind the discontinuation of the research related to the compounds. The only information in the public domain about the characteristics of the compound under development can be found in the patent(s).

Next we take into account the average number of citations received in the early years after patent application, before the termination of the project. A time frame of five years from the application year is considered. Since the time elapsed from discovery is limited, only few information are available about the properties of the protected compound/process, leading us to expect no significant difference from the comparison of the distributions characterizing discontinued patents and marketed patents. Indeed, on average, within the first five years from the application date, discontinued patents receive 0.88 self-citations, whereas the figure is 0.87 for marketed patents. If citations by firms/institutions other than the original assignee are considered, discontinued patents receive an average of 1.68 citations versus 1.95 citations for marketed patents. A Kolmogorov-Smirnov test does not allow to reject the null hypothesis that the two distributions are drawn from the same underlying population, even though a larger share of discontinued patents receive zero citations. This result, coupled with previous findings (Figure 1), suggests that differences in citation behavior between successes and failures are driven by post-outcome behavior, i.e. from the

citations received by the patent after the launch of the associated product on the market[15].

The disclosure of the information about the compound under study in patents and the advances in science sets the ground for a "race" for reaching the market, where competitors start exploring the new research arena pursuing parallel research trajectories even though the outcome is still highly uncertain. Competition on the R&D side in the pharmaceutical industry is substantial and firms entering the new research arena build both on failures and successes.

Does the outcome of the cited references provide information about the future outcome of current research efforts? Put it differently, do patents citing a success (failure) have a higher (lower) probability of success?

Table 2 reports the average probability of success of patents classified according to the outcome of their backward references[16]. We distinguish four categories: (i) 1,579 patents with no information about cited patents; (ii) 167 patents citing at least one previous failure; (iii) 264 patents citing at least one previous success; and (iv) 47 patents citing both previous successes and previous failures. For each category we computed the average probability of success by looking at the share of successes, i.e. at the ratio between marketed patents and the total number of patents in each cell.

	Citing previous failure?		
Citing previous success?	No	Yes	Total
No	38.32	13.77	35.97
	(1,579)	(167)	(1,746)
Yes	66.29	42.55	62.70
	(264)	(47)	(311)
Total	42.32	20.93	40.00
	(1,843)	(214)	(2,057)

Table 2: Probability of success of patent-projects building on failures/successes (number of patent-projects considered for computation in parenthesis)

On average, patents citing previous failure have a lower probability of

[15]Further research is needed to properly address this issue. Moreover, in order to distinguish "real" knowledge spillovers, it would be useful to distinguish the citations added by the patent examiner.

[16]We further restrict our sample and discard patents applied for before the year 1980, and patents assigned to individual inventors, leaving us with a sample of 2,057 patent-projects.

success, whereas patents citing a previous success have a higher probability of success. The patents included in the analysis have a probability of reaching the market of 40.00 percent. The rate of success changes substantially if the patent cites at least one previous success (62.70%), whereas it almost halves if the patent builds on at least one previous failure (20.93%). On average, 38.32 percent of patents with no information about the outcome of their backward references has reached the market, and the figure increases to 66.29 percent for patents citing at least one previous success and no previous failures. On the contrary it is equal to 13.77 percent for patents citing previous failures and no previous success.

Results in Table 2 seems to suggest that patents building on previous success have a higher probability of success, and the probability of failure substantially decreases if the patent builds on a previous failure. However, we need to take into account the fact that the innovator selects the basis of his/her innovation and we expect "good" compounds to build on "good" research, whereas "bad" compounds are more likely to build on "bad" research. Therefore, results in Table 2 might be driven only by the selection capabilities on the side of the innovator. In order to disentangle this issue we estimate a model where the dependent variable is an indicator of the success of the project (i.e. it is equal to 1 if the patent/project is successfully marketed, and 0 if the patent/project is discontinued), and the independent variables aim at capturing the relevant characteristics of the cited patents, i.e. of the research the patent builds upon. Explanatory variables included in the analysis are described in Table 3, along with their mean and standard deviation.

The main variables of interest are pt_succ and pt_fail, two dummy variables indicating respectively patents building on previous successes and failures. The dummies pt_succpst and pt_failpst further distinguish those patents by taking into account the timing of the citations: they only consider citations to previous successes or failures made when the outcome of the cited patent has already been disclosed (i.e. the compound is either launched on the market or announced to be discontinued). The variables pt_selfc, pt_orig, pt_science, and pt_timeb aim at capturing major characteristics of the innovation and are computed on the basis of backward citations. These are built as described in Trajtenberg et al. (1997) but only taking into account pharmaceutical citations.

A higher share of self citations (i.e. a higher pt_selfc) indicates a higher level of appropriability of the research on the side of the innovator, as self-citations are indicative of the cumulative nature of the technology and a measure of the extent to which innovators are able to reap the benefits of

Variable	Description	Mean	StdD.
Explanatory variables			
pt_succ	1 if the patent-project cites a previous success, 0 otherwise	0.15	0.36
pt_fail	1 if the patent-project cites a previous failure, 0 otherwise	0.10	0.31
pt_succpst	1 if the patent-project cites a previous success (after the outcome is known), 0 otherwise	0.06	0.24
pt_failpst	1 if the patent-project cites a previous failure (after the outcome is known), 0 otherwise	0.03	0.16
pt_selfc	share of self citations[a]	0.15	0.30
pt_orig	index of originality of the patent[a]	0.44	0.37
pt_science	share of references to non-patent literature[a]	0.29	0.33
pt_timeb	average time lag (computed on backward citations)[a]	5.09	4.32
ass_coree	share of assignee's patents within the same IPC of the patent	0.11	0.21
trend	time trend (1 if application year is 1980 to 21)	9.83	4.71
Instruments			
pt_importbi	importance (in terms of received citations) of cited patents[a]	75.95	505.4
Nfailure	number of previous failed patents	646.0	441.3
Nsuccess	number of previous successful patents	580.2	239.6

[a] Please refer to Trajtenberg et al. (1997) for a detailed description.

Table 3: Description of the variables included in the regressions

their own research (Hall et al. 2001). The index of originality of the patent (pt_orig) measures the breadth of the technological roots of the innovation: higher value of the index indicate that the patent under analysis builds on previous patents spanning many different IPC classes. pt_science aims at capturing the reliance of research on scientific sources: this is the share of references to non-patent literature, as a percentage of the total number of references (both patents and articles in scientific journals). As it is well documented in the literature, the innovation process in the biopharmaceutical industry relies heavily on the advances in "basic science", i.e. in the basic understanding of the mechanisms characterizing the targeted diseases. A higher share of non-patent literature is associated with a higher degree of basicness of the innovation and, likely, a stronger linkage with public re-

search organizations, whose advances are typically disclosed trough scientific publications.

The last variable built on the basis of patent information, pt_timeb, measures the time distance between the citing and the cited patents. The higher pt_timeb, the older the sources the patent builds upon, therefore the wider the knowledge base available to the innovating firm.

In order to control for firm capabilities in the technology class of the innovation, ass_core measure the share of assignee's patents within the same IPC of the patent under study.

Finally a time trend is included in order to take into account the increase in attrition rates that has been characterizing pharmaceutical research over the Nineties (Mervis 2005).

Results are reported in Table 4. A probit model is applied, and we also consider an instrumental variable (IV) approach in order to solve the endogeneity issue concerning pt_succ and pt_fail[17].

As expected the time trend is negative: younger patents exhibits lower probability of success, coherently with the increase in attrition rates that is characterizing the industry.

Coherently with the results in Table 2, standard probit analysis highlight a positive association between the success of previous cited patent references and the patent under study. Results change substantially when the IV approach is considered. The estimated "IV-Probit" coefficient of pt_succ is no longer significant, pointing to the fact that success does not breed success. With regard to failures, no clear-cut picture emerges from our analysis. When all citations are considered, building on a previous failure substantially reduces the probability of success, even though the coefficient is not precisely estimated. If post-outcome citations to failures are considered, the coefficient is positive but only statistically significant at the 10 percent level. All in all, building on a previous success does not assure a higher probability of success, whereas evidence about patents that build on previous failures is mixed.

As far as the characteristics of the research the patent builds upon, only pt_science exerts a positive and significant effect on the probability of success. Not surprisingly, given the characteristics of the innovation process in pharmaceuticals, stronger linkage with basic science increases the probability of success.

[17]The IVs selected for pt_succ and pt_fail are a measure of importance of the backward citations (pt_importb), the number of previous failed patents (Nfailure), and number of previous successful patents (Nsuccess).

Variable	Probit	IV-Probit	IV-Probit
pt_succ	0.7740***	1.0615	
	(.0876)	(1.090)	
pt_fail	-0.6217***	-4.3346**	
	(.1103)	(1.879)	
pt_succpst			1.2966
			(1.3984)
pt_failpst			5.6629*
			(2.9359)
pt_selfc	0.0230	0.4413	-0.2414
	(.1104)	(.3738)	(.2054)
pt_orig	-0.2806***	-0.4316*	0.0713
	(.0996)	(.2537)	(.2399)
pt_science	0.3653***	0.6195***	0.2778*
	(.1091)	(.2251)	(.1468)
pt_timeb	0.0146*	0.0261	-0.0160
	(.0084)	(.0267)	(.0240)
ass_coree	0.6345***	0.5734**	0.8123***
	(.1454)	(.2048)	(.2091)
trend	-0.0550***	-0.0302**	-0.0902***
	(0.0071)	(.0152)	(.0162)
constant	-0.0333	-0.0235	0.1702
	(.0856)	(.1686)	(.1338)

*** 1% - ** 5% - * 10% statistical significance

Table 4: Estimation results

Also ass_core exerts a positive and statistically significant effect: a higher probability of success characterizes patents where the innovator has previous substantial knowledge as measure by the share of assignee's previous patents within the same IPC of the patent, and this might be the result of wider experience in the developed technology.

Summing up, the analysis reveals that patents represent an important source of information for monitoring the R&D activities undertaken by competitors and provide a spur to innovative efforts by other firms in related fields or in the same area of application of the original patent. Nonetheless building on a previous success does not assure a higher probability of reaching the market.

The case studies presented in the next Section will allow us to further in-

vestigate the dynamics of technological competition arising from knowledge disclosure in patents.

4 Case Studies

Two case studies in the fields of inflammation and cancer are discussed. The analysis presented in this Section are based upon literature search and patent text analysis. Certainly, results of these studies cannot be extended beyond their field of application and generalized to pharmaceutical innovation. Nonetheless, they provide some hints on the nature of R&D competition in pharmaceuticals. Typically pharmaceutical research advances through a process of trial-and-error, where both successes and failures play a role in guiding subsequent innovative efforts. Parallel research trajectories are pursued as the result of fierce competition driven by science in molecular biology and by the disclosure of relevant information through patents and scientific publications.

In the MAPK p38 case, a vast array of research efforts were building on pioneer compounds, highly cited, that never reached the market. In the case of DNA topoisomerase I inhibitors, knowledge disclosed in patents of compounds that never reached the market contributed to successful development of follow-on drugs by firms other than the original innovator. Next, more effective drugs with less side effects have been developed based on knowledge about first-in-class compounds.

The distinction between druglike and leadlike compounds is useful in understanding the pattern described, especially in the p38 MAPK case. The two set of compounds differ in terms of chemical characteristics (Oprea et al. 2001). The analysis of these characteristics is of little interest for our purposes, nonetheless the distinction between leads and drugs can help us to understand the patterns of R&D competition highlighted in this paper. Lead structures typically do not exhibit optimal target binding affinity, nonetheless they have characteristics that make them starting points in medicinal chemistry efforts, and patent citation data offer a way to track subsequent research efforts. Indeed, the pioneer compound analyzed in the p38 case study, i.e. SB-203580, identified as a lead by Oprea et al. (2001), never got to the market but is highly cited.

4.1 MAPK p38 inhibitors

The p38 mitogen activated protein kinase (MAPK) is a serine-threonine kinase which regulates the production of pro-inflammatory cytokines, such

as IL-1 and TNF-α. These cytokines play a central role in the body's inflammatory response. Excess levels of IL-1 and TNF are associated with a broad range of acute and chronic inflammatory diseases such as rheumatoid arthritis, osteoarthritis, osteoporosis, inflammatory bowel disease, asthma, atheroscleresis, and cachexia. As a whole, the inflammatory disease US market is expected to reach a value of 8 billions of USD in 2007. The potential market size, high competition and rapid advances in this field have urged companies to base their decisions uniquely on the most recent discoveries and projections, even if no compound targeting p38 has reached the market yet. This has led to a vast array of drug candidates targeting p38, which have failed to reach the market but whose associated patents are highly cited.

The most widely studied class of p38 MAPK inhibitors are the vicinal aryl/ pyridine-4-yl heterocycles. Anti-inflammatory activity for this structural class was first reported in the early Eighties in two patents[18] by GlaxoSmithKline (GSK), the first mover in this field. The molecular target has been unravelled only later in 1994 by GSK researchers (Lee et al. 1994).

Since then, as demonstrated by the number of R&D projects and submitted patents, several big pharmaceutical companies and emerging biotech firms have undertaken research programs focused on the p38 MAPK. At present, more than 60 research programs have been started though a large share has been discontinued before last phases of clinical trials and no compound has reached the market so far.

GSK has been leading the way by reporting 2,4,5-triarylimidazole inhibitors and filing an extensive array of patents claiming compounds based on optimization of the pyridylimidazole template represented by SB-203580, the first p38 inhibiting compound in preclinical trials. Even if the lead compound SB-203580 showed to be a very selective inhibitor of p38 MAPK occupying the ATP-binding site, its development was discontinued in 1998 probably due to potent inhibition of hepatic cytocrome P450 isoenzymes, posing toxicological problems in drug development.

In the following years, numerous replacements for the imidazole scaffold of the original lead were disclosed in the patent literature[19]. Most of the research projects based on these compounds have been discontinued.

After several compounds were patented with structural homology to SB-203580, structurally different p38 inhibitors were disclosed by different com-

[18] US4175127, WO8801169.

[19] Oxazole: WO9513067, US5559137. Pyrazole: WO9705877, WO9705878. Substituted pyrazole: WO9852937, WO9852940, WO9852941, WO9852558.

panies, e.g. Vertex and Boheringer Ingelheim, both citing patents by GSK claiming imidazole as the central ring.

Vertex was the first company to develop a molecule belonging to the class of nitrogen containing heterocyclic compounds. VX-745, the first product to enter clinical trials and to reach phase 2, showed good anti-inflammatory efficacy but was discontinued due to concerns regarding possible neurological adverse effects. Research by Vertex is still active and a back-up compound has been selected (VX-702), now in phase II clinical trial for cardiovascular disorders, inflammation, and rheumatoid arthritis.

In 1999, Boehringer Ingelheim was granted a patent claiming a diaryl urea inhibitor (BIRB-796) which binds a different allosteric binding pocket of p38 MAPK that is spatially distinct from the ATP pocket. This compound seemed to overcome the toxicity problems related to SB-203580 and structural homologues. BIRB-796 moved into phase III clinical trial for psoriasis, and was the first p38 MAPK inhibitor to reach this stage of development. However, Boehringer Ingelheim announced the discontinuation of BIRB-796 R&D project in 2005.

The citation pattern for BIRB-796 depicts an interesting trend: no citations came from other firms, while six new patents citing the original one were applied for by Boehringer Ingelheim few years before discontinuing the BIRB project, indicating active research around the original compound.

Even if GSK disclosed the first anti-inflammatory compounds targeting p38 MAP kinase almost ten years ago, the race for the first drug based on p38 MAPK inhibition is still open. Other companies have entered the field building on the pioneer patents and the information disclosed therein. This dynamic is reflected in the pattern of citations. The patents related to SB-203580 are highly cited, and still continue to receive citations both from GSK and its rivals, even though their development has been discontinued. Figure 2 reports the number of patents covering p38 MAPK compounds applied for in the US in the year 1998-2002, soon after the time the original SB compound has been discontinued (top panel). The bottom panel report the share of patents citing the original imidazole compounds by GSK: about 20% of patents still cite the original research even though the compound is known to exert toxicological effects.

4.2 DNA topoisomerase inhibitors

Camptothecin (CPT) was first discovered in a National Cancer Institute drug screen of naturally occuring agents in 1966 from the bark of the Chinese tree Camptotheca Acuminata (Wall et al. 1966). Potent cytotoxic

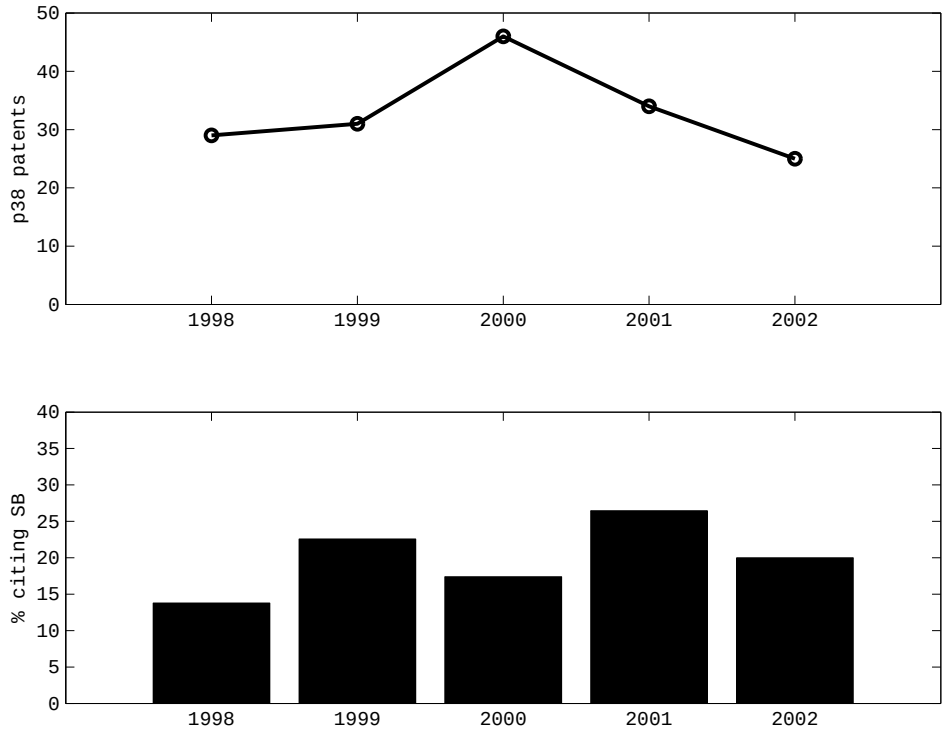

Figure 2: p38 patents: (top) number of patents applied for in the US, (bottom) share of patents citing the original SB compound

activity was immediately noticed. However, early clinical trials with the sodium salt of CPT in the 1970's were soon discontinued because severe and often unpredictable side effects (myelosuppression and haemorrhagic cystitis). Further clinical work was suspended until, in 1985, researchers of GSK in Philadelphia (US) in collaboration with the John Hopkins University discovered that human DNA topoisomerase I (Topo I) is the molecular target of CPT (Hsiang et al. 1985)[20]. Unfortunately, the lactone ring of CPT, which is necessary for a proper fit into the active size of Topo I, is readily hydrolyzed to the carboxylate inactive form at physiological pH. In addition, CPT is fairly insoluble. Thus, since then, CPT has become the prototype Topo I

[20]One of the more exciting developments in oncology has been the identification of Topo I as a molecular target of a variety of anticancer drugs. First discovered in 1971, the enzymes are referred to as topoisomerases because they are able to change the topology of DNA molecules without changing the underlying chemical structure of the DNA.

specific inhibitor and several institutions have started research projects to discover more stable and soluble CPT analogues compounds. CPT derivatives specifically target Topo I leading to premature termination of DNA replication and inhibition of transcription. Cells can repair DNA breaks caused by low doses of CPT, whereas higher doses lead to cell death. Since many neoplastic cells are characterized by high levels and activities of Topo I, this enzyme has become one of the main cellular targets for anticancer therapy. Although the pathway which leads from Topo I drug target DNA damage to cell death is not entirely clear yet, the number of patents granted each year concerning CPTs has increased steadily from 1995. Although these numbers testify the ramping interest in this field, a large number of them concern formulations, drug delivery systems and combinations with other drugs rather than innovative compounds (Dallavalle et al. 2002).

Figure 3 represents the pattern of citations received by patents that have been matched to the main camptothecin analogue compounds: Irinotecan, Topotecan, and 9-aminocamptothecin (9AC), whose chemical structures are reported in Figure 4 along with the chemical structure of camptothecin lead compound. On the x-axis time is reported from first patent grant to 2004, whereas the y-axis reports the cumulated number of citations received by the patent associated to each compound. The line markers identify the development stage in each year.

Irinotecan and Topotecan are the only two candidate drugs that have been launched worldwide. Irinotecan has been introduced in 1994 by the Japanese company Yakult Honsha KK for colon carcinoma and Topotecan by GSK for lung and ovarian carcinoma in 1996. Irinotecan is one of the most active agents gastrointestinal (GI) tumours; however, it has heavy side effects such as concurrent GI and hematological toxicities. On the contrary, the toxicity of topotecan is principally hematological and patients with renal insufficiently require significant dose reduction. In both cases, there is room for more effective and better tolerated drugs. At present, more than 90 R&D projects focused on the development of Topo I inhibitors have been started and more than 60% of them are still active.

The year 1994, when 9AC entered into human clinical trials and irinotecan reached the market, reports a sharp increase in the citation trend of the three compounds in the Figure. After this year the two series or 9AC and Irinotecan proceed paired until 2001, when the citations to Irinotecan seem to stabilize, whereas 9AC still continues to receive citations, even if the development has been discontinued. Even if not general, the pattern is interesting since the patents associated to the three compounds are deeply

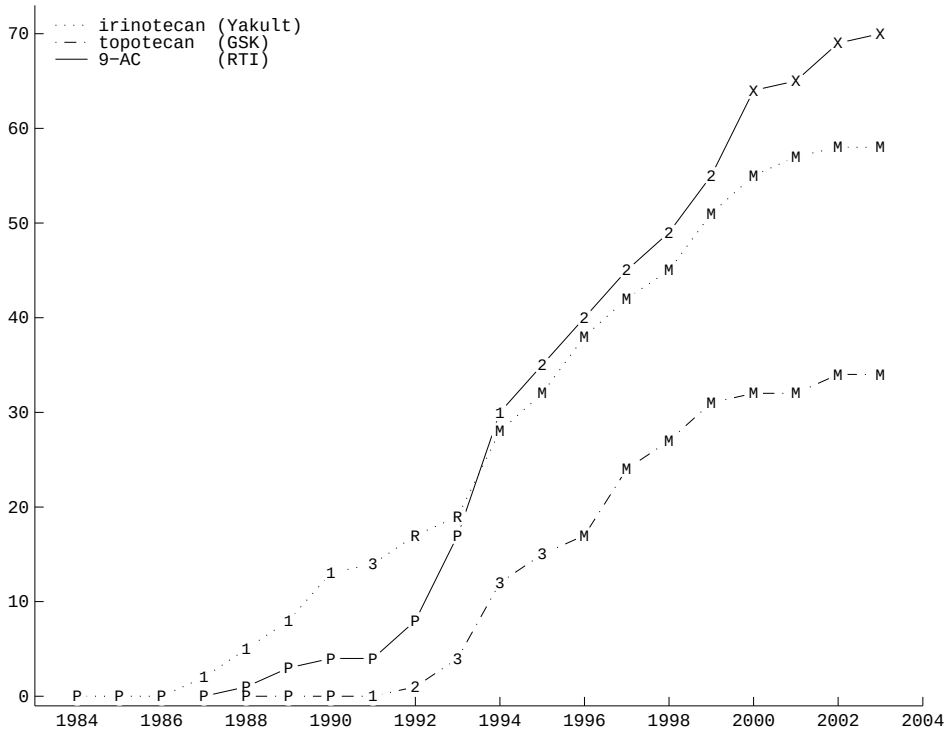

Figure 3: Cumulated number of non-self forward citations for selected Topo I inhibitors. Markers correspond to phases of development (P: discovery/preclinical; 1-2-3: clinical phases 1-2-3; R: (pre)registration; M: marketed; X: discontinued.

interwined[21]. In fact, Irinotecan is cited by 9AC, and both are cited by Topotecan.

After the commercialization of the successful compound (irinotecan), the (unsuccessful) molecule (9AC) still continues to receive citations, even after the decision to discontinue its development. Moreover, the analysis shows the high uncertainty of the drug discovery activity. Building on marketed products does not assure success. Indeed, the 9AC patent benefits from the information in the Irinotecan patent, even though the associated research has been discontinued. On the contrary, the Topotecan patent, citing both 9AC and irinotecan led to a marketed compound.

All in all, the pharmaceutical R&D process characterizes itself as a trial

[21]Check the similarity in molecular structures reported in Figure 4.

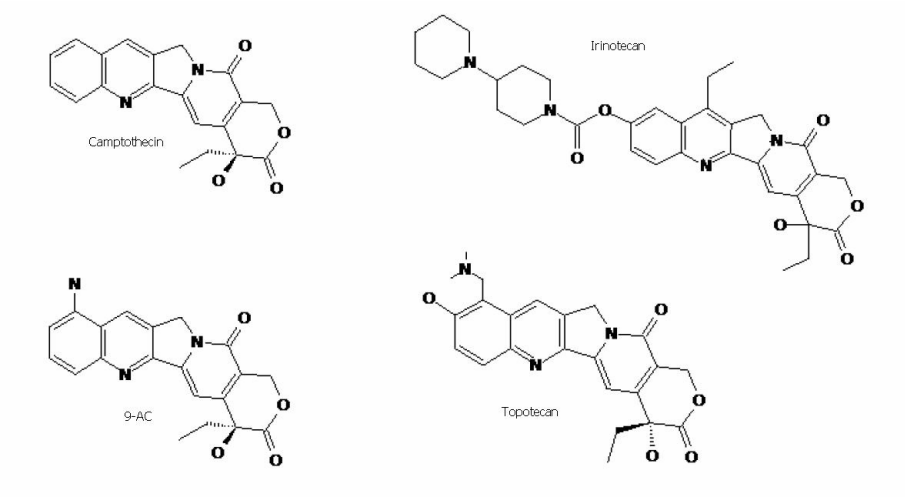

Figure 4: Chemical structure of camptothecin and selected analogues.

and error process, where both successes and failures contribute to the advances of the technological frontier. This is testified by the size and direction of knowledge spillovers, as measured by patent citations, since both successful and failed compounds still receive citations after the termination of the associated project.

5 Concluding Discussion

The evidence presented in this paper contributes to enrich our understanding of the dual nature of patents and of their fundamental contribution to information disclosure (Arrow 1962), unraveling the complexity of dynamics of knowledge production and competition in pharmaceutical R&D. The information disclosed through patents leads to an expansion of the knowledge frontier. The increase in knowledge further stimulates R&D efforts, both in terms of new patents and new firms entering the research arena, therefore stimulating competition within the industry, and fostering research in the new field.

Building on a integrated and comprehensive dataset about the innovative activities of the firms and institutions operating within the pharmaceutical domain, we are able to link a sample of patent to the development history of the protected compound, allowing us to distinguish successful patents, i.e.

patents protecting compounds that are successfully marketed, from failed patents, i.e. patents protecting compounds that are discontinued either in preclinical or clinical trials. By taking into account patent citations, we are able to assess knowledge transfers and linkages between different research trajectories. Understanding the rationale behind the citing behavior is important to understand the characteristics of R&D competition in pharmaceuticals.

Our analysis shows that a large share of citations to marketed patents take place after the compound is commercialized, proving to be safe and effective in targeting a selected disease. This is not surprising, however building on a previous success does not assure higher probability of success.

Differently from previous studies, our analysis also takes into account the role of failures in providing the ground for subsequent innovation. While information on biological and therapeutic properties of a pharmaceutical product can be easily obtained when it reaches the market, in most cases firms do not publicly disclose the reasons behind their failures, making the associated patent the sole source of information available.

Citation patterns to discontinued compounds reveal interesting differences between innovator's and rivals' behavior. A small share of self-citation is made by the original innovator to failed compound after the outcome has been disclosed, nonetheless failed patents do provide information to firms and institutions other than the original innovator also after the related research is abandoned, as a large share of subsequent citations is made to discontinued patents also after the time the outcome is disclosed.

Results about the effect on success probabilities of building on a failure are mixed, but overall our empirical analysis suggests the existence of a social value associated to information disclosure on discontinued patents, in terms of new (and sometimes better) research trajectories exploring new therapies for treating a disease. Information contained in patents would be greatly enriched by the disclosure of the information gathered during the drug development process. Once a product is approved for marketing, a wide variety of information is available about its properties. On the contrary, firm usually do not disclose information about their failures or about compound toxicity or lack of efficacy and scientific journals prefer to publish positive outcomes and, with few recent notable exceptions in the case of open-access online journals, it is extremely difficult to find negative results in peer reviewed scientific journals.

Nonetheless, these would be valuable information for avoiding waste of resources and duplication of efforts. In 2007, the U.S. Congress enacted the FDA Amendments Act (FDAAA) which expand the scope of required

registrations at ClinicalTrials.gov and provides for the first federally funded trial results database[22]. It has been argued that disclosure of clinical results could undermine competitive advantage. In this paper we show that the social value of failures in drug development is much higher than the private value and a full disclosure of clinical trial results is an important complement to the patent disclosure function.

References

Alcacer, J. and Gittelman, M.: 2004, How do I know what you know? Patent examiners and the generation of patent citations, *New York University, mimeo* .

Anton, J. and Yao, D.: 2003, Patents, Invalidity, and the Strategic Transmission of Enabling Information, *Journal of Economics & Management Strategy* **12**(2), 151–178.

Arora, A., Gambardella, A. and Fosfuri, A.: 2001, *Markets for Technology: The Economics of Innovation and Corporate Strategy*, MIT Press.

Arrow, K.: 1962, Economic Welfare and the Allocation of Resources for Invention, *in* R. Nelson (ed.), *The Rate and Direction of Inventive Activity: Economic and Social Factors*, Princeton, NJ: Princeton University Press.

Baker, S., Lichtman, D. and Mezzetti, C.: 2005, Disclosure and investment as strategies in the patent race, *Journal of Law and Economics* **48**, 173–194.

Bessen, J.: 2005, Patents and the Diffusion of Technical Information, *Economics Letters* **86**(1), 121–128.

Chan, A.-W., Hrobjartsson, A., Haahr, M., Gotzsche, P. and Altman, D.: 2004, Empirical Evidence for Selective Reporting of Outcomes in Randomized Trials, *Journal of the American Medical Association* **291**(20), 2457–2465.

Cohen, W., Nelson, R. and Walsh, J.: 2000, Protecting Their Intellectual Assets: Appropriability Conditions and Why US Manufacturing Firms Patent (or Not).

[22] See Section 801.

Dallavalle, S., Merlini, L., Penco, S. and Zunino, F.: 2002, Perspective in camptothecin development, *Expert Opinions on Therapeutic Patents* **12**, 837–844.

David, P., Mowery, D. and Steinmuller, E.: 1992, Analyzing the Economic Payoffs from Basic Research, *Economics of Innovation and New Technology* **2**, 73–90.

DiMasi, J.: 1995, Success rates for new drugs entering clinical testing in the United States, *Clinical Pharmacology & Therapeutics* **58**(1), 1–14.

Gambardella, A.: 1995, *Science and Innovation: The US Pharmaceutical Industry in the 1980 s*, Cambridge University Press.

Grabowski, H., Vernon, J. and DiMasi, J.: 2002, Returns on Research and Development for 1990s New Drug Introductions, *Pharmacoeconomics* **20**(3), 11–29.

Hall, B., Jaffe, A. and Trajtenberg, M.: 2001, The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools.

Harhoff, D., Narin, F., Scherer, F. and Vopel, K.: 1999, Citation Frequency and the Value of Patented Inventions, *The Review of Economics and Statistics* **81**(3), 511–515.

Heller, M. and Eisenberg, R.: 1998, Can Patent Deter Innovation? The Anti-Commons in Biomedical Research, *Science* **280**, 698–701.

Henderson, R. and Cockburn, I.: 1996, Scale, Scope, and Spillovers: The Determinants of Research Productivity in Drug Discovery, *Rand Journal of Economics* **27**(1), 32–59.

Henderson, R., Orsenigo, L. and Pisano, G.: 1999, The Pharmaceutical Industry and the Revolution in Molecular Biology: Exploring the Interactions between Scientific, Institutional and Organizational Change, *in* D. Mowery and R. Nelson (eds), *Sources of Industrial Leadership*, Cambridge University Press: Cambrigde, MA.

Horstmann, I., MacDonald, G. and Slivinski, A.: 1985, Patents as Information Transfer Mechanisms: To Patent or (Maybe) Not to Patent, *The Journal of Political Economy* **93**(5), 837–858.

Hsiang, Y., Hertzberg, R., Hecht, S. and Liu, L.: 1985, Camptothecin induces protein-linked DNA breaks via mammalian DNA topoisomerase I, *Journal of Biological Chemistry* **260**(27), 14873–14878.

Jaffe, A. and Trajtenberg, M.: 1996, Flows of Knowledge from Universities and Federal Labs: Modeling the Flow of Patent Citations Over Time and Across Institutional and Geographic Boundaries.

Jaffe, A. and Trajtenberg, M.: 2002, *Patents, citations and innovations: a window on the knowledge economy*, MIT Press.

Jaffe, A., Trajtenberg, M. and Fogarty, M.: 2000, The Meaning of Patent Citations: Report on the NBER/Case-Western Reserve Survey of Patentees.

Janis, M.: 2000, On Courts Herding Cats: Contending with the Written Description Requirement (and Other Unruly Patent Disclosure Doctrines), *Washington University Journal of Law and Policy* **2**, 55–108.

Johnson, R. and Dickersin, K.: 2007, Publication bias against negative results from clinical trials: three of the seven deadly sins, *Nature Clinical Practice Neurology* **3**(11), 590–591.

Landes, W. and Posner, R.: 2003, *The economic structure of intellectual property law*, Belknap Press of Harvard University Press Cambridge, Mass.

Lanjouw, J. and Schankerman, M.: 1999, The Quality of Ideas: Measuring Innovation with Multiple Indicators.

Lee, J., Laydon, J., McDonnell, P., Gallagher, T., Kumar, S., Green, D., McNulty, D., Blumenthal, M., Keys, J., Land Vatter, S. et al.: 1994, A protein kinase involved in the regulation of inflammatory cytokine biosynthesis, *Nature* **372**(6508), 739–746.

Levin, R., Klevorick, A., Nelson, R., Winter, S., Gilbert, R. and Griliches, Z.: 1987, Appropriating the Returns from Industrial Research and Development, *Brookings Papers on Economic Activity* **Microeconomics**(3), 783–831.

Lichtman, D., Baker, S. and Kraus, K.: 2000, Strategic Disclosure in the Patent System, *Vanderbilt Law Review* **53**, 2175–2217.

Machlup, F. and Penrose, E.: 1950, The Patent Controversy in the Nineteenth Century, *The Journal of Economic History* **10**(1), 1–29.

Mansfield, E.: 1986, Patents and Innovation: An Empirical Study, *Management Science* **32**(2), 173–181.

Mervis, J.: 2005, Productivity Counts–But the Definition Is Key.

Murray, F. and Stern, S.: 2007, Do Formal Intellectual Property Rights Hinder the Free Flow of Scientific Knowledge? An Empirical Test of the Anti-Commons Hypothesis, *Journal of Economic Behavior and Organization* **63**(4), 648–687.

Oppenheim, C.: 1998, How SMEs use the patent literature.

Oprea, T., Davis, A., Teague, S. and Leeson, P.: 2001, Is There a Difference between Leads and Drugs? A Historical Perspective, *Journal of Chemical Information and Computer Sciences* **41**(5), 1308–1315.

Orsenigo, L., Pammolli, F. and Riccaboni, M.: 2001, Technological change and network dynamics: lessons from the pharmaceutical industry, *Research Policy* **30**(3), 485–508.

Parchomovsky, G.: 2000, Publish or Perish, *Michigan Law Review* **98**(4), 926–952.

Plant, A.: 1934, The Economic Theory Concerning Patents for Inventions, *in* R. . K. Paul (ed.), *Selected Economic Essays and Addresses*, London, pp. 35–36.

Reichman, J.: 1995, Universal Minimum Standards of Intellectual Property Protection under the TRIPS Component of the WTO Agreement, *International Lawyer* **29**, 345–388.

Scotchmer, S. and Green, J.: 1990, Novelty and Disclosure in Patent Law, *The RAND Journal of Economics* **21**(1), 131–146.

Sutton, J.: 1998, *Technology and market structure: theory and history*, The MIT Press Cambridge, Mass.

Tang, P., Adams, J. and Pare, D.: 2001, Patent protection of computer programmes, Report submitted to the European Commission - Directorate-General Enterprise.

Trajtenberg, M.: 1990, A Penny for Your Quotes: Patent Citations and the Value of Innovations, *The RAND Journal of Economics* **21**(1), 172–187.

Trajtenberg, M., Henderson, R. and Jaffe, A.: 1997, University vs. Corporate Patents: A Window on the Basicness of Innovations, *Economics of Innovation and New Technology* **5**(1), 19–50.

Wall, M., Wani, M., Cook, C., Palmer, K., McPhail, A. and Sim, G.: 1966, Plant Antitumor Agents. I. The Isolation and Structure of Camptothecin, a Novel Alkaloidal Leukemia and Tumor Inhibitor from Camptotheca acuminata1, 2, *Journal of the American Chemical Society* **88**(16), 3888–3890.

WHO: 2006, Public Health: Innovation and Intellectual Property Rights.

Zarin, A. and Tse, T.: 2008, Moving Toward Transparency of Clinical Trials, *Science* **319**(5868), 1340–1342.