Scientometrics (2011) 88:805–818
DOI 10.1007/s11192-011-0400-3

Tracking R&D behavior: bibliometric analysis of drug patents in the Orange Book

Ming-Chao Huang · Shih-Chieh Fang · Shao-Chi Chang

Received: 13 January 2011 / Published online: 8 May 2011

Abstract The Publication *Approved Drug Products with Therapeutic Equivalence Evaluations* (commonly known as the Orange Book) identifies drug products approved by the United States Food and Drug Administration (USFDA) for safety and effectiveness, and provides substantial information on new drug applications (NDAs) with patent data. To explore the patterns among drug patents in the Orange Book, this study used patent bibliometric analysis. The productivity and impact are presented at the assignee level and applicant level, respectively, and the applicant's patent portfolio is further discussed. 2,033 drug patents are identified in this current study. Our findings indicate that the applicant's patent portfolio in the Orange Book is helpful in revealing the technological capability and patent strategy of the pharmaceutical incumbents. By linking drug data and patent information, this current study sheds light on patent research in the pharmaceutical industry.

Keywords Bibliometric analysis · Pharmaceutical industry · Orange Book · NDAs · Patent sourcing

Introduction

The pharmaceutical industry is heavily R&D intensive, with incumbents requiring ten to 15 years on average to develop new drug products from initial discovery to final FDA approval. The cost of a new drug development increased from $231 million in 1987 to $802 million in 2000 (DiMasi 2001). To measure the R&D activities and performance in the pharmaceutical industry, studies have adopted various proxies, e.g., R&D expenditures, patent and citation count, number of claims, new molecular entities, and new drug approval (Chandy et al. 2006; Chen and Chang 2009, 2010a, b; Dominguez-Lacasa 2006; Koenig

M.-C. Huang
Graduate Institute of Business Administration, National Cheng Kung University, Tainan, Taiwan

S.-C. Fang (✉) · S.-C. Chang
Department of Business Administration, National Cheng Kung University,
No. 1, University Road, Tainan City 701, Taiwan
e-mail: fangsc@mail.ncku.edu.tw

and Mezick 2004; McMillan and Hamilton 2000). Following previous studies, we explore drug patents based on the approval of new drugs from the Orange Book using the bibliometric method.

The approval of drugs with additional patent information from the Orange book is used as the sample in the current study for two reasons. First, the Orange Book identifies drug products approved on the basis of safety and effectiveness by the USFDA and offers considerable information on the listed drugs, e.g., the active ingredient, proprietary name, drug applicant name, and drug applicant number (Hill 2005). The drug applicant provides a distinct unit of analysis comparable to the assignee, which is commonly used in patent-based research. Second, for the purpose of competition on generic drug price, the Hatch and Waxman Act requires all newly submitted drug applications file patent information. Hence, any patents held by the applicant that claim the drug or a method of using the drug must be submitted (USFDA 2010). As a result, a patent is listed in the Orange Book when the new drug application is approved by the USFDA, and can serve as a basis of a patent infringement claim (Hill 2005). Therefore, we can connect drug products with their corresponding patents and address related research issues using the Orange Book.

We use the patent bibliometric method to analyze the characteristics of drug patents in the Orange Book. Accordingly, this study may be divided into three parts. First, the overviews of our data are presented by annual patent counts (including application year and granted year), distribution of patents across countries, and distribution of impact on patent-level. Second, we analyze the productivity and impact by assignee and applicant, respectively. Finally, because external patents occupy a portion of an applicant's patent portfolio, and some studies of the industry have noted an increasing proportion of firm revenue derived from technology discovered outside the firm (Ceccagnoli et al. 2010), we address technological capability and patent source by analyzing the focal applicant's patent portfolio.

This paper is organized as follows. In the first section, we review the literature on patent bibliometric analysis in the pharmaceutical industry. Next, we present an overview of drug patents in the Orange Book. In the following sections, the productivity and impact are presented by assignee and applicant, respectively. We next analyze the applicant's patent portfolio by technological capabilities and patent sources, and further summarize patent sourcing strategies. Finally, we discuss the findings of this paper.

Literature review

Patents play an important role in protecting a firm's innovation in an R&D intensive industry (Chen and Chang 2009, 2010a, b; McMillan and Hamilton 2000, 2007). Firms in the pharmaceutical industry are critically dependent on the patent system for appropriation of returns from R&D. Narin (1994) applied bibliometric methods, including patent citation analysis, to establish the use of patent bibliometric methods, and patent bibliometrics have been used to investigate several technology fields (Acosta et al. 2009; Bhattacharya and Meyer 2003; Lo 2007, 2008; Van Looy et al. 2007; Wang et al. 2010; Yuan et al. 2010). Most studies use patent bibliometrics primarily to address issues of patent or citation count to obtain an overall view of the sample (Lo 2007, 2008; McMillan and Hamilton 2007). For the patent count, the number of patents indicates the productivity. This indicator may be applied to country-, assignee-, inventor- and technology-level analysis (Acosta et al. 2009; Lo 2007, 2008; Yuan et al. 2010). For the citation count, the forward citation counts indicate the level of impact or the quality of a firm's patents, since the frequent citation of a

patent provides useful technological knowledge for other patents (Acosta et al. 2009; Lo 2007, 2008; McMillan and Hamilton 2007). Drawing on previous studies, it is appropriate to depict the pattern of drug patents in the Orange Book by measuring patent and citation count.

Over the last decade several studies have examined drug patents using patent bibliometric methods (Chandy et al. 2006; Chen and Chang 2009, 2010a, b; Koenig and Mezick 2004; McMillan and Hamilton 2000). McMillan and Hamilton (2000) assert that bibliometrics provide a quality measure of a company's scientific and technological knowledge. Their study explored the publishing and patenting citation patterns of 12 US pharmaceutical companies and suggested that the pharmaceutical industry is appropriate for bibliometrics due to its extensive use of knowledge-based property and patent enforceability (McMillan and Hamilton 2000). Chen and Chang conducted a series of studies that concerned the relationships between patent indicators and firm performance in the pharmaceutical industry by artificial neural network technique (Chen and Chang 2009, 2010a, b). The findings of their studies demonstrate the importance of patent-related indicators. They also found that non-linear relationships exist between patent traits and firm performance.

The drug patents listed by the USFDA provide a well-defined linkage between technologies and products. Chandy et al. (2006) used the FDA drug data and USPTO patent data to investigate the conversion ability of pharmaceutical companies. They measured conversion by converting a drug patent into an actual drug product (binary variable). Their findings suggest that the pharmaceutical industry is appropriate for exploration of the relationship between patent (technology) and new drug (product). They excluded patents that are acquired from outside, whether through licensing or M&A (Chandy et al. 2006, p. 497). However, the phenomenon of M&A and licensing is an important issue in the pharmaceutical industry at present. To investigate the impact of M&A activities on R&D performance, Koenig and Mezick (2004) examined drug data from the USFDA. The cost of new molecular entity (NME) and patents were evaluated as the R&D performance while the pre-merger company performance was compared to post-merger performance. Their conclusion suggested that post-merger products might be examined to determine the source of the patents used to develop each new product (Koenig and Mezick 2004).

Based on the studies above, the current study addresses the following issues. First, we analyze drug patents in the Orange Book using patent bibliometric methods, to present an overview of our sample. Though previous studies have used the Orange Book data as the research sample, the patterns among drug patents in the Orange Book are remain unknown. Second, we identify the top 10 assignees and applicants with the most patents in the Orange Book to explore their productivity and impact, respectively. We especially focus on the analysis of applicants, the major players in the Orange Book who apply for NDAs with attached patent information. Previous studies in bibliometrics have primarily emphasized the level of assignee or inventor. However, for the pharmaceutical industry, analysis of drug applicants is more practical. Third, the Orange Book offers a fine-grained platform to explore pharmaceutical firms' patent portfolios and their patent sources. Hence, we can evaluate the applicant's technological capabilities and patent sources as well. Additionally, patent sourcing strategies can be clarified through the lens of licensing and M&A. The intent of this study is to link USFDA drug data and USPTO patent information by patent bibliometrics and sheds light on patent research in the pharmaceutical industry.

Data and methods

In this study, the NDA data is collected from the USFDA Electronic Orange Book (EOB) and the patent data is taken from the USPTO patent database. The EOB primarily contains three parts: 'Product', 'Patent', and 'Exclusivity'. We capture the patent number and drug application number from the file 'Patent', and match the drug application number to the corresponding drug applicant information from the file 'Product'. Each NDA applicant may have several applications (approved drugs) and added patents, we compare the patents submitted by the applicants to identify the unique patent information for this study. In addition, review of the various types of patents shows that utility patents can best describe R&D capability (Chen et al. 2005), we follow this rule and only collect utility patents. Consequently, 1,011 different applications with 2,033 unique patents are identified in the EOB database. Furthermore, we collect patent data from USPTO patent database to capture the relevant information (e.g., assignee name, application date, issued date, and nationality of inventors).

Quantitative methods are used to measure patent counts as a proxy of productivity, while the number of forward citations is the proxy of impact in this study. Although using forward citation data may not entirely reflect the value of patents, recognizing the importance of a given patent (technology) may require other forms of data, such as expert assistance. However, patent bibliometrics is still an efficient way to evaluate bundles of patents and draw general conclusions. McMillan and Hamilton (2000, p. 466) argue that bibliometrics is based on a simple premise: the greater the merit, influence or importance of a particular paper or patent, the more frequently it will be cited in the future. Given this, bibliometrics may thus provide a quality measure of a company's technological knowledge by examining the citation patterns of its patents (McMillan and Hamilton 2000). A number of studies also adopted this rationale and have demonstrated a positive relation between patent citation and patent value (Acosta et al. 2009; Chen and Chang 2009, 2010b; Hall et al. 2005; McMillan and Hamilton 2007). Therefore, we follow previous studies and evaluate a company's patents using forward citations. The patent forward citations are collected from 1987 (the year that the first patent issued in our sample) to June 2010.

An applicant's patent portfolio may comprise patents from a variety of sources (patentees). To further analyze the applicant's technological capability and independence, we count self-owned patents (in-house patents) on the one hand, and on the other, measure the concentration of patent sources within the portfolio. The concentration of patent sources within the portfolio can reflect a firm's technological independence, and we use the Herfindahl–Hirschman Index (HHI) to measure the concentration of patent sources. HHI is generally applied in Industrial Organization studies to measure the level of industry concentration. Studies have used the HHI to measure the concentration of patents across technological classes (Chen and Chang 2009, b; Hill 2005) or the distribution of paper citations (Yang et al. 2010). In line with the concept of concentration, we use HHI to measure the concentration of patentees within the applicant's portfolio as a proxy of technological independence. We calculate the HHI using the following expression:

$$\mathrm{HHI} = \sum_{i=1}^{p} \left(\frac{N_i}{N}\right)^2$$

where N is the total number of patents within a given applicant's patent portfolio, P denotes the number of distinct patentees within the portfolio, and N_i is the number of patents held by each distinct patentee within the portfolio. When an applicant's patent

portfolio consists largely of a single patentee, the HHI is close to 1. Conversely, if there are a large number of patentees within a portfolio, the HHI is close to 0.

Results

Overall analysis

Annual productivity

2,033 unique pharmaceutical patents are listed in the Orange Book (Fig. 1). The first filed patent is patent 5162362 (1983), while the first issued patent is patent 4634697 (1987) of our sample. 1995 saw the most patent applications (245 patents) while 1999 is the year in which the most patents were granted (179 patents). Based on the number of patents granted annually, our data may be divided into two periods, 1987–1997, and 1998–2010. During the period 1987–1997, the number of patents granted annually increased, peaking in 1997. The period 1996–2004 was the period of highest productivity, with over 100 patents issued annually. Since 1998, the number of patents granted has begun to fall.

Productivity analysis by country

We measured patent count as the indicator of productivity (Lo 2007). For the 2,033 patents, 34 countries are identified by assignee country. Table 1 presents the top 10 countries with the highest number of patent applications in the Orange Book. The United States is the most productive country, with 1,290 patents (63.45%), followed by Japan with 115 (5.66%). Great Britain, Switzerland and Germany round out the top five countries. The United States submitted more than half the patents in Orange Book, and the top five countries were granted over 83% of patents.

Impact analysis

Forward citation is an indicator of one patent's impact on others. The 2,033 patents have been cited 28,059 times by 14,798 distinct patents. On average, each patent has received

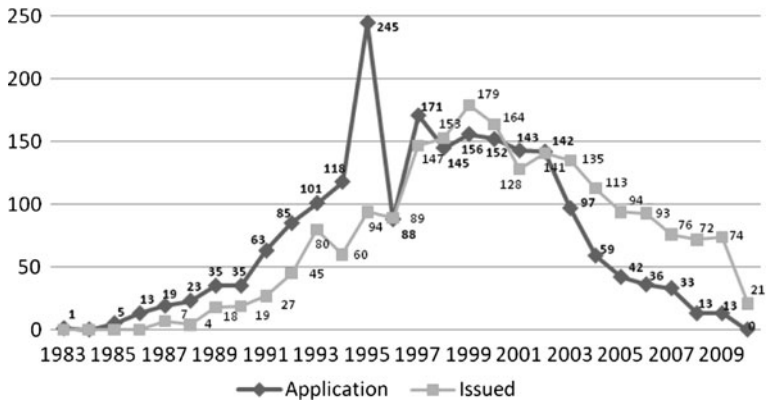

Fig. 1 Annual patent count of our sample

Table 1 Distribution of countries and assignees

Country	Count	%	No. of assignees	%
United State	1290	63.45	314	56.17
Japan	115	5.66	35	6.26
Great Britain	109	5.36	24	4.29
Germany	93	4.57	33	5.90
Switzerland	89	4.38	15	2.68
Sweden	66	3.25	8	1.43
Canada	46	2.26	21	3.76
France	43	2.12	17	3.04
Belgium	28	1.38	6	1.07
Denmark	26	1.28	5	0.89
Others	173	8.51	81	14.49
Total	2033	100.00	559	100.00

13.80 forward citations. We treat this indicator as a benchmark to compare assignees and applicants. Patent 5585089 received the most forward citations (314), while 354 patents (17.41% of the sample) received only one patent citation, and 124 patents (6.10% of the sample) have no forward citations. 52.78% of our sample has at least six patent citations, while only 28 patents (1.38%) are cited more than 100 times. Table 2 shows the distribution of forward citations of our sample.

Analysis by assignees

We identify the top 10 assignees with the most patents in Orange Book (Table 3). The assignee name is presented as original information in the USPTO, even if the assignee has

Table 2 Distribution of forward citations

Forward Citations	Count	%	Cumulative Count	%
0	124	6.10	124	6.10
1	354	17.41	478	23.51
2	197	9.69	675	33.20
3	120	5.90	795	39.10
4	104	5.12	899	44.22
5	87	4.28	986	48.50
6	87	4.28	1073	52.78
7	68	3.34	1141	56.12
8	59	2.90	1200	59.03
9	53	2.61	1253	61.63
10	55	2.71	1308	64.34
11-100	697	34.28	2005	98.62
101-200	23	1.13	2028	99.75
Over 200	5	0.25	2033	100.00

Table 3 Ten assignees with the highest number of patents

Assignee	Number of Patents	%
Merck and Co	68	3.34
SmithKline Beecham	60	2.95
Novartis	58	2.85
Glaxo	55	2.71
Eli Lilly and Co	50	2.46
Schering	47	2.31
Abbott Lab.	43	2.12
ALZA	43	2.12
Pfizer	43	2.12
Takeda	36	1.77
Total	503	24.40

been merged or changed. For example, SmithKline Beecham has merged with Glaxo Wellcome to form GSK in 2000, but we still list these two firms separately.

Productivity analysis

503 patents are granted to the top 10 assignees, which account for 24.4% of our sample. Merck and Co is the most productive assignee, with 68 patents. SmithKline Beecham owns 60 patents, followed by Novartis (58), Glaxo (55), and Eli Lilly and Co (50). Among the top 10 companies, Takeda is the only Japanese company, with all others being either US or European.

Impact analysis

Table 4 presents forward citations of the top 10 assignees. Patents of the top 10 companies received a total of 5,511 forward citations, an average of 10.96 each. Merck and Co received the most forward citations among the top 10 assignees (819), or 12.22 each, followed by Glaxo (738, 13.42). Eli Lilly and Co (703, 14.06), ALZA (615, 14.30), and Pfizer (612, 14.23) complete the top five firms. Although Eli Lilly and Co, ALZA and Pfizer have fewer citations than Merck and Co and Glaxo, the average impact of the patents of these three firms is highest among the top 10 companies. Schering, ranked sixth in this study, received 547 citations whose average impact is 11.64. Takeda received 406 citations with an impact of 11.28. The last three assignees among the top 10 companies are Novartis (404), SmithKilne Beecham (335), and Abbott (332), all averaging less than eight forward citations per patent (6.97, 5.58, and 7.72 respectively).

In conclusion, the average number of forward citations of the top 10 assignees (10.96) is less than that of the total sample (13.8). Obviously, quantity of citation counts does not guarantee the company's patent impact. For instance, Merck and Co has the most patents and forward citations, but its impact is ranked fifth. SmithKline Beecham has the second most patents, but has only received 335 forward citations (ranked 9th) with a mere 5.58 average patent impact (ranked 10th). The lower impact of the top 10 assignees shows that internal technology development is insufficient, and firms in the industry must have access to external technological resources for new product development. Thus, to show the distribution of technology it is necessary look at applicants rather than assignees.

Table 4 Forward citations of the top 10 assignees

Assignee	No. of Forward Citations	Avg. of Forward Citations
Merck and Co	819	12.22 (5)
SmithKline Beecham	335	5.58 (10)
Novartis	404	6.97 (9)
Glaxo	738	13.42 (4)
Eli Lilly and Co	703	14.06 (3)
Schering	547	11.64 (6)
Abbott Lab.	332	7.72 (8)
ALZA	615	14.30 (1)
Pfizer	612	14.23 (2)
Takeda	406	11.28 (7)
Total	5511	10.96

Analysis by applicants

The USFDA requires pharmaceutical firms to submit relevant patent information when they are in the NDA process. Hence the applicant becomes an important unit of analysis in bibliometric analysis of the pharmaceutical industry. Applicant names vary in the Orange Book. For example, the applicant name 'Abbott' and 'Abbott Labs' are both listed in the Orange Book. M&A activities in the pharmaceutical industry further complicate applicant names, as when GlaxoSmithKline was established after the merger of Glaxo Group and SmithKilne Beecham. To confirm and unify applicant names, we consulted experts in the pharmaceutical industry. Most applicants have been adjusted, except for Wyeth. We separate Wyeth from Pfizer even after Wyeth merged with Pfizer in 2009 since there are no Wyeth NDA applications by Pfizer, and it is helpful in clarifying the conclusions of this study. 286 unique applicants with patent information are listed in the Orange Book. We then identified the top 10 applicants by patents in the Orange Book (Table 5).

Comparing the list of top 10 applicants (Table 5) with assignees (Table 3), we can find that some company names have changed. A number of assignees have been merged or acquired. For example, Glaxo and SmithKline Beecham have merged to form GSK, and

Table 5 Ten applicants with the highest number of patents

Applicant	No. of Patents	%	Country
GlaxoSmithKline	128	6.03	GB
Novartis	108	5.31	CH
Johnson and Johnson	85	4.18	US
Abbott	77	3.79	US
Pfizer	67	3.30	US
Merck and Co	66	3.25	US
AstraZeneca	61	3.00	GB
Sanofi-Aventis	56	2.75	FR
Eli Lilly and Co	49	2.41	US
Wyeth	34	1.67	US
Total	731	35.96	

J&J acquired ALZA. Schering AG was acquired by Bayer in 2007, and is not listed among the top 10 applicants. The disappearance of Takeda from the list of top 10 applicants is because Takeda's patents were submitted by other applicants for new drug applications (Abbott and AstraZeneca submit eight and six of Takeda's patents respectively). In brief, SmithKline Beecham, Glaxo, Schering, ALZA, and Takeda are replaced by GSK, J&J, AstraZeneca, and Sanofi-Aventis in the list of top 10 applicants.

Productivity analysis

The top 10 applicants submitted a total of 731 distinct patents (35.96% of our sample). The applicant GlaxoSmithKline (abbreviated GSK) is denominated by GlaxoSmithKline, Glaxo GRP LTD, GlaxoSmithKline Cons, SmithKline Beecham, and SB Pharmco in the Orange Book. GSK applied for the most patents for new drug applications in the Orange Book (128 patents). Novartis, Novartis and Novartis Pharms in the Orange Book, is second with 108 patents. Johnson and Johnson (abbreviated J&J) has multiple pharmaceutical manufacturing divisions. We combine Johnson and Johnson, Ortho McNeil Janssen, Janssen Pharma, McNeil, and Ortho Dermatologics to represent J&J's applications (85 patents).

Abbott consists of Abbott, Abbott Labs, and Knoll, for a total of 77 patents. Pfizer is active in M&A activity in the pharmaceutical industry and several entities have been merged or acquired by Pfizer, including Wyeth, Agouron, Warner Lambert, GD Searle, and Pharmacia and Upjohn. To focus on the linkage between patentees and applicants, we only identify Pfizer, Pfizer Inc, Pfizer Pharms, Pfizer Global, and Pfizer Ireland as Pfizer (67 patents). Merck is sixth among the top 10 applicants. It is represented by Merck, Merck and Co Inc, Merck Sharp Dohme, and Merck Co Inc in the Orange Book. The other four applicants in the Orange Book are AstraZeneca (61), Sanofi-Aventis (56), Eli Lilly and Co (49), and Wyeth (34).

Impact analysis

Similar to the impact analysis of assignee, we also count forward citations of each applicant and measure the average number of forward citations as the indicator of patent impact. The top 10 applicants received 11,695 forward citations (Table 6), and the average number of forward citations is roughly 16, higher than that of the top 10 assignees (Table 4). Clearly, the top 10 applicants acquire high impact patents from the outside to enhance new drug development.

GSK submitted 128 patents and received 1,590 forward citations, ranking second, for a patent impact of 12.42, ranking ninth. Novartis received the most forward citations among the top 10 applicants (1,963 times) and has the third ranking impact, 18.18. J&J has 1,297 forward citations for a patent impact of 15.26. Both the forward citation number (695 times) and impact (9.03) of Abbott are ranked last, but these two indicators are higher than assignee Abbott (see also Table 4). Pfizer submitted 1,481 patents to the USFDA for a patent impact of 22.10 (ranked 2nd).

The remaining five applicants, Merck and Co, AstraZeneca, Sanofi-Aventis, Eli Lilly and Co, and Wyeth, have 1,013, 795, 811, 753, and 1,297 forward citations, respectively. It is interesting that Wyeth has an outstanding performance on patent impact (38.15), ranked first. We investigated the 34 patents submitted by Wyeth in the Orange Book, and found that four patents have more than 100 forward citations each (patent number 5100899, 5145684, 5585089, and 5693762). The technologies in these four patents were not invented

Table 6 Impact of the top 10 applicants

Applicant	No. of Forward Citations	Avg. of Forward Citations
GlaxoSmithKline	1590	12.42 (9)
Novartis	1963	18.18 (3)
Johnson and Johnson	1297	15.26 (6)
Abbott	695	9.03 (10)
Pfizer	1481	22.10 (2)
Merck and Co	1013	15.35 (5)
AstraZeneca	795	13.03 (8)
Sanofi-Aventis	811	14.48 (7)
Eli Lilly and Co	753	15.37 (4)
Wyeth	1297	38.15 (1)
Total	11695	16.00

by Wyeth, however. Patent 5100899 is invented by an individual, patent 5145684 is granted to Sterling Drug, and patents 5585089 and 5693762 are granted to Protein Design Labs. Through licensing from other assignees, Wyeth has enlarged its patent portfolio and enhanced its new drug production process. In the following section, we further explore the top 10 applicants' patents portfolios to reveal the technological capabilities and the distribution of patent sources.

Analysis of applicant's patent portfolio

From the viewpoint of the applicant, we define the patent portfolio as the combination of patents that is submitted with a NDA by the focal applicant. As mentioned above, both internally developed and externally acquired patents are submitted to the Orange Book by focal applicants. Hence, we attempt to explore the applicant's patent portfolio by evaluating its technological capabilities and clarifying its patent sources. The applicant's technological capability is measured by counting the number of patents that belong to the focal applicant in the portfolio (in-house patents). The ratio of in-house patents to all patents of each applicant is presented. The patent sources are identified by the number of unique patentees (patent owners) within the portfolio. Further, we measure the HHI of patent sources to demonstrate the distribution of patentees.

Before counting in-house patents and number of patentees identically, we must first define the boundary of the applicants. Based on the concept of business group, both the subsidiaries and the companies that have been acquired or merged by the focal applicant are treated as the same applicant. For instance, Laboratories Fournier is a subsidiary of Abbott (Abbott submitted six patents from Laboratories Fournier), Knoll Pharmaceutical and Kos Pharmaceuticals are merged with Abbott (Abbott submitted seven patents from Knoll and Kos). Hence we combine these three companies' patent data with the applicant Abbott. By contrast, patents obtained through licensing from the outside are not combined with those of the focal applicant. For example, we separate three patents from Abbott's patent portfolio because they were acquired from Elan Pharma. Finally, patents granted to individuals are grouped into one category. Table 7 shows the results of in-house patents and patent sources of the top 10 applicants.

Analysis of in-house patents and patent sources

GSK has 96 in-house patents, which means that GSK holds 96 of the 128 patents (75.00%) itself. The sources of GSK's patents are derived from 18 different entities (patentees), and the concentration of GSK's patents on patentees is 0.5673, ranked third among the top 10 applicants. Novartis submits 108 patents, 69 of which (63.89%) are invented by Novartis. Novartis's 108 patents are from 15 different patentees, and the concentration of patentees is 0.4206. J&J owns 52 of 85 patents (61.18%), the number of patentees is 16, and the degree of concentration is 0.3954 (ranked 10th), meaning the sources of J&J's patents are the most diverse.

Abbott submits 77 patents, 54 of which are in-house patents (70.13%). Abbott's patents are derived from nine different patentees, and the concentration is 0.5089. Pfizer submits 67 patents, 43 of which (64.18%) are in-house patents. Pfizer obtains patents from eight different patentees for a concentration is 0.4462. Merck and Co owns 48 of 66 patents (72.73%). Merck's 66 patents are from 11 different entities, and the concentration is 0.5381.

AstraZeneca submits 61 patents to the USFDA and holds 46 of the in-house patents (75.41%). AstraZeneca's application patents can be divided into six different entities for a concentration of 0.5856 (ranked 2nd). Sanofi-Aventis submits 56 patents, 40 of which are in-house patents (71.43%). The 56 patents are from 10 different entities for a concentration of 0.5254. Eli Lilly and Co submits 49 patents, with just four sourced externally, yielding the highest concentration among the top 10 applicants (0.8459). Wyeth submits 34 patents to the Orange Book, 22 of which are in-house patents (64.71%). Wyeth's patents are derived from six different patentees, for a concentration of 0.4516.

In conclusion, the ratio of in-house patents exceeds 60% for all firms, while half of the top 10 applicants have more than 10 patentees (patent sources) among their patents. Among the top 10 applicants, GSK has the most in-house patents (96) and patentees (18). By contrast, Sanofi-Aventis has the fewest in-house patents (40), while Eli Lilly and Co has the smallest number of patentees (4). The results of the ratio of in-house patents and HHI analysis show that Eli Lilly and Co has the highest ratio of in-house patents (91.84%) and HHI of patentees (0.8459), while J&J has the lowest ratio of in-house patents (61.18%) and HHI of patentees (0.3954).

Table 7 In-house patents and patent sources of the top 10 applicants

Applicant	No. of patents	No. of in-house patents	%	No. of patentees	HHI of patentees
GlaxoSmithKline	128	96	75.00	18	0.5673 (3)
Novartis	108	69	63.89	15	0.4206 (9)
Johnson and Johnson	85	52	61.18	16	0.3954 (10)
Abbott	77	54	70.13	9	0.5089 (6)
Pfizer	67	43	64.18	8	0.4462 (8)
Merck and Co	66	48	72.73	11	0.5381 (4)
AstraZeneca	61	46	75.41	6	0.5856 (2)
Sanofi-Aventis	56	40	71.43	10	0.5242 (5)
Eli Lilly and Co	49	45	91.84	4	0.8459 (1)
Wyeth	34	22	64.71	6	0.4516 (7)

Summary of patent sourcing strategies

In this section, we pay special attention to the distribution of patent sources (patentees) within the applicants' patent portfolios. The phenomenon of diversified patent sources reflects the multidisciplinary properties of the pharmaceutical industry, where pharmaceutical firms source needed patents from other companies. The investigation of drug patents in the Orange Book can help to disclose the applicants' patent sources and depict the patent strategy as well. By analyzing the applicants' patent portfolios, we can divide applicant's patent sourcing strategies into three types: (1) applicants invent patents internally; (2) applicants acquire patents through M&A; and (3) applicants obtain necessary patents by licensing from the patentees.

The first type of patent source represents an R&D intensive strategy, and our findings indicate that all top 10 applicants source more than 60% of their patents in-house (Table 7), reflecting the importance of the internal technological capability of the pharmaceutical incumbents. Furthermore, the volume of in-house patents is a good proxy for the stock of technological capability, which can deepen the absorption capacity (Bhattacharya and Meyer 2003; Cohen and Levinthal 1990) and consequently help in the acquisition and use of external technology.

The second type of patent source requires heavy capital investment to carry out M&A activity. Pfizer spent $68 billion for Wyeth, and Merck and Co spent $41 billion for Schering Plough. The primary purpose of M&As, common in the pharmaceutical industry during the last decade, is to maintain a healthy pipeline of products (Koenig and Mezick 2004). Though the performance of M&A is hard to estimate, pharmaceutical firms constantly look for new targets.

Licensing from external sources is also a key source of patents. Through licensing, a firm can access the patent immediately, reducing R&D costs and M&A investments. Patent licensing reflects an emerging concept of open innovation in which firms may seek opportunities to license their own patents or acquire necessary patents under permit. For example, Novartis licenses patent 6596260 to GSK, and AstraZeneca licenses patent 7342005 to Merck and Co.

In brief, all three types of patent sourcing strategies summarized herein are used by the major applicants. An applicant may pursue more than one patent sourcing strategy simultaneously, but internal R&D capability still plays the dominant role during the process of new drug development.

Conclusions

This current study uses patent bibliometric method to explore the patent data in the USFDA Orange Book. We first investigate the overall characteristics of our sample, and then we present the productivity and impact at the assignee level. We emphasize the role of new drug applicants and address issues of productivity and impact using the applicant's patent portfolio. The internal technological capabilities (in-house patents) and patent sources of the applicants are also discussed.

The results of the patent productivity analysis show that the US is the most active country. Nine of top 10 countries with the most pharmaceutical patents in our sample are US or European countries. From the viewpoint of assignees, Merck and Co has the most patents, and almost all top 10 assignees are US-based or Euro-based companies, the only

exception being a Japanese company, Takeda. All top ten applicants are US or European firms, with GlaxoSmithKline submitting the most patents for NDAs.

The analysis of patent impact on assignee shows that Merck and Co received the most forward citations, while ALZA has the highest index of impact. On the other hand, the applicant Novartis' patent portfolio received the most forward citations, while applicant Wyeth has the highest index of impact. We further discuss the patent sources of top 10 applicants, and categorize the patent sourcing strategies into three types: internal, M&A, and licensing. Based on these findings, the three patent sources are simultaneously used to construct applicant patent portfolios. Therefore, pharmaceutical industry incumbents rely on multiple technology sources to develop new drugs and enhance their competitive advantage.

In general, drugs filed in the USFDA are intended for the US market. This study used the Orange Book to link drugs filed with the USFDA and corresponding patents filed with the USPTO, and consequently presented the pattern of drug patents in the US. Hence, based on certain institutional characteristics, such as the time-lag of patent application, an over-representation of larger firms, and English-language bias in citation (Iversen 2000), there are several limitations when we generalize our findings. First, our findings reflect R&D behaviors of the companies that act in the US pharmaceutical market. Although, the US is the largest single market in the world and each country (company) has the same propensity to patent in the US (Chen and Chang 2009; Iversen 2000), the context here should be clarified. Further studies could investigate drug and patent data from other countries.

Second, we present the top ten pharmaceutical firms with the most patents, a situation which may not generalize across firms with fewer patents. Third, based on the constraint of monthly updated patent data in the Orange Book, this study used only drug patents that have not expired. Further study may collect patent data from the Orange Book continuously to carry out a longitudinal analysis. Finally, the methodology using in this study has been adopted in several studies, but we suggest that this methodology may help us to reveal and understand the pattern of the Orange Book. Further study may use other bibliometric methodologies to analyze the relationship between drug applicants and patent owners.

References

Acosta, M., Coronado, D., & Feranadez, A. (2009). Exploring the quality of environmental technology in Europe: Evidence from patent citations. *Scientometrics, 80*(1), 131–152.

Bhattacharya, S., & Meyer, M. (2003). Large firms and the science-technology interface: Patents, patent citations, and scientific output of multinational corporations in thin films. *Scientometrics, 58*(2), 265–279.

Calero, C., Leeuwen, T. N., & Tijssen, R. J. W. (2007). Research cooperation within the bio-pharmaceutical industry: Network analyses of co-publications within and between firms. *Scientometrics, 71*(1), 87–99.

Ceccagnoli, M., Graham, S. J. H., Higgins, M. J., & Lee, J. (2010). Productivity and the role of complementary assets in form's demand for technology innovations. *Industrial and Corporate Change, 19*(3), 839–869.

Chandy, R., Hopstaken, B., Narasimhan, O., & Prabhu, J. (2006). From Invention to innovation: Conversion ability in product development. *Journal of Marketing Research, 43*(3), 494–508.

Chen, Y. S., & Chang, K. C. (2009). Using neural network to analyze the influence of the patent performance upon the market value of the US pharmaceutical companies. *Scientometrics, 80*(3), 637–655.

Chen, Y. S., & Chang, K. C. (2010a). Analyzing the nonlinear effects of firm size, profitability, and employee productivity on patent citations of the US pharmaceutical companies by using citations of the US pharmaceutical companies by using artificial neural network. *Scientometrics, 82*(1), 75–82.

Chen, Y. S., & Chang, K. C. (2010b). The nonlinear nature of the relationships between the patent traits and corporate performance. *Scientometrics, 82*(1), 201–210.
Chen, D. Z., Huang, H. W., Huang, M. H., & Fu, F. C. (2005). Core technologies and key industries in Taiwan from 1978 to 2002: A perspective from patent analysis. *Scientometrics, 64*(1), 31–53.
Cohen, W. M., & Levinthal, D. A. (1990). Absorptive capability: A new perspective on learning innovation. *Administration Science Quarterly, 35*(1), 128–152.
Dimasi, J. A. (2001). Risks in new drug development: Approval success rates for investigational drugs, *Clinical Pharmacology and Therapeutics*, 29(May), 297–307.
Dominguez-Lacasa, I. (2006). Capturing the changes in the knowledge base underlying drug discovery and development in the 20th century and the adjustment of Bayer, Hoechst, Schering AG and E. Merck to the advent of modern biotechnology. *Scientometrics, 66*(2), 345–364.
Hall, B. H., Jaffe, A., & Trajtenberg, M. (2005). Market value and patent citations. *Rand Journal of Economics, 36*(1), 16–38.
Hill, L. L. (2005). Patent primer, The Orange Book. *Nature Reviews Drug Discovery, 4*(8), 261.
Iversen, E. J. (2000). An excursion into the patent-bibliometrics of Norwegian patenting. *Scientometrics, 49*(1), 63–80.
Koenig, M. E. D., & Mezick, E. M. (2004). Impact of mergers and acquisitions on research productivity within the pharmaceutical industry. *Scientometrics, 59*(1), 157–169.
Lo, S. C. (2007). Patent analysis of genetic engineering research in Japan, Korea and Taiwan. *Scientometrics, 70*(1), 183–200.
Lo, S. C. (2008). Patent coupling analysis of primary organizations in genetic engineering research. *Scientometrics, 74*(1), 143–151.
Mcmillan, G. S., & Hamilton, R. D. (2000). Using Bibliometrics to measure firm knowledge: An analysis of the US pharmaceutical industry. *Technology Analysis and Strategic Management, 12*(4), 465–475.
Mcmillan, G. S., & Hamilton, R. D. (2007). The public science base of US biotechnology: A citation-weighted approach. *Scientometrics, 72*(1), 3–10.
Narin, F. (1994). Patent bibliometrics. *Scientometrics, 30*(1), 147–155.
USFDA (2010) *Approved drug products with therapeutic equivalence evaluations.* 30th edn. http://www.fda.gov/downloads/Drugs/DevelopmentApprovalProcess/UCM071436.pdf. Accessed 19 April 2010.
Van Looy, B., Magerman, T., & Debbakere, K. (2007). Developing technology in the vicinity of science: An examination of the relationship between science intensity (of patents) and technological productivity within the field of biotechnology. *Scientometrics, 70*(2), 441–458.
Wang, M. H., Yu, T. C., & Ho, Y. S. (2010). A bibliometrics analysis of the performance of water research. *Scientometrics, 84*(3), 813–820.
Yang, S., Ma, F., Song, Y., & Qiu, J. (2010). A longitudinal analysis of citation distribution breadth for Chinese scholars. *Scientometrics, 85*(3), 755–765.
Yuan, J. P., Yue, W. P., Su, C., Wu, Z., Ma, Z., Pan, Y. T., et al. (2010). Patent activity on water pollution and treatment in China- a scientometric perspective. *Scientometrics, 83*(3), 639–651.