We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900
Open access books available

185,000
International authors and editors

200M
Downloads

154
Countries delivered to

Our authors are among the

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Towards More Inclusive IP Analysis by Frontier Tools

Yoshiyuki Osabe and Mari Jibu

Additional information is available at the end of the chapter

Abstract

This chapter introduces multilateral analysis on IP rights: (1) a new indicator "Innovation Front" and its use, (2) analysis of patent quality, and (3) future prospect in pharmaceutical field. Through these items, more inclusive IP analyses have been conducted. We introduce the origin, trajectory, and destination of knowledge spillovers in the science and technology system, especially in pharmaceutical field. "Innovation Front" is also covered, where it is possible to find major hotspots in basic research, which give a great influence to technologies. Readers of this chapter will find (1) the major hotspots in basic research, (2) patent quality analysis ranging from basic research to application research, and (3) an overview of drug R&D and future competitiveness in the pharmaceutical field.

Keywords: patents, non-patent literature, knowledge spillovers, knowledge flows, patent quality, pharmaceutical, drug pipelines

1. Introduction

There are no doubts about the importance of knowledge spillovers for creation of intellectual properties (IPs) and furthermore economic growth. Because the spillovers allow a better penetration and diffusion of innovation and stimulate cooperation in R&D for innovators to try to internalize knowledge flows, the role of knowledge exchange and dissemination is often as important as the role of direct investment in a knowledge- and technology-driven economy. Various reports have been published to study the origin, pass way, and end products of knowledge spillovers in the science and technology ecosystem. Patents and citations between patents and non-patent literature (NPL) are analyzed to make comprehensive grasp of knowledge spillovers [1] or to estimate patent quality [2].

Recently, we have developed a new indicator, named "Innovation Front" by calculating papers cited in patents by co-citation analysis [3]. Since papers cited in patents are close to

technology, it is possible to find major hotspots in basic research, which give a great influence to technologies.

Apart from our analysis, a new indicator "Patent-Science Link" is developed by the Organisation for Economic Co-operation and Development (OECD). This is able to understand how the technical knowledge has been flowing from the science-based study to the innovation activity [4]. The Patent-Science Link indicates that pharmaceutical patents reckoned the large part of citations made from patents to scientific papers. That is, in the fields of pharmaceutical, the science-based study is much closer to the innovation activity comparing with other fields of technology.

Therefore, we analyzed the intellectual properties in pharmaceutical science using Innovation Front and OECD's Patent-Science Link. Based on these indicators, we clarified how the knowledge flows on each pharmaceutical R&D stage and how a drug has been created as a final end product. In this study, papers, patents, and drug pipelines represent each pharmaceutical R&D stage. Especially, the indicators of the drug pipelines also showed an overview and future prospects of pharmaceutical industries.

This chapter introduces multilateral analysis on IP rights: (1) a new indicator "Innovation Front" and its use, (2) analysis of patent quality, and (3) future prospect in the pharmaceutical field.

Note that the opinions expressed in this chapter are those of the authors and do not represent those of the institutions that the authors belong to.

2. Innovation Front

The citations between scientific papers and patents have been analyzed since Narin started the study of science linkage, which is seen in the administrative process as patent examiners refer scientific papers in examining a patent [5, 6]. Japan Science and Technology Agency (hereafter JST) is one of the funding agencies in Japan with aims of promoting technology transfers and technology-based innovation. As such, it is important to investigate the JST's contribution as a result of its funding. JST has a point of view that non-patent literature (NPL) in patents is seemed to provide a hint of knowledge flows between science-based study and innovation ecosystem. JST has developed as such an indicator, the "Innovation Front" in order to show specific research areas where science significantly influences technology by calculating co-citation between patents and research papers [7].

2.1. Innovation process

For the purpose of stimulating innovation, the innovation process needs to be carefully analyzed. **Figure 1** shows that the innovation develops on the bases of the two internal actions, "knowledge embodiment (in other words, development part in R&D)" and "knowledge creation (in other words, research part in R&D)" [8]. Yamaguchi developed his idea, so-called innovation diagram that shows the way of visualizing an innovation process using two intellectual elements. They represent the element of the "knowledge embodiment (development)" and the element of the "knowledge creation (research)" (see **Figure 1**). Once the
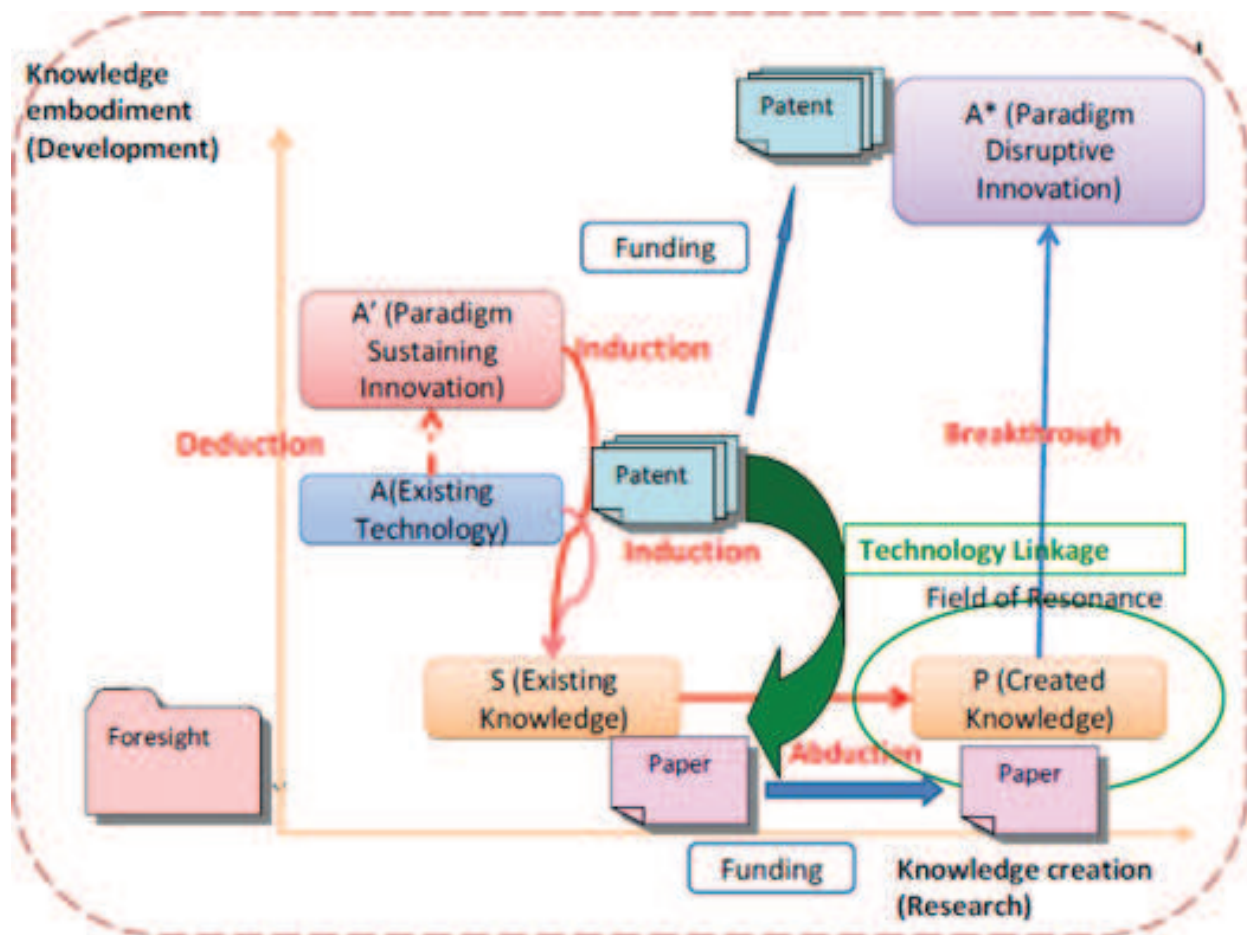
Figure 1. Innovation process and scientometrics.

technological exploitation come to a total deadlock (A″ in **Figure 1**), the R&D need going back to the basic scientific knowledge ("Induction" in **Figure 1**) and force it to move with scientific exploitation ("Abduction" in **Figure 1**) for the purpose of creating another innovation by overcoming the deadlock above (A* in **Figure 1**). Making a breakthrough and creating a new innovation will be possible by the actions above. As you see, by linking science-based knowledge to technological information, we can see a whole way of the innovation process, as shown in **Figure 1**.

Figure 1 indicates an overview of how to understand and visualize the innovation process by connecting the results such as patent analysis, bibliometric analysis, and clustering analysis by the following innovation process. In this attempt, **the linkage between patents and papers** is a key to obtain an understanding of the innovation process, which is evidence-based, such as the review of science-based knowledge from technology stage ("induction" in **Figure 1**). Because patent examiners generally cite patent and paper information as a reference, we can develop an index that is called "the science linkage index," that is, the number of forward citations and the NPL share (share of non-patent literature) citations, and the index may show whether a patent is technology-oriented or science-oriented. Many studies show that backward citations to the non-patent literature (NPL) relate the closeness between a patented

invention and science-based knowledge. On the other hand, the importance of a patent for the R&D is related to forward patent citations.

In order to create a new indicator "Innovation Front," we utilized the above-mentioned relationship between patents and papers. As shown in **Figure 2**, pairs of papers are co-cited by different patents. For example, Paper A and Paper B are co-cited by both Patent X and Patent Y. So, we can recognize that Paper A and Paper B are close to each other **in the point of technical (patent) view**. On the other hand, the relationship between Paper A and Paper C or Paper B and Paper C is not technically close because they"do not share a group of patents which co-cite them. In this way, we can create a set of clusters consisted of papers which are technically close to each other.

"Technical closeness" is the keyword in this research. Innovation front is the first indicator who can show the technical closeness between papers by calculating patent co-citations.

2.2. Methodology

2.2.1. Series of database

The series of databases used for the Innovation Front are as follows:

Papers: "Essential Science Indicators," "Web of Science" by Thomson Reuters.

Patents: "Derwent World Patents Citation Index," "Derwent World Patents Index" by Thomson Reuters.

2.2.2. Classification

Papers:

Note that 22 category codes from Thomson Reuters "Essential Science Indicators" for Innovation Front.
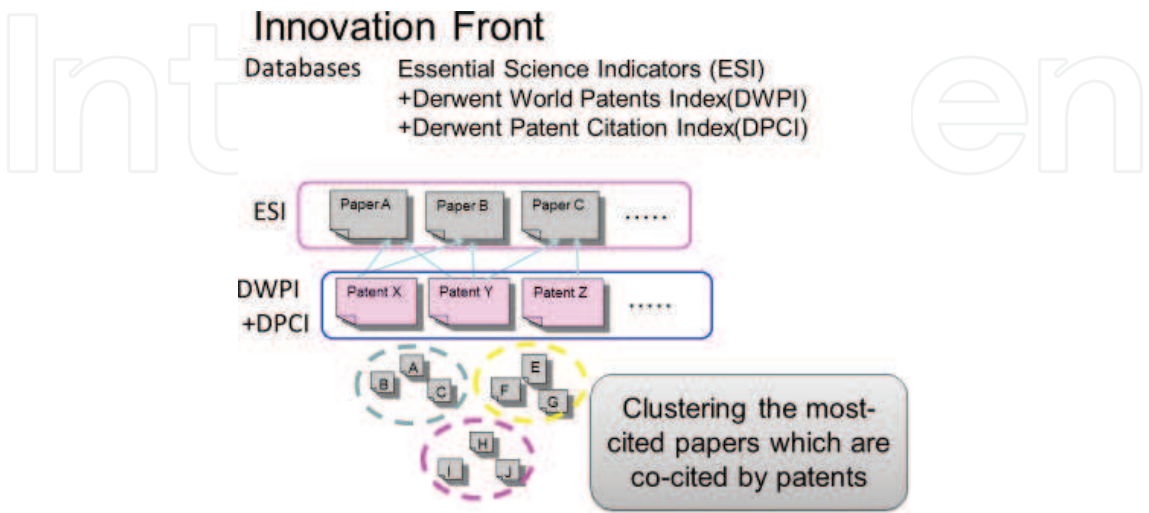

Figure 2. Structure of Innovation Front.

Patents:

(1) International Patent Classification (IPC) codes for Information and Communication Technology (ICT), biotechnology, environment-related technologies, nuclear energy, and fuel cells [1].

(2) ECLA codes for nanotechnology.

2.2.3. Calculation

(1) 1. Extract two papers arbitrarily (the papers were published between 2006 and 2010. These papers were extracted from "Essential Science Indicators."). 2. Calculate the frequency of forward co-citation by patents. This calculation is done with "Derwent World Patents Citation Index." 3. Then, calculate how frequent forward citation patents are in the two arbitrary papers.

(2) Derive cosine coefficient N from Eq. (1).

$$N = \frac{F(A, B)}{\sqrt{F(A) \times F(B)}}$$
$$F(A, B) \geq 2$$

(1)

where *A* and *B* of formula (1) show arbitrary papers, $F(A)$ and $F(B)$ show the cited frequency of the arbitrary papers, and $F(A, B)$ also shows the co-citation frequency of the arbitrary papers. Note that definition of $F(A, B)$ is larger than 2 or equal to 2.

(3) Then define $N \geq 0.3$ (note that N is larger than 0.3 or equal to 0.3).

(4) 1. Compile the papers that are extracted under the condition (3) above as nodes. 2. Connect the linkage among papers as the edge function. 3. Visualize the network of nodes and linkages, by using Cytoscape Web.

2.3. Analytical results

"Innovation Front" shows scientific hotspots by means of making clusters by co-citation analysis between scientific papers and patents. Since scientific papers cited by patent examiners are approximate to technology in the patent, we can find who are the major researchers and important scientific specialties in terms of papers having an influence on technology. Thomson Reuters publishes a "research front," that is the papers which have a strong impact on Science. It is based on the calculation of co-citation frequency of papers. JST has a similar approach in terms of patent. JST clusters the top 1% in terms of the most frequently cited papers that patent co-cite and show them as a cluster indicating an impact in terms of innovation. It has named as "Innovation Front." The clustering by the Innovation Front is shown in **Figure 3**.

Each node represents a paper and the nodes' colors express 22 kinds of fields. The node's size represents how many citations are there in papers. The line width between nodes represents
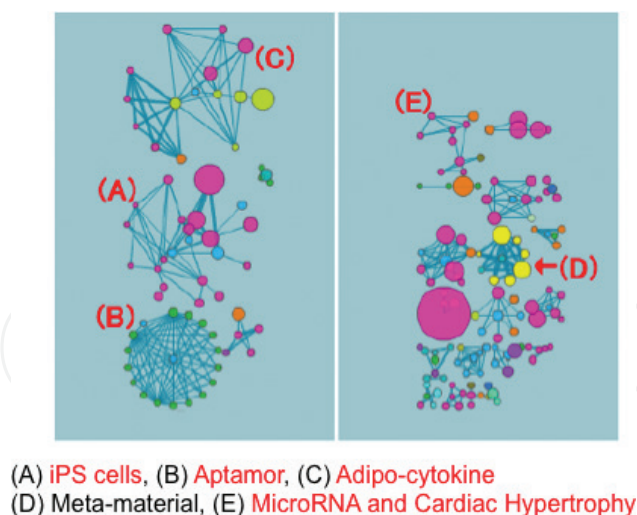
(A) iPS cells, (B) Aptamor, (C) Adipo-cytokine
(D) Meta-material, (E) MicroRNA and Cardiac Hypertrophy

Figure 3. Result of the Innovation Front.

the number of co-citations by patent families. As a result, generated 24 clusters composed of 183 papers in total. The total citations of patent families and those of papers are 1095 and 46,038, respectively. The most frequently published journal is *Science*, with 18 papers and the second top journal is *Nature* with 17 papers. As to subjects of papers in clusters of Innovation Front: the largest subject is clinical medicine (45.4%) and the second top is chemistry (14.2%).

Cluster (A) is the field of induced pluripotent stem (iPS) cells composed of 22 papers. The iPS cells represent "induced pluripotent stem cells" which was discovered by Shinya Yamanaka in Kyoto University in 2006. Among the 22 papers, the title of the core paper is "Induction of Pluripotent Stem Cells from Adult Human Fibroblasts by Defined Factors," written by Shinya Yamanaka et al. This paper was published in *Cell* in 2007.

Cluster (B) is the field of aptamor composed of 18 papers. Aptamor is a peptide or oligonucleotide that binds to a specific target molecule. The main fields of the 18 papers are engineering and chemistry. This is slightly different from the main field of the definition of aptamar above. In terms of 18 papers composing the cluster (B), the title of the core paper is done by Lee et al. in Northwestern University and the title is "Colorimetric Detection of Mercuric Ion (Hg^{2+}) in Aqueous Media using DNA-Functionalized Gold Nanoparticles." This paper was published in *Angewandte Chemie International Edition* in 2007.

Cluster (C) is the field of adipocytokine (cytokine secreted by adipose cells) composed of 15 papers. The main field is clinical medicine, immunology, and biology/biochemistry.

Cluster (D) is the field of meta-material, composed of nine papers. The main field is physics.

Cluster (E) is the field of microRNA and cardiac hypertrophy composed of nine papers. The main fields are clinical medicine and neuroscience and behavior.

Figure 4 shows the detail mapping of cluster A that represents the field of iPS cells composed of 22 papers and the number of citation papers are 7517. Squares indicate papers written by Yamanaka group and the largest square is titled "Induction of Pluripotent Stem Cells from
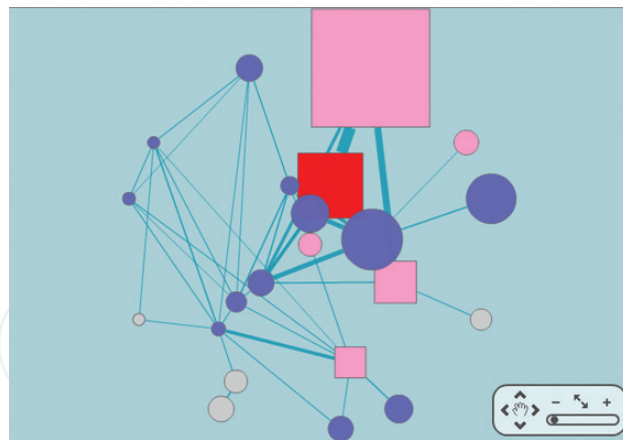
Figure 4. Innovation Front: a cluster of iPS cells.

Adult Human Fibroblasts by Defined Factors," written by Shinya Yamanaka, published in *Cell* in 2007. The aim of the core paper by our indicator is the iPS cells derived from **human fibroblast**. On the other hand, the Nobel Prize in Physiology or Medicine 2012 was awarded jointly to Sir John B. Gurdon and Shinya Yamanaka, and the Press Release by the Nobel Assembly announced" Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors" by Shinya Yamanaka as one of the key publications. This paper's aim is to develop iPS cells derived from **mouse embryonic and adult fibroblast cultures**. As the study advances from basic to practical research, subjects for study are changing from *in vitro*, *in vivo*, mice, apes, and to humans. Our indicator is able to find the papers and fields (clusters) that affect the **technology, which is closer to applied stage than science study**. This represents the result of iPS cell cluster above. In other words, Innovation Front showed the application study papers that target humans instead of *in vitro* study or mouse in the analysis of iPS cell field. Other papers of iPS cells also target humans in cluster (A). The results of Innovation Front composed of roughly six groups are presented below.

As to papers of Node 1, this is by a group of the International Consortium of Stem Cell Networks, which is a consortium of human embryonic stem (ES) cell researchers from around the world. As to Node 2, this is by Yamanaka group, which made a success in converting mice and human skin cells into iPS cells. As to Node 3, this is written by Thomson J. A. (University of Wisconsin) group. Thomson group succeeded in isolating the human embryonic stem (ES) cells in 1998 and also known as a study of human iPS cells. They published the paper related to iPS cells at almost the same time as Yamanaka group, 2007. As to Node 4, this is the group of Schoeler H. R. (Max Planck Institute) group. Schoeler groups succeeded in reprogramming of adult mouse neural stem cells by introducing on 4 October. As to Node 5, this is the group of Jaenisch R. (MIT) group. Jaenisch group had a success in terms of reprogramming from mature, differentiated mouse B cells. Finding alternatives instead of the cancer-causing retroviruses for making iPS cells is also their study. As to Node 6, this is the group of Hochedlinger, K. (Harvard University) group. Hochedlinger group used an adenovirus in order to transport the transcription factors into the DNA.

As shown above, Innovation Front shows "new geography of innovation hotspots." According to analysis by Innovation Front, the hotspots are the science fields of induced pluripotent stem

cells (iPS cells), aptamor, microRNA, cardiac hypertrophy and other fields of pharmaceutical science. Apart from Innovation Front, the OECD also introduced a new indicator called "Patent-Science Link" and indicated that patented inventions in pharmaceuticals account for the majority of citations to scientific papers in patents [4]. Therefore, we have moved forward with our aims to the analysis of intellectual properties in pharmaceutical science fields.

3. Patent quality on pharmaceutical field

Various reports have been published to show the study the origin, pass way, and end products of knowledge flows and the delays in the science and technology system. Patents and citations between patents and non-patent literature (NPL) are analyzed to make comprehensive grasp of knowledge spillovers [1] or to measure patent quality [2].

Pharmaceutical innovation is particularly important for drug discovery. There is a steady decrease in R&D productivity of drugs over the last number of years [10]. According to Scannell and Bosley, inflation-adjusted industrial R&D costs per novel drug increased nearly 100 fold between 1950 and 2010 [11]. R&D efficiency per billion US dollars of R&D spending has declined fairly steady, measured simply in terms of the number of new drugs brought to market by the global biotechnology and pharmaceutical industries. They call this trend "Eroom's Law in Pharmaceutical R&D" [12]. On the other hand, several cases of success have been found recently. For example, the reports present that drugs sourced via open innovation have a higher chance of later-phase clinical success, among 281 biopharma companies, between 1988 and 2012 [13].

We present an analysis of knowledge flows in the pharmaceutical innovation process. Backward citations, citations to NPL and forward citations that link patents, scientific papers and pharmaceutical pipelines data are analyzed and visualized to provide a more holistic understanding in hotspots of R&D. Because new drug discovery is the global issue and based on science knowledge, like biotechnology and chemistry, the analysis related to this field is eligible for science and innovation for global challenges.

3.1. Methodology

3.1.1. Dataset and its preparation

The datasets below by Thomson Reuters, which is covered from 1981 to 2011, are prepared. (1) For patent data, the Derwent World Patents Index (DWPI). (2) For papers' data, the Web of Science (WoS) database. (3) For drug pipeline data, the Thomson Reuters Cortellis for Competitive Intelligence database (hereinafter "Cortellis") including detailed information of drugs. (3) For citations data, the Derwent Patents Citation Index (DPCI). (4) For linkages between patents and papers, the WoS-DPCI Linktable computed by Thomson Reuters and JST that prepares backward citation data from patents to the NPL (non-patent literature) from the DPCI. Data were prepared on December 11, 2013.

Because of interest to find patents and their relationship to the non-patent literature in phar-maceutical fields, we prepared all 833,376 patents having one of the IPC (International Patent Classification) codes A61P that represents "specific therapeutic activity" from the Derwent Patents Citation Index and also we extracted their citations from DPCI, and we named them as "Pharma_Patents." Then, we prepared 57,800 patents that are linked to drug pipeline data from the Cortellis, named as "Drug_Patents (DP)." Then, the DP were subtracted from the patents having A61P resulting in a dataset of 325,576 "Non-Drug Pharma Patents (NDPP)." In other words, NDPP is a patent that has the code "A61P" but is not linked to drug pipelines. **Figure 5** shows the relationship between DP and NDPP.

Finally, all 115,252 NPL for Drug_Patents (DP) and 718,269 Non-Drug_Pharma_Patents (NDPP) were retrieved using the WoS-DPCI Linktable.

3.1.2. Calculation

As to patent quality analysis, patent family size, IPC counts, forward citations, backward citations, and citations to NPL were compared between 701 DP of random sampling and 701 NDPP for logistic regression analysis.

The citation lag: it can be calculated for forward citations to identify the speed by which patents are cited by future patens. It can also be used for backward citations to identify how prompt the existing works are cited by patents, see **Figure 6**. The citation lag is calculated as the average time gap of the years when the focus patent published (see **Figure 6**, patent A) minus the publication years of all cited works (see **Figure 6**, patents B, C, and NLP D). Similarly, the citation lag of all forward citations is defined as the average time gap of the publication years from all citing works (patents E–G in **Figure 6**) to the publication year of the patent A.

The generality index G_x: This is a quantitative index that represents the technical diversity of patents that are cited by a given focal patent (patent A in **Figure 6**). It also represents the technical diversity in terms of a group of patents that cite one focal patent. The count of dif-ferent IPCs, which are associated with citing and cited patents, is used for calculation of the diversity. The generality index will be high if citing and cited patents occupy a wide spread
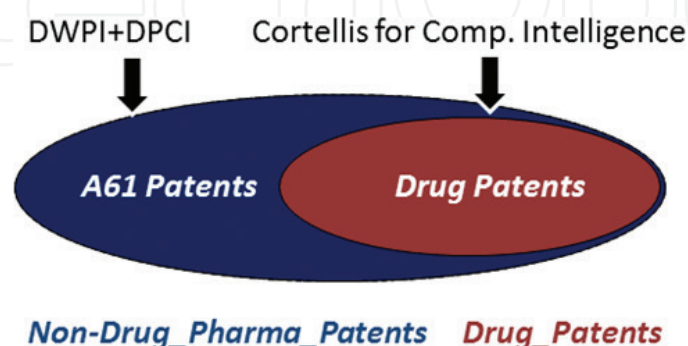

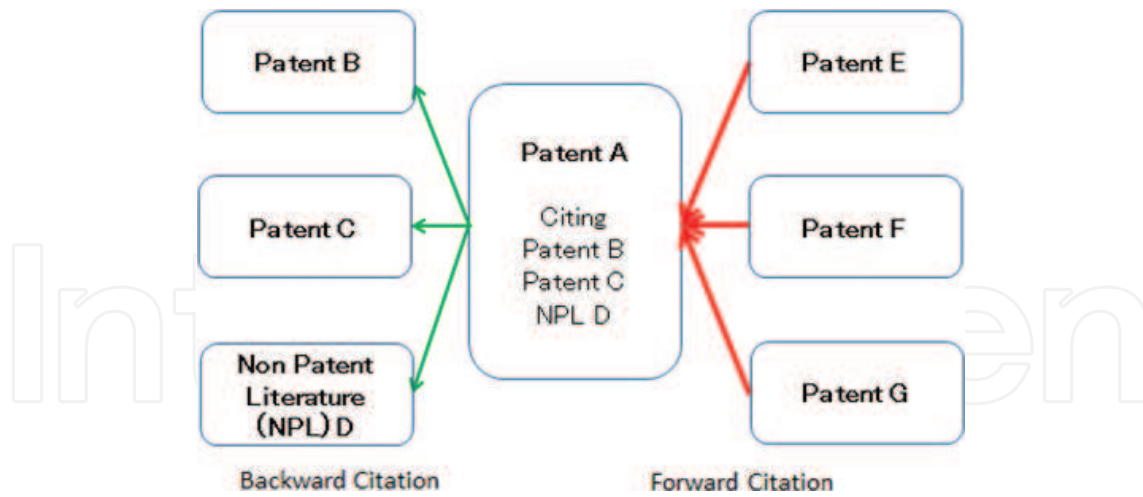**Figure 5.** Relationship between Drug_Patents (DP) and Non-Drug_Pharma_Patents (NDPP).

Figure 6. Backward citations to patents and NPL and forward citations to patents.

of technology fields. For example, when x is the focal patent and y_i patents are citing the focal patent x, with $i = 1, …, N$, then G_x can be calculated by following formula:

$$G_x = 1 - \sum_{j=1}^{M_i} \left(\frac{1}{N} \sum_{i=1}^{N} \frac{T_{ji}^n}{T_i^n} \right)^2 \tag{2}$$

Where T_i^n is the total number of IPC n-digit classes in y_i, T_{ji}^n is the total number of IPC n-digit classes in the jth IPC 4-digit class in y_i, and $j = 1… M_i$ is the cardinal of all IPC 4-digit classes in y_i.

The index was calculated using all different 6-digit IPC subclasses for all patents in Drug_Patents (DP) and Non-Drug_Pharma_Patents (NDPP).

The subject index S_x: this is a new indicator we proposed. It is based on the generality index above, but the difference is that it is computed for NPL. For example, when x is the focal patent which cites y_i, $i = 1, …, N$ scientific papers (NLP), then S_x can be calculated as following formula:

$$S_x = 1 - \sum_{j=1}^{j} \left(\frac{N_{ij}}{N_i} \right)^2 \tag{3}$$

Where N_i is the total number of subject code in y_i and N_{ij} is the total number of subject code in jth the subject code in y_i.

The subject codes attached in each scientific paper are counted using basic element 1. The subject index was calculated for all non-patent literatures cited by patents in DP and NDPP.

The patent scope "SCOPEp": this is often associated with the technological and economic values of patents. It is said that broad scope patents tend to have a higher value (Lerner, 1994). For each patent P, the patent scope is defined as:

$$SCOPE_p = n_p; n \in \left\{ IPC_1^4; … IPC_i^4, … IPC_j^4; …; IPC_n^4 \right\} \text{ and } IPC_i^4 \neq IPC_j^4 \tag{4}$$

where n_p denotes the number of distinct 4-digit IPC subclasses listed in the patent P and is normalized. The patent scope was calculated for all distinct 6-digit IPC subclasses for all patents in Drug_Patents and Non-Drug_Pharma_Patents.

3.2. Analytical results

3.2.1. Patent quality

Table 1 shows the results of logistic regression analysis. As a result, forward citations ($P < 0.001$), IPC count ($P < 0.001$) and also citations to NPL ($P < 0.05$) are significantly associated with patent quality. Therefore, forward citations, IPC count, and citations to NPL are emerged as new indicators for distinguish genuine patents that have strong linkage with R&D processes from other patents related to drug.

3.2.2. Citation lag: technology delays

Comparison of citation lags for Drug_Patents and Non-Drug_Pharma_Patents shows the dynamics of knowledge spillovers. **Table 2** represents that the forward citation lag of Non-Drug_Pharma_Patents is 2.17 years later on average while Drug_Patents are cited faster—after 1.89 years on average. Here, we can see the high-quality patent (Drug_Patents) tend to be referenced faster than other patent (Non-Drug_Pharma_Patents) The backward citations lag of

	Coefficient	Std. error	z-value	P-value	Sig.
Intercept	−0.05474	0.0824	−0.664	0.50638	
IPC count	−0.009	0.0018	−4.89	1.01E−06	0.1%
Forward cites by patent	0.020355	0.0031	6.645	3.03E−11	0.1%
Backward cites to patent	0.003728	0.0025	1.516	0.12958	
Backward cites to NPL	0.003718	0.0014	2.626	0.00864	1%
Patent family size	−0.00062	0.0041	−0.153	0.87859	

Table 1. Logistic regression analysis.

	NDPP	DP
Forward cites by patents	2.17	1.89
Backward cites to patents	3.40	5.64
Backward cites to NPL	1.69	2.50

Table 2. Forward and backward citation lags.

Non-Drug_Pharma_Patents to patents is 3.4 years earlier on average and they go to much more recent non-patent literature (NPL)—published on average only 1.69 years earlier. Surprisingly, Drug_Patents cite older works comparing with Non-Drug_Pharma_Patents : cited NPL are 2.5 years old on average and cited patents are 5.64 years old. All values are statistically significant at the 1% level. In resume, we can see that Drug_Patents cover wider ranges and are cited more quickly comparing with Non-Drug_Pharma_Patents.

3.2.3. Generality index, subject index: knowledge diversity

The generality index was calculated for 4- and 6-digit IPCs. It was analyzed for backward and forward citations, also for Non-Drug_Pharma_Patents and Drug_Patents, see **Table 3**. Drug_ Patents have higher generality index and subject index than Non-Drug_Pharma_Patents. That is, in general, Drug_Patents is based on wider "technologically based knowledge" and is cited by a wider range of set of patents that have more diversified IPCs. All values are statistically significant at the 1% level.

3.2.4. Scope: technology value

The patent scope was calculated for NDPP and DP, see **Table 4**. Contrary to expectation, in the pharmaceutical fields, the scope of DP (Drug_Patents) tends to be narrower than that of NDPP (Non-Drug_Pharma_Patents). This is unexpected as, in general, patents linked to drugs are seemed more valuable than those not linked to drugs.

Some possibilities are conceivable below: (1) NDPPs are used for protect peripheral technologies surrounding one core DP and broader patents are useful to protect an inconsequential broad area, (2) because it takes long time to get an approval of drug, a patent owner has to obtain the other patent which is narrower than original patent, like second use patent or formulation patent, in order to extend a patent term, and (3) pharmaceutical companies might hide their core patent, therefore, Thomson Reuters cannot link drug pipelines to an appropriate patent and so on.

		NDPP	DP
Generality index (4-digits)	Forward citations	0.36	0.37
	Backward citations	0.40	0.54
Generality index (6-digits)	Forward citations	0.46	0.50
	Backward citations	0.52	0.73
Subject Index	Backward citations to NPL	0.22	0.28

Table 3. Generality index.

Scope	NDPP	DP
4-digit	0.13	0.11
6-digit	0.16	0.15

Table 4. Scope for Non-Drug_Pharma_Patents and Drug_Patents.

3.3. Knowledge flows in pharmaceutical innovation

Our study compared and contrasted patents that are linked or not linked to drugs to understand knowledge flows and delays in pharmaceutical innovation. Results are summarized in **Figure 7**. As can be seen, Drug_Patents ("patents linked with drugs" in **Figure 7**) draw from a more diverse set of technologies and are cited more widely across the technology landscape. However, they tend to be more technically specialized (lower scope) than Non-Drug_Pharma_Patents ("patents linked without drugs" in **Figure 7**). Concerning citation lag, Drug_Patents tend to refer to older patents and scientific papers and are cited faster than Non-Drug_Pharma_Patents.
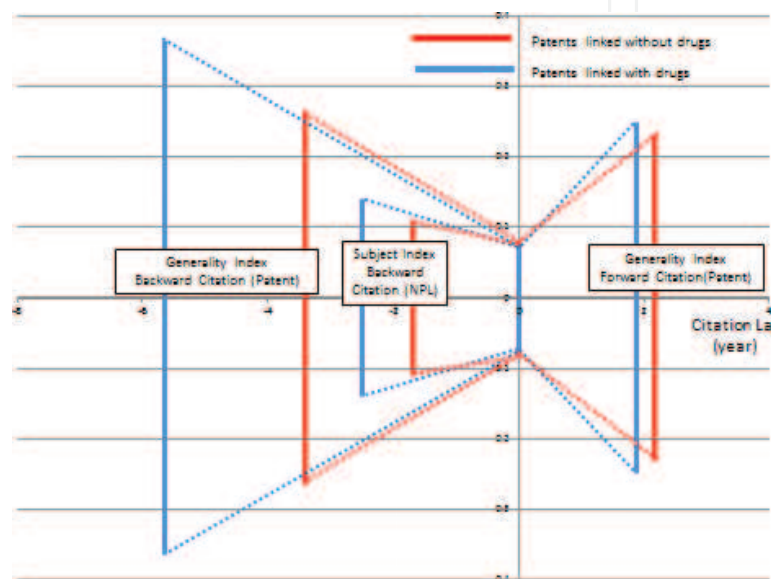

Figure 7. Comparison of Non-Drug_Pharma_Patents and Drug_Patents.

4. An overview and future prospects of pharmaceutical industry

For the sake of providing evidence that contribute to policy making or strategy planning in national governments and pharmaceutical companies, we tried to show an overview and future prospects of the pharmaceutical industry based on data related to science papers, patents, drug pipelines and other various data of enterprises.

As to the evidence for policy making, there have been many analyses on current status, based on drug sales and R&D expense. For example, OECD has published the report overviewing the pharmaceutical industry in a global market based on the indicators of R&D expense, trade balance and term of drug approval in September 2009 [9, 14]. In addition, many reports on drugs of top sales by research companies and publications of in-house products by pharmaceutical companies have been published. What is more, the global competitiveness ranking of pharma companies is announced by collecting and arranging the information. Although these indicators represent the R&D capacity or competitiveness of the past and the present, they cannot foresee the future status.

We aimed to focus on the R&D pipelines that pharmaceutical company in each country have. The results indicated the R&D potential of each country in the current stage and in the future, by compiling the pipeline data in each R&D step, comparing the number of marketed pharmaceuticals and so on.

4.1. Drug development and pipelines

Drug development is the process of bringing a new pharmaceutical lead compound to the market. There are several stages in drug development. It starts from "Basic research" stage, then moves to "Pre-clinical," "Clinical Phase," "Filed and Approved," and finally it goes to "Marketed" stage. Sometimes, it starts from intermediate stage like "Clinical Phase" when a company installs a lead compound from other company by licensing or M&A. Lead compounds in these stages are called as "R&D pipelines" or simply as "pipelines."

"Pre-clinical" is a stage of research that begins before clinical trials. Typically, animal testing and *in vitro* testing (test with microorganisms or cells) will be performed.

"Clinical Phase" is a stage of research on human participants, tested for safety and effectiveness in humans in order to be validated for a therapeutic use by a ruling authority of a government. "Clinical Phase" is composed of three stages. "Phase I" trials determine safety and dosing, usually in healthy volunteers. "Phase II" trials are the test in order to obtain an initial reading of efficacy and presumable safety in small groups of patients suffered by the disease targeted by the lead compound. "Phase III" trials are large and pivotal tests to decide its safety and also efficacy in sufficiently large groups of patients suffered by the targeted disease.

When these trials prove its adequate safety and efficacy, drug development goes to "Filed" stage, where a new drug application is filed to the ruling authority. After safety and efficacy are adequately confirmed by the authority, the application is approved. It means that drug development goes to "Approved" stage. When sales of the drug starts, the stage goes to "Marketed." Because the several stages above are need to be achieved in order to bring a lead compound to the market, it needs immense of R&D expense and dozen years by selecting an appropriate one from several tens of thousands of compounds and bring it to upper stages.

4.2. Methodology

4.2.1. Data acquisition and preparation

Two datasets of Evaluate and Thomson Reuters are used in this analysis. The dataset of Evaluate we used is "EvaluatePharma," which includes the pipelines data (about 45,000), licensing data, and M&A data from big pharma and biotech companies (about 7560 companies) in the world. It includes the data from 1986 to 2012. Also, the dataset of Thomson Reuters we used is "Cortellis for Competitive Intelligence," which includes timely global information on over 61,000 drugs, 6,000,000 patents, and 44,000 deals from big pharma to biotech companies.

4.2.2. Pipelines

Pipelines datasets from the EvaluatePharma of Evaluate are used in this analysis. The EvaluatePharma is the database containing R&D pipelines data, licensing data and M&A data, etc. in about 7560 pharma or biotechnology companies. We extracted 43,057 products, which are drug products and pipelines, related to small molecule drug or biomedicine covering 1986–2012 on May 2013. As to licensing data, we extracted in-licensed and out-licensed data covering January 2006–May 2013 from the EvaluatePharma. The Cortellis for Competitive Intelligence database by Thomson Reuters is also used in the pipelines analysis. Note that 21,086 pipelines are extracted and compiled on 11 December 2013.

4.2.3. Categories of business entities

Focusing on the sizes and categories of the business entities possessing pipelines, we classified 43,057 products data from the EvaluatePharma into eight categories, which are SMEs&VBs (Ventures), Majors, Generics, Specialties, Universities, Government, NGOs, and Others.

4.3. Analytical results

Figure 8 represents the R&D pipelines and marketed drugs covering from 1986 to 2012 in terms of small molecule drugs. **Figure 8** reveals that the number of US "Marketed" is immense, Japan is the second, followed by Germany, Switzerland, the United Kingdom, and France. The same goes for bio-medicines that include recombinant product, bioengineered vaccine, monoclonal antibodies, cell therapy, gene therapy, and DNA&RNA therapeutics. In terms of bio-medicines, the USA is the most competitive among the countries [15]. Later analysis revealed that the R&D pipelines in Korea are also high, however, they include many generic drugs [16].

Then we focused on the types of business entities that have R&D pipelines.

The conventional R&D process was "**closed innovation"** which happened in closed environment, like inside of laboratories in major pharmaceutical companies, where they discovered a lead compound and brought it into the market by themselves. This was because they had strength in chemical synthesis and took advantage of the strength to R&D of small molecular drugs. By bringing a blockbuster to the market, they got huge income to make an investment to the R&D of next drug. However, a new business model has recently been necessary for the pharmaceutical industry in order to cope with higher risk of drug development caused by longer term of R&D, increasing R&D expense. Examples are the capital expansion by a merger of major pharma companies and the role-sharing between major pharma companies and ventures. Recently "**open innovation"** has been attention-getting, which looks for the R&D pipelines outside of companies in order to reduce the higher risk. A report said that the success rate of drug R&D with open innovation would be three times as one with conventional R&D [13].

Therefore, we tried to find what kind of business entities have an important role in the drug R&D process, by analyzing the types, scales, and licensing activities of business entities having
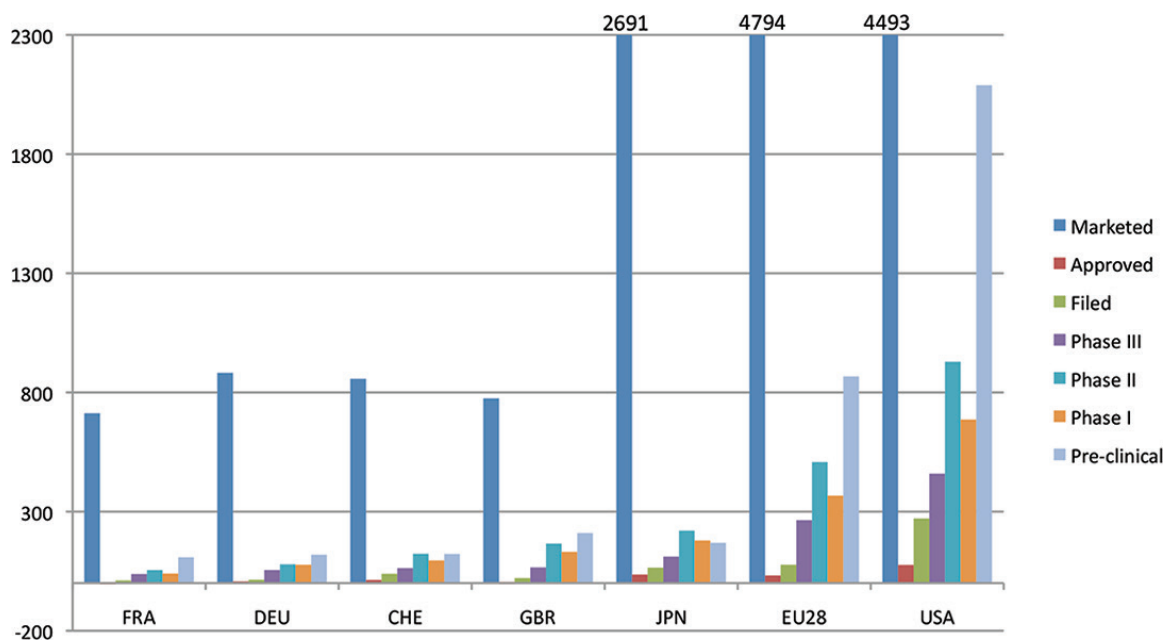
Figure 8. R&D pipelines and marketed drugs by country.

R&D pipelines. The business entities having R&D pipelines are divided into eight categories such as "SMEs, VBs," "Majors," "Generics," "Specialties," "Universities," "NGOs," "Governments," and "Others" by using the database EvaluatePharma. **Figure 9** shows licensing activity by country. The number of licensing in the USA is predominantly large. This is because the USA has the largest number of pipelines and its market is the biggest in the world. Following to the USA, the number of licensing activities in Japan is second large, but about one-third of the USA.

Then, **Figure 10** shows licensing activities by categories of business entities by country. Although the largest number of licensing in the USA is "SMEs, VBs," the largest number in Japan and EU countries are "Majors." As we show the predominance of the USA in **Figure 1**, it seems that a key role of the predominance is strong activity of SMEs and ventures. Later analysis revealed that the pipelines in the USA flowed not only from a university to a major via SME&VB, but also they flowed from a major to a SME&VB and from an SME&VB to other SME&VB. The multilateral flows of pipelines constitute, so to speak, "the roundabout of drug R&D" and that is the strength of the USA [16].

4.4. Drug R&D status by IPC

The International Patent Classification (IPC) has a subclass "A61P" which represents "specific therapeutic activity of chemical compounds or medicinal preparations." We analyzed the R&D status by providing IPC subclass to the therapy field that each pipeline has.

Figure 11 shows the origin countries of marketed drugs by IPC. We prepared the drug data of the countries (the USA, Japan, the UK, Switzerland, Germany, France, and Korea) that have many pipelines. As to the marketed drugs, the fields of infections and cancers are hotspots.
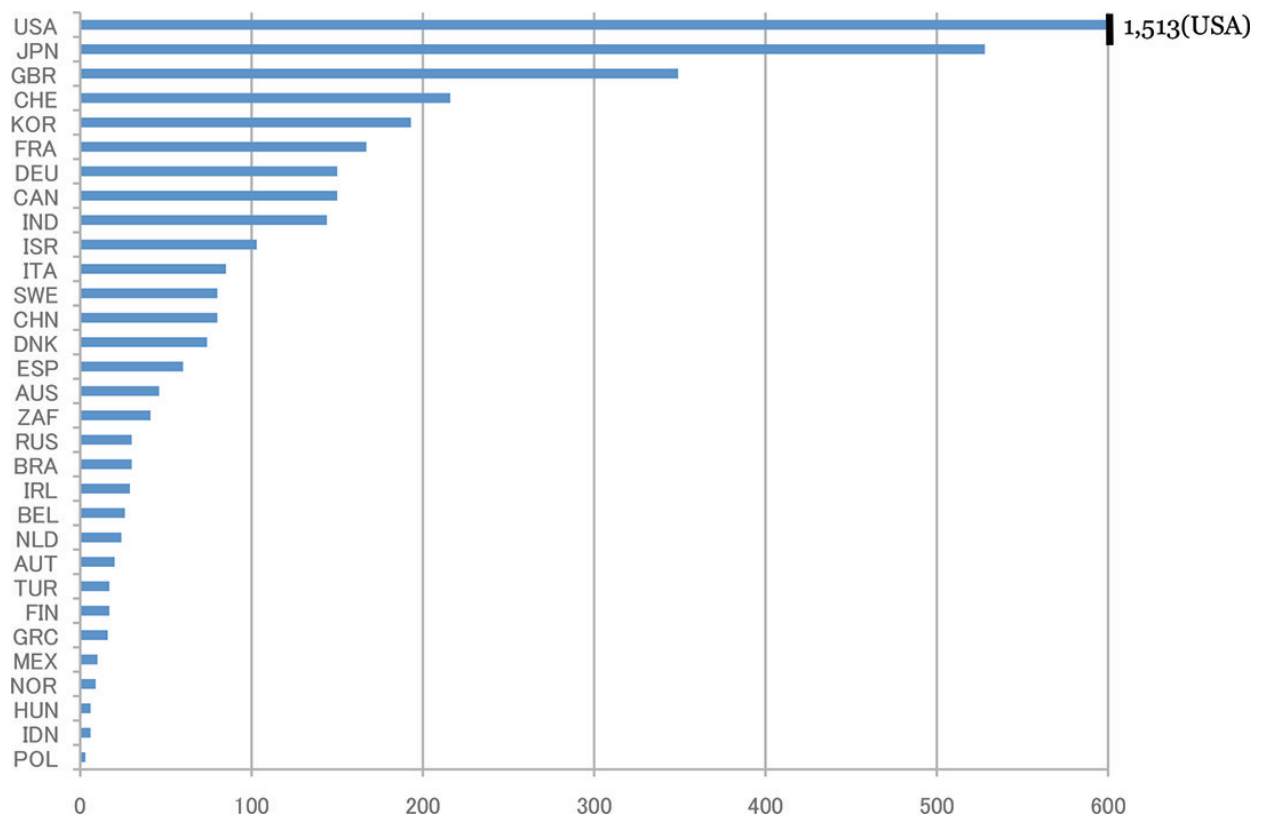
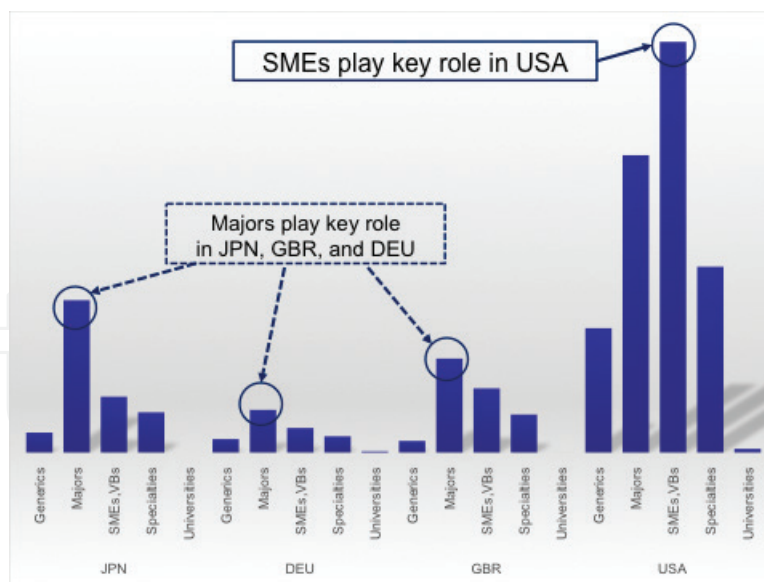**Figure 9.** Licensing activities by country.

Figure 10. Licensing activity by category of business entities.

The number of Japan's original drugs is also relatively large and the ratio to total is about 20%, although the number of the US origins is the largest (about 44%). **Figure 12** shows the origin countries of R&D pipelines by IPC. In order to overview the present R&D status, **Figure 12**
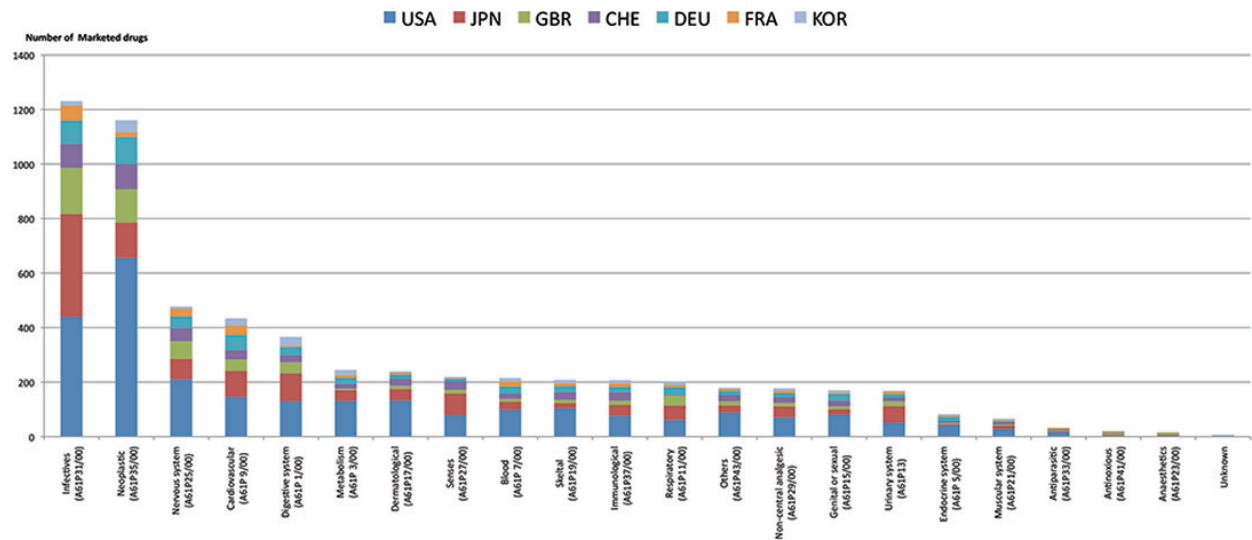
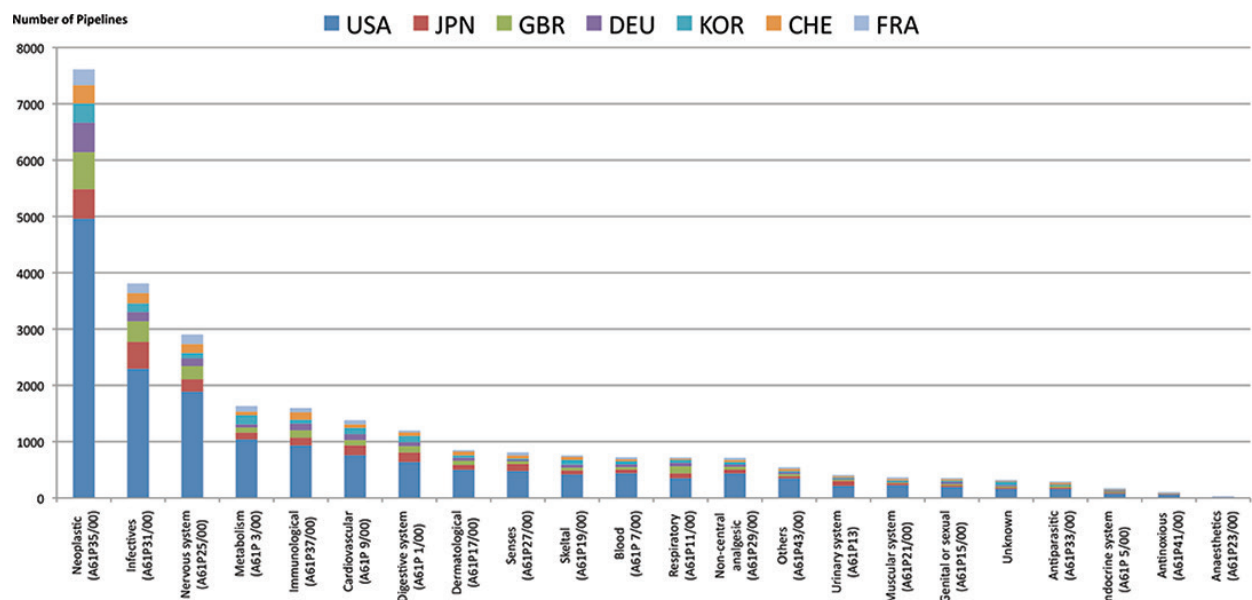**Figure 11.** Origin countries of marketed drugs by IPC.

Figure 12. Origin countries of pipelines by IPC.

does not include the marketed drugs. It reveals that, as to drug candidates that are in process of R&D, the fields of cancers, infections, and mental disorders are largest. It also reveals that the pipelines of the US origin are largest.

Cancer is the top rank in cause of mortality among developed countries. It is also the top rank in the DALY (disability-adjusted life year; a new metric based on the sum of years of potential life lost due to premature mortality and the years of productive life lost due to disability) and in the YLD (years lived with disability) [17, 18]. We can see that anti-cancer drugs have been under development by many countries.

As to infection, IPC code A61P31/00 includes antivirus for HIV, influenza, antimicrobials, and antifungal. Infections are a major cause of high mortality and low DALY among developing countries. Disorders of the nervous system (A61P 25/00) include neurodegenerative disorders like Alzheimer's disease or Parkinson's disease, antidepressants, and antipsychotics like schizophrenia. Comparing to "Marketed," the number of pipelines is larger, especially in "discovery" stage that represents basic research and pre-clinical stage. It is noted that this is the field with high future growth potential.

5. Final discussion

This chapter discussed multilateral analysis on IP rights: (1) a new indicator "Innovation Front" and its use, (2) analysis of patent quality, and (3) future prospect in the pharmaceutical field.

In the part (1), a new indicator "Innovation Front" showed scientific hotspots by focusing on "Technical closeness," that is to say, by means of making clusters by co-citation analysis between scientific papers and patents. It shows that the science fields of induced pluripotent stem cells (iPS cells), aptamor, adipocytokine, meta-material, and microRNA are the hotspots. In the part (2), we prepared patent datasets that are linked or not linked to drug pipelines in order to understand knowledge flows in the drug R&D field. The results indicate that DPs (Drug_Patents) are based on a wider range of technologies and are cited wider technology landscape. Contrary to expectation, they have a narrower scope. It means that they tend to be more technically specialized than NDPP (Non-Drug_Pharma_Patents). Concerning citation lag, DPs seem to refer to older documents and are cited faster than NDPPs. In the part (3), we showed an overview and future prospects of pharmaceutical industries, focusing on drug pipelines, size of business entities, and the International Patent Classification (IPC). The intricacies of patenting of pharma products by country-based and business entity-based were discussed.

Since many years, diverse studies have been conducted to study the origin, trajectory, and destination of knowledge spillovers in the science and technology system. The work presented here also contributes to a study for innovation analysis by showing a newly developed indicator (Innovation Front) and a new way of analysis, like comparison of Drug_Patents to Non-Drug_Pharma_Patents. Our next approach will be a linkage of "trademark" database to other database like patents and drug pipelines and also "financial analysis" linked to IP rights.

Author details

Yoshiyuki Osabe[1]* and Mari Jibu[2]

*Address all correspondence to: osabe-yoshiyuki@jpo.go.jp

1 Japan Patent Office (JPO), Tokyo, Japan

2 Japan Science and Technology Agency (JST), Tokyo, Japan

References

[1] Lukach R, Plasmans J. Measuring knowledge spillovers using patent citations: Evidence from the Belgian firm's data. CESifo Working Paper NO.754 Category 9: Industrial organization; July 2002

[2] Squicciarni M, Dernis H, Criscuolo C. Measuring patent quality: Indicators of Technological and Economic Value. Vol. 3. OECD/DSTI/DOC;2013.

[3] Jibu M, Osabe Y. Knowledge flows and delays in the pharmaceutical innovation. Business Management Studies. 2016;**2**(1):88-94

[4] OECD. OECD Science, Technology and Industry Scoreboard 2013: Innovation for Growth. Paris, France: OECD Publishing; 2013

[5] Narin F, Pozek RP. Bibliometric analysis of US pharmaceutical industry research performance. Research Policy. 1988;**17**:317-330 http://dx.doi.org/10.1016/0048-7333(88)90039-X

[6] Narin F, Hamilton K, Olibastro D. The increasing linkage between US technology and public science. Research Policy. 1997;**26**:317-330. http://dx.doi.org/10.1016/S0048-7333(97)00013-9

[7] Jibu M. Analysis achievements of JST operations through scientific patenting: Linkage between patents and scientific papers. In: Atlanta Conference on Science and Innovation Policy, IEEE; 15-17 Sept. 2011, Atlanta, GA, USA. pp. 214-220. http://dx.doi.org/10.1109/acsip.2011.6064473

[8] Yamaguchi E. Innovation theory leading to breakthrough—How will the memory device industry face the "post-Moore" future? Oyo Buturi. 2010;**79**(12):1077-1083

[9] OECD. Compendium of Patent Statistics, 2008. Organization for Economic Co-operation and Development; 2008.

[10] Booth B, Zemmel R. Prospects for productivity. Nature Reviews Drug Discovery. 2004;**3**:451-456

[11] Scannell JW, Bosley J. When quality beats quantity: Decision theory, drug discovery, and the reproducibility crisis. PLoS ONE. 2016;**11**(2):e0147215

[12] Scannell JW, Blanckley A, Boldon H, Warrington B. Diagnosing the decline in pharmaceutical R&D efficiency. Nature Reviews Drug Discovery. 2012;**11**:191-200

[13] Deloitte Development LLC. Executing an Open Innovation Model: Cooperation is Key to Competition for Biopharmaceutical Companies [Internet]. 2015. Available from: https://www2.deloitte.com/content/dam/Deloitte/us/Documents/life-sciences-health-care/us-lshc-open-innovation.pdf (Accessed: January 19, 2017)

[14] Lerner, J. (1994). The Importance of Patent Scope: An Empirical Analysis. The RAND Journal of Economics. **25**(2):319-333.

[15] OECD. OECD Health Policy Studies: Pharmaceutical Pricing Policies in a Global Market. OECD [Internet]. 2008-09-24. Available from: http://www.oecd.org/els/pharmaceutical-pricing-policies-in-a-global-market.htm (Accessed: January 19, 2017)

[16] Osabe Y, Jibu M. Development of new indicators for the launch of a Japanese version of the NIH (1): An overview and future prospects of pharmaceutical industry based on new indicators, Journal of Information Processing and Management. 2013;**56**(7):448-458

[17] Osabe Y, Jibu M. Development of new indicators for the launch of a Japanese version of the NIH (3) : Analysis of entities conducting pharmaceutical R&D activities, Journal of Information Processing and Management. 2013;**56**(10):685-696

[18] World Health Organization (WHO). About the Global Burden of Disease (GBD) Project [Internet]. Available from: http://www.who.int/healthinfo/global_burden_disease/about/en/index.html, (Accessed: January 19, 2017)