Knowledge Flows and Delays in the Pharmaceutical Innovation System

Mari Jibu[1], Yoshiyuki Osabe[2], and Katy Börner[3]
[1] OECD, Paris (France) and Japan Science and Technology Agency, Tokyo (Japan)
[2]Japan Patent Office, Ministry of Economy, Trade and Industry, Tokyo (Japan)

[3] *katy@indiana.edu,*
CNS, ILS, SOIC & IUNI, Indiana University, Bloomington, IN (USA)

Introduction

This paper presents an analysis of knowledge flows in the pharmaceutical innovation process. Backward citations, citations to non-patent literature (NPL), and forward citations that link patents, scientific publications, and pharmaceutical pipelines data on drug developments are analyzed and visualized to provide a more holistic understanding. Results show that patents linked to drugs tend to be technically specialized when compared to patents without linkages to drugs. Moreover, patents linked to drugs tend to cite older patents and scientific publications and impact wider technological and scientific fields than pharmaceutical patents not linked to drugs.

Diverse studies have been conducted to study the origin, trajectory, and destination of knowledge flows and the delays in the science and technology system. Patents and citations between patents and to non-patent literature (NPL) are analyzed to understand knowledge spillovers (Lukach & Plasmans, 2002) or to measure patent quality (Squicciarni et al., 2013). The OECD Science, Technology and Industry Scoreboard 2013 (OECD, 2013) uses comprehensive and up-to-date data to report on knowledge flows via collaboration networks (e.g., derived from co-authored publications and co-inventors on patents), international migration of researchers (e.g., estimated from changes in author's addresses on publications), but also flows of royalty and license fees for technologies. Recently, the OECD introduced a new indicator, called "Patent-Science Link," that aims to measure knowledge flows between the science base and the innovation system (OECD, 2013). According to this new indicator, patented pharmaceutical inventions account for the majority of citations made from patents to scientific publications. That is, the distance between the science base and the innovation system is much closer in pharmaceutical fields than it is in other technological fields. Pharmaceutical innovation is particularly important for drug discovery, as research and development (R&D) costs are huge and major challenges exist for arriving at cost-effective new drugs. In fact, there is a steady decrease in R&D productivity over the last number of years (Booth & Zemmel, 2004).

The structure of the paper is as follows: The next Section details data acquisition and preparation. This is followed by a description of the methodology and results. The paper concludes with a discussion of key insights and their comparison to prior work.

Data Acquisition and Preparation

Five datasets by Thomson Reuters covering 1981 to 2011 are used in this analysis. (1) Publication data from the *Web of Science (WoS)* database. (2) Patent data from the *Derwent World Patents Index (DWPI)* and associated citations from the (3) *Derwent Patents Citation Index (DPCI)*. (4) Linkages between publications and patents come from the *WoS-DPCI Linktable* computed by Thomson Reuters and JST that provides information on backward citations from patents and to the non-patent literature (NPL), i.e., scholarly publications, derived from the *DPCI*. (5) Drug pipeline data was retrieved from the *Cortellis for Competitive Intelligence* database including detailed information of exactly drugs a patent is associated with. Data was compiled on December 11, 2013.

Interested to identify patents and their linkages to the NPL in pharmaceutical fields, we extracted all 833,376 patents with the International Patent Classification (IPC) code "A61P: Specific therapeutic activity of chemical compounds or medicinal preparations" from the *DWPI* with their citations from *DPCI*, called "Pharma_Patents." Then, we extracted 57,800 patents linked to pipeline data from the *Cortellis for Competitive Intelligence* database, called "Drug_Patents." Next, the *Drug-Patents* were subtracted from the *A61P-Patents* resulting in a dataset of 325,576 "Non-Drug Pharma Patents" that have the A61P code but are not linked to drugs.

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

Methodology

Four metrics were computed: (1) *citation lag*; (2) *generality index* computing the diversity of patents that are cited by a given focal patent as well as the diversity of patents that are citing the focal patent;

(3) *subject index*, a new indicator based on the generality index but computed for NPL; (4) *patent scope***,** often associated with the technological and economic value of patents with broad scope patents having a higher value (Lerner, 1994).

Results

Using the four metrics, a number of novel results can be computed.

Technology Delays: Citation Lag

Comparing citation lag data for *DP* and *NDPP* reveals the temporal dynamics of knowledge flows. Table 1 shows that forward citations from *NDPP* come from patents that were published on average 2.17 years later while *DP* are cited faster—after 1.89 years on average. Backward citations from *NDPP* go to patents that were published on average 3.4 years earlier and they go to much more recent NPL—published only 1.69 years earlier on average. Interestingly, *DP* cite older works than *NDPP*: Cited patents are 5.64 years old and cited NPL are 2.5 years old on average. All values are statistically significant at the 1% level. In sum, they show that *DP* cover larger temporal ranges and are cited more quickly than *NDPP.*

Table 1. Forward and Backward Citation Lags.

	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

Technology Diversity: Generality & Subject Index

The generality index was calculated for 4- and 6-digit IPCs for forward and backward citations for *NDPP* and *DP*, see Table 2. *DPs* have higher generality index and subject index than *NDPP*. That is, on average, *DP* draw on more diverse technology "base knowledge" and are cited by a more diverse set of patents that have more varied IPCs. All values are statistically significant at the 1% level.

Table 2. Generality Index for Forward Citations (FC) and Backward Citations (BC).

		NDPP	DP
Generality Index (4-Digits)	FC	0.36	0.37
	BC	0.40	0.54
Generality Index (6-Digits)	FC	0.46	0.50
	BC	0.52	0.73
Subject Index	BC to NPL	0.22	0.28

Technology Value: Scope

The patent scope was computed for *NDPP* and *DP*, see Table 3. The scope of *DP* is lower than that of *NDPP*. This is unexpected as patents linked to drugs are presumably more valuable than those not linked to drugs.

Table 3. Scope.

	NDPP	DP
Scope (4-Digits)	0.13	0.11
Scope (6 Digits)	0.16	0.15

Conclusions

This paper compared and contrasted patents that are linked or not linked to drugs to understand knowledge flows and delays in pharmaceutical innovation. The results indicate that *Drug_Patents* 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*. Concerning citation lag, *Drug_Patents* tend to refer to older patents and scientific publications and are cited faster than *Non-Drug_Pharma_Patents*.

In our prior work, we introduced new drug-patent indicators for identifying patents related with pharmaceutical entities' R&D progress (Jibu & Osabe, 2014) and that IPC count, forward citations, and citations to NPL are efficient drug-patent-indicators. The work presented here is novel is that it shows that citation lags and the generality of backward citations are statically significantly different for *Non-Drug_Pharma_Patents* and *Drug_Patents*.

Acknowledgments

We would like to thank Fernando Galindo-Rueda for his expert comments and support of this research. This work was partially funded by the National Institutes of Health under awards P01AG039347, U01GM098959, and U01CA198934.

References

Booth, B., & Zemmel. R. (2004). Prospects for Productivity. *Nature Reviews Drug Discovery 3*, 451-456.

Jibu, M. & Osabe, Y. (2014). Refined R&D Indicators for Pharmaceutical Industry. *Future Information Technology, Lecture Notes in Electrical Engineering*, 309, 549-554.

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

Lukach, R. & Plasmans. J. (2002). Measuring knowledge spillovers using patent citations: evidence from the Belgian firm's data. CESifo Working Paper NO.754 Category 9: Industrial organization,

OECD, (2013). *OECD Science, Technology and Industry Scoreboard 2013: Innovation for Growth.* Paris, France: OECD Publishing.

Squicciarni, M., Dernis, H. & Criscuolo, C. (2013). Measuring Patent Quality: Indicators of Technological and Economic Value. *OECD/DSTI/DOC* 3.