feature

Potential effects of increased openness in pharma: the original knowledge behind new drugs

Francesca Bignami[1], francesca.bignami@ki.se and Pauline Mattsson[1,2]

This study seeks to determine potential changes in the degree of openness in pharmaceutical R&D by investigating where the knowledge behind new molecular entities (NMEs) comes from in terms of type of organization, geography and time. We find that the organizations granted NMEs increasingly rely on external knowledge sources but that these are increasingly shared among NME grantees. Universities are the most important indirect knowledge contributor and their relative importance has increased with time. NME grantees are increasingly relying on knowledge from different countries and the age of the knowledge sources confirms that recent NMEs are mostly follow-on drugs. This work provides evidence of the increasing openness of pharma to new knowledge sources as a means to improving the drug discovery and development process.

Introduction

Historically, most R&D was carried out internally in pharmaceutical companies. The internal R&D department performed the upstream, basic research to elucidate the underlying mechanisms and pathways of disease to identify promising points of intervention as well as the applied research to discover compounds that could be used to treat diseases and to test them in the clinical phase [1]. Over the past decades, however, the pharmaceutical industry has seen increasing regulatory demands and a low outcome rate of new drug approvals despite increasing R&D investments [2,3]. These challenges have resulted in pharmaceutical companies becoming more open to external influences with the aim of changing the traditional drug development process by increasing

its speed, lowering barriers and searching for knowledge from public and private actors that are geographically dispersed [4]. Market-leading companies such as Pfizer and Novartis have set up research units closely located to elite universities and premier hospitals [5,6]. Eli Lilly has used crowdsourcing platforms to screen for new compounds [7]. In addition, public–private partnership consortia such as the Innovative Medicine Initiative (IMI) have been established with the intention to boost pharmaceutical innovation in Europe [9]. Here, we estimate the degree of openness in pharmaceutical R&D over time by studying where the knowledge behind new drugs comes from. We propose a proxy metrics of openness based on an analysis of the prior art associated with patents behind new molecular entity (NME) filings.

Previous studies have already analyzed the knowledge contribution of public versus private actors in the drug discovery and development process. For example, Stevens et al. [10] looked at NMEs that were licensed from public sector research institutions to companies. Kneller [11] also examined the patents behind NMEs but expanded the analysis and differentiated between universities, biotechnology companies and pharmaceutical companies. Patridge et al. [12] used a similar classification but, instead of using patents, they identified the earliest peer-reviewed publication to report the NME in a synthesis or purified formulation, and then extracted all authoring institutions. These studies focused on the direct role of organizations in the drug development process. However, the organizations involved directly

Features • PERSPECTIVE

might be present mainly for marketing reasons, having little to do with the actual scientific knowledge contribution. The development of drugs relies on underlying knowledge and tools from a wide variety of sources that go beyond the direct production of patents. Few studies have explored the indirect knowledge contribution of organizations in the drug development process. An exception is the study by Sampat and Lichtenberg [13], who used prior art references in patents behind NMEs to study the indirect effect that public actors might have. These previous studies have investigated the overall contribution of the public sector but not how its importance might have changed. Hence, a more comprehensive analysis over time of the knowledge sources behind NMEs is missing.

This study investigated the degree of openness in pharma by analyzing the direct and indirect contribution of internal versus external organizations and how the relative importance might have changed. It goes beyond earlier studies that differentiated between public (governmental laboratories and universities) and private actors by also distinguishing between the owner of the drug, industry competitors, academia, hospitals and research institutes. In addition, we studied the role of geography behind knowledge sources since globalization has facilitated interactions with geographically distant actors. Finally, we analyzed the age of the knowledge behind the drugs, to understand how the degree of openness influences the access to recent knowledge sources.

Approach

The time-period 1998–2017 was selected to cover the changes of openness in the pharmaceutical industry in the past decades. First, we collected information about all NMEs and their main patents for each year in the time-period using FDA Administrative Correspondence and the hard-copy versions of the FDA Orange Book. Based on a trend analysis, used to extract the underlying pattern of behavior of the number of NMEs and main patents, the time period was divided into four time series. The median number of NMEs and main patents were calculated for each time series (see Box S1 and Box S2 for more details in the supplementary material online). We selected the four years: 2000, 2005, 2011 and 2015, that were closest to the median number of NMEs and patents for each time series (Figs. S1–S3, see supplementary material online). This resulted in 102 NMEs, where 90 were covered by at least one main patent. In total, 322 main patents were found. The direct knowledge sources were identified as the assignees of main patents behind the NMEs.

In the second step, the cited prior art in the main patents was used as a proxy to identify the indirect impact of knowledge. The citations were identified from the full text of the main patents and included patents and scientific publications. In this way, the citations added by the patent applicants were captured and not the extrapolated citations modified by the examiner that appear in the reference list on the front page of the patent application. For the 322 main patents, we extrapolated 3607 citations. The indirect knowledge sources were identified as the assignees of cited patents or the affiliations of authors in cited scientific publications in the main patents.

Direct and indirect knowledge sources were classified according to whether they were internal or external to the company granted the NME, the type of organization (industry, academia, hospitals and research institutes), the geographic locations of the organizations, inventors and authors. Finally, the age of the prior art for each NME was calculated (additional details related to the dataset, method and limitations can be found in the supplementary material online).

Analysis
Type of knowledge sources
Table 1 shows the number of knowledge sources according to the classifications: direct and indirect, internal and external, and type of organization. The number of unique knowledge sources behind the 90 NMEs increased between 2000 and 2015. The increase was higher for indirect knowledge sources than for direct knowledge sources. There is not much difference between the numbers of internal versus external knowledge sources at the direct level. However, at the indirect level, the number of external knowledge sources increased 3.4 times (from 312 in 2000 to 1064 in 2015), whereas internal knowledge sources only increased 2.3 times (from 10 in 2000 to 23 in 2015).

The data in Table 1 further show that universities are the most important indirect knowledge source behind NMEs. The number of unique universities increased by a factor of four between 2000 and 2015 (from 123 universities in 2000 to 490 in 2015). The number of hospitals and external companies as knowledge sources more than tripled between 2000 and 2015 (from 44 hospitals in 2000 to 152 in 2015 and from 88 external companies in 2000 to 271 in 2015).

Star hubs for new targets
The results above show that the number of indirect knowledge sources has increased. However, to determine the extent to which the organizations granted NMEs rely on the same knowledge sources or new knowledge sources outside the traditional knowledge network, we analyzed the degree of interconnectedness of

TABLE 1

Number of knowledge sources per NME according to direct versus indirect, internal versus external and type of organization

|	2000	2005	2011	2015
NMEs that includes at least one main patent	22	16	20	32
Main patents	62	48	50	162
Direct knowledge sources	28	17	29	49
Indirect knowledge sources	322	535	1088	1087
Direct knowledge sources				
Internal	13	7	18	24
(median number of knowledge sources per NME)	(1)	(1)	(1)	(1)
External	15	10	11	25
(median number of knowledge sources per NME)	(1)	(0)	(0)	(1)
Indirect knowledge sources				
Internal	10	17	19	23
(median number of knowledge sources per NME)	(0)	(1)	(1)	(1)
External	312	518	1069	1064
(median number of knowledge sources per NME)	(7)	(11)	(21.5)	(22)
Type of indirect external knowledge sources				
Universities	123	244	472	490
(median number of knowledge sources per NME)	(3)	(7)	(7.5)	(11.5)
Hospitals	44	72	134	152
(median number of knowledge sources per NME)	(0)	(1.5)	(1.5)	(1)
Institutes	57	80	177	151
(median number of knowledge sources per NME)	(0)	(1)	(2.5)	(3)
Companies	88	122	286	271
(median number of knowledge sources per NME)	(2.5)	(2.5)	(10)	(7.5)
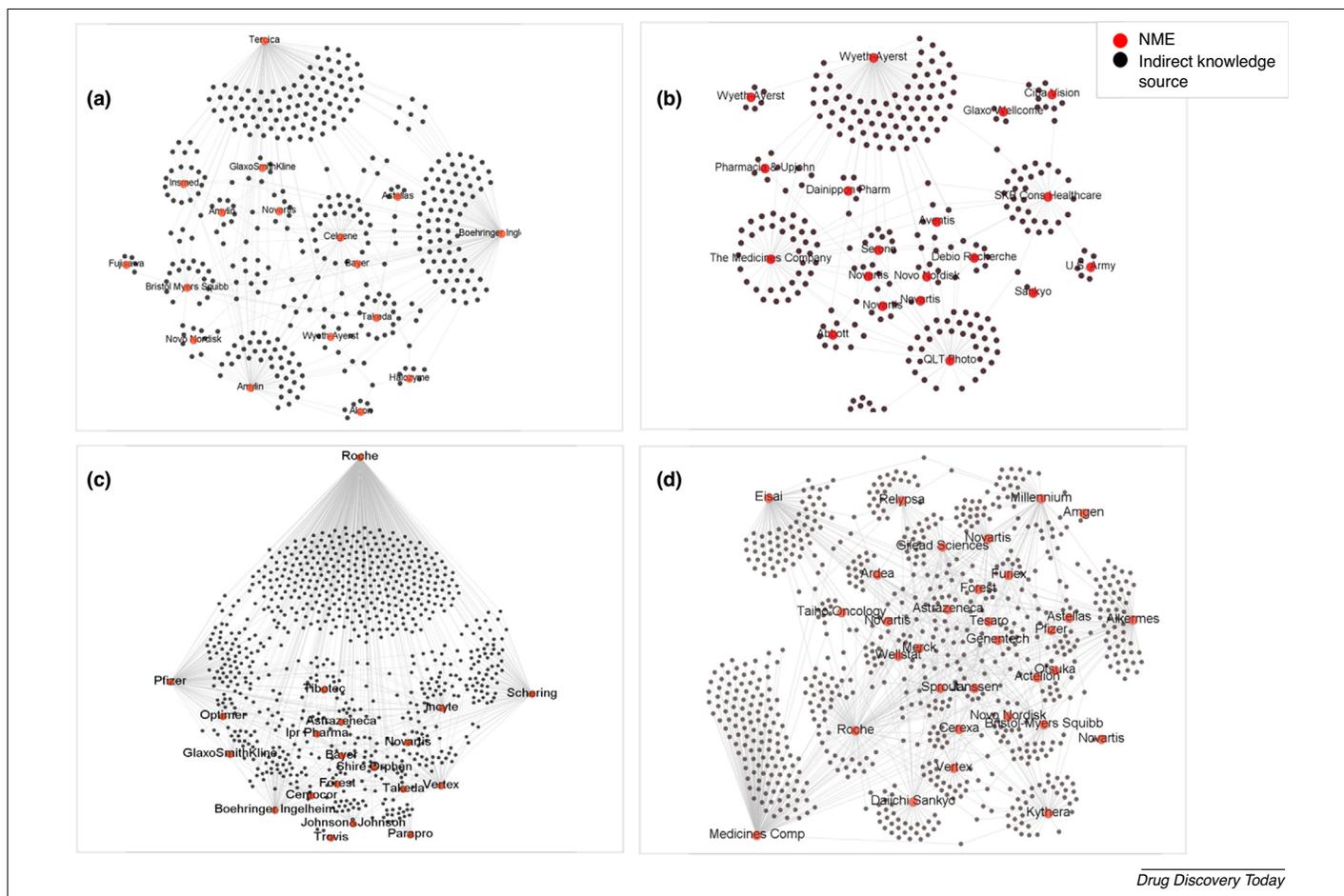
FIGURE 1

Indirect knowledge source networks for the 4 years. **(a)** 2000, **(b)** 2005, **(c)** 2011 and **(d)** 2015. Each black node is an external knowledge source (organization) and the red nodes are the new molecular entities (NMEs). The labels of the red nodes are the organizations granted the NMEs. Each line represents a knowledge source contributing to an NME. Network graphs were produced in Gephi, using the Force Atlas layout (with gravity 10), a variant of the Fruchterman-Reingold algorithm with stronger clustering. The Force Atlas layout is used with the option 'adjust by sizes' to avoid nodes overlapping, depending on the size of each node. Nodes that are highly similar are clustered close together; nodes with low similarity or no similarity are positioned further apart.

indirect knowledge sources between NMEs. Fig. 1 shows the networks of the indirect knowledge sources for each of the 4 years where each red node represents the NMEs and the black nodes represent the indirect knowledge sources. The edges illustrate the knowledge contribution from the indirect knowledge sources to NMEs. The network characteristics of the networks shown in Fig. 1 are presented in Table 2.

The average degree, measuring the structural cohesion, of the knowledge source network has increased with time, suggesting that NME-granted organizations share more knowledge sources (from 1.13 in 2000 to 1.38 in 2015). However, the network density decreased between 2000 and 2015, which can be explained by the higher number of nodes in 2011 and 2015 (Table 2). Network density measures the proportion of all connections

made, out of all that could be made, for the network as a whole.

The number of communities was higher in 2015 than in 2000 (Table 2). A community is defined as an NME with its indirect sources of knowledge. However, some NMEs share the same indirect sources of knowledge. Therefore, the number of

communities is less than the number of NMEs. Because the number of NMEs is different in the 4 years, the number of communities must be normalized according to the number of NMEs. When normalized, the number of communities in 2015 was lower than in any other year, which suggests that the 2015 network is more interconnected

TABLE 2

Network characteristics of the indirect knowledge sources

	2000	2005	2011	2015
NMEs	22	16	20	32
Knowledge sources	267	401	784	754
Nodes	289	417	804	786
Average degree	1.13	1.27	1.36	1.38
Modularity	0.71	0.63	0.56	0.64
Density	0.004	0.003	0.002	0.002
Communities	15	11	12	18
Communities normalized	0.68	0.69	0.60	0.56

Features • PERSPECTIVE

TABLE 3

Number of unique countries of knowledge sources according to direct versus indirect, internal versus external and type of organization

	2000	2005	2011	2015
NMEs	7	6	5	6
Direct knowledge sources	8	6	7	12
Indirect knowledge sources	17	32	36	45
Direct knowledge sources				
Internal	6	5	4	7
External	4	3	4	9
Indirect knowledge sources				
Internal	4	6	8	9
External	17	32	36	45
Europe	11	19	23	26
North America	2	2	2	2
Central and South America	1	4	1	3
Africa	0	3	2	2
Asia and Oceania	3	4	8	12
Type of indirect external knowledge sources				
Universities	13	26	32	38
Hospitals	9	15	16	28
Institutes	10	19	21	27
Companies	10	13	18	18

and less clustered (0.68 in 2000 and 0.56 in 2015). This is illustrated in the networks of Fig. 1 where the 2000 network is visibly more modular with tight community structures and fewer knowledge sources shared between NMEs than in the 2015 network. Modularity is a measure of the extent to which a network is divided into densely connected communities.
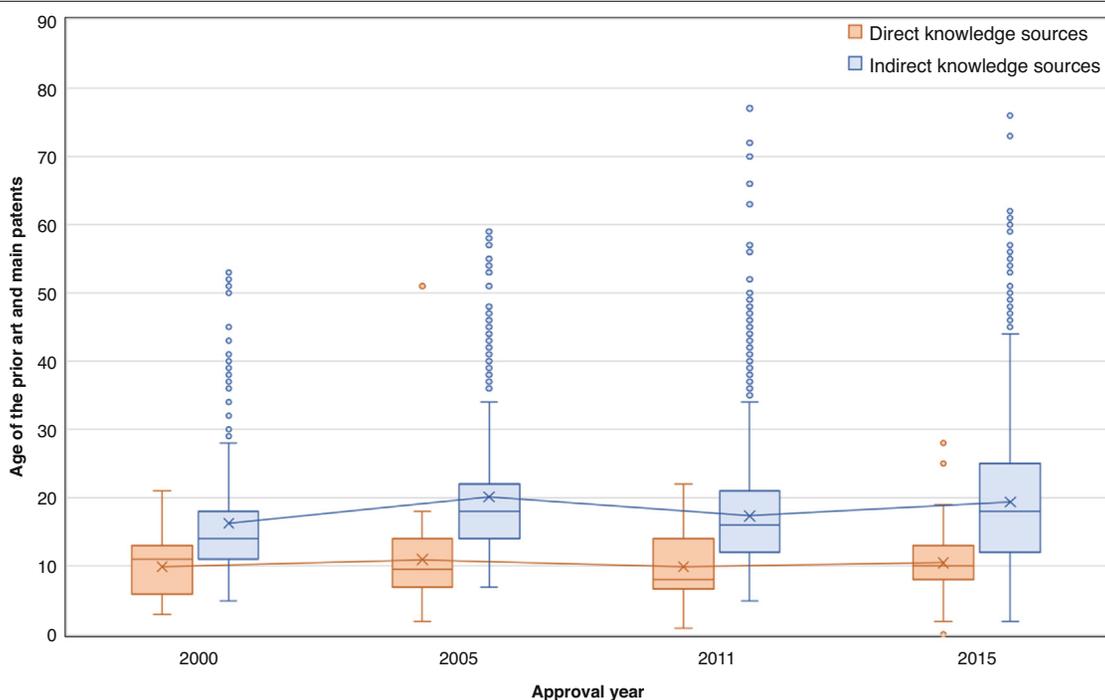
We also analyzed the shared indirect knowledge sources between NMEs. A shared knowledge source is an organization that appears as a knowledge source in more NMEs in the same year. An example is Utah University that appeared in three NMEs in 2000. The number of shared knowledge sources increased until 2011 and then slightly decreased in 2015. The most commonly shared organization type in 2015 was other companies, followed by universities, institutes and hospitals (Fig. S4 and Table S4, see supplementary material online).

Geographic distribution of knowledge sources

We investigated the geographic distribution of the knowledge sources. Table 3 shows that the number of countries increased at all levels of analysis. The increase was especially pronounced in the indirect external knowledge sources. The number of countries in 2015 was almost three-times the number in 2000. In particular, the number of organizations from countries within Europe (11 countries in 2000 to 26 in 2015), Asia and Oceania (3 countries in 2000 to 12 in 2015) has increased. Universities, hospitals and institutes from different countries increased threefold, whereas companies from different countries increased twofold. One explanation behind these results could be that the organizations granted NMEs are increasingly

FIGURE 2

Box plot of the age of knowledge sources. The box plot shows the first quartile, median and third quartile. The x in the box represents the mean. The ends of the 'whiskers' represent the 1.5 interquartile range (IQR) of the lower quartile and the 1.5 IQR of the upper quartile. The dots outside the 'whiskers' are outliers that do not fall within the IQR.

coming from more countries. However, where the companies behind NMEs come from has not changed and the leading drug discovery countries are still the USA, UK, Switzerland, Japan, Germany and other Northern European countries (Table S3, see supplementary material online).

Age of knowledge sources

To determine how the degree of openness influences the access to recent knowledge, we calculated the age of the main patents and citations behind the NMEs. The age is calculated as the difference between the approval year of the NME and the year of the priority date or publication year of the main patents or prior art. The mean and the age dispersion of the direct prior art are similar in the four years. For the indirect prior art, the age dispersion is comparable in 2000, 2005 and 2011 but increased in 2015, meaning that drugs in 2015 relied on indirect prior art from a wider range of years than in 2000 (Fig. 2). In other words, cited prior art in 2015 is partly from articles published or patents filed in the same years as cited prior art in 2000.

Concluding remarks

Increased openness has been presented as a possible solution to the pharmaceutical industry's productivity crisis. The idea is that, to stay competitive and innovative, companies have to search for knowledge externally because it is impossible to find all relevant knowledge internally. Increased openness could facilitate greater access to novel disease targets and enable getting closer to subject matter experts, facilitating the translation of these novel findings into drugs. This study provides specific insights into the openness of the pharmaceutical industry over time by studying where the knowledge behind drugs come from.

We found that organizations behind NMEs are relying more on external knowledge sources. It is mainly the number of indirect knowledge sources that have increased whereas direct knowledge sources have only increased marginally. Universities are the most important indirect knowledge contributor behind NMEs and their relative importance has increased. They increasingly assume much of the greater risk associated with discovering highly innovative drugs and bringing them to the proof-of-concept stage. However, academia plays a passive part in drug creation by only contributing indirectly to basic science knowledge from which biotech and pharmaceutical companies can top up their decreasing investment in R&D.

Although the pharmaceutical industry has become more open to external knowledge sources, we still see a traditional division in drug development roles, with the main patents being owned by private companies.

While studying the drug development knowledge network, we found that indirect knowledge sources are increasingly shared between NMEs. These shared knowledge sources are mainly 'star hubs' located in the leading drug discovery countries. As part of the increasing openness, companies have established small R&D sites inside 'elite institutions' to stay ahead of competitors by accessing pre-publication knowledge from the best clinical and basic science researchers. Companies are providing funding, human resources, infrastructure, compound libraries and technologies, whereas the academic and clinical partners bring in new drug target ideas. The results reflect that open publicly supported academic research is valuable for companies, as are the scientists trained in the course of academic research.

As part of increasing globalization, pharma companies also aim to attract organizations outside the traditional network from smaller specialized universities and emerging countries such as India and China. These trends can be observed in the increasing number of knowledge sources from Europe and Asia. Globalization has resulted in offshoring of clinical trials to nontraditional clinical trial countries far away from headquarters and traditional R&D sites of pharmaceutical companies [14,15]. In addition, there has been an increasing investment in R&D in countries such as China, India and South Korea [16]. It is important to underline that the increase in knowledge sources involved in the drug development process cannot be directly linked to the open innovation platforms that were established in the early 2000s because it is too early to observe these effects. However, the results could indicate a first step toward increasing openness in the pharmaceutical industry.

Finally, even though we see an increase in new knowledge sources, we found that NMEs in 2015 partly rely on knowledge from the same years as NMEs in 2000. This result suggests that the increase in openness has not focused on accessing recent knowledge. Instead, the number of breakthrough discoveries that recent NMEs rely on has been rather limited and new NMEs are follow-on drugs that are rather incremental in nature. This explanation is further strengthened by the rather low number (36%) of first-in-class

drugs in 2015. This number has not changed in subsequent years with 33% (15 out of 46 NMEs) first-in-class drugs in 2017. The pharmaceutical industry is still far from filling the drug pipeline with innovative medicine. In the interest of public health and to ensure public benefits from publicly funded biomedical research, it is therefore important that countries continue to improve their environment for basic science and the subsequent translation of the resulting discoveries, because reduced research funding could lead to a slowdown in the pipeline for treating morbid disease.

References

1 Petrova, E. *et al.* (2014) Innovation in the pharmaceutical industry: the process of drug discovery and development. In *Innovation and Marketing in the Pharmaceutical Industry* (Ding, M., ed.), pp. 19–81, Springer, New York
2 Kessel, M. (2011) The problems with today's pharmaceutical business – an outsider's view. *Nat. Biotechnol.* 29, 27–33
3 Munos, B. (2009) Lessons from 60 years of pharmaceutical innovation. *Nat. Rev. Drug Discov.* 8, 959–968
4 Hunter, J. and Stephens, S. (2010) Is open innovation the way forward for big pharma? *Nat. Rev. Drug Discov.* 9, 87–88
5 Pfizer's Centers for Therapeutic Innovation. Available at: https://www.pfizercti.com/.
6 Novartis Institutes for BioMedical Research. Available at: https://www.novartis.com/our-science/novartis-institutes-biomedical-research.
7 Carroll, G.P. *et al.* (2017) Measuring the effectiveness and impact of an open innovation platform. *Drug Discov. Today* 22, 776–785
9 Innovative Medicine Initiative. Available at: https://www.imi.europa.eu/about-imi/mission-objectives.
10 Stevens, A. *et al.* (2011) The role of public-sector research in the discovery of drugs and vaccines. *N. Engl. J. Med.* 364, 535–541
11 Kneller, R. (2010) The importance of new companies for drug discovery: origins of a decade of new drugs. *Nat. Rev. Drug Discov.* 9, 867–882
12 Patridge, E.V. *et al.* (2015) An analysis of original research contributions toward FDA-approved drugs. *Drug Discov. Today* 20, 1182–1187
13 Sampat, B. and Lichtenberg, F.R. (2011) What are the respective roles of the public and private sectors in pharmaceutical innovation? *Health Aff.* 30, 332–339
14 Glickman, S.W. *et al.* (2009) Ethical and scientific implications of the globalization of clinical research. *N. Engl. J. Med.* 360, 816–823

Features • PERSPECTIVE

15 Haeussler, C. and Rake, B. (2017) The changing geography of clinical research: a critical analysis of its drivers. *Ind. Corp. Chang.* . http://dx.doi.org/10.2139/ssrn.2846672

16 Hughes, B. (2010) Evolving R&D for emerging markets. *Nat. Rev. Drug Discov.* 9, 417–420

Francesca Bignami[1,*]
Pauline Mattsson[1,2]

[1]Department of Learning, Informatics, Management and Ethics, Karolinska Institutet, Stockholm, Sweden

[2]Department of Business Administration, Lund University, Lund, Sweden

*Corresponding author.