From the DEPARTMENT OF LEARNING, INFORMATICS,
MANAGEMENT AND ETHICS (LIME)
Karolinska Institutet, Stockholm, Sweden

EXTERNAL COLLABORATIONS IN MULTINATIONAL PHARMACEUTICAL COMPANIES

Francesca Bignami

Karolinska Institutet

Stockholm 2020

Original cover illustration by Bettina Schwalm
All previously published papers were reproduced with permission from the publisher.
Published by Karolinska Institutet.
Printed by E-Print AB 2020

ISBN 978-91-7831-630-4

External Collaborations in Multinational Pharmaceutical Companies

THESIS FOR DOCTORAL DEGREE (Ph.D.)

By

Francesca Bignami

The public defence of this thesis will take place:

At Karolinska Institutet, Inghensalen, Tomtebodavägen 18a, Solna.
On Friday, the 17th of April 2020 at 10:00.

Principal Supervisor:
Pauline Mattsson (Ph.D.)
Lund University
Department of Business Administration

Co-supervisor(s):
Jarno Hoekman (Ph.D.)
Utrecht University
Department of Sustainable Development

Professor Carl Johan Sundberg
Karolinska Institutet
Department of Learning, Informatics, Management and Ethics (LIME)

Opponent:
Professor Emeritus Finn Valentin
Copenhagen Business School
Department of Strategy and Innovation

Examination Board:
Professor Mats Engwall
KTH Royal Institute of Technology
Department of Industrial Economics-Management

Professor Olof Zaring
University of Gothenburg
Department of Economy and Society

Associate professor Shai Mulinari
Lund University
Department of Sociology

To my family

ABSTRACT

Traditionally, the internal research and development (R&D) departments of multinational companies (MNCs) have served as a main driver of MNCs innovative capacity. Today's high pace of change and competitive landscape have forced MNCs however to look beyond their organizational boundaries and to involve external organizations in their R&D for technological advancement and innovation. In particular, MNCs are using R&D collaborations as a means to create and access new knowledge. Collaborations are particularly relevant in science-based sectors such as the pharmaceutical industry where R&D mainly relies on complex and basic scientific knowledge. In this sector, the sources of expertise are widely dispersed and drug discovery and development requires coordination between different actors. The globalization trend has facilitated collaborations across long distances and companies have adopted a combination of long and short distance collaborations in their innovation process.

The role of geographic proximity in collaboration continues to puzzle researchers. In fact, it may be more complex than previously addressed in the literature, most of which has considered external collaborations in a MNC as a homogenous entity in terms of knowledge. A MNC consists of different R&D units that specialize in different research areas and are active at different stages of the innovation process. Furthermore, a MNC collaborates with a large variety of external organizations and individuals. Thus, R&D collaborations in a MNC can be considered heterogeneous in terms of the knowledge and actors involved. The various types of knowledge and actors well differ in how important the role of geographic proximity is for successful collaborations.

This thesis studies the role of geographic proximity in the R&D collaborations of MNCs when creating, accessing and embedding different types of knowledge. In particular, I examine these aspects by differentiating between (1) the nature of knowledge (basic science vs clinical science, core vs explorative knowledge), and (2) the actors involved on the organization level (university, hospital, research institute and company) as well as on the individual level (star scientists). The different levels of analysis describe different aspects of the R&D collaboration and how these affect the internal knowledge of MNCs.

I used co-publications and patents as a proxy for R&D collaborations and analysed the role of geographic proximity using descriptive, social network and econometric analysis. The results show an increasing openness of pharmaceutical MNCs to collaborate over the past 20 years in terms of the organizations and countries involved in drug discovery and development. While the main patents behind innovative drugs are still mainly owned by companies themselves, external organizations increasingly contribute indirectly to knowledge creation, as visible from an increasing proportion of cited patents and publications from external organizations. This substitutes for biotech and pharmaceutical companies decreasing investments in R&D. Furthermore, considering the nature of knowledge, the results show that collaborations in basic science and core knowledge areas are more positively affected by geographic proximity than collaborations within clinical science and knowledge exploration of the MNCs. I also find that different types of actors embed different natures of knowledge. The knowledge accessed by

MNCs from universities is more positively affected by geographic proximity during the collaborative process, compared to hospitals or companies. However, highly skilled individual scientists who work at MNCs (star scientists), can help to maintain local collaborations.

Based on these findings I conclude that the role of geographic proximity in R&D collaborations of MNCs varies between the types of collaboration and must be more precisely assessed distinguishing between each R&D collaboration between a MNC and another organization or actor. This thesis underlines the crucial role of R&D collaborations for MNCs and emphasizes the importance of geography for the R&D management of MNCs to create and access knowledge effectively in collaborations. From a policy perspective, the importance of different knowledge types in R&D collaborations should be kept in mind when facilitating the development of R&D collaborations, particularly when local actors are trying to attract foreign MNCs.

Keywords: *R&D collaborations; geographic proximity; knowledge; multinational companies; pharmaceutical industry.*

LIST OF SCIENTIFIC PAPERS

I. **Bignami, Francesca &** Mattsson, Pauline. (2019). Potential effects of increased openness in pharma: The original knowledge behind new drugs. Drug Discovery Today. Volume 24, Issue 10, 1957-1962. DOI: https://doi.org/10.1016/j.drudis.2019.06.015

II. **Bignami, Francesca**, Mattsson, Pauline, & Hoekman, Jarno. (2019). The importance of geographical distance to different types of R&D collaboration in the pharmaceutical industry. Industry and Innovation., 1–25. DOI: https://doi.org/10.1080/13662716.2018.1561361

III. **Bignami, Francesca** & Mattsson, Pauline**.** Should they stay or should they go? How the closure of MNC R&D sites affects regional collaborations. (submitted to *Research Policy*)

CONTENTS

LIST OF ABBREVIATIONS

DID	Difference-in-difference model.
FDA	Food and Drug Administration (US regulatory agency).
MNC	Multinational company.
NDA	New drug application.
NME	New molecular entity.
R&D	Research and development.
SNA	Social network analysis.
US	United States.
USPTO	United States Patent and Trademark Office.

LIST OF DEFINITIONS

R&D collaboration	Active participation by two or more partners in a joint R&D or innovation project (Cassiman and Veugelers, 2002).
External organization	An organization that collaborates in R&D with a MNC. This organization can be a university, hospital, research institute, industry, agency or foundation.
Knowledge source	An organization that contributes to the knowledge behind a NME. It can be the assignee of a main or cited patent or the affiliation of an author in a cited scientific publication.
Direct knowledge source	An organization involved directly in the drug development process of a NME, as the assignee of the main patent behind a NME.
Indirect knowledge source	An organization involved indirectly as the assignee of a cited patent or the affiliation of an author in a cited scientific publication of the main patent behind a NME.
Organization granted NME	An organization that sponsors the development or grants the FDA approval of a NME.
Internal knowledge source	An organization that grants the NME or an organization that merges or is acquired by an organization that grants the NME.
External knowledge source	An organization other than the one that grants the NME.
Shared knowledge source	An organization that appears as a knowledge source for multiple NMEs in the same year.
Age of the knowledge source	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 (for patents this is the priority year and for scientific publications the publication year).
Basic science knowledge	Knowledge type refers to the understanding of a phenomenon. This often takes place during the first phases of the drug discovery process.
Clinical science knowledge area	Knowledge type also called applied knowledge refers to the knowledge for a specific end-use. In the drug discovery and development process is related to clinical trials, where

	compounds are tested on humans to assess the safety and efficacy of a drug.
Core knowledge areas	Defined by March (1991) as exploitative knowledge: the use and development of things already known. This knowledge is the main source of innovation for companies.
Exploration knowledge areas	Defined by March as ‘a pursuit of new knowledge for the company’ (1991). It represents the knowledge in a new therapeutic area in which the company decides to diversify its market offer.
Local collaborator	An external organization located less than 100 km from a closed or control R&D site.

1 INTRODUCTION

Knowledge has long been regarded as the most critical asset for an organization and one of the principal sources behind companies' competitive advantage (Teece, 1992; Grant, 1996b). Firms innovate by combining existing and new knowledge (Schumpeter, 1942; Nelson and Winter, 1982; Henderson and Clark, 1990). In large multinational companies (MNCs) that operate in science-based industries, such as pharmaceutical and biotechnology firms, the combination and creation of knowledge mostly occur as part of research and development (R&D) activities (Cohen and Levinthal, 1989; Rosenberg, 1990). However, with the increasing complexity of products and technologies, in-house R&D is not enough to create innovation because knowledge is distributed across different fields and organizations (Nooteboom, 2009). Companies therefore increasingly access and use external knowledge to complement their in-house expertise and increase their competitive advantage (Freeman, 1991; Powell and Grodal, 2005).

R&D collaborations – the main focus of this thesis – are considered to be one of the most important conduits for accessing and using external knowledge. R&D collaborations are defined as "active participation by both partners in a joint R&D or innovation project" (Cassiman and Veugelers, 2002). These partnerships allow for mutually beneficial exchanges in arrangements where both sides make long-term investments (Hagedoorn and Duysters, 2002; Nooteboom, 2004). They are not only about the creation of new knowledge but also about firms accessing existing knowledge from a variety of individuals and external organizations such as universities, research institutes, other firms and competitors (Duysters and Lokshin, 2011).

Previous studies have shown that R&D collaborations have become an increasingly critical element in the innovation strategies of firms (Powell, Koput and Smith-Doerr, 1996; Chesbrough, 2003). The knowledge from R&D collaborations contributes to different performance outcomes including innovation, technological advances (Powell, Koput and Smith-Doerr, 1996; Duysters and Lokshin, 2011), patenting (Gittelman and Kogut, 2003) and organizational growth or failure (Mitsuhashi and Greve, 2009). R&D collaborations are also a constitutive part of open innovation models that companies use to access external knowledge such as recruitment, acquisition and formal as well as informal exchanges with other organizations (Chesbrough, 2003). Considering the crucial importance of R&D collaboration for innovation strategies, a key question for firms is how to create and assimilate knowledge effectively from R&D collaborations (Arora and Gambardella, 1990; Powell, Koput and Smith-Doerr, 1996).

In an attempt to answer this question, many scholars have debated the role of geographic proximity in a firm's ability to access external knowledge (Boschma, 2005; Ponds, van Oort and Frenken, 2007; Balland, Boschma and Frenken, 2015). One stream of literature has argued that geographic proximity and clustering facilitates collaborations because face-to-face and social interactions create the trust that is important for collaborations (Porter, 1998; Zucker,

Darby and Armstrong, 1998; Rosenfeld, 2005). In contrast, several authors have argued that geographic proximity is neither a necessary nor sufficient condition for collaboration and innovative performance. Local collaborations may become more valuable if the company uses also distant interactions that facilitate the variety of ideas and decrease the problem of spatial knowledge lock-in (Bathelt, Malmberg and Maskell, 2004; Owen-Smith and Powell, 2004; Boschma, 2005). Moreover, advances in information and telecommunication technologies and the integration of global markets have contributed to the adoption of a combination of local and distant collaborations by MNCs. Thus, the spatial aspect of knowledge involved in R&D collaborations appears to be more complex than what previously has been indicated.

The starting point of this thesis is therefore that the proper role of the geographic dimension in R&D collaborations must be more precisely assessed in relation to different types of knowledge. This is particularly relevant within the context of MNCs (Buckley and Carter, 2004). MNCs are characterized by different R&D units that specialize in different research areas and that are active at different stages of the innovation process. Thus, the knowledge MNCs maintain is heterogeneous. Furthermore, the R&D collaborations of MNCs tend to involve a variety of external organizations and individuals from which the company aims to access specific types of knowledge. So far the literature looking at geographic proximity has not considered this heterogeneity of knowledge acquired by MNCs through their R&D collaborations.

One way of filling this gap in the context of MNCs is to focus on the diversity of each R&D collaboration and consider the differences in the type of knowledge created, accessed and embedded within these partnerships. Thus, the different research areas that the MNCs are active in are characterized by different natures of knowledge. Also, the variety of actors from which the MNCs access knowledge need to be taken into consideration, since they may embed different natures of knowledge as well as different institutional norms and values with regard to R&D collaboration and knowledge creation. In this context, the pharmaceutical sector – the research setting for this thesis – provides a particular case in which to study collaborations as pharmaceutical MNCs are considered one of the most research-based sectors, carrying out a high proportion of collaborative R&D activities (Kessel, 2011).

In the geography literature, different types of knowledge have only been indirectly taken into account when studying R&D collaborations. Some have argued that geographic proximity is important in knowledge transfer because of the difficulty of transferring tacit knowledge (Polanyi, 1966). However, a more comprehensive study at the R&D collaboration level, looking at the types of knowledge in each collaboration, is required.

To fill these gaps, this thesis investigates the role of geographic proximity in the R&D collaborations of pharmaceutical MNCs when they create, access and embed different types of knowledge. In particular, it studies the role of geographic proximity in relation to R&D collaborations, looking at (1) the nature of knowledge referring to the type of knowledge created from R&D collaboration; (2) the actors involved at the organization level, called external organizations, referring to the type of knowledge that MNCs access from external

organizations in R&D collaboration; and on individual level referring to the type of knowledge that is embedded within individuals in the MNCs in the R&D collaborations. These different levels of analysis describe different aspects of R&D collaborations. My main argument is that the role of geographic proximity varies depending on the type of knowledge created, accessed and embedded by MNCs in R&D collaborations.

2 BACKGROUND

2.1 R&D COLLABORATIONS

For firms in science-based industries, R&D collaboration is a crucial path to innovation. Accordingly, the number of R&D partnerships has risen steadily in firms, while internal R&D has declined (Chesbrough, 2003; Tijssen, 2009; Simpson and Reichman, 2013; Rafols *et al.*, 2014; Crescenzi, Nathan and Rodríguez-Pose, 2016). For example, Crescenzi, Nathan, and Rodríguez-Pose (2016) studied patents in the UK and observed that the 'co-invented patents' across all major technology fields made up around 25% of all patents in 1978 and over 67% in 2007. Rafols *et al.*, (2014) reviewed the publication activities of the R&D laboratories of major European and US pharmaceutical firms during the period 1995–2009. They observed a growing trend of collaborations and the overall reduction of in-house knowledge production. Furthermore, this study has shown that in addition to an increase in collaborations, pharmaceutical companies allow their partners to serve as the first authors of publications, taking the lead in an increasing proportion of projects (Rafols *et al.*, 2014).

The importance of collaborations has generated a lot of attention in the literature, with several motives for the establishment of collaborations in companies listed. The most widely cited reason for a company to work with others is the creation of and access to knowledge from external organizations (Hamel, 1991; Hagedoorn, 1993; Mowery, Oxley and Silverman, 1996). Firms are also heavily engaged in R&D collaborations to share risks, obtain access to new markets and technologies, speed up the process of product development and pool complementary skills (Cantwell, 1995; Pittaway *et al.*, 2004; Dahlander and Gann, 2010; Cantner and Rake, 2014). Furthermore, the literature shows that the innovation advantage of R&D collaborations is a signal to the market, as well as to potential partners, of both the firm's activities and products (Pittaway *et al.*, 2004).

In combination with increased collaborations, globalization has facilitated collaborations at greater geographic distances. Firms and in particular MNCs are increasingly combining local and long-distance collaborations in their R&D activities (Gertler, 2003; Bathelt, Malmberg and Maskell, 2004; Owen-Smith and Powell, 2004; Boschma, 2005). Moreover, companies have also invested in their ability to integrate knowledge dispersed across different locations to translate the advantage of the geographic dispersion of R&D activities into innovation (Singh, 2008). These trends have drawn attention to the geographic proximity of R&D collaborations in the literature and have developed into a debate on its role.

2.2 GEOGRAPHIC PROXIMITY IN R&D COLLABORATIONS

Authors have typically argued that innovation is a highly localized phenomenon and geographic proximity factors into the success of collaborations (Jaffe *et al.*, 1993; Audretsch and Feldman, 1996; Ponds, van Oort and Frenken, 2007). In particular, the importance of geographic proximity in R&D collaborations is most often explained by the fact that short geographic distances facilitate face-to-face and social interactions that form a basis to build

trust between companies and external organizations (Porter, 1998; Boschma, 2005; Rosenfeld, 2005; Ponds, van Oort and Frenken, 2007)

Further, geographic proximity has been considered important for knowledge access and creation. First, it facilitates access to highly skilled people, who can be recruited from local environments (Tijssen, 2009). Second, short distances are crucial for knowledge creation between organizations in clusters. Clusters are defined by Porter (1998) as 'a geographically proximate group of interconnected companies and associated institutions in a particular field'. Following the seminal insight by Marshall (1927), the literature on clustering argues that 'there is something in the air' meaning that the knowledge creation is promoted when firms and individual are located in clusters (Maskell and Malmberg, 1999; Gertler, 2003; Martin and Moodysson, 2011). Authors suggest that clusters allow geographic proximity between organizations, beneficial for the creation and the rapid diffusion of new knowledge (Krugman, 1998). Hence, in these clusters, tacit knowledge can be easily diffuse through social and spontaneous meetings without collaborations, because it is 'in the air'. The importance of this is evidenced by the observation that companies located in knowledge clusters have a higher innovation performance (Jaffe *et al.*, 1993; Acs, Audretsch and Feldman, 1994).

In contrast to these streams of literature, Boschma (2005) does not consider geographic proximity a necessary or sufficient condition for facilitating innovative activities. In particular, he proposes a proximity framework exploring the role of geographic proximity in combination with non-spatial dimensions such as cognitive, organizational, social and institutional proximity. The presence of such proximities between two organizations can compensate for the lack of geographic proximity between them (Boschma, 2005).

Another body of literature declares the 'end to the tyranny of distance' (Cohen and Cairncross, 1997; Stark and Castells, 1997) due to increasing advances in information and communication technologies. As a result of such advances, firms and in particular MNCs increasingly combine local and long-distance R&D collaborations(Gertler, 2003; Bathelt, Malmberg and Maskell, 2004; Owen-Smith and Powell, 2004; Boschma, 2005).

Local collaborations may be more valuable for a company if they are coupled with geographically distant partnerships that facilitate the merger of ideas and reduce the problem of spatial knowledge lock-in (Boschma, 2005). One example is the case of the Boston biotechnology industry where companies combine local and regional interactions with strategic partnerships with interregional and international actors (Owen-Smith and Powell, 2004). Several authors demonstrate that papers' citations, often used as an indicator of scientific quality, rise due to international collaboration (Narin, Stevens and Whitlow, 1991; Katz and Hicks, 1997; Iorio *et al.*, 2012). Malmberg and Maskell (2002) argue that long-distance interactions facilitate a mixture in the knowledge base of companies. Furthermore, distant collaborations allow the firms to learn about market trends and the newest technologies (Bathelt, Malmberg and Maskell, 2004). Therefore, relationships over longer geographic distances may entail higher costs, but this can be repaid by innovation, which might happen

accessing to both new knowledge and collaboration with highly competent researchers in long-distance interactions.

2.3 TYPES OF KNOWLEDGE AND GEOGRAPHIC PROXIMITY

Despite being a well-researched topic, the role of geographic distance in R&D collaborations continues to puzzle researchers. Because firms and, in particular, MNCs, have adopted a combination of local and distant collaborations, it appears that the role of geographic proximity is more complex and cannot be generalized for all collaborations in a firm. One limitation of previous studies on geographic proximity is that they do not study the collaboration behind the firm-level and thus they treat external collaborations as a black box. More specifically, the majority of studies assume that internal collaborations and the actors involved are similar, ignoring the fact that external collaborations are context-dependent and can be characterized internally by types of knowledge (Mattes, 2012). As a consequence, it remains unclear what type of R&D collaborations are organized on a more local level and which collaborations tend to be more globally organized. Thus, to move beyond a firm-level analysis, a more comprehensive study is required.

The study of the role of geographic proximity in each collaboration depending on the type of knowledge involved is particularly relevant for MNCs. These companies combine different types of knowledge from various disciplines within the same institution and their collaborations involve a variety of external organizations and individuals. R&D collaborations can be studied taking the following different aspects into account: (1) the nature of knowledge (e.g. the type of knowledge created from R&D collaboration); (2) the actors involved (a) at the organizational level, called external organizations (e.g. the type of knowledge that MNCs access from external organizations through R&D collaboration); (b) individual actors (e.g. the type of knowledge that is embedded within individuals in the MNCs R&D collaborations). These three aspects address the role of geographic proximity in R&D collaborations in more depth.

In the next section, the different types of knowledge in R&D collaborations and their geographic aspect are explained in detail.

2.3.1 Nature of knowledge

The nature of knowledge refers to the characteristics of the knowledge created from R&D collaboration in a firm. For example, in the pharmaceutical sector, MNCs divide their R&D structure into therapeutic areas that produce different types of knowledge that are often located at separate R&D sites.

The most common distinction of the nature of knowledge can be found in the work of philosopher Michael Polanyi (1966) that differentiates between two categories: tacit and codified knowledge. The first is associated with experience, embedded in practice, skills, emotions and human interactions, and cannot be adequately articulated by verbal means (Polanyi, 1966). It can hardly be formalized and transmitted because it is closely embedded in individuals. Codified knowledge, by contrast, is the ‘knowledge about’, easily understood

because it can be transferred by written documents (David and Foray, 1995). Looking at the knowledge in R&D collaborations, tacit knowledge is mostly spread through face-to-face meetings and personal relations that need partners at geographic proximity (von Hippel, 1994; Feldman and Lichtenberg, 2000). Codified knowledge, on the other hand, can more easily be carried on via long distances (David and Foray, 1995).

The dichotomy between tacit and codified knowledge has often been used indirectly in the literature to study the role of geographic proximity in collaborations. However, this approach has been criticized as difficult to measure. In particular, the implicitness of knowledge is conceptually vague (Cowan, David and Foray, 2000). Numerous scholars, therefore, have presented a number of alternatives. For example, Asheim and Coenen (2005) and Asheim and Gertler (2009) classify knowledge into three categories: analytical, synthetic and symbolic. The first mainly refers to the scientific knowledge employed to understand and explain empirical phenomena. Synthetic knowledge is defined as know-how and is more tacit and problem-driven. Symbolic knowledge is related to cultural meaning such as texts, films, and fashion designs. Although this classification has been used recently in a few studies (Mattes, 2012; Davids and Frenken, 2017), it does not describe different types of knowledge relevant to R&D. In this context, Broström (2010) suggests an alternative classification of knowledge, dividing the R&D process into three phases: learning and impulses (benefits for the first phase of the R&D cycle), short-term projects (benefits for the last phase of the R&D cycle) and long-term projects. Although this classification is focused on R&D, it does not directly consider the nature of knowledge. Furthermore, Broström (2010) adopts the perspective of the user, rather than the knowledge producer, the firm.

In studies that acknowledge the diversity of knowledge in the R&D departments of MNCs, in particular for the science-based sector, it has been observed that certain R&D activities are highly tacit while others are more codified. This distinction has been used to characterize basic versus applied science knowledge, which is the most common way to classify scientific knowledge and R&D activities (Stokes, 1997). Basic science knowledge is the understanding of a scientific phenomenon based on experience accumulated by individuals (Lim, 2004). More precisely, the Organization for Economic Cooperation and Development defines it as ‘an experimental or theoretical work undertaken primarily to acquire new knowledge of the underlying foundation of phenomena and observable fact (OECD, 2011). Basic science knowledge is relatively tacit: it cannot be articulated verbally and is based on experience accumulated by individuals (Lim, 2004). It follows that basic science knowledge needs face-to-face interaction and geographic proximity to be transferred. Applied science knowledge, on the other hand, seeks to produce knowledge for a specific end-use (NSF, 2006). It is more codified than basic science and can be more easily transmitted through written documents that allow it to travel long distances (Castellani and Zanfei, 2006).

Next to the distinction between basic and applied science knowledge areas, this thesis argues that another way to distinguish between tacit and codified knowledge is by identifying areas that are core and areas that are exploratory to the firms. Companies generally have a few core

knowledge areas in which they have a high level of internal R&D knowledge they can exploit. In contrast, firms also have a number of knowledge areas, called exploration knowledge areas, where the level of internal R&D knowledge is lower than in core areas and they aim to develop new knowledge. This distinction is related to the internal knowledge of the firm and in particular to the choice of the firm to access new knowledge outside its core areas and leverage existing internal knowledge with external knowledge (Koza and Lewin, 1998; Rothaermel, 2001; Beckman, Haunschild and Phillips, 2004).

In literature, the dichotomy between core and explorative knowledge areas can be linked to the concepts of exploitation and exploration of knowledge in processes of organizational learning and technological innovation which was first introduced by March (1991). Exploitative as 'the use and development of things already know' and exploration is defined as 'a pursuit of new knowledge' (March, 1991; Levinthal and March, 1993). Looking at the role of geographic proximity, the exploration of knowledge, linked to the exploration knowledge areas in firms, is considered more codified and easy to transfer across long distances compared to the exploitation of knowledge (Miller, Zhao and Calantone, 2006). However, the exploitation of knowledge, defined as the core knowledge areas in firms, is the main source of innovation in firms and its knowledge transfer is closely controlled. Thus, these core knowledge areas consequently benefit from short distances to be transferred (Kale, Singh and Perlmutter, 2000; Cantwell and Mudambi, 2011). Explorative knowledge requires flexibility to assimilate knowledge from a diversity of skills and continuous scanning for new technological opportunities globally dispersed (Hansen, Nohria and Tierney, 1999).

Thus, the use of the dichotomies basic versus applied science knowledge, and core vs explorative knowledge to describe the knowledge created in R&D collaborations gives possibilities to look at the difference in the nature of knowledge in research-based sectors.

2.3.2 Collaborators at the organizational level

R&D collaborations in MNCs can involve a diversity of external organizations such as universities, research institutes and companies. Each of these collaborative actors is characterized by a particular type of knowledge that MNCs entering into R&D collaborations aim to access.

The geography literature has mostly focused on university-industry collaborations. Academia is the most important external organization that collaborates with science-based MNCs. Universities are considered the 'holy grail' of basic scientific knowledge and the base of breakthrough innovations. The knowledge embedded in universities is mostly tacit in nature and transferred verbally by individual researchers (Arora and Gambardella, 1990; Owen-Smith and Powell, 2004). Anselin, Varga and Acs, (2000) find that universities increase the innovation output of firms within a distance of about 50–75 miles. Tijssen (2009) observes that firms tend to collaborate locally with universities to gain access to research infrastructure and recruit highly skilled academics. Though most of the geography literature is limited to university-industry collaboration, academia is only one of the actors involved in R&D

collaborations with MNCs. The literature suggests that the diversity of external organizations contributes to a company's innovative performance (Duysters and Lokshin, 2011; De Leeuw, Lokshin and Duysters, 2014). Thus, it is important to look at the knowledge that MNCs access from hospitals, research institutes and companies. For example, in the case of the pharmaceutical sector, hospitals are the primary source of applied knowledge (Cockburn and Henderson, 1996, 2001). MNCs collaborate with hospitals at long distances to access this type of knowledge (Glickman *et al.*, 2009; Hoekman *et al.*, 2012).

Additionally, the difference between external organizations is often linked to the concept of institutional proximity. Institutional proximity is defined by Boschma (2005) as the distance between two organizations in terms of norms and values. For example, universities are primarily focused in the creation of new knowledge and education. Private companies are driven by commercialization, focusing on knowledge that can be used for competitive advantage and profit (Dasgupta and David, 1994). Thus, collaborations between companies and universities cannot be considered institutionally proximate. Focusing on the geographic dimension of collaboration, Boschma (2005) and the related literature argue that institutional proximity compensates for the lack of geographic proximity and vice-versa (Boschma, 2005; Ponds, van Oort and Frenken, 2007; D'Amore *et al.*, 2013). R&D collaborations between organizations that lack institutional proximity are more likely over short distances. This lack is compensated by the creation of mutual trust through face-to-face and informal interactions (Ponds, van Oort and Frenken, 2007; Balland, Boschma and Frenken, 2015).

In sum, most scholars have assessed the role of geographic proximity in R&D collaboration between firms and external organizations using the concept of institutional proximity. It is important when studying geographic proximity to consider the type of knowledge that a MNC aims to access through the collaboration with that external organization.

2.3.3 Collaborators at the individual level, star scientists

Finally, another way to describe the different types of knowledge in R&D collaborations is to study the heterogeneity of the knowledge embedded in individuals.

During R&D collaborations, knowledge is created through the interactions of individuals. Individuals are the locus of knowledge (Powell, 1998) and knowledge is created by them and through their interactions, not by the organization itself (Nonaka, 1991). Company employees and their embedded knowledge are critical ingredients to gain a competitive advantage for an organization. Firms create value through the selection, development and use of human capital (Grant, 1996a). The distribution of knowledge is heterogeneous and extremely skewed in organizations. Only a small group of individuals contributes to the production of knowledge in an organization (Lotka, 1926; Price, 1963). More specifically, these highly skilled individuals are commonly referred to as 'star scientists' (a term coined by Zucker and Darby, 1996). Star scientists play a primary role in the scientific knowledge creation of a firm and thus are a unique source of innovative competence (Henderson and Cockburn, 1994). Furukawa and Goto (2006) find that stars in science were responsible for a disproportionally large number of publications

in scientific journals and were thus engaging in the creation of new knowledge. In particular, they embed tacit knowledge and combine it with the genius and vision in the most promising areas of research. Beyond their knowledge, they also provide access to large networks that have extraordinary value for access to external knowledge (Henderson and Cockburn, 1994; Zucker and Darby, 1996; Lacetera, Cockburn and Henderson, 2004; Azoulay, Graff Zivin and Wang, 2010). Star scientists act as boundary spanners, bridging organizational and environmental boundaries (Hess and Rothaermel, 2011).

In the context of R&D collaborations, most of the literature has focused on the role of academic star scientists. Authors have shown that university star scientists involved in R&D collaborations with firms influence positively the number and average quality of firm innovations (Zucker and Darby, 1996, 2001; Baba, Shichijo and Sedita, 2009). Furthermore, Azoulay, Zivin and Sampat (2011) discuss the labour mobility of elite academic life scientists. They suggest that the interactions between academic scientists and industry professionals may require more face-to-face meetings than those involving only academics. As mentioned before, star scientists hold tacit knowledge and thus their collaborations require geographic proximity between the organizations involved (Audretsch and Feldman, 1996; Paci and Usai, 2000).

In sum, the literature acknowledges the role of geographic proximity in collaborations involving star scientists. However, very few authors study the role of firms and star scientists in collaborations (for exception Hess and Rothaermel, 2011).

3 RESEARCH SETTING: MULTINATIONAL PHARMACEUTICAL COMPANIES

The research setting for my thesis focuses on pharmaceutical MNCs for both personal and scientific reasons. Prior to starting my Ph.D. research, I worked as a lean manufacturing engineer for a multinational medical technology company. Over this experience, the lean manufacturing methodology taught me how to uncover what adds value to a project by reducing everything else by a better understanding of the innovation behind the R&D process of the industry. In particular, I had the possibility to look closely at how R&D is conducted inside a firm and became interested in how knowledge production is dependent on external organizations from different sectors. After that, I wanted to investigate more deeply the concept of innovation systems and I started working as a research assistant at the Karolinska Institutet. During this project, I researched the university innovation system and conducted several interviews with university professors. I noted how their experiences and importance of R&D collaborations with the industry, and in particular with pharmaceutical companies, have changed in the last 20 years. From these two different personal experiences, I developed the idea behind my Ph.D. project. My experience inside the life sciences sector has inspired me to look at R&D collaborations from a firm perspective.

From a scientific perspective, the pharmaceutical sector is a suitable arena in which to study R&D collaboration because in no other industry are science, research and development so crucial for the innovation process (Munos, 2009; Kessel, 2011). Furthermore, R&D mainly relies on complex and basic science knowledge, the sources of expertise are widely dispersed and development requires coordination between different actors (Calero, Van Leeuwen and Tijssen, 2007; Plotnikova and Rake, 2014).

Thus, the research setting of this thesis is the pharmaceutical industry. In particular, Paper I is focused on pharmaceutical innovation, looking at new molecular entities (NMEs) approved by the regulatory authority of the United States (US), the Food and Drug Administration (FDA). Papers II and III are centred around six pharmaceutical MNCs (Pfizer, AstraZeneca, Eli Lilly, GlaxoSmithKline, Novartis and Takeda) which belong to the top-20 largest companies according to revenue and R&D expenditures in 2012 (they were still part of this list in 2019). These MNCs have adopted different innovation and collaboration strategies, represent different cultures and are geographically distributed.

In the next sections, I explain the knowledge and collaboration dynamics in the pharmaceutical industry.

3.1 KNOWLEDGE IN THE DRUG DISCOVERY AND DEVELOPMENT PROCESS

The innovation process of pharmaceutical companies is commonly called the ‘drug discovery and development process’ and can be divided into two main phases: discovery and development (Figure 1). The distinction between the discovery and development stages reflects the different knowledge characteristics, professional training and expertise required in the

process. The discovery stage establishes the theoretical base of the new drug. In the clinical development stages, all the results from basic science, preclinical pharmacology and safety are applied to see whether the scientific theory can be used into a valuable new drug for patients. Drug discovery is characterized by basic science knowledge, an invariability exploration of the unknown, and the resulted drug may end up with compounds quite different from the expected results. The development stage, in contrast, needs to follow the guidelines from the regulator using standard experimental protocols. It is characterized by applied science knowledge that in this case is called clinical science knowledge (Cockburn and Henderson, 1996, 2001; Cockburn, Henderson and Stern, 1999; Moodysson, Coenen and Asheim, 2008).

In particular, the discovery stage aims to understand the mechanism and the process of the disease. First, the disease area is chosen and the therapeutic need to be addressed is delineated. Next, the biochemical, cellular or pathological mechanism that will be targeted are identified, and if it results in positive outcomes, a molecular drug target is identified. Then, the lead structure is established, followed by design, testing and the optimization of a small number of drug candidates judged suitable for the development process. After the identification of drug candidates, the preclinical test starts, where the mechanism of absorption, distribution, metabolism, excretion, and toxicology of the drug candidates are studied and tested on animals. At the end of the preclinical stage, the first approval is required by the authorities to start human clinical trials. In the case of the US, considered the leading drug-discovering country in the world, the first approval required by the FDA to start human clinical trials is called the Investigational New Drug application. In this first approval step, the results from the preclinical testing are checked by the authorities. They analyse the side effects and other safety features of a drug candidate, the drugs' chemical structure. In addition, the manufacturing process of the drug is considered (Hill, Rang and Vallance, 2012).

If the drug candidate receives approval for human clinical trials, the development stage starts, which is focused on proving the effectiveness and safety of the drug for humans. It includes all the steps from a drug candidate to the approval of the drug for marketing by the appropriate regulatory authorities, for example, the FDA in US. Clinical development involves four phases. Phase I examines the clinical pharmacology, involving a relatively small group of healthy people, to study how the body takes in and discards the drug. Then it looks at the side effects and the outcomes of the drug. Phase II focuses on exploration, proof of concept, confirmatory efficacy and dose range finding. This phase is mostly concentrated on the safety of the drug, looking both at the short-term side effects and the efficacy of the drug against a specific condition. Phase III concentrates on confirmatory, large-scale efficacy and safety. It involves more than a few hundred patients and is the most expensive and time-consuming phase of the drug development process. An NME that successfully passes through all these stages finally enters the approval stage when an application is filed by the drug sponsor to obtain sales and marketing approval from the regulatory authority. In the US, the application is called a New Drug Application (NDA) and is submitted to the FDA (Hill, Rang and Vallance, 2012).

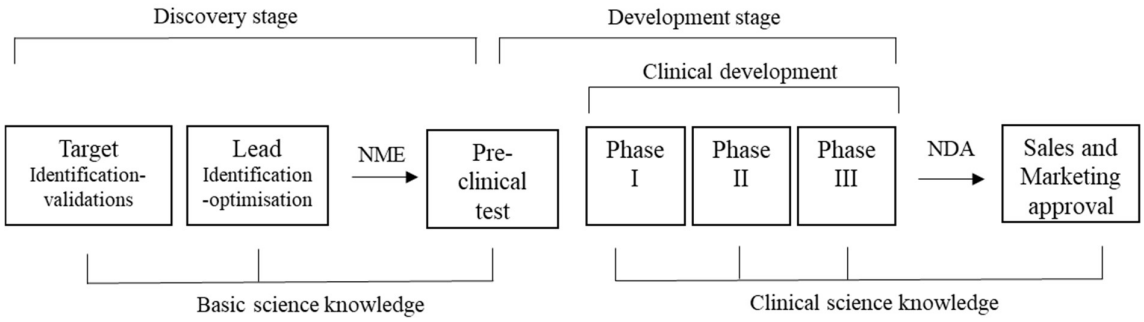

Figure 1| Drug discovery and development process in the US. NME is a New Molecular Entities and NDA is a New Drug Application.

3.2 TRENDS IN THE PHARMACEUTICAL INDUSTRY

In recent decades, the pharmaceutical sector has faced a productivity crisis (Pammolli, Magazzini and Riccaboni, 2011). The number of NMEs approved has decreased and the cost of bringing a drug to the market has grown constantly. The cost to put a drug on the market was more than $2 billion and the failure rate approximates 95% in 2013 (DiMasi, Grabowski and Hansen, 2016). Furthermore, the long time to market is due to the complexity of clinical studies and a relatively stringent regulatory environment (Munos, 2009). Secondly, the patents behind several blockbuster drugs have expired, triggering competition in the generics market. In particular, between 2009–2013, of the top 20 best-selling drugs in the world, 18 went off patent (Munos, 2009; Kessel, 2011).

Looking at the six pharmaceutical MNCs selected for this thesis, Figure 2 shows that R&D expenditures between 2000 and 2010 increased, while revenues started to decrease around 2010. Novartis, for example, in the period 2003–2010, increased its research and development investment by 171%, thus becoming the biggest R&D spender in the entire pharmaceutical industry. However, between 2010 and 2011, the revenues decreased in all six MNCs. These trends can be associated with the productivity crisis in the pharmaceutical industry.

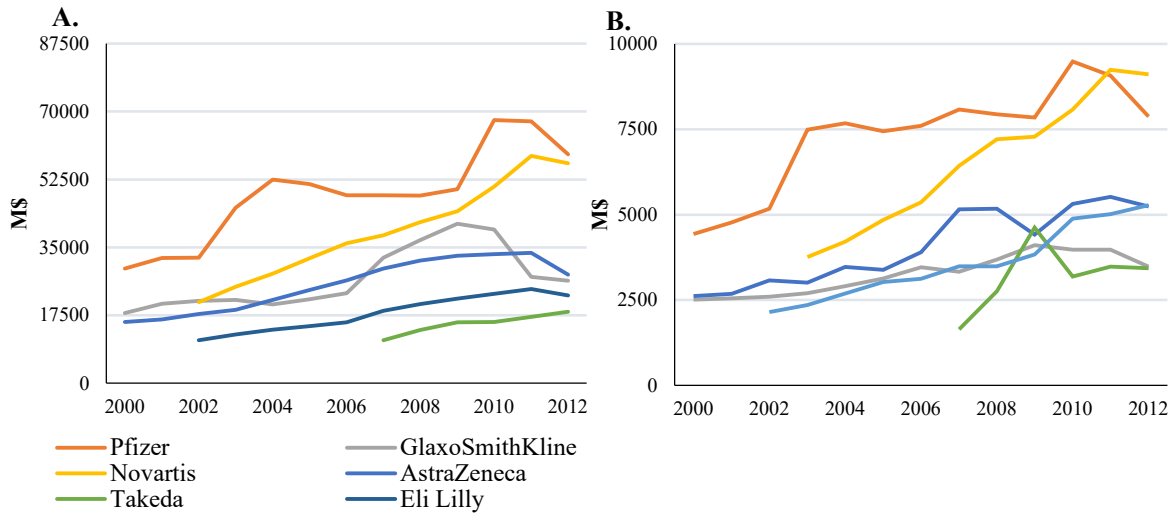

Figure 2| R&D investment between 2000 and 2012**; B. Revenues** between 2000 and 2012 of six pharmaceutical MNCs. Source: Company Annual Reports.

In this challenging and dynamic environment, the drug discovery and development process is more distributed and pharmaceutical MNCs are heavily engaged in R&D collaborations. Rafols et al. (2014) explore the R&D of the major European and US pharmaceutical firms by examining their publication activities during the period 1995–2009. They observe a relative increase in the publications of these firms with external organizations, suggesting a tendency to outsource. Furthermore, different external organizations are involved depending on the phase of the drug discovery process. Figure 3 depicts the involvement of external organizations in various phases of the R&D process. The discovery stage involves mostly collaborations with academia, research institutes and biotech companies, for example. These organizations are crucial to the generation of basic science knowledge in firms (Arora and Gambardella, 1990; Owen-Smith and Powell, 2004). Hospitals and contract research organizations collaborate more frequently with pharmaceutical companies in the development stage. Clinical trials are conducted in hospitals and frequently managed by contract research organizations. Thus, through these collaborations, MNCs have access to clinical science knowledge.

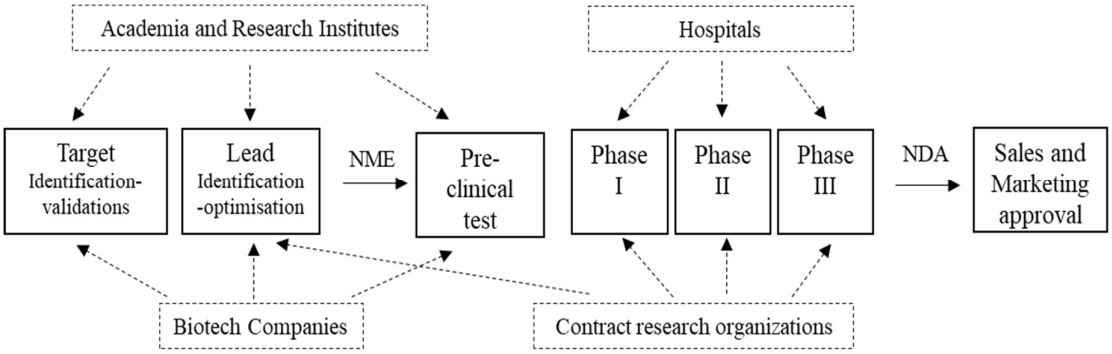

Figure 3| Drug discovery and development process in pharma and critical partners. NME: New Molecular Entities; NDA: New Drug Applications (in the US).

In addition, globalization has facilitated collaborations across long distances and MNCs are using a combination of local and long-range collaborations. For example, in the development phase, pharmaceutical MNCs have started to offshore clinical trials to non-traditional clinical research countries far from their R&D sites or coordinated from distant R&D sites (Glickman *et al.*, 2009). Furthermore, the globalization trend has increased competition as firms compete on a global scale and focus more on core knowledge areas where they have a stronger knowledge base to compete and create a more flexible structure for the exploration knowledge areas.

As a consequence of this trend, a change in the structure of pharmaceutical MNCs can be observed. Traditionally pharmaceutical MNCs were organized with one or more large central R&D laboratories, but in the last ten years, they have shifted to decentralized R&D structures,

with several R&D locations around the world to access local knowledge. A new trend is emerging that expresses the view that 'smaller is better', with the creation of small and autonomous discovery entrepreneurial units active in the interaction with external partners (Douglas *et al.*, 2010). To this end, pharmaceutical companies have adopted different solutions. For example, they create small R&D hubs inside a university to work side by side with academic researchers or organizations that are more science-driven and less bureaucratic, close to the entrepreneurial model of the biotechnology industries. Another example is the small units tasked with identifying and developing external drug discovery opportunities with shared costs, risks and rewards (Gassmann and Von Zedtwitz, 2003; Munos, 2009; Kessel, 2011).

The rising interest in external collaboration and the globalization trend of MNCs is illustrated in the positive growth of the number of publications of the six selected pharmaceutical MNCs between 2000 and 2012 (Figure 4). The overall number of publications increased as well as publications with external organizations, while the number of internal written publications remained constant (Figure 4a). Further, looking at the geographic location of the organizations involved in the publications (Figure 4b), the number of countries involved has increased in the time period studied. This rise is also generated by the change of the structure of MNCs as they have increased the number of their R&D locations spread around the world to access local specialized knowledge.

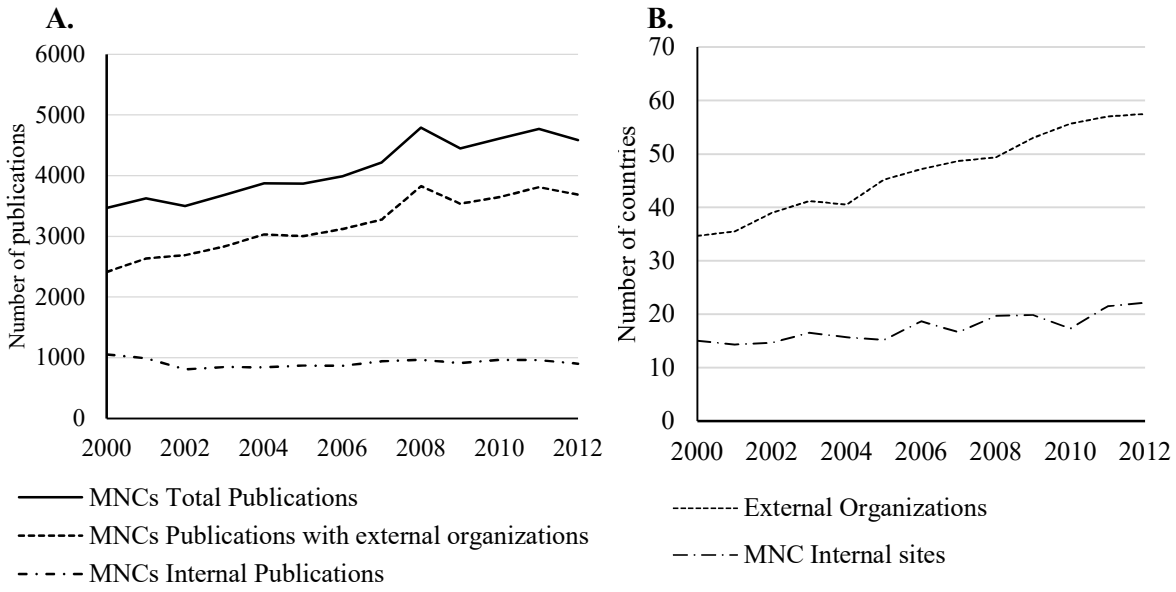

Figure 4| A. Number of MNCs publications, **B. Number of countries** involved in the publications of the six MNCs selected between 2000 and 2012. The line related to external organizations shows the number of countries related to the external organizations that collaborate with the MNCs selected. The line related to MNC internal sites depicts the number of countries of the R&D sites of the MNCs. Source: publications from the Scopus database.

4 AIM AND CONTRIBUTION OF THE THESIS

The aim of this thesis is to investigate the role of geographic proximity in R&D collaborations of MNCs when creating, accessing and embedding different types of knowledge. The research setting is the pharmaceutical industry.

My starting point is that the role of geographic proximity may vary according to the nature of knowledge and the type of actors who are involved in the R&D collaboration. I examine these aspects by:

1. Differentiating between the nature of knowledge (basic vs. clinical science knowledge, and core vs explorative knowledge).
2. Differentiating between actors
    a. on the organizational level (external organizations such as a university, hospital, research institute or company).
    b. on the individual level (star scientists vs other individuals).

As to the nature of knowledge, I refer to the knowledge that is created within the collaboration. For actors on the organization level, I refer to the knowledge that is accessed from external organizations. Concerning actors on the individual level, I refer to the knowledge that is embedded within individuals in the firm that form the loci of collaborations. These three levels address the role of geographic proximity in R&D collaborations in more depth.

The main contribution of this thesis to the existing literature on geographic proximity and R&D collaborations is to show that the role of geographic proximity in collaborations needs to be assessed in relation to the type of knowledge involved. To my knowledge, this is one of the first studies that goes beyond the firm-level and takes as the unit of analysis an R&D collaboration of a firm based upon knowledge type.

The thesis consists of three papers. Their aims and contributions are explained in detail in the next sections and an overview is presented in Table 1.

4.1 AIM AND CONTRIBUTION OF THREE PAPERS

In my first paper entitled “Potential effects of increased openness in pharma: The original knowledge behind new drugs“, I seek to provide a general descriptive analysis of how the degree of openness and the importance of geographic proximity in R&D collaborations are changing the drug discovery and development process. This paper contributes to the overall aim of the thesis, by looking at the changing role of geographic proximity in terms of the location of the organizations behind new drugs. Furthermore, I look at the different types of actors - on the organizational level – that are involved in the creation of new drugs. Previous studies have analysed the organizations in the drug discovery and development process (Kneller, 2010; Stevens *et al.*, 2011). These studies have only focused on the organizations that are granted new drugs and have studied the patents behind the new drugs to identify these organizations. However, the development of drugs is the result of a wide range of organizations

involved beyond the direct production of patents. To have a better overview of the organizations involved, this article used the prior art references in the patents behind new drugs to identify the organizations that contribute to the knowledge underpinning new drugs. In addition, this article looks at the type of organizations that contribute to the development of new drugs change over time.

The second paper of my thesis, entitled 'The importance of geographical distance to different types of R&D collaboration in the pharmaceutical industry', deals with the role of geographic proximity in R&D collaborations of MNCs based upon types of knowledge. This paper contributes to the overall aim of this thesis by studying the nature of knowledge and the actors involved at the organization level. As mentioned earlier, the geography literature to date has considered knowledge to be a homogenous phenomenon, without looking into differences in the access and creation of knowledge and their geographic logic (Mattes, 2012). This paper fills this gap by providing a better understanding of the factors that influence the role of geographic proximity in collaboration based upon types of knowledge. The collaborations are divided by the nature of knowledge into basic science vs applied science, and core vs exploration knowledge. Furthermore, this paper also studies the role of geographic proximity in collaboration distinguishing between different actors such as universities, hospitals, and companies.

Finally, Paper III, entitled 'Should they stay or should they go? How the closure of MNC R&D sites affects regional collaboration', focuses on the effect that the closure of MNCs R&D sites has on access to local knowledge. Collaborations with actors are affected differently depending on the type of organizations involved. Further, the involvement of certain individuals such as star scientists is also shown to influence how well MNC can continue to collaborate with local actors. The existing literature on closures and their regional consequences is limited, and most of the studies have focused on the consequences of closures for employees (Boschma, 2005; Holm and Østergaard, 2015). However, this paper provides a better understanding of how the role of geographic proximity in collaborations is affected by the closure of R&D sites depending upon the type of knowledge involved. To my awareness, this study is one of the first empirical attempts to quantitatively identify the impact a closure may have on local R&D collaborations and thereby access to knowledge. The study provides a more nuanced understanding of the importance of geographic proximity for local collaborations. Second, it gives insights on how to minimize the negative loss of knowledge due to closure for companies and regions.

Paper Title	Objective	Paper- Methodology
Paper 1- Potential effects of increased openness in pharma: The original knowledge behind new drugs.	Determine the change in the degree of openness and the importance of geographic proximity in terms of location of the organizations behind new drugs in pharmaceutical R&D.	Method: Bibliometrics and Patentometrics Sample: NMEs approved by FDA between 2000 and 2015 Data analysis: Descriptive statistics and social network analysis.
Paper 2- The importance of geographical distance to different types of R&D collaboration in the pharmaceutical industry.	Investigate the factors that influence the role of geographic proximity in collaborations, looking at the nature of knowledge (basic science, clinical science, core knowledge and exploration knowledge) and the types of actors at the organizational level (university, hospital and company) involved.	Method: Bibliometrics Sample: Publications of six pharmaceutical MNCs between 2000 and 2012. Data analysis: Regression analysis using gravity models.
Paper 3- Should they stay or should they go? How the closure of MNC R&D sites affects regional collaboration.	Investigate the effect of the closure of MNC R&D sites on local external collaborations, distinguishing between types of actors at the organizational (university, hospital and company) and individual level (star scientists).	Method: Bibliometrics Sample: Publications of five pharmaceutical MNCs between 2000 and 2014. Data analysis: Regression analysis using difference-in-difference models.

Table 1| Details of the three papers.

5 METHODOLOGY

5.1 RESEARCH PROCESS

The research process of the thesis is illustrated in Figure 5. I first conducted exploratory interviews with high-level managers at the MNCs to examine the role of geographic proximity in R&D collaborations in MNCs. These interviews served to identify how companies are working with R&D collaborations and provided a basic understanding of MNCs collaboration strategies.

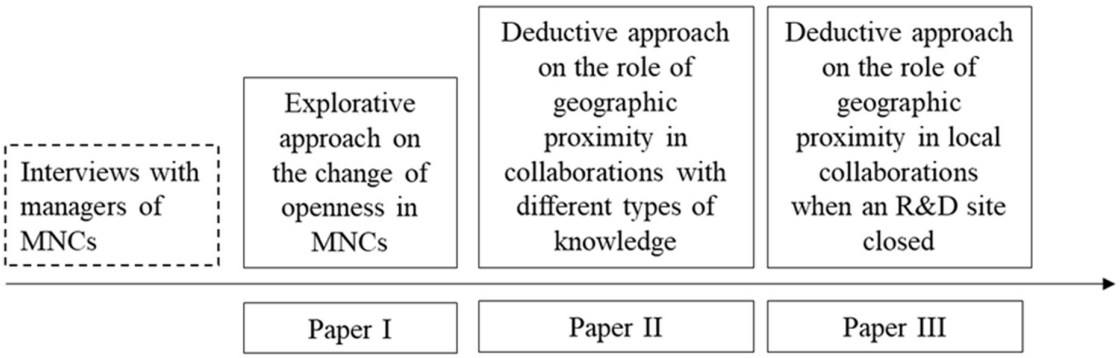

Figure 5| Research process

The interviews triggered me to quantify the changing role of openness and geographic proximity in the drug discovery and development process. To do that, I decided to use an explorative approach. Specifically, geographic proximity is analysed with attention to globalization and the increased collaborations between countries. The results of this first explorative phase in the research process are found in Paper I.

Further into the research process and based on the results of the Paper I, I employed a deductive approach to test the existing theory of the importance of geographic proximity in R&D collaborations. A deductive approach focuses on 'developing a hypothesis (or hypotheses) based on existing theory and then designing a research strategy to test the hypothesis' (Wilson, 2014). I began with the theory on the importance of geographic proximity in R&D collaborations. Then new hypotheses were formulated to test this theory and I considered how the role of geographic proximity may differ depending on the knowledge involved in collaborations. In particular, I distinguished between the nature of the knowledge and the types of actors involved in collaborations. Thus, the hypotheses were based on the relationship between geographic distances, nature of the knowledge and type of actors involved in R&D collaborations. They were then tested using multiple regression analysis. The results of this deductive approach are found in Paper II.

Based on the initial interviews, it emerged that MNCs are characterized by continuous changes in their R&D structure. These changes may include the opening, closure, and relocation of R&D sites due to for example mergers and acquisitions, access to market and local

competences, political pressure, access to the specific target population and governmental incentives. The closures of R&D sites by MNCs are not unusual. Furthermore, the interviewers underline that when a MNC decides to close an R&D site, this may have implications for local collaborations and star scientists. Based on this information, I decided to test the role of geographic proximity on R&D collaborations when a local R&D site is closing using a deductive approach similar to Paper II. Looking at the knowledge type involved in R&D collaborations, I distinguished according to the actors at organizational as well as individual level. The results of this study are found in Paper III.

I use a quantitative approach in all papers. Quantitative methods are used to test the theoretical hypotheses and measure the relationship between variables to explain a phenomenon (Denzin and Lincoln, 1998). One reason for using quantitative methods in this thesis is that this approach provides me with the opportunity to obtain an overview of each collaboration of a MNC and to compare this across MNCs. In addition, the choice to use quantitative methods is related to the research setting of this thesis: MNCs. MNCs are large companies that collaborate intensively. Thus, quantitative methods can provide broad and general conclusions of the use of R&D collaborations in MNCs. The quantitative methods used in this thesis are bibliometric and patentometric methods. In the next section, the reasons for this choice are explained.

5.2 BIBLIOMETRIC AND PATENTOMETRIC AS METHODS

Different methods have been used in the literature to analyse R&D collaborations and knowledge exchange. The most common methods to measure the outcome of collaborations in science, and in particular R&D collaborations, are bibliometric and patentometric methods.

The bibliometric method applies to scientific publications. It is based on the concept that the core of scientific research is the production of 'knowledge' and scientific literature is the constituent presentation of that knowledge creation process, the output of R&D activity. Publications are one way for researchers to make new knowledge available to the scientific community and to claim it as intellectual property. Price (1963), in his seminal work 'Little Science-Big Science', underlines that 'a scholarly publication is not a piece of information but an expression of the state of a scholar or a group of scholars at a particular time'. Furthermore, according to Cockburn and Henderson (Cockburn and Henderson, 1996) co-authorship of papers is evidence of a significant, sustained and productive interaction between researchers in different organizations. In particular, publications involving multiple organizations serve as a proxy for collaborations between individuals and organizations. In this case, the article includes more than one author and location, suggesting that the authors come from different organizations.

Similar to publications, the patentometric method identifies patents taken by multiple organizations as a proxy for collaborations. Patents have been the most frequently used indicator of knowledge transfer and technological change (Narin and Noma, 1987; Agrawal and Henderson, 2002). Patents have been used to evaluate the innovation performance of organizations and individual researcher. Furthermore, patent analysis has become an important

tool to explore technological collaboration and measure the output of the innovation system (McAleer and Slottje, 2005; Gao, Guan and Rousseau, 2011). Since patent data is publicly available and regularly updated, it provides a specific and detailed indicator for analysing how public research is developed by industries (Cohen, Nelson and Walsh, 2002).

A number of authors have identified reasons to support publications or patents as representations of a robust method to measure R&D collaborations. First, publications and patents are public information, accessible to everyone, and the results of their analyses can be reproduced using the same method. Second, they are scalable. Analyses can be performed at individual, organizational, national or international levels (Katz and Hicks, 1997). Third, among other indicators used to map collaborations, analyses of publications or patents are methods that can be uniformly applied to all organizations. They are also commonly used in science-based sectors such as the pharmaceutical industry (Tijssen, 2009, 2012).

However, there are several significant limitations in using patents and publications as a proxy for collaborations. First, not all the collaborations between organizations result in a patent or a co-publication. In the context of this thesis, the complexity of the knowledge generation process in the pharmaceutical industry requires the involvement of many different actors across sectors and disciplines that might not publish or patent. Therefore, those involved, for example, in technology development and clinical trials are overlooked using the above methods. Second, collaboration is the result of individual interactions and therefore equates to a fuzzy phenomenon, about which publications or patents only reveal aspects of the actual collaboration. Furthermore, collaborations can take place in various forms and for a number of reasons, for example giving general advice, providing materials, sharing ideas and data, and the production of a patent or publication. A bibliometric or patentometric study is not able to reveal these characteristics. Patents and publications may also overemphasize the contribution of certain actors. Not all the names on a publication are responsible for the work. Indeed, some publications list authors for purely social reasons. This is the case of academia in publications or industry in patents. On the contrary, university researchers might, for example, publish results of joint efforts without listing in the publication the involvement of industry organizations. Thus, these collaborations cannot be identified using a bibliometric method. Finally, the date when a paper or patent is published does not reflect the time when the research activities were performed. Further, citations both in patents and publications may be added because of strategic reasons even when they do not reflect actual knowledge sources (Katz and Martin, 1997; Moed, Glänzel and Schmoch, 2004).

The difference between publications and patents as a proxy of collaborations is the difference between science and technology. Patents are related to a technical invention and publication to science. Furthermore, universities are focused on publishing and they want to publish their results quickly to increase their (citation) impact. However, the industry is oriented to commercialize knowledge. Firms seek to keep their knowledge internally, rather than to publish in scientific papers (Dasgupta and David, 1994). As a result, they may want to limit the disclosure of funding until they apply for a patent. However, with the increasing importance of

R&D collaborations, industries are increasingly publishing in particular in science-based industries like the pharmaceutical sector (Rafols *et al.*, 2014). Furthermore, comparing publications and patents, the knowledge in R&D collaborations is covered more widely by scientific publications than patents because knowledge must be novel to qualify for a patent.

In sum, publications and patents do not capture all the collaborations and their aspects. However, I assumed that significant R&D collaborations lead to these outputs in most cases in the pharmaceutical MNCs.

5.3 SAMPLE

The research setting of this thesis is the pharmaceutical sector. I have selected two different samples in order to first exploratively investigate the openness of pharmaceutical industry and then to focus on the role of the geographic proximity in R&D collaborations. First, I have selected 102 NMEs between 2000 and 2015, focusing on four time periods (2000, 2005, 2011 and 2015). Second, I chose six pharmaceutical MNCs (Pfizer, AstraZeneca, Eli Lilly, GlaxoSmithKline, Novartis and Takeda) that belong to the top-20 largest companies according to revenue and R&D expenditure in 2012. They continue to top the list for 2019. These companies have been selected because they have adopted strategic, cultural and geographic differences. Based on interviews with R&D managers of these MNCs, I also checked that these firms adopted different collaboration strategies. Selecting companies with non-identical collaboration strategies, cultures and geographic distribution allow me to generalize my results.

This thesis uses publications and patents in both samples as the main proxy of R&D collaborations. There are several databases that index scientific publications, such as for example, Scopus, Web of Science and Google Scholar. The dataset for this project has been built on publications extracted from Scopus, a database of scientific documents edited by Elsevier and launched in 2004. Scopus contains 33 million records of which 16 million include references going back to 1996; 17 million pre-1996 records go back as far as 1841. Scopus lists scientific articles from a wide collection of publishers and includes information such as authors´ name, title, publication date, abstract, etc. I chose Scopus rather than Web of Science as it links author names with their affiliations in the period analysed in this thesis (this was introduced by Web of Science around 2008–2009). This is important in the disambiguation of individuals across the publications and, for this thesis, in the study of the actors at the individual level. Moreover, Scopus includes broader and robust information than Google Scholar (Yang and Meho, 2006). For these reasons, I selected Scopus as the publication database of choice in this thesis.

For the patent, this thesis used United States Patent and Trademark Office (USPTO) data because the US market is the leading market for pharmaceutical drugs and Paper I is focused on NMEs granted by Food and Drug Administration (FDA) that is the agency that approved new drugs in the United States.

In particular, Paper I analyses the main patents of the 102 NMEs approved in 2000, 2005, 2011 and 2015 by the FDA. The patents behind these drugs were extracted from the New Drug

Application (NDA) documents and Orange books of the Food and Drug Administration (FDA). The NDA is the application by the drug sponsor to formally propose the approval of a new drug by the FDA in the US and it is collected by the same agency in the public database DRUG@FDA. The drug sponsor applying to the NDA must include all active patents covering the NMEs in the NDA document to ensure that no other company holds the rights to that drug. However, some of the online NDA documents in the DRUG@FDA did not include patent information. In these cases, the Orange book hard copy version was used to search for patents. Orange Book is the common name of the yearly publication called "Approved Drug Products with Therapeutic Equivalence Evaluations". From the 102 NMES, 322 patents were identified. From these patents we collected the citations as publications and patents. The total sample consists of 322 main patents and 3,607 citations (publications and patents).

Paper II uses publications between 2000 and 2012 derived from the selected six pharmaceutical MNCs in collaboration with at least one external organization. The database includes 24,561 publications and 17,631 external organizations.

Finally, Paper III reviews the publications of the selected pharmaceutical MNCs and identifies 605 local external organizations affected by the closure of five R&D sites and 860 local external organizations of five control sites between 2000 and 2014.

5.4 DATA COLLECTIONS

I built up a database including both the patents and publications.

Publications and patents contain information describing many characteristics of collaborations. In particular, a bibliometric record includes author names and affiliation(s) (including institution, country and city), year of publication, citations and keywords. The inclusion of author addresses enables the identification of interactions both inside (different R&D locations and departments) and outside of a company. Similar to a publication, a patent document includes a variety of information, such as the name of the inventor(s), address(es) and the name of the organization receiving the patent. In addition, citations are included in the patents.

In order to investigate the change of openness in R&D collaborations, the main patents behind new drugs were analysed looking at the name of the organizations receiving the patent. However, these organizations might be present mainly for marketing reasons, having little to do with the actual scientific knowledge contribution. Thus, the citations in the main patents were also analysed to have a better understanding of the knowledge behind new drugs. A citation can be a patent or a publication.

In order to analyse the geographic proximity between MNC and collaborators, the address of the authors and inventors of all publications and patents of the MNCs were cleaned. Next, the cities of the collaborating organizations and MNC R&D sites were used to determine the geographical coordinates (longitude, latitude) of each location using Google Maps. In addition, looking at the effect on collaborations after the closure of an R&D site, I defined local collaboration as a collaboration between MNCs and external local organizations within 100

km from a closed or control R&D site. I identified those local collaborations with the publications of MNCs that were co-published with a local organization.

The information contained in the publications and patents also allowed for them to be differentiated by the nature of knowledge and the type of actors involved in R&D collaboration.

5.4.1 Categorization of the nature of knowledge

To categorize the publications according to the nature of knowledge (basic science vs clinical science knowledge; core vs exploration knowledge), the keywords of the publication were identified. Keywords identify the major themes of a publication and typically consist of Medical Subject Heading (MeSH) terms, a comprehensive controlled vocabulary that is used by the National Library of Medicines to classify publications and books on life sciences and medicine. In other words, they are labels assigned to each article in Medline in order to describe what the article is about. The use of MeSH terms captures the heterogeneity among papers within a journal; it is a more specific approach compared to the traditional CHI index, based on the classification of the journal in which an article is published (Hamilton, 2003). Keywords are used in this thesis to classify the publications into basic vs clinical science knowledge and core vs exploration knowledge.

In addition, it is important to underline that citations from the main patents can be considered basic science knowledge because the organizations filed the patents for the new drugs when the drug is at the discovery phase

5.4.2 Categorization of the types of actors at the organizational level

Based on the address(es) of authors and inventors, an actor was classified according to type at the organizational level, i.e. university, hospital, research institute, company and other (governmental agency, foundations, etc.).

5.4.3 Categorization of the types of actors at the individual level

Finally, to categorize the publications on the individual level, I have adopted the definition of star scientists by Rothaermel and Hess (2007). A star scientist is defined as a MNC researcher who publishes at a rate of three standard deviations above the average MNC scientists over the previous ten years. Star scientists were identified in the database of the total publications for each MNC looking at the names of authors in the publications. In this way, a list of star scientists was created and matched with the names of authors of the local publications.

5.5 ANALYSIS

For the analysis, I used descriptive statistics, social network analysis (SNA) and multiple regression models.

5.5.1 Social network analysis

In Paper I, I used social network analysis (SNA) to quantify the changing role of openness and geographic proximity in the drug discovery and development process. Many scholars use SNA to map and measure the relationships and flows between people, groups and organizations. In Paper I, SNA is applied to investigate and visualize the connection between the actors involved in patents and publications. The descriptive statistics were also used in Paper II and III for the base of the regression models.

5.5.2 Econometric approach

In Papers II and III, I used multiple regression models to test the importance of geographic proximity in R&D collaborations for different types of knowledge. Multiple regression models allow for the prediction of the number of collaborations between two organizations based on the geographic distances between collaborators and the type of knowledge involved.

In particular, in Paper II, I used the gravity model as a multiple regression model. In the geography literature, the gravity model has been used to approximate spatial interactions between locations (Ponds, van Oort and Frenken, 2007; Hoekman, Frenken and van Oort, 2009). This has enabled me to control for other determinants – types of knowledge – that affect MNC collaborations in addition to geographic distance. The model is based on the 'gravity equation', an analogy to the Newtonian theory of gravity. In Paper II, the gravity model has been used to estimate the effect of geographical distance on the number of collaborations of a MNC. Generally, a gravity model assumes that the number of R&D collaborations between a MNC R&D site and an external organization will be directly proportional to the product of their economic mass (measured by the number of total publications of the MNC R&D site and the external organization) and inversely proportional to the distance between them. Thus, the dependent variable of the gravity model is the number of publications as a proxy of R&D collaborations of the MNCs. The geographic distance and the nature of the knowledge of each publication are the independent variables. The types of actors at the organizational level were included as control variables.

In Paper III, I tested hypotheses regarding the effects of closures of MNC R&D sites on the number of local collaborations and the involvement of star scientists. Building a regression using a difference-in-difference model (DID), I estimated the difference in the number of collaborations between a MNC and a local organization affected by the closure, assessed against a control group. Following the work of Ashenfelter and Card (1985), the DID model has become a popular tool to estimate the effect of a policy change or treatment in the medical field (Meyer, 1995; Heckman, Lalonde and Smith, 1999; Blundell and Costa Dias, 2009; Lechner, 2010). This model is used to estimate outcomes before and after the policy or treatment intervention. In Paper III, the treatment intervention is the closure of the R&D site. The dependent variable is the number of collaborations. The types of actors at the organizational and individual levels are analysed using two independent variables.

6 RESULTS

In the next section, I first present the explorative results on the changing role of openness in the drug discovery and development process (section 6.1). Then, the results on the changing role of geographic proximity in relation to the types of knowledge in R&D collaborations are outlined (section 6.2).

6.1 INCREASING OPENNESS OF PHARMACEUTICAL MNCS

In accordance with earlier studies, I observe that the number of organizations involved in the main patents of newly approved drugs increased between 2000 and 2015. However, this increased number is more pronounced for organizations that are involved in the citations of the main patents. The main patents of newly approved drugs are still granted to few organizations. Instead, the citations in the main patents involved an increasing number of organizations between 2000 and 2015 that indirectly contributed to drug discovery (Table 2). In addition, the organizations involved in the citations are more external organizations than the companies granted new drugs.

Using SNA, I have analysed the degree of interconnectedness of the organizations involved in cited patents to determine the extent to which those granted new drugs share the same or new organizations in the citations. Figure 6 illustrates the networks of the organizations involved in the citations in 2000, 2005, 2011, 2015. Each red node represents the NMEs and the black nodes represent the organizations involved in the citations. The edges illustrate the knowledge contribution from the organization in the citations to NMEs. It can be observed that the 2000 network has fewer organizations (black nodes) shared between NMEs than in 2015. Thus, organizations applying for drugs are increasingly relying on knowledge from the same organizations. This is, for example, the case of California University, which contributed to 16 new drugs in 2015.

	2000	**2005**	**2011**	**2015**
NMEs that includes at least one main patent	22	16	20	32
Main Patents	62	48	50	162
Organizations in the main patents	28	17	29	49
Organizations in the citations	322	535	1088	1087
Internal	10	17	19	23
External	312	518	1069	1064

Table 2| Number of the organizations behind the new drug approved.

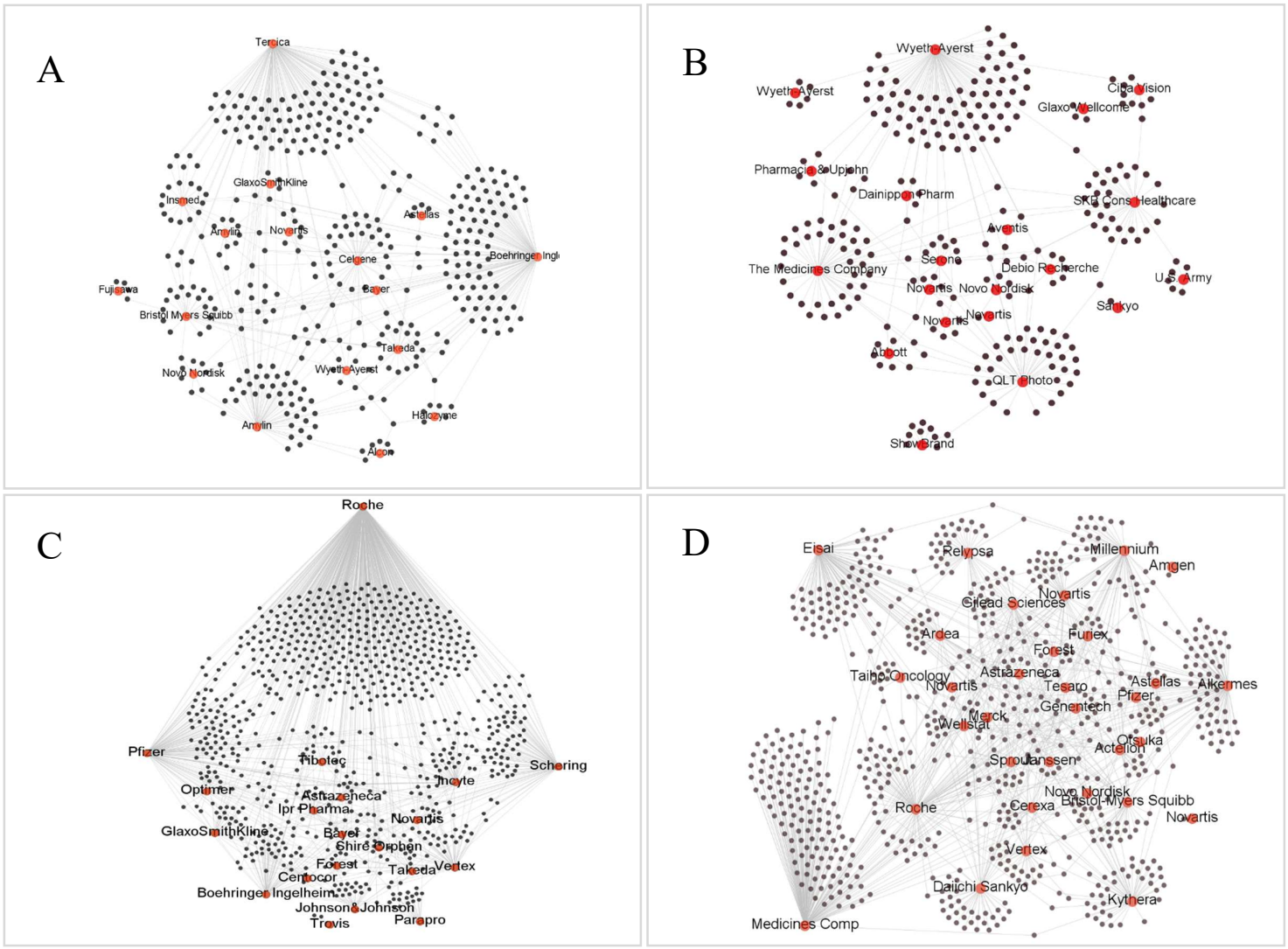

Figure 6| External organization networks over the sample years: (**A**) 2000, (**B**) 2005, (**C**) 2011 and (**D**) 2015. Each black node identifies an external organization; red nodes signify NMEs. The labels of the red nodes are the organizations granted new drugs. Each line represents an external organization contributing to a new drug.

6.2 GLOBALIZATION AND THE DIVERSE GEOGRAPHIC DISTANCES IN COLLABORATIONS

The effect of globalization can be observed with the increased number of countries in the organizations behind a new drug, in particular the external organizations included in citations. The number of countries of the organizations involved in the citations in 2015 is almost three times the number in 2000 (Table 3).

Considering that globalization and informational communication technologies have facilitated collaborations at long distances, I first tested with gravity and DID models whether geographic proximity still matters in collaborations. The results show that geographic distances negatively affect R&D collaborations. This is also the case for local collaborations affected by the closure of an R&D site.

Next, to further investigate the role of geographic proximity in R&D collaboration in MNCs, I analysed the relationship between geographic distance and the types of knowledge (nature of the knowledge, type of actors at the organizational and individual level) in collaborations.

	2000	2005	2011	2015
NMEs	7	6	5	6
Organizations in the main patents	8	6	7	12
Organizations in the citations	17	32	36	45
External Organizations in the citations				
Universities	13	26	32	38
Hospitals	9	15	16	28
Research Institutes	10	19	21	27
Companies	10	13	18	18

Table 3| Number of unique countries of organizations involved in the main patents and in their citations (Paper I).

6.2.1 Geographic proximity matters for collaboration in basic science and core knowledge

The nature of the knowledge involved in the R&D collaborations is investigated to distinguish between basic science, clinical science, core and exploration knowledge areas. The results show that clinical science and core knowledge collaborations are the highest in number between the R&D collaborations of the selected MNCs (Table 4).

The relationship between the mean number of collaborations and the geographical distance between MNC R&D sites and external organizations looking at the nature of the knowledge (basic science, clinical science, core and explorative) is presented in Figure 7. The geographical distances of R&D collaborations involving a MNC range from 0 km to almost 20,000 km. The distribution of the distance is twisted as 25% of collaborations take place within 400 km. In particular, Figure 7(a) shows that the mean number of collaborations in the range of 100 km is higher for basic science knowledge areas than for clinical science knowledge areas. Beyond

100 km, the mean number of collaborations in clinical science knowledge area is always higher than that of basic science areas. Furthermore, Figure 7(b), adding the distinction core and exploration to the basic and clinical science knowledge areas, shows that the mean number of R&D collaborations is higher in clinical science-core knowledge than in clinical science-exploration knowledge at all geographic distances. However, the opposite result is observed for the R&D collaborations that create basic science knowledge.

I subsequently tested the effect of the relationships between geographic distances and the nature of knowledge on the number of R&D collaborations with gravity models. Geographic proximity displayed greater importance for R&D collaborations in basic science compared to clinical science knowledge areas. Finally, in both basic science and clinical science knowledge areas, geographical distance affected more negatively the number of collaborations in core knowledge areas than in exploration knowledge areas.

	N	Mean	SD	Min	Max
Total number of collaborations	41 650	2.311	4.391	1	274
Basic science	5 721	1.864	2.411	1	72
Clinical science	37 759	2.273	3.945	1	202
Basic science-core	3 415	1.545	1.667	1	40
Basic science-exploration	3 497	1.571	1.711	1	40
Clinical science-core	26 155	1.943	3.101	1	162
Clinical science-exploration	21 280	1.667	2.157	1	106

Table 4| Descriptive statistics of the nature of knowledge in R&D collaborations of the six selected MNCs (Paper II).

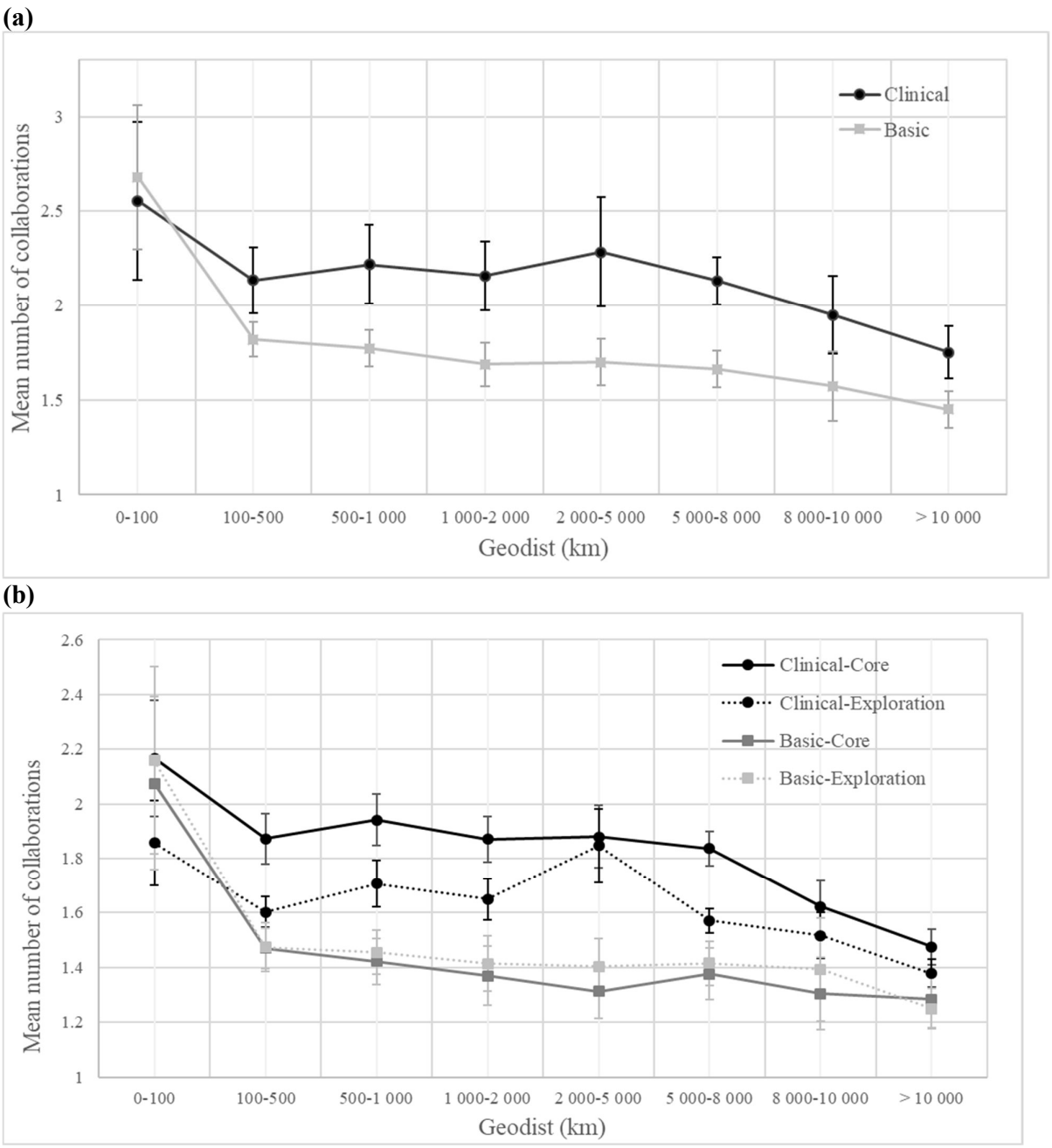

Fig. 7| Mean number of collaborations for geographical distance intervals between MNC R&D sites and external organizations with confidence intervals according to the nature of the knowledge (Paper II).

6.2.2 Geographic proximity matters for collaborations with universities

Next, the type of actors at the organization level involved in the R&D collaborations are studied, separating the external organization into universities, hospitals, research institutes and companies. Universities are the most important external organizations involved indirectly in the discovery and development of a new drug and their importance increased between 2000 and 2015 (Table 5). The results in Table 5 show that the role of universities increases by a factor of four, while hospitals and companies grow by a factor of three between 2000 and 2015. It is important to underline that the citations from the main patents can be considered basic science knowledge because the organizations filed the patents for the new drugs when the drug is at the discovery phase.

Furthermore, looking at the increased number of countries in organizations indirectly involved in the creation of new drugs, this increase is less pronounced for companies compared to other types of actors, such as universities, hospitals and research institutes (Table 3). One explanation could be that companies behind new drugs are still based in the leading countries of the pharmaceutical industry such as the US, UK, Switzerland, Japan and Germany. In contrast, universities, hospitals and research institutes are instead increasingly spread around the globe, in line with the globalization trend.

In the R&D collaborations, we can observe that hospitals are the most important collaborator in terms of the number of collaborations with MNCs and in particular in clinical science knowledge areas. Instead, universities- MNC collaborations are the most frequently in basic science knowledge.

Based on these results, I tested the relation between the types of actors at the organization level and the geographic distances in collaborations with the gravity models. I found that the role of geographic proximity differs between collaborations that involve universities, hospitals and companies. More specifically, if an R&D collaborator is a university, the role of geographic proximity is more important compared to other types of organizations. In contrast, R&D collaborations that involve hospitals are positively affected by geographic distances. This means that MNCs and hospital can successfully collaborate at long distances.

Further, similar results were shown by the DID models looking at the type of actors when R&D collaborations are affected by the closure of a local MNC R&D site. Local R&D collaborations when an R&D site closes are more important if the collaboration involves a university compared to a hospital or a company.

	2000	2005	2011	2015
Universities	123	244	472	490
(median number of external organizations per NME)	(3)	(7)	(7.5)	(11.5)
Hospitals	44	72	134	152
(median number of external organizations per NME)	(0)	(1.5)	(1.5)	(1)
Research Institutes	57	80	177	151
(median number of external organizations per NME)	(0)	(1)	(2.5)	(3)
Companies	88	122	286	271
(median number of external organizations per NME)	(2.5)	(2.5)	(10)	(7.5)

Table 5| Number of actors at the organizational level in the citations of the main patents behind new drugs (Paper I).

6.2.3 Importance of star scientists in the retention of local collaborations

Finally, the relationships between geographic distances and types of actors at the individual level were studied with DID models, focusing on the R&D collaborations involving star scientists.

In accordance with earlier studies, the involvement of star scientists with MNCs positively affects the number of collaborations. Furthermore, local R&D collaborations are less negatively affected by the closure of R&D sites if they are linked to star scientists of MNCs before the closure compared to if star scientist were not involved. Finally, I compared the effect of the location of star scientists distinguishing between local (located at the R&D site) versus non-local (located elsewhere in the company). The local R&D collaborations involving a MNCs star scientist are less affected by the closure if the star scientist was not located in that R&D site that is closing. These results suggest that star scientists are important for MNC to retain R&D collaborations and to continue to be able to have access to local knowledge after the closure of an R&D site.

7 DISCUSSION

The aim of this thesis was to investigate the role of geographic proximity in R&D collaborations of pharmaceutical MNCs when creating, accessing and embedding different types of knowledge. I find that the role of geographic proximity is dependent on the nature of the knowledge and type of actors that are involved in the collaboration.

In the next section, I discuss the changing role of openness in the drug discovery and development process and in R&D collaborations. Then, the role of geographic proximity in R&D collaborations is discussed.

7.1 INCREASING OPENNESS OF PHARMACEUTICAL MNCS

In accordance with earlier studies (Chesbrough, 2003; Tijssen, 2009; Simpson and Reichman, 2013; Rafols *et al.*, 2014; Crescenzi, Nathan and Rodríguez-Pose, 2016), the results of my thesis point to an increasing openness of the pharmaceutical industry with a high number of R&D collaborations and different organizations involved in the drug development process. Pharmaceutical innovation is no longer a stand-alone activity conducted by a few big pharmaceutical companies in isolation.

The openness of the pharmaceutical industry was investigated looking at how the knowledge behind new drugs has changed over time. Pharmaceutical MNCs are increasingly relying on external organizations to create new drugs. However, this increased openness was found to be mostly indirect. This means that the patents and the ownership of the NMEs are still internal to the MNCs; external organizations are used only indirectly to access new knowledge. Thus, although the pharmaceutical industry is using more external organizations, I still observed a traditional division of the drug discovery and development process. The private sector is mainly responsible to test new drug compounds in the clinical phase, and they still own the main patents behind new drugs. The public sector is the main driver of basic research to clarify the underlying mechanisms and pathways of disease and detect promising drug compounds. Thus, despite the increasing involvement of universities, companies sponsoring drug development still own the main patents behind them.

Collaborators are increasingly shared between MNCs developing drugs. Thus the same external organizations contribute indirectly to the creation of several NMEs. These are mainly ‘elite institutions’ and MNCs collaborate with them regularly to access new knowledge from the best clinical researchers and scientists. Firms arrange funding, human resources, infrastructure, compound libraries and technologies, whereas academic and clinical partners supply in new ideas for the drug targets.

Looking deeper into how MNCs use external collaborations, this thesis shows that MNCs use R&D collaborations to create, access and embed different types of knowledge that contribute to an internal heterogeneous knowledge pool. For example, R&D collaborations of MNCs are more common in the clinical science and core knowledge areas compared to basic science and exploration types of knowledge. Looking at the type of actors at the organizational level, I

found that universities are the most important knowledge contributor to get access to basic science in collaborations for MNCs. This is in line with earlier literature. For example, Cohen, Nelson and Walsh (2002) conclude that a third of industrial R&D projects relied on research carried out in academia. Additionally, the results show that universities are also crucial actors in the drug discovery and development process and their relative importance has increased between 2000 and 2015. This can be explained by Bayh-Dole Act (Mowery and Sampat, 2005), the growing number 'translational research' activities by goverments (Zerhouni, 2005), and public-private research partnerships (Stiglitz and Wallsten, 1999) that have promoted the growing number of interactions between universities and companies in the last 40 years. On the other side, MNCs collaborate with universities to share costs and access the knowledge of highly skilled academic researchers (Rothaermel and Hess, 2007).

Finally, looking at the type of actors at the individual level, this thesis shows that the knowledge embedded in star scientists internal to the MNCs influences positively the number of collaborations. To this point, prior literature has suggested that collaborations with top researchers have a positive effect on the innovative performance of firms (Baba, Shichijo and Sedita, 2009). In fact, star scientists are considered an important source of knowledge for a company not only through their own research but also by being part of a large scientific network (Lacetera, Cockburn and Henderson, 2004; Rothaermel and Hess, 2007).

7.2 DIFFERENT TYPES OF KNOWLEDGE DETERMINE THE ROLE OF GEOGRAPHIC PROXIMITY

In combination with the increased openness, globalization has facilitated collaborations at long geographic distances and pharmaceutical MNCs have adopted a combination of local and distant collaborations. This trend is observed in the study of the organizations that contributed to new drugs. The results show that the number of involved organizations from different countries has increased. But also that MNCs choose to open R&D locations throughout the world to harness locally external capabilities and to exploit local knowledge (Gassmann and Reepmeyer, 2005; Rusu, Kuokkanen and Heier, 2011).

Due to an increasing number of global collaborations, considerable attention has been paid in the literature to the (changing) role of geographic proximity. In accordance with earlier studies, this thesis shows that – despite the globalization trend and the development of information communication technologies – geographic proximity still matters in R&D collaborations (Ponds, van Oort and Frenken, 2007; Douglas *et al.*, 2010; Abramovsky and Simpson, 2011; Plotnikova and Rake, 2014). In particular, geographic proximity positively affects the number of collaborations of MNCs. Being located at short distance facilities the number of collaboration between an MNC and an external organization. By studying the effect of closures of R&D sites, I find that MNCs partly loose access to local knowledge. Thus, local R&D sites act as temporary co-location spaces for collaborations between MNCs and local organizations.

This thesis also shows that the role of geographic proximity in R&D collaborations is complex and, that the nature of knowledge created and the types of actors involved in the collaborations have to be taken into account.

7.2.1 Basic science and core knowledge collaborations are more sensitive to geographic proximity

Considering the nature of knowledge, collaborations are less affected by geographic distance in clinical science than in basic science knowledge areas. Basic science knowledge is highly tacit and characterized by its serendipitous nature. This can be explained by looking at the different platforms adopted by the major pharmaceutical MNCs under the umbrella of 'open innovation' to facilitate R&D collaborations. For example, Pfizer established the Center for Therapeutic Innovation, a network for academic collaborators. It aims to bridge the gap between early scientific discoveries and translation into new drugs, building flexible and small Pfizer locations inside major universities in the US. Another example is the open innovation initiative by GlaxoSmithKline called Open lab at 3 Cantos, a research centre concentrating on diseases common in developing countries such as malaria and tuberculosis. The centre supports visiting researchers from universities to carry out research there. In 2002, Novartis launched its Institute for BioMedical Research in Cambridge (Massachusetts) to gain access to local universities and research hubs.

In contrast, clinical science knowledge can be considered more codified knowledge and thus can more easily be transferred through collaborations at geographic distances. The number of clinical trials carried out in so-called emerging regions, especially in Eastern European, Latin American and Asian countries, increased in the last 20 years. Thiers, Sinskey, and Berndt (2008) showed that between 2002 and 2006, the number of biopharmaceutical clinical trial sites in the emerging regions grow with 21.3%. The reasons behind this shift include lower costs, faster patient recruitment, establishment of contract research organizations focused on global clinical trials, and access to new markets (Drain *et al.*, 2018). Additionally, Haeussler and Rake (2017) have shown that the knowledge base of emerging regions plays an important role in the location decision for clinical trials by pharmaceutical companies. Emerging countries see the opportunities to enhance their national innovative capacity with the active participation of domestic scientists in clinical research projects together with scientists for developed countries.

Finally, in both basic and clinical science, R&D collaborations in core knowledge areas need closer geographic proximity than in exploration knowledge areas. Companies are exploring potential new areas to invest in and using distant collaborations to learn about market trends and the newest technologies (Bathelt, Malmberg and Maskell, 2004). Instead, collaborations in core areas rely on long-term relationships and trust. These knowledge areas are the main income sources for organizations and require close proximity (Kessel, 2011).

7.2.2 Universities are more sensitive collaborators concerning geographic proximity

Looking at the actors at the organizations level, this thesis shows that R&D collaborations are more negatively influenced by distance if the interaction involves a university compared to a hospital. A similar result is found in my analysis of the effect of the closure of R&D sites on local collaborations. The presence of a local MNC R&D site is more important when it comes to setting up and maintaining collaborations with universities or research institutes compared to other types of organizations. These results are in line with those of the previous section about the nature of the knowledge. In fact, academia embeds mostly basic science knowledge that due to its tacit nature depends on short distances and face-to-face interactions to be transferred (Arora and Gambardella, 1990; Owen-Smith and Powell, 2004). When MNCs access knowledge from hospitals, it is mainly with the aim to access clinical science knowledge which is more accessible over distance.

These results are also in line with the role of institutional proximity introduced by Boschma (2005). He argued that geographic proximity is neither a prerequisite nor sufficient for successful collaboration and it can compensate for the lack of institutional proximity (when actors are far apart in norms, practices and/or incentives) in collaborations. Hence, MNC-university collaborations require geographic proximity for their lack of institutional proximity.

7.2.3 Star scientists compensate for the lack of geographic proximity

Finally, this thesis looked at the role of geographic proximity when the collaborations involve particular individuals, such as star scientists. The results suggest that star scientists of MNCs play a positive role in the retention of local R&D collaborations after the closure of an R&D site. When an R&D site is closing, the risk that the MNC may lose out on local knowledge is mitigated by the presence of star scientists in the collaborations. The results show that even though the star scientists are located elsewhere within the MNC they have a positive effect on local collaborations.

Thus, geographic proximity in R&D collaboration is less important when it involves a star scientist. This result highlights the boundary-spanning roles of star scientists and their importance for collaborations at geographic distances. Star scientists often have a large network and can serve as a channel for the firm to reach the scientific community (Arora and Gambardella, 1990). These individuals act as a bridge between organizations (Lacetera, Cockburn and Henderson, 2004) and "creators of knowledge road" between different units and geographic areas (Maier, Kurka and Trippl, 2007).

8 CONCLUSIONS

Increased openness in the innovation process of pharmaceutical MNCs has been presented as a solution to the productivity crisis faced by these companies. Pharmaceutical MNCs are increasingly collaborating with a variety of organizations, individuals and expanding into different knowledge areas. Considering this increased openness, a key strategic question for MNCs is how to create, assimilate and embed knowledge effectively in R&D collaborations. This thesis highlights the role of geographic proximity in R&D collaborations and how it is dependent on the knowledge types involved. MNCs are characterized by high internal diversity and the collaborations of a MNC differ depending on the stages of the innovation process and research and therapeutic areas involved. Thus, companies need to develop different strategies for collaborations depending on the type of knowledge involved. The importance of knowledge types in R&D collaboration supports the view that an organization or individual knowledge base involved determines the success of R&D collaborations. This suggests that MNCs need to attract and involve the right organizations and individuals in their collaborations.

Nevertheless, my analysis shows that there is room for universities to play a more active role in the drug innovation process given their unique knowledge base. It is therefore important to continue to promote 'translational research' and public-private partnerships not only to increase collaborations but also to encourage a more active role for academia and organizations from other sectors in the creation of innovation. Furthermore, the analysis of the role of geographic proximity and star scientists when an R&D site of MNCs is closing highlights the crucial role of star scientists for successful external collaborations of MNCs. Star scientists are important for MNCs not only because of their expertise and knowledge base but also because of their ability to span collaborations over longer distances and compensate for the absence of local R&D sites and mitigate the negative effect of the closure of a MNC R&D site.

Finally, the link between knowledge type and geography emphasizes the importance of considering geography in R&D management. Managers, when they need to close or open a MNC R&D site, must consider that the importance of geographic proximity and face-to-face interactions depends on the nature of the knowledge created and accessed in that specific location. Thus, the knowledge type of the MNC's R&D site can be considered a location driver as well as a strategic characteristic to consider when a MNC decides to close or open a R&D site. From a policy standpoint, these arguments are important to keep in mind when local actors are trying to attract foreign MNCs and when facilitating and supporting the development of collaboration.

8.1 FUTURE WORKS

This thesis has mainly contributed to increasing our understanding of the role of geographic proximity in MNCs R&D collaborations. However, R&D collaboration is only one of the mechanisms of the so-called open innovation strategy of a firm. Companies vary considerably in their modes and strategies regarding how to access external knowledge. These different modes, take many forms, such as joint-equity, ventures, recruitment and acquisitions (Powell,

Koput and Smith-Doerr, 1996). The strategy towards external knowledge can affect why some organizations gain more from external interactions than others (Dahlander and Gann, 2010). Thus, it could be interesting to compare the role of R&D collaborations and the importance of geographic proximity therein with other modes and strategies used by MNCs to access external knowledge.

Another interesting area for future research could be an in-depth investigation into knowledge creation, looking specifically at each R&D collaboration of an MNC. This can provide insights for firms on how to engage in successful collaborations. For example, the in-depth study might look at how the knowledge being created in the R&D collaborations involving hospitals differs from those with universities.

Furthermore, the role of individuals, and in particular star scientists, in collaborations and teams have received a lot of attention recently in the literature (for example Bennett and Gadlin, 2012; Bozeman and Boardman, 2014). However, a key question remains: What is the role of individuals in R&D collaboration? Additionally, are they increasingly becoming network orchestrators?

Finally, the main focus of this thesis is on geographic proximity. Institutional proximity has been considered only marginally in relation to geographic proximity. However, proximity is a multi-dimensional concept that includes geographic, institutional, organizational, social and cognitive proximity. To be able to further understand R&D collaborations, other dimensions of proximity need to be considered. Thus, it could be interesting to investigate how the different proximity dimensions vary depending on the type of knowledge involved in R&D collaborations. For example, I would expect that cognitive proximity is more important for collaborations in basic science knowledge due to its tacit nature compared to clinical science knowledge. In fact, if the cognitive distance is too large, the actors involved will not understand each other and will not be able to interpret and assimilate knowledge; this is particularly difficult for tacit knowledge (Nooteboom, 1992). In addition, several authors have shown that the dimensions of proximity may be complementary to each other or they may also act as substitutes (Boschma, 2005; D'Amore *et al.*, 2013; Davids and Frenken, 2017). A future research question might ask how does complementarity vary according to different dimensions of proximity and knowledge?

9 ACKNOWLEDGEMENTS

Collaboration is possible only with challenges and growth – quite similar to the experience of my Ph.D., which encompasses many scientific as well as personal developments. I am very grateful for this experience. It would not have been possible and enjoyable without the help of friends and family. In particular, I would like to thank the following people.

First and foremost, I would like to express my sincere gratitude to my main supervisor **Pauline**. Thank you for rewarding discussions, for the long travels and for sharing your perspective. You gave me the freedom to develop my own idea, you pushed me beyond my limits and you made me step up. I have gained a lot of knowledge from our collaboration, which has been fun and inspiring. You have been without doubts an excellent researcher and guide in this journey.

Jarno, my co-supervisor, thank you for encouraging support during all stages of my Ph.D. Utrecht was a critical learning step to form me as a researcher as well as an enjoyable and cultural enrichment. I have really appreciated your insightful comments and professional guidance as well as the fun time.

Carl Johan, my co-supervisor, your energy has been an inspiration for me. Even with your busy schedule, you have always found time for few words. Thank you for always being ready to help in all research and non-research related challenges. Your encouragement and kind words have meant a lot to me.

Hanna, my manager at UBE, thanks for introducing me to UBE, for your countless advice, for supporting me and for not allowing me to quit– you always supported me emotionally as well as professionally during this last year.

I would also like to thank all the members of the UBE. Special thank you to **Maddis**, **Ulrika**, **Sanna**, **Lena** for all the discussions about entrepreneurship during the fika.To **Anna S.** and **Patrik** to help me at the beginning of my Ph.D. To **Liisa** who provided unreserved help whenever needed. To **Andreas**, **Cecilia**, **Anna B.**, **Pu** and **Samer** for bringing a lot of fun in UBE. To **Charlotta** and **Danielle**, to share part of the Ph.D´s journey. A thank you to **Max** and **Linda** for the nice evenings and tough training sections together. **Anneliese** for the Italian vacation and our supportive friendship. To **Jill** and **Annika O.** for being my first Swedish friends and make me fall in love with Stockholm.

For all the colleagues at **MMC** and **LIME** to make my Ph.D. a fun journey. To **Pamela** and **Carl**, for welcoming me in your research group. To **George, Rafiq, Mairi, Mimmi, Carolina**, **Filip** and **Anton** for sharing the ups and the downs of my Ph.D.

To all my colleagues and friends at **CARTANA**. It has been an honour to be part of the CARTANA journey since the beginning. What a journey! It is very inspiring to work with you every day! A special thanks to **Malte** for believing in me. You have been a true source of inspiration for your passion, your values, and your intelligence. We have come so far already and yet we have just started. I'm only looking forward to the years to come!

To **Elisabeth, Shahul, Tiago**, my Swedish family, thanks for the great memories together over these past six years. I feel extremely lucky to have you in every life challenges. A special thanks to **Elisabeth** for being my dearest friend, for caring so much. I am looking forward to our adventure with your family, **Julian** and **David**. **Shahul**, thank you for enriching me with your culture and your amazing friendship. Thank you for the dinners, trips and for being patience when I was late. India is still by far my favourite trip and I miss living in the same city. **Tiago**, for sharing the fun times, the travelling, the parties, the conferences.. and thanks for bringing in our group of friends Agnes. **Agnes**, thanks for your smile and your positivity. It is always so fun to spend time with you.

To **Alvaro**, for being so honest, so special and open in our friendship. To **Raquel**, first friend and then also a UBE colleague. Thanks for your kind words and to always find time to listen. To **Alice**, **Marco** and **Giulia**, you have been a surprise in Stockholm, thanks for the many Sunday hike and many more good moments. Mi sento sempre in famiglia con voi! To **Davide**, **Justyna** and **Giulia**, you have changed my life and I will never be enough grateful to you. I wish you will soon move back to Stockholm. To **Goncalo** and **Simona**, I always enjoy spending time with you. I hope your future will be in Stockholm. To **Viktoria** for being always happy. To **Bettina** for the fantastic cover of this thesis, for your creativity and for thinking outside the box. To **Annika B**. thanks for bringing me serenity. To **Claudia and Fede**, for many more good moments between Italy and Sweden . Thank all the other friends (**Elira** and **Paolo**, **Manu, Marion, Pim, Veronica, Jonathan, Eniko, Aron, Alberto, Anastasia, Lucio, Francesco, Susanna**) I met in Sweden, thanks for making me love the time in Sweden.

To **Fredrik P.**, thanks for the support during my life challenges journey. I have learned a lot in biology and clinical pharmaceutical from our stimulating discussion. Life is very surprising and thanks to you, I have discovered a different perspective of the pharmaceutical industry.

During my Ph.D., I have been involved in several extra activities that make the journey even more interesting. Thanks to CHASE team (**Tiago, Cécile, Christian, Daniel, Emelie, Florian, Lizan, Michael, Moritz, Susanna, Thibaud**), Ssidelinecityrunclub and TEDxStockholm team.

Per gli amici in Italia lontani ma sempre presenti nella mia vita! **Marco B.**, you convinced me to move in Stockholm and our first year in Stockholm was very intense and unforgettable: work, science, star up competition, hilarious discussions about life and science. You have taught me to follow my passions and to be the exception in society. It has been an honour to be part of your journey. To **Joy, Maddy** and **Betta** for our lifelong friendship. Even if we meet only once a year, I love our time together. Per la **CC** per essere ”personi reali” e farmi ritornare alla realta´. E´ sempre bellissimo tornare a Romano e adoro la semplicita´ con cui affrontate la vita!

Alla famiglia Stucchi, alle **nonne, zii** e **cugini di Daniele**, per essere sempre pronti ad accogliermi a Romano. To **Enrico**, avermi fatta subito sentire parte della famiglia. To **Elia**, per

essere il miglior cognato e per le nostre discussioni imprenditoriali. Per **Leti e Neva** grazie per farmi sentire un po´ come una sorella maggiore.

To my family in Cuneo, thank you for always being by my side. To **my parents**, thank you for creating the perfect family to grow up in. You gave me unconditional love and provided me with everything I need to follow my dreams. You always push me to challenge myself and you were always close to me, despite many miles between. To **my Dad**, thank you for triggering my curiosity in science and for encouraging me to never stop learning. Alla **mia mamma**, per essere la spina dorsale della famiglia. Grazie per la tua pazienza e il tuo incoraggamento nel mio viaggio imprenditoriale! To my sisters for being my best friends, I know that you will always be there for me. A **Martina e Davide** grazie per avermi prottetto in tutti questi anni e per la vostra bellissima famiglia. To **Maria Piera**, my youngest sister, I'm so proud of what you have become; you are my role model for being so young yet so ambitious in life. A **Olivia e Jacopo**, voi siete stati un gioia pura nella mia vita durante questo viaggio, riempidoni di energia e sorrisi!

My final words of acknowledgement are addressed to **Daniele**, who came into my life halfway through this journey and has enriched my life with fun, joyful, and unforgettable moments. Thank you for being patient, for your constant encouragement and for celebrating all of the big and small milestones of this Ph.D. You mean more than words could ever say. You are my rock and this Ph.D would not have been possible without you. Thank you for making my every day so special.

10 REFERENCES

Abramovsky, L. and Simpson, H. (2011) 'Geographic proximity and firm-university innovation linkages: Evidence from Great Britain', *Journal of Economic Geography*, 11(6), pp. 949–977. doi: 10.1093/jeg/lbq052.

Acs, Z. J., Audretsch, D. B. and Feldman, M. P. (1994) 'R & D spillovers and recipient firm size', *Review of Economics & Statistics*, 76(2), pp. 336–340. doi: 10.2307/2109888.

Agrawal, A. and Henderson, R. (2002) 'Putting Patents in Context: Exploring Knowledge Transfer from MIT', *Management Science*, 48(1), pp. 44–60. doi: 10.1287/mnsc.48.1.44.14279.

Anselin, L., Varga, A. and Acs, Z. (2000) 'Geographical spillovers and university research: A spatial econometric perspective', *Growth and Change*. Blackwell Publishing Inc., 31(4), pp. 501–515. doi: 10.1111/0017-4815.00142.

Arora, A. and Gambardella, A. (1990) 'Complementarity and External Linkages : the Strategies of the Large Firms in Biotechnology', *The Journal of Industrial Economics*, 38(4), pp. 361–380. doi: 10.2307/2098345.

Asheim, B. T. and Coenen, L. (2005) 'Knowledge bases and regional innovation systems: Comparing Nordic clusters', *Research Policy*, 34(8), pp. 1173–1190. doi: 10.1016/j.respol.2005.03.013.

Asheim, B. T. and Gertler, M. S. (2009) 'The Geography of Innovation: Regional Innovation Systems', in *The Oxford Handbook of Innovation*. Oxford University Press. doi: 10.1093/oxfordhb/9780199286805.003.0011.

Ashenfelter, O. and Card, D. (1985) *Using the Longitudinal Structure of Earnings to Estimate the Effect of Training Programs*, *The Review of Economics and Statistics*.

Audretsch, D. B. and Feldman, M. P. (1996) 'Innovative clusters and the industry life cycle', *Review of Industrial Organization*, 11(2), pp. 253–273. doi: 10.1007/BF00157670.

Azoulay, P., Graff Zivin, J. S. and Wang, J. (2010) 'Superstar Extinction', *The Quarterly Journal of Economics*, 125(2), pp. 549–589. doi: 10.1162/qjec.2010.125.2.549.

Azoulay, P., Zivin, J. S. G. and Sampat, B. (2011) *The Diffusion of Scientific Knowledge Across Time and Space: Evidence from Professional Transitions for the Superstars of Medicine*. Cambridge, MA. doi: 10.3386/w16683.

Baba, Y., Shichijo, N. and Sedita, S. R. (2009) 'How do collaborations with universities affect firms' innovative performance? The role of "Pasteur scientists" in the advanced materials field', *Research Policy*, 38(5), pp. 756–764. doi: 10.1016/j.respol.2009.01.006.

Balland, P. A., Boschma, R. and Frenken, K. (2015) 'Proximity and Innovation: From Statics to Dynamics', *Regional Studies*. Taylor & Francis, 49(6), pp. 907–920. doi: 10.1080/00343404.2014.883598.

Bathelt, H., Malmberg, A. and Maskell, P. (2004) 'Clusters and knowledge: Local buzz, global pipelines and the process of knowledge creation', *Progress in Human Geography*. Arnold, 28(1), pp. 31–56. doi: 10.1191/0309132504ph469oa.

Beckman, C. M., Haunschild, P. R. and Phillips, D. J. (2004) 'Friends or Strangers? Firm-Specific Uncertainty, Market Uncertainty, and Network Partner Selection', *Organization*

Science, 15(3), pp. 259–275. doi: 10.1287/orsc.1040.0065.

Bennett, L. M. and Gadlin, H. (2012) 'Collaboration and Team Science: From Theory to Practice', *Journal of investigative medicine : the official publication of the American Federation for Clinical Research*. NIH Public Access, 60(5), p. 768. doi: 10.231/jim.0b013e318250871d.

Blundell, R. and Costa Dias, M. (2009) 'Alternative approaches to evaluation in empirical microeconomics', *Journal of Human Resources*, pp. 565–641. doi: 10.3368/jhr.44.3.565.

Boschma, R. (2005) 'Proximity and innovation: A critical assessment', *Regional Studies*, 39(1), pp. 61–74. doi: 10.1080/0034340052000320887.

Bozeman, B. and Boardman, C. (2014) *Research Collaboration and Team Science: A State-of-the-Art Review and Agenda*. doi: 10.1007/978-3-319-06468-0.

Broström, A. (2010) 'Working with distant researchers - Distance and content in university-industry interaction', *Research Policy*, 39(10), pp. 1311–1320. doi: 10.1016/j.respol.2010.09.002.

Buckley, P. J. and Carter, M. J. (2004) 'A formal analysis of knowledge combination in multinational enterprises', *Journal of International Business Studies*, pp. 371–384. doi: 10.1057/palgrave.jibs.8400095.

Calero, C., Van Leeuwen, T. and Tijssen, R. (2007) 'Research cooperation within the bio-pharmaceutical industry: Network analyses of co-publications within and between firms', *Scientometrics*, 71(1), pp. 87–99. doi: 10.1007/s11192-007-1650-y.

Cantner, U. and Rake, B. (2014) 'International research networks in pharmaceuticals: Structure and dynamics', *Research Policy*. Elsevier B.V., 43(2), pp. 333–348. doi: 10.1016/j.respol.2013.10.016.

Cantwell, J. (1995) 'The globalisation of technology : what remains of the product cycle model?', *Cambridge Journal of Economics*, 19, pp. 155–174. doi: http://cje.oxfordjournals.org/.

Cantwell, J. and Mudambi, R. (2011) 'Physically attraction and the geography of knowledge sourcing in multinational enterprises', *Global Strategy Journal*, 232, pp. 206–232. doi: 10.1111/j.2042-5805.2011.00024.x.

Cassiman, B. and Veugelers, R. (2002) 'R&D cooperation and spillovers: Some empirical evidence from Belgium', *American Economic Review*, 92(4), pp. 1169–1184. doi: 10.1257/00028280260344704.

Castellani, D. and Zanfei, A. (2006) *Multinational Firms, Innovation and Productivity*, *Edward ElgardPublishing*. doi: 10.1016/j.respol.2007.08.002.

Chesbrough, H. W. (2003) *Open Innovation - The New Imperative for Creating and Profiting from Technology*. Harvard Business School Publishing Corporation.

Cockburn, I. and Henderson, R. (1996) 'Public-private interaction in pharmaceutical research', *Proceedings of the National Academy of Sciences of the United States of America*, 93(23), pp. 12725–12730. doi: 10.1073/pnas.93.23.12725.

Cockburn, I., Henderson, R. and Stern, S. (1999) 'The Diffusion of Science Driven Drug Discovery: Organizational Change in Pharmaceutical Research', *NBER Working Paper*

Series, 7359. doi: 10.3386/w7359.

Cockburn, I. M. and Henderson, R. M. (2001) 'Scale and scope in drug development: Unpacking the advantages of size in pharmaceutical research', *Journal of Health Economics*, 20(6), pp. 1033–1057. doi: 10.1016/S0167-6296(01)00108-4.

Cohen, E. A. and Cairncross, F. (1997) 'The Death of Distance: How the Communications Revolution Will Change Our Lives', *Foreign Affairs*. Harvard Business School Press, 76(6), p. 157. doi: 10.2307/20048298.

Cohen, W. and Levinthal, D. (1989) 'Innovation and learning: the two faces of R & D', *The economic journal*, 99(397), pp. 569–596.

Cohen, W., Nelson, R. and Walsh, J. (2002) 'Links and Impacts: The Influence of Public Research on Industrial R&D', *Management Science*, 48(1), pp. 1–23. doi: 10.1287/mnsc.48.1.1.14273.

Cowan, R., David, P. and Foray, D. (2000) 'The explicit economics of knowledge codification and tacitness', *Industrial and Corporate Change*, 9(2), pp. 211–253. doi: 10.1093/icc/9.2.211.

Crescenzi, R., Nathan, M. and Rodríguez-Pose, A. (2016) 'Do inventors talk to strangers? on proximity and collaborative knowledge creation', *Research Policy*. Elsevier B.V., 45(1), pp. 177–194. doi: 10.1016/j.respol.2015.07.003.

D'Amore, R. *et al.* (2013) 'Research Collaboration Networks in Biotechnology: Exploring the Trade-Off Between Institutional and Geographic Distances', *Industry & Innovation*, 20(3), pp. 261–276. doi: 10.1080/13662716.2013.791127.

Dahlander, L. and Gann, D. M. (2010) 'How open is innovation?', *Research Policy*. Elsevier B.V., 39(6), pp. 699–709. doi: 10.1016/j.respol.2010.01.013.

Dasgupta, P. and David, P. (1994) 'Toward a new economics of science', *Research Policy*, 23(5), pp. 487–521. doi: 10.1016/0048-7333(94)01002-1.

David, P. and Foray, D. (1995) 'Accessing and expanding the science and technology knowledge base', *STI review*, pp. 13–68.

Davids, M. and Frenken, K. (2017) 'Proximity, knowledge base and the innovation process', *Regional Studies*. doi: 10.1080/00343404.2017.1287349.

Denzin, N. Y. and Lincoln, Y. (1998) *Strategies of Qualitative Inquiry*, *SAGE Publications*. doi: 10.21225/d5001g.

DiMasi, J. A., Grabowski, H. G. and Hansen, R. W. (2016) 'Innovation in the pharmaceutical industry: New estimates of R&D costs', *Journal of Health Economics*, 47, pp. 20–33. doi: 10.1016/j.jhealeco.2016.01.012.

Douglas, F. L. *et al.* (2010) 'The case for entrepreneurship in R&D in the pharmaceutical industry', *Nature Reviews Drug Discovery*. Nature Publishing Group, pp. 683–689. doi: 10.1038/nrd3230.

Drain, P. K. *et al.* (2018) 'Global migration of clinical research during the era of trial registration', *PLoS ONE*. Public Library of Science, 13(2). doi: 10.1371/journal.pone.0192413.

Duysters, G. and Lokshin, B. (2011) 'Determinants of alliance portfolio complexity and its

effect on innovative performance of companies', *Journal of Product Innovation Management*, 28(4), pp. 570–585. doi: 10.1111/j.1540-5885.2011.00824.x.

Feldman and Lichtenberg (2000) 'The Impact and Organization of Publicly-Funded Research and Development in the European Community', *Annales d'Économie et de Statistique*, (49/50), p. 199. doi: 10.2307/20076115.

Freeman, C. (1991) 'Networks of innovators: A synthesis of research issues', *Research Policy*, 20(5), pp. 499–514. doi: 10.1016/0048-7333(91)90072-X.

Furukawa, R. and Goto, A. (2006) 'The role of corporate scientists in innovation', *Research Policy*, 35(1), pp. 24–36. doi: 10.1016/j.respol.2005.07.007.

Gao, X., Guan, J. and Rousseau, R. (2011) 'Mapping collaborative knowledge production in China using patent co-inventorships', *Scientometrics*. Kluwer Academic Publishers, 88(2), pp. 343–362. doi: 10.1007/s11192-011-0404-z.

Gassmann, O. and Reepmeyer, G. (2005) 'Organizing Pharmaceutical Innovation: From Science-based Knowledge Creators to Drug-oriented Knowledge Brokers', *Creativity and Innovation Management*, 14(3), pp. 233–245. doi: 10.1111/j.1467-8691.2005.00344.x.

Gassmann, O. and Von Zedtwitz, M. (2003) 'Trends and determinants of managing virtual R&D teams', *R and D Management*. Blackwell Publishing Ltd, pp. 243–262. doi: 10.1111/1467-9310.00296.

Gertler, M. S. (2003) 'Tacit knowledge and the economic geography of context, or The undefinable tacitness of being (there)', *Journal of Economic Geography*, 3(1), pp. 75–99. doi: 10.1093/jeg/3.1.75.

Gittelman, M. and Kogut, B. (2003) 'Does Good Science Lead to Valuable Knowledge? Biotechnology Firms and the Evolutionary Logic of Citation Patterns', *Management Science*, 49(4), pp. 366–382. doi: 10.1287/mnsc.49.4.366.14420.

Glaeser, E. L. (1998) 'Are Cities Dying?', *Journal of Economic Perspectives*. American Economic Association, 12(2), pp. 139–160. doi: 10.1257/jep.12.2.139.

Glickman, S. W. *et al.* (2009) 'Ethical and scientific implications of the globalization of clinical research', *New England Journal of Medicine*. Massachussetts Medical Society, 360(8), p. 816. doi: 10.1056/NEJMsb0803929.

Grant, R. M. (1996a) 'Prospering in Dynamically-Competitive Environments: Organizational Capability as Knowledge Integration', *Organization Science*, 7(4), pp. 375–387. doi: 10.1287/orsc.7.4.375.

Grant, R. M. (1996b) 'Toward a Knowledge-Based Theory of the Firm', *Strategic Management Journal*, 17, pp. 93–107.

Haeussler, C. and Rake, B. (2017) 'The changing geography of clinical research: A critical analysis of its drivers', *Industrial and Corporate Change*, 26(2), pp. 285–310. doi: 10.1093/icc/dtx002.

Hagedoorn, J. (1993) 'Understanding the rationale of strategic technology partnering: Nterorganizational modes of cooperation and sectoral differences', *Strategic Management Journal*. Wiley, 14(5), pp. 371–385. doi: 10.1002/smj.4250140505.

Hagedoorn, J. and Duysters, G. (2002) 'External sources of innovative capabilities: The

preference for strategic alliances or mergers and acquisitions.', *Journal of Management Studies*, 39(2), pp. 167–188. doi: 10.1111/1467-6486.00287.

Hamel, G. (1991) 'Competition for competence and interpartner learning within international strategic alliances', *Strategic Management Journal*, 12(1 S), pp. 83–103. doi: 10.1002/smj.4250120908.

Hamilton, K. S. (2003) *Subfield and Level Classification of Journals*.

Hansen, M. T., Nohria, N. and Tierney, T. (1999) 'What's your strategy for managing knowledge?', *Harvard business review*, 77(2), pp. 106–16. Available at: www.hbr.org (Accessed: 27 January 2020).

Heckman, J. J., Lalonde, R. J. and Smith, J. A. (1999) 'Chapter 31 The economics and econometrics of active labor market programs', in *Handbook of Labor Economics*. Elsevier, pp. 1865–2097. doi: 10.1016/S1573-4463(99)03012-6.

Henderson, R. and Clark, K. B. (1990) 'Architectural Innovation: The Reconfiguration of Existing Product Technologies and the Failure of Established Firms', *Administrative Science Quarterly*. JSTOR, 35(1), p. 9. doi: 10.2307/2393549.

Henderson, R. and Cockburn, I. (1994) 'Measuring competence? Exploring firm effects in pharmaceutical research', *Strategic Management Journal*, 15, pp. 63–84. doi: 10.1016/0024-6301(95)94304-H.

Hess, A. M. and Rothaermel, F. T. (2011) 'When are assets complementary? Star scientists, strategic alliances, and innovation in the pharmaceutical industry', *Strategic Management Journal*, 32(8), pp. 895–909. doi: 10.1002/smj.916.

Hill, R. G., Rang, H. P. and Vallance, P. (2012) *Drug Discovery and Development: Technology in Transition*. 2nd ed., *Drug Discovery and Development: Technology in Transition*. 2nd ed. Edited by H. R. RG Hill. Edinburgh: Churchill Livingstone/Elsevier.

von Hippel, E. (1994) '"Sticky Information" and the Locus of Problem Solving: Implications for Innovation', *Management Science*. Institute for Operations Research and the Management Sciences (INFORMS), 40(4), pp. 429–439. doi: 10.1287/mnsc.40.4.429.

Hoekman, J. *et al.* (2012) 'The Geographical Distribution of Leadership in Globalized Clinical Trials', *PLoS ONE*. Public Library of Science, 7(10). doi: 10.1371/journal.pone.0045984.

Hoekman, J., Frenken, K. and van Oort, F. (2009) 'The geography of collaborative knowledge production in Europe', *Annals of Regional Science*, 43(3 SPEC. ISS.), pp. 721–738. doi: 10.1007/s00168-008-0252-9.

Holm, J. R. and Østergaard, C. R. (2015) 'Regional Employment Growth, Shocks and Regional Industrial Resilience: A Quantitative Analysis of the Danish ICT Sector', *Regional Studies*. Routledge, 49(1), pp. 95–112. doi: 10.1080/00343404.2013.787159.

Iorio, R. *et al.* (2012) 'Research Collaboration and Publication's Citation Impact: Evidence in the Italian Biotech Sector', *Rivista italiana degli economisti*. Società editrice il Mulino, XVII(1/2012), pp. 75–100. doi: 10.1427/36440.

Jaffe, A. B. *et al.* (1993) 'Geographic localization of knowledge spillovers as evidence by patent citations', *Quarterly Journal of Economics*, (August), pp. 577–598. doi: 10.2307/2118401.

Kale, P., Singh, H. and Perlmutter, H. (2000) 'Learning and protection of proprietary assets in strategic alliances: Building relational capital', *Strategic Management Journal*, 21(3), pp. 217–237. doi: 10.1002/(SICI)1097-0266(200003)21:3<217::AID-SMJ95>3.0.CO;2-Y.

Katz, J. S. and Hicks, D. (1997) 'How much is a collaboration worth? A calibrated bibliometric model', *Scientometrics*. Springer Netherlands, 40(3), pp. 541–554. doi: 10.1007/BF02459299.

Katz, J. S. and Martin, B. R. (1997) 'What is research collaboration?', *Research Policy*, 26(1), pp. 1–18. doi: 10.1016/S0048-7333(96)00917-1.

Kessel, M. (2011) 'The problems with today ' s pharmaceutical business — an outsider ' s view', *Nature Publishing Group*. Nature Publishing Group, 29(1), pp. 27–33. doi: 10.1038/nbt0111-27.

Kneller, R. (2010) 'The importance of new companies for drug discovery: origins of a decade of new drugs', *Nature reviews. Drug discovery*. Nature Publishing Group, 9(11), p. 156. doi: 10.1038/nrd3251.

Koza, M. P. and Lewin, A. Y. (1998) 'The Co-evolution of Strategic Alliances', *Organization Science*, 9(3), pp. 255–264. doi: 10.1287/orsc.9.3.255.

Krugman, P. (1998) 'Space: The Final Frontier', *Journal of Economic Perspectives*. American Economic Association, 12(2), pp. 161–174. doi: 10.1257/jep.12.2.161.

Lacetera, N., Cockburn, I. and Henderson, R. (2004) 'Do firms change capabilities by hiring new people? A study of the adoption of science-based drug discovery.', *Advances in Strategic Management*, pp. 133–159. doi: 10.1016/S0742-3322(04)21005-1.

Lechner, M. (2010) 'The Estimation of Causal Effects by Difference-in-Difference Methods', *Foundations and Trends R in Econometrics*, 4(3), pp. 165–224. doi: 10.1561/0800000014.

De Leeuw, T., Lokshin, B. and Duysters, G. (2014) 'Returns to alliance portfolio diversity: The relative effects of partner diversity on firm's innovative performance and productivity', *Journal of Business Research*. Elsevier Inc., 67(9), pp. 1839–1849. doi: 10.1016/j.jbusres.2013.12.005.

Levinthal, D. A. and March, J. G. (1993) 'The myopia of learning', *Strategic Management Journal*, 14(2 S), pp. 95–112. doi: 10.1002/smj.4250141009.

Lim, K. (2004) 'The relationship between research and innovation in the semiconductor and pharmaceutical industries (1981-1997)', *Research Policy*, 33(2), pp. 287–321. doi: 10.1016/j.respol.2003.08.001.

Lotka, A. J. (1926) 'The frequency distribution of scientific productivity', *Journal of the Washington Academy of Science*, 16, pp. 317–323. doi: -.

Maier, G., Kurka, B. and Trippl, M. (2007) 'Knowledge spillover agents and regional development: spatial distribution and mobility of star scientists', *DYNREG (Dynamic Regions in a Knowledge-Driven Global Economy)*, 17, p. 35.

Malmberg, A. and Maskell, P. (2002) 'The Elusive Concept of Localization Economies: Towards a Knowledge-Based Theory of Spatial Clustering', *Environment and Planning A: Economy and Space*, 34(3), pp. 429–449. doi: 10.1068/a3457.

March, J. (1991) 'Exploration and exploitation in organizational learning', *Organization*

Science, 2(1), pp. 71–87. doi: 10.1287/orsc.2.1.71.

Marshall, A. (1927) *Industry and Trade: A Study of industrial technique and businnes organization; and of their influences on the condition of various classes and nations*. 3rd editio. Edited by London: Macmillan.

Martin, R. and Moodysson, J. (2011) 'Innovation in symbolic industries: The geography and organization of knowledge sourcing', *European Planning Studies*. Taylor & Francis, 19(7), pp. 1183–1203. doi: 10.1080/09654313.2011.573131.

Maskell, P. and Malmberg, A. (1999) 'The competitiveness of firms and regions. "Ubiquitification" and the importance of localized learning', *European Urban and Regional Studies*. SAGE Publications Ltd, 6(1), pp. 9–25. doi: 10.1177/096977649900600102.

Mattes, J. (2012) 'Dimensions of Proximity and Knowledge Bases: Innovation between Spatial and Non-spatial Factors', *Regional Studies*, 468, pp. 1085–1099. doi: 10.1080/00343404.2011.552493.

McAleer, M. and Slottje, D. (2005) 'A new measure of innovation: The patent success ratio', *Scientometrics*. Springer Netherlands, pp. 421–429. doi: 10.1007/s11192-005-0222-2.

Meyer, B. D. (1995) *Natural and Quasi-Experiments in Economics*, *JBES Symposium on Program and Policy Evaluation*.

Miller, K. D., Zhao, M. and Calantone, R. J. (2006) 'Adding interpersonal learning and tacit knowledge to March's exploration-exploitation model', *Academy of Management Journal*. Academy of Management, pp. 709–722. doi: 10.5465/AMJ.2006.22083027.

Mitsuhashi, H. and Greve, H. (2009) 'A matching theory of alliance formation and organizational success: Complementarity and compatibility', *Academy of Management Journal*, 52(5), pp. 975–995. doi: 10.5465/AMJ.2009.44634482.

Moed, H. F., Glänzel, W. and Schmoch, U. (2004) *Handbook of quantitative science and technology research: the use of publication and patent statistics in studies of S &T systems*, *Book*. Kluwer Academic Publishers. doi: 10.1007/1-4020-2755-9_20.

Moodysson, J., Coenen, L. and Asheim, B. (2008) 'Explaining spatial patterns of innovation: Analytical and synthetic modes of knowledge creation in the Medicon Valley life-science cluster', *Environment and Planning A*, 40(5), pp. 1040–1056. doi: 10.1068/a39110.

Mowery, D. C., Oxley, J. E. and Silverman, B. S. (1996) 'Strategic alliances and interfirm knowledge transfer', *Strategic Management Journal*, 17(S2), pp. 77–91. doi: 10.1002/smj.4250171108.

Mowery, D. C. and Sampat, B. N. (2005) 'The bayh-dole act of 1980 and university-industry technology transfer: A model for other OECD governments?', in *Essays in Honor of Edwin Mansfield: The Economics of R&D, Innovation, and Technological Change*, pp. 233–245. doi: 10.1007/0-387-25022-0_18.

Munos, B. (2009) 'Lessons from 60 years of pharmaceutical innovation.', *Nature reviews. Drug discovery*. Nature Publishing Group, 8(12), pp. 959–68. doi: 10.1038/nrd2961.

Narin, F. and Noma, E. (1987) 'Patents as indicators of corporate technological strength', *Research Policy*, 16(1), pp. 143–155.

Narin, F., Stevens, K. and Whitlow, E. S. (1991) 'Scientific co-operation in Europe and the

citation of multinationally authored papers', *Scientometrics*. Kluwer Academic Publishers, 21(3), pp. 313–323. doi: 10.1007/BF02093973.

Nelson, R. and Winter, S. G. (1982) 'The schumpeterian tradeoff revisited', *American Economic Review*. American Economic Association, 72(1), pp. 114–132. doi: 10.2307/1808579.

Nonaka, I. (1991) 'The Knowledge-Creating Company:How Japanese Companies Create the Dynamics of Innovation', *Harvard Business Review*, 69(November-December), pp. 96–104.

Nooteboom, B. (1992) 'Towards a dynamic theory of transactions', *Journal of Evolutionary Economics*. Springer-Verlag, 2(4), pp. 281–299. doi: 10.1007/BF01200127.

Nooteboom, B. (2004) *Inter-firm collaboration, learning and networks: An integrated approach*. doi: 10.4324/9780203392102.

Nooteboom, B. (2009) *A cognitive theory of the firm: Learning, governance and dynamic capabilities*, *A Cognitive Theory of the Firm: Learning, Governance and Dynamic Capabilities*.

NSF (2006) '2006: Year in Review'. Available at: https://www.nsf.gov/discoveries/disc_summ.jsp?cntn_id=108277.

OECD (2011) *Development Co-operation Report 2011 50th Anniversary Edition*, *Development Co-operation Report 2011 : 50th Anniversary Edition*. doi: 10.1787/dcr-2011-en.

Owen-Smith, J. and Powell, W. (2004) 'Knowledge Networks as Channels and Conduits: The Effects of Spillovers in the Boston Biotechnology Community', *Organization Science*, 15(1), pp. 5–21. doi: 10.1287/orsc.1030.0054.

Paci, R. and Usai, S. (2000) 'Technological enclaves and industrial districts: An analysis of the regional distribution of innovative activity in Europe', *Regional Studies*, 34(2), pp. 97–114. doi: 10.1080/00343400050006032.

Pammolli, F., Magazzini, L. and Riccaboni, M. (2011) 'The productivity crisis in pharmaceutical R&D.', *Nature reviews. Drug discovery*. Nature Publishing Group, 10(6), pp. 428–38. doi: 10.1038/nrd3405.

Pittaway, L. *et al.* (2004) 'Networking and innovation: A systematic review of the evidence', *International Journal of Management Reviews*, 5–6(3–4), pp. 137–168. doi: 10.1111/j.1460-8545.2004.00101.x.

Plotnikova, T. and Rake, B. (2014) 'Collaboration in pharmaceutical research: Exploration of country-level determinants', *Scientometrics*, 98(2), pp. 1173–1202. doi: 10.1007/s11192-013-1182-6.

Polanyi, M. (1966) 'The logic of tacit inference', *Philosophy*, 41(155), pp. 1–18. doi: 10.1017/S0031819100066110.

Ponds, R., van Oort, F. and Frenken, K. (2007) 'The geographical and institutional proximity of research collaboration', *Papers in Regional Science*, 86(3), pp. 423–443. doi: 10.1111/j.1435-5957.2007.00126.x.

Porter, M. E. (1998) 'Clusters and the new economics of competition.', *Harvard business review*, 76(6), pp. 77–90. doi: 10.1042/BJ20111451.

Powell, W. (1998) ‘Learning from Collaboration: Knowledge and Networksin the Biotechnology and Pharmaceutical Industries’, *Readings in Economic Sociology*, (April), pp. 262–273. doi: 10.1002/9780470755679.ch14.

Powell, W. and Grodal, S. (2005) ‘Networks of Innovators’, in *The Oxford Handbook of Innovation*, p. 78. doi: 10.1093/oxfordhb/9780199286805.003.0003.

Powell, W., Koput, K. and Smith-Doerr, L. (1996) ‘Interorganizational Collaboration and the Locus of Innovation: Networks of learning in biotechnology’, *Administrative Science Quarterly*, 41(1), pp. 116–145. doi: 10.2307/2393988.

Price, D. J. de S. (1963) *Little Science, Big Science*. New York: Columbia University Press.

Rafols, I. *et al.* (2014) ‘Big Pharma , little science ? A bibliometric perspective on Big Pharma ’ s R & D decline’, *Technological Forecasting & Social Change*, 81(February), pp. 22–38. doi: 10.1016/j.techfore.2012.06.007.

Rosenberg, N. (1990) ‘Why Do Firms Do Basic Research (with Thier Own Money)?’, *Research Policy*, 19, pp. 165–174.

Rosenfeld, S. (2005) ‘Industry Clusters : Business Choice , Policy Outcome , or Branding Strategy ?’, *Journal of New Business Ideas and Trends*, 3(2), pp. 4–13.

Rothaermel, F. T. (2001) ‘Incumbent’s advantage through exploiting complementary assets via interfirm cooperation’, *Strategic Management Journal*, 22(6–7), pp. 687–699. doi: 10.1002/smj.180.

Rothaermel, F. T. and Hess, A. M. (2007) ‘Building Dynamic Capabilities: Innovation Driven by Individual-, Firm-, and Network-Level Effects’, *Organization Science*, 18(6), pp. 898–921. doi: 10.1287/orsc.1070.0291.

Rusu, A., Kuokkanen, K. and Heier, A. (2011) ‘Current trends in the pharmaceutical industry--a case study approach.’, *European journal of pharmaceutical sciences : official journal of the European Federation for Pharmaceutical Sciences*. Elsevier B.V., 44(3), pp. 437–40. doi: 10.1016/j.ejps.2011.07.008.

Schumpeter, J. A. (1942) *Capitalism, Socialism and Democracy.* doi: 10.2307/2518575.

Simpson, P. B. and Reichman, M. (2013) ‘Opening the lead generation toolbox.’, *Nature reviews. Drug discovery*. Nature Publishing Group, 13(1), pp. 3–4. doi: 10.1038/nrd4202.

Singh, J. (2008) ‘Distributed R&D, cross-regional knowledge integration and quality of innovative output’, *Research Policy*, 37(1), pp. 77–96. doi: https://doi.org/10.1016/j.respol.2007.09.004.

Stark, D. and Castells, M. (1997) ‘The Rise of the Network Society’, *Contemporary Sociology*. Edited by M. : B. P. Malden, 26(6), p. 725. doi: 10.2307/2654643.

Stevens, A. J. *et al.* (2011) ‘The Role of Public-Sector Research in the Discovery of Drugs and Vaccines’, *New England Journal of Medicine*, 364(6), pp. 535–541. doi: 10.1056/NEJMsa1008268.

Stiglitz, J. and Wallsten, S. (1999) ‘Public-Private Technology Partnerships: Promises and Pitfalls’, *American Behavioral Scientist*, 43(1), pp. 52–73. doi: 10.1177/00027649921955155.

Stokes, D. E. (1997) ‘Pasteur’s Quadrant: Basic Science and Technological Innovation’,

Technology and Culture, 40(2), pp. 390–392. doi: 10.1109/MSPEC.1999.738233.

Teece, D. J. (1992) 'Competition, cooperation, and innovation', *Journal of Economic Behavior & Organization*, 18(1), pp. 1–25. doi: 10.1016/0167-2681(92)90050-L.

Thiers, F. A., Sinskey, A. J. and Berndt, E. R. (2008) 'Trends in the globalization of clinical trials', *Nature Reviews Drug Discovery*, 7(1), pp. 13–14. doi: 10.1038/nrd2441.

Tijssen, R. (2009) 'Internationalisation of pharmaceutical R&D: how globalised are Europe's largest multinational companies?', *Technology Analysis & Strategic Management*, 21(7), pp. 859–879. doi: 10.1080/09537320903182330.

Tijssen, R. (2012) 'Co-authored research publications and strategic analysis of public-private collaboration', *Research Evaluation*, 21(3), pp. 204–215. doi: 10.1093/reseval/rvs013.

Wilson, J. (2014) *Essentials of business research: a guide to doing your research project*. 3rd edn. Edited by S. Publications.

Yang, K. and Meho, L. I. (2006) 'Citation Analysis: A Comparison of Google Scholar, Scopus, and Web of Science'. Richard B. Hill.

Zerhouni, E. A. (2005) *Translational and Clinical Science-Time for a New Vision*, *n engl j med*. Available at: http://grants.nih.gov/grants/guide/ (Accessed: 9 March 2020).

Zucker, L. and Darby, M. (1996) 'Star scientists and institutional transformation: patterns of invention and innovation in the formation of the biotechnology industry.', *Proceedings of the National Academy of Sciences of the United States of America*, 93(23), pp. 12709–12716. doi: 10.1073/pnas.93.23.12709.

Zucker, L., Darby, M. and Armstrong, J. (1998) 'Geographically localized knowledge: Spillovers or markets?', *Economic Inquiry*, 36(1), pp. 65–86. doi: 10.1111/j.1465-7295.1998.tb01696.x.

Zucker, L. G. and Darby, M. R. (2001) 'Capturing Technological Opportunity Via Japan's Star Scientists: Evidence from Japanese Firms' Biotech Patents and Products', *Journal of Technology Transfer*, 26(1/2), pp. 37–58. doi: 10.1023/a:1007832127813.