

Article

Long-Term Collaboration Network Based on ClinicalTrials.gov Database in the Pharmaceutical Industry

Heyoung Yang * and Hyuck Jai Lee

Future Information Research Center, Korea Institute of Science and Technology Information, 66 Hoegi-go, Dongdaemun-gu, Seoul 02456, Korea; hlee@kisti.re.kr

* Correspondence: hyyang@kisti.re.kr; Tel.: +82-2-3299-6069

Received: 12 December 2017; Accepted: 23 January 2018; Published: 26 January 2018

Abstract: Increasing costs, risks, and productivity problems in the pharmaceutical industry are important recent issues in the biomedical field. Open innovation is proposed as a solution to these issues. However, little statistical analysis related to collaboration in the pharmaceutical industry has been conducted so far. Meanwhile, not many cases have analyzed the clinical trials database, even though it is the information source with the widest coverage for the pharmaceutical industry. The purpose of this study is to test the clinical trials information as a probe for observing the status of the collaboration network and open innovation in the pharmaceutical industry. This study applied the social network analysis method to clinical trials data from 1980 to 2016 in ClinicalTrials.gov. Data were divided into four time periods—1980s, 1990s, 2000s, and 2010s—and the collaboration network was constructed for each time period. The characteristic of each network was investigated. The types of agencies participating in the clinical trials were classified as a university, national institute, company, or other, and the major players in the collaboration networks were identified. This study showed some phenomena related to the pharmaceutical industry that could provide clues to policymakers about open innovation. If follow-up studies were conducted, the utilization of the clinical trial database could be further expanded, which is expected to help open innovation in the pharmaceutical industry.

Keywords: ClinicalTrials.gov; collaboration network; collaborative strategy; pharmaceutical industry; clinical trial study

1. Introduction

Due to social issues such as an improvement in quality of life and the global aging phenomenon, the need for new drug development is growing rapidly, and many pharmaceutical companies and research institutes are making efforts to develop new drugs. However, the cost of R&D investment in the pharmaceutical industry is increasing, and productivity is stagnant. For this reason, the demand for innovation is increasing. A recent study estimated that the cost of new drug development increased from 800 million United States dollars (USD) (in 2000 dollars) to 2.87 billion USD (in 2013 dollars) (DiMasi et al., 2016) [1]. Healthcare expenditure in the US also increased from 27.4 billion USD in 1960 to 2.9 trillion USD in 2013 (Rho and Kim, 2017) [2]. A key part of the cost increase is clinical trial studies; particularly, the cost of performing phase II and phase III is growing rapidly (Orloff et al., 2009) [3]. Clinical trials are not only cost intensive; they are also time-consuming in new drug development. According to Kaitin (2010) [4], it takes 7.2 years on average from the start of clinical testing to regulatory approval in the case of FDA-approved new drugs. In an extreme case, the drug is supplied to the market more than nine years after new drug-candidate development. This means that clinical trials could be a threat to the management of pharmaceutical companies, and it could remove opportunities

for the extension of life for patients. Productivity crisis is a frequent research topic in the pharmaceutical industry in this decade (Pammolli et al., 2011) [5]. Some studies supposed that the stagnancy may reflect the limitation of the current R&D model (Munos, 2009) [6]. There is a widespread recognition that pharmaceutical R&D, despite increasing R&D costs, has stagnated, and is faced with a productivity crisis. As one of the ways to overcome this productivity crisis, open innovation is getting attention in the pharmaceutical industry, and has been studied in various ways.

Open innovation has been actively studied in various perspectives and various industrial fields. Yun (2015) [7] proposed the dynamic model of an open innovation economy system as a way to overcome the growth limit of capitalism. The dynamics of an open innovation economy system explains the positive interaction among the open innovation economy, closed innovation economy, and social innovation economy. Witt (2016) [8] suggested what kind of innovation is needed for a sustainable future. Gupta et al. (2017) [9] described the mutual empowerment of corporations and communities by studying the relationship between market and social forces and social open innovation. Powell et al. (1996) [10] argued that innovation is created in networks of interorganizational collaboration, rather than in the interior of an individual company, in industrial fields where the knowledge base is complex and sources of expertise are disperse. Dougherty (2017) [11] and Leydesdorff and Ivanova (2016) [12] pointed out the importance of “interorganizational collaboration and innovation”, and proposed the concepts of a “complex innovation ecosystem” and a “triple helix model”. These suggest new frames of open innovation. DeMets and Califf (2011) [13] explained the importance of leadership in the collaboration of clinical trials from a historical point of view. The results may vary depending on who provides the leadership in the collaboration of clinical trials. They emphasized the importance of academics if the public character of the new drugs were to be maintained.

Analyzing co-authorship in journal articles or the co-assignees in patents is a general methodology for examining interorganizational collaborations from an academic or technological point of view. However, it is not unusual in the pharmaceutical industry for journal articles to not be published when technological details need to be hidden for strategic reasons, and the generic drugs do not have any related patent. For this reason, those two types of documents have severe limitations as the probe for examining interorganizational collaborations in the pharmaceutical industry. Clinical trials information, on the contrary, can show a wider range of relationships between organizations. It makes clinical trials a more appropriate source of information when examining interorganizational collaborations in the pharmaceutical industry.

Even though humans are the research subject in clinical trials, they have been considered valuable activities in which a cure and/or medical treatment is determined (Dickersin and Rennie, 2003) [14]. However, concerns of bias in clinical trials, such as excluding the unexpected or unwanted result from the report or publishing, have been made. In order to prevent this bias, and secure the objectivity of the clinical trials, a mandatory registration of the clinical trial studies, even at the early stages, has been proposed (Simes, 1986) [15]. Hence, many countries and funding agencies nowadays operate clinical trials registration systems. These can be found at the webpage of the World Health Organization (WHO) [16]. The United States (U.S.) is the first country that regulated the registrations of clinical trials through the Food and Drug Administration Modernization Act (FDAMA) [17]. Based on FDAMA National Institute of Health (NIH) developed the database ClinicaTrials.gov, where information on clinical trials to develop drugs is gathered and provided to the public. ClinicaTrials.gov, which has been in service since 2000, is the database on clinical trials with the longest history and the largest amount of data, including the trial registration date, funding sources, the primary sponsor, collaborators, conditions, interventions, the study type, study phase, location, and so on.

A clinical trials database would help us understand the current state of the pharmaceutical sector in several ways. First, in order for a new drug to be introduced into the market, it must undergo a clinical trial study beforehand. Whether it is successful or not, information on the clinical trial study should be registered to the database prior to the first patient enrollment. Therefore, the clinical trials

database is one of the information sources with the widest coverage for new drug development. Second, the clinical trials database includes information on the development of biosimilar drugs, which are replication medicines. The importance of biosimilar drugs is increasing in the pharmaceutical industry. Original drugs are protected by patents on discovery and development, and the patent holders have exclusive rights. A replication medicine is basically a medicine that has the same effect as the existing drug, but available at a lower price, so the academic importance is low. The market importance of biosimilar drugs, on the contrary, is increasing. Recently, biosimilar drugs have exceeded the market share of original drugs. Even drug makers of original drugs foster the development of biosimilar drugs. This means that, without the biosimilar drugs, the current activity of the pharmaceutical industry cannot be understood properly. Third, clinical trial study information can provide insight for improving the productivity of new drug development. Since clinical trial studies, especially phase II and phase III, are the largest parts of the cost increase in new drug development, these stages are where the greatest effort needs to be made in order to improve productivity and reduce risk. Fourth, the clinical trials database is 'an underutilized source' of information (Glass et al., 2015) [18]. For the reasons described above, the clinical trials information is a very important source of information, with a large amount of data accumulated: there are data from over 230,000 clinical trials from 195 countries in ClinicalTrials.gov. Nevertheless, it has not been adequately analyzed and examined. There are some previous studies, such as the correlation analysis between ClinicalTrials.gov and publications in biomedical journals, and an analysis on the characteristics of trends in the clinical trials database. There have been few studies on collaborations in clinical trials, despite their importance.

Social network analysis (SNA) is used in this study to examine the interorganizational collaborations that can be found in clinical trial studies. SNA was originally developed to study the relationships between people, and is most frequently used in disciplines such as sociology and anthropology. Due to the development of information communication technology, SNA is now applied to complex social systems and natural phenomena. SNA itself is forming a transdisciplinary field covering economics, organization studies, business management, public health, information science, biology, complexity, and so on (Knoke and Yang, 2008) [19]. SNA has been applied to many cases to analyze the relationships between researchers and institutions, and find out the important players within the network (Newman (2004) [20], Kretschmer (1994) [21], Wang et al. (2012) [22], Liu et al. (2017) [23], Almgren and Lee (2016) [24]). Interorganizational collaboration is a key to understanding open innovation, and thus, many studies have been carried out to understand open innovation and improve productivity (Powell et al. (1996) [10], Fritsch and Kauffeld-Monz (2010) [25]).

In a nutshell, an open innovation strategy is needed to address productivity problems in the pharmaceutical industry, where R&D costs and risks are increasing. In order to look for an open innovation strategy, it is first necessary to understand past and present collaborative relations in the pharmaceutical industry. Therefore, the research question of this study is as follows: Does the clinical trial database help us explore interorganizational collaboration and understand open innovation in the pharmaceutical industry? Information on participating organizations is extracted from the clinical trials database, to which SNA methodology was applied. Analyzing about 40 years of clinical trials information is expected to help understand the past and current collaboration status, and eventually the open innovation trend in the pharmaceutical industry.

2. Literature Review

Glass et al. (2015) [18] argued that ClinicalTrials.gov is underutilized, despite it being an important source of information for understanding the entire clinical trials. Research on ClinicalTrials.gov data has been limited to investigations of registration trends and trial attributes in clinical trial, publications, and understanding the taxonomy of ClinicalTrials.gov. In particular, ClinicalTrials.gov contains a wealth of data on partnerships and clinical trial data from over 195 countries. Glass et al. derived yearly statistical data for ClinicalTrials.gov data, and mentioned that incomplete site information was not more

than 3%, and hence ClinicalTrials.gov data could be a valuable research source. They also presented various research topics for in-depth research studies and explained the need for future studies.

Califf et al. (2012) [26] examined the characteristics of 96,346 interventional clinical study data in ClinicalTrials.gov between 2004 and 2010. Various topics, such as primary purposes, intervention types, regions, anticipated enrollments, and so on, were examined for 2004–2007 and 2007–2010. They also studied clinical trial attributes by therapeutic areas, i.e., oncology, cardiovascular, and mental health. They concluded that the clinical trials were dominated by small trials, and contained a significant heterogeneity in methodological approaches.

Roumiantseva et al. (2013) [27] studied clinical trial design characteristics according to sponsorship in ClinicalTrials.gov. They analyzed 108,315 trials registered in ClinicalTrials.gov as of 11 June 2011, and compared interventions and medical conditions by sponsor type. They found that industry-sponsored studies differed systematically from government-sponsored ones in study type, choice of interventions, and study conditions.

Thiers et al. (2008) [28] showed globalization trends by statistically classifying countries and regions where clinical trials were conducted. As expected, most clinical trials were conducted in wealthy countries. However, in recent years, clinical trials in emerging regions were also on the rise. This study also showed the global status of clinical trials. If a pharmaceutical company plans to enter the global market with new drugs, it will need to grasp the global network.

Apart from the previously mentioned studies, there are research examples that applied SNA to observe the innovation status of the firms. For example, Powell et al. (1996) [10] collected interorganizational agreements data in 1990–1994 from the Bioscan database, and analyzed the collaboration relationship between dedicated biofirms (mostly U.S. firms). They found that the density of the biotech firm network increased, and claimed that the companies tended to enlarge the interorganizational agreement in order to obtain more specific information, resources, and products. Another example is Fritsch and Kauffeld-Monz (2010) [25]. They analyzed the network structure of 300 firms and research organizations in 16 German regions. Data were gathered by postal questionnaires in 2004 asking the names of the most important partner firms and research organizations. They argued that strong ties were more effective in conveying knowledge and information than the weak ones.

According to the above literature review on previous studies, it is reasonable to consider that the clinical trials database and SNA is an appropriate combination of information source and analytical methodology to examine past and current interorganizational collaboration and open innovation in the pharmaceutical industry.

3. Data and Method

In this study, the ClinicalTrials.gov database was used to analyze the pharmaceutical network of collaborations. To the best of our knowledge, this study is the first long-term analysis of a clinical trials database. The whole dataset of ClinicalTrials.gov, containing approximately 230,000 trial studies as of January 2017, was downloaded from its Downloading Contents for Analysis page (<https://clinicaltrials.gov/ct2/resources/download>). SNA was applied to the sponsor–collaborator relationship data, which was also acquired from ClinicalTrials.gov (Figure 1). The collaboration network in each decade was constructed, and network properties and characteristic trend were examined in order to conduct a time-series analysis. The characteristics of the collaboration network of clinical trials were analyzed, and major players of a network for each decade were identified. The characteristics of the four types of sponsors of clinical trials, such as groups of national institutes, universities, companies, and others, were examined. The findings, such as the differences in their roles in clinical trials, could provide some implications for the understanding of open innovation in the pharmaceutical industry.

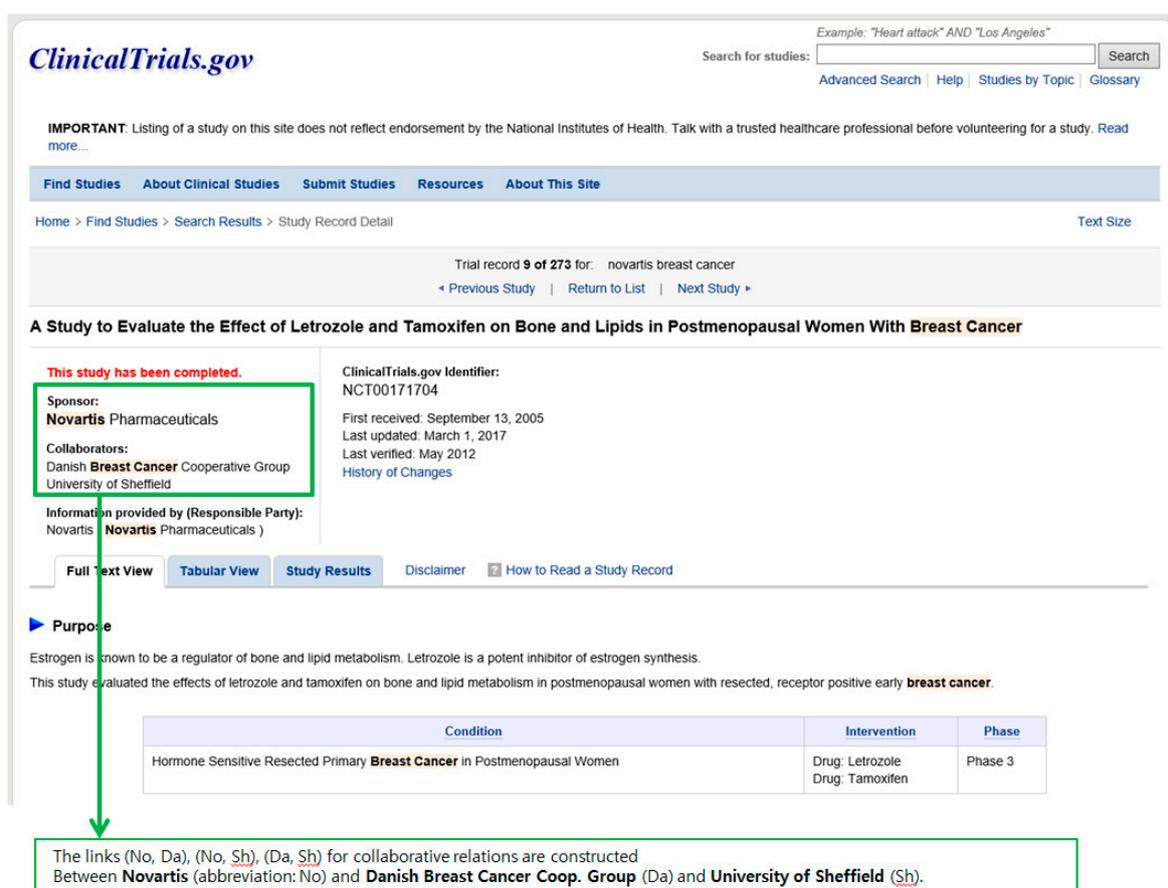


Figure 1. An example of clinical trial data and collaborative relationships between sponsors and collaborators.

Four time periods were set: 1980–1989 (1980s), 1990–1999 (1990s), 2000–2009 (2000s), and 2010–2016 (2010s), and the collaboration network in each period was constructed. Table 1 shows the basic parameters for the dataset. The number of participants for each time period increased from 181 (1980s) to 28,227 (2010s). Since the 1990s, as the numbers of participating agencies exceeds 1000, node reduction was required to carry out the network analysis. In this study, the top 77% agencies were selected based on the number of clinical trials per agency, and the number of nodes of collaboration networks for each decade increased from 222 (1990s) to 1467 (2000s) to 3087 (2010s). The number of edges (the collaborative relationships) of each network also increased drastically, from 563 (1980s) to 40,577 (2010s). Figure 2 shows the trends in the number of clinical trials, and the number of participating agencies. From Figure 2, it can be seen that the number of clinical trials increased sharply since the 2000s, and the number of participating agencies is not as high as the number of clinical trials. This shows that as the number of agencies participating in clinical trials increases, the number of clinical trials participating for each agency also increases.

Table 1. Parameters of data from ClinicalTrials.gov.

Parameters	2010s	2000s	1990s	1980s	1970s
Numbers of Clinical Trials	115,518	85,440	5963	468	96
Numbers of Agencies	28,227	16,559	1593	181	37
Nodes (77% of all agencies)	3087	1467	222	181	-
Edges (Collaborative relationships)	40,577	17,901	1023	563	-

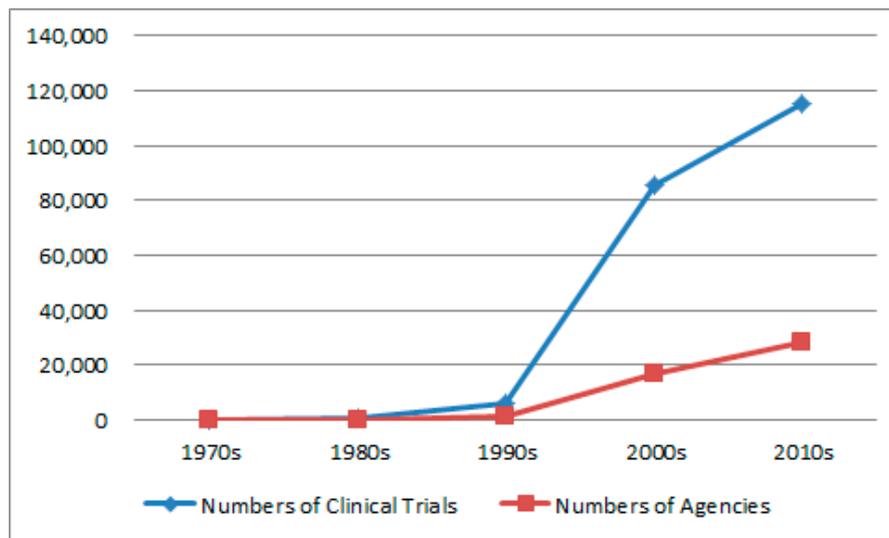


Figure 2. Trends in the numbers of clinical trials and participating agencies.

4. Results

4.1. Distributions of Number of Clinical Trials and Degrees

Figure 3 shows the distribution of the number of clinical trials per agency with the number of clinical trials for the 2010s (black circles), 2000s (red circles), 1990s (blue circles), and 1980s (green circles). It is clear that the number of clinical trials per agency increased with time in the 1980s, 1990s, and 2000s. However, it can be seen that distributions in the 2000s and 2010s significantly overlapped. Until the 2000s, the number of clinical trials greatly increased, but it can be interpreted that the number of clinical trials was maintained in the 2010s. Figure 4 shows the degree of distributions in the clinical trial collaboration networks in the four time periods. The degree is the number of links attached to a node (agency, in this case), and it could be interpreted as the number of collaborators of agencies. As with the number of clinical trials, it can be seen that the degree increases in recent time periods. Particularly from the 1990s to the 2000s, the degree of distribution increases noticeably. Though the high-degree regions in the 2000s and 2010s overlap each other, the degree in the 2010s is higher than the 2000s, as a whole. Figures 3 and 4 illustrate collaboration trends in clinical trials. By the 2000s, the agencies have increased in both the number of clinical trials and collaborators. The quantitative and cooperative expansion of clinical trials has been achieved. However, by the beginning of 2010, the number of clinical trials was stagnant, and the number of collaborators was steadily increasing. This is believed to be a correction period. It is difficult to continuously increase the number of clinical trials due to the expansion of clinical trial costs and productivity issues. Instead, they seek to improve productivity by accompanying more partnerships. This is a clue showing that companies and agencies look for chances of open innovation in the pharmaceutical industry.

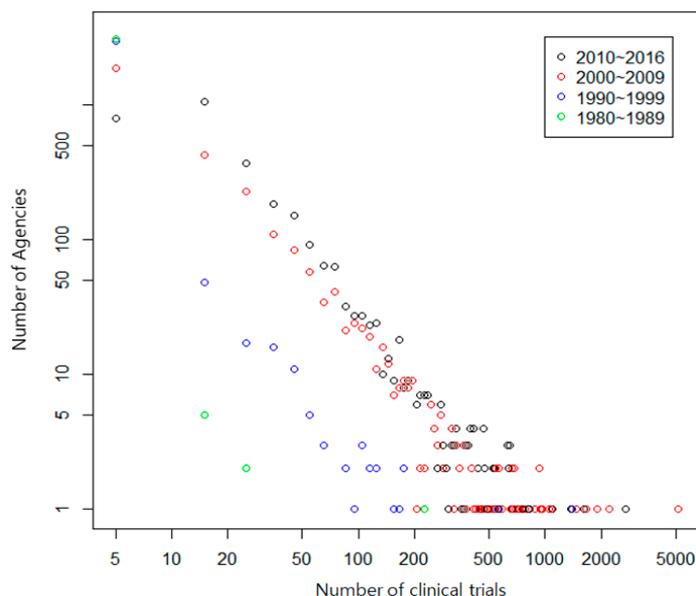


Figure 3. Distribution of the number of clinical trials per agency for each time period.

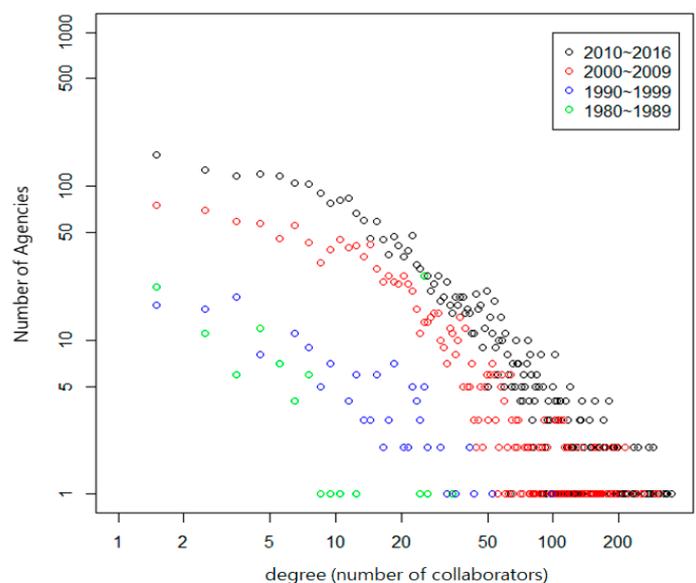


Figure 4. Distribution of the degrees (number of collaborators) per agency for each time period.

4.2. Major Players on Clinical Trial Networks

The top 20 agencies in each time period, with respect to the number of clinical trials and degrees, were extracted. Table 2 shows statistics on the top 20 agencies with respect to the number of trials. According to Table 2, it is apparent that the (global) companies lead the clinical trials from the 2000s (the numbers of companies are 10 for the 2010s and 13 for the 2000s), while national institutes did those roles in the 1980s and the 1990s (the numbers of national institutes are 11 for the 1990s and 12 for the 1980s). Table 3 shows lists of the top 20 agencies for the 2010s and the 2000s with respect to the number of clinical trials. A large number of clinical trials are carried out by companies such as GlaxoSmithKline (Brentford, UK), Pfizer (New York, NY, USA), AstraZeneca (Cambridge, UK), Merck Sharp & Dohme Corp. (Whitehouse Station, NJ, USA), Novartis Pharmaceuticals (Basel, Switzerland), and so on. National institutes such as the National Cancer Institute, the National Heart, Lung, and Blood Institute, the National Institute of Health, and the National Institute of Allergy and Infectious

Diseases are also major agencies with large numbers of clinical trials. (The National Cancer Institute is the biggest player in regard to the number of clinical trials, with 5190 trails in the 2000s, and 2705 trials in the 2010s.) However, universities are rare: the University of California, San Francisco is the only university in the top 20 rank in the 2010s, and there is no university ranked in the 2000s.

Table 2. Statistics of the top 20 agencies per type for each time period (number of clinical trials).

Type	Number of Top 20 Agencies in Number of Clinical Trials ¹			
	2010s	2000s	1990s	1980s
University	1	0	0	6
National institute	4	5	11	12
Company	10	13	1	2
Others	5	2	8	2

¹ It includes all agencies of equal rank.

Table 3. Top 20 agencies for the 2010s and the 2000s (the number of clinical trials).

Type	Agencies (Top 20 in 2010s)	Agencies (Top 20 in 2000s)
University	University of California, San Francisco, CA, USA (675)	-
National institute	National Cancer Institute, MD, USA (2705) National Heart, Lung, and Blood Institute, MD, USA (649) National Institute of Health, MD, USA (636) National Institute of Allergy and Infectious Diseases, MD, USA (629)	National Cancer Institute, MD, USA (5190) National Heart, Lung, and Blood Institute, MD, USA (1088) National Institute of Mental Health, MD, USA (962) National Institute of Allergy and Infectious Diseases, MD, USA (886) National Institute on Drug Abuse, MD, USA (697)
Company	GlaxoSmithKline, Brentford, UK (1623) Pfizer, New York, USA (1374) AstraZeneca, Cambridge, UK (1.096) Merck Sharp & Dohme Corp., NJ, USA (1088) Novartis Pharmaceuticals, Basel, Switzerland (1056) Hoffmann-La Roche, Basel, Switzerland (871) Eli Lilly and Company, Indiana, USA (718) Bayer, Leverkusen, Germany (662) Boehringer Ingelheim, Ingelheim, Germany (641) Bristol-Myers Squibb, New York, USA (639)	GlaxoSmithKline, Brentford, UK (2197) Pfizer, New York, USA (1910) Merck Sharp & Dohme Corp., NJ, USA (1656) AstraZeneca, Cambridge, UK (1480) Sanofi, Paris, France (1058) Boehringer Ingelheim, Ingelheim, Germany (960) Novartis, Basel, Switzerland (946) Eli Lilly and Company, Indiana, USA (927) Hoffmann-La Roche, Basel, Switzerland (781) Novartis Pharmaceuticals, Basel, Switzerland (772) Bayer, Leverkusen, Germany (722) Bristol-Myers Squibb, New York, USA (709) Genentech, Inc., CA, USA (689)
Others	Massachusetts General Hospital, MA, USA (829) Mayo Clinic, Minnesota, USA (812) Assistance Publique—Hôpitaux de Paris, Paris, France (760) M.D. Anderson Cancer Center, TX, USA (747) VA Office of Research and Development, DC, USA (642)	M.D. Anderson Cancer Center, TX, USA (923) VA Office of Research and Development, DC, USA (690)

Note: The numbers in parentheses are the numbers of clinical trials for each agency.

Table 4 shows statistics of the top 20 agencies with respect to degrees. Degree means the number of collaborators, so the agencies with high degrees could be interpreted as intermediators in the collaboration networks. According to Table 4, universities are the main players, which is contrary to the result in Table 2. In 2010s, 12 universities were ranked in the top 20 with high degrees, which is twice the amount of that in the 2000s (rank 6). The universities, such as Johns Hopkins University, the University of California, San Francisco, Duke University, the University of Pittsburgh, and Columbia University are the major universities for 20 years (in the 2000s and 2010s), as shown in Table 5. The Washington University School of Medicine, Stanford University, the University of Washington,

the University of Michigan, the University of Pennsylvania, and Emory University are emerging in the 2010s. Meanwhile, only two national institutes and three companies are ranked in the top 20 list.

Table 4. Statistics of top 20 agencies per type for each time period (the degrees).

Type	Number of Agencies in Top 20 Rank of Degrees ¹			
	2010s	2000s	1990s	1980s
University	12	6	11	0
National institute	2	4	7	1
Company	3	8	1	21
Others	3	2	4	0

¹ It includes all agencies of equal rank.

Table 5. Top 20 agencies for the 2010s and 2000s (the degrees).

Type	Agencies (Top 20 in the 2010s)	Agencies (Top 20 in the 2000s)
University	Johns Hopkins University (352)	
	University of California, San Francisco (337)	
	Duke University (331)	
	Washington University School of Medicine (303)	University of California, San Francisco (216)
	Columbia University (280)	Johns Hopkins University (203)
	Stanford University (278)	University of Pittsburgh (196)
	University of Washington (278)	University of Alabama at Birmingham (192)
	University of Michigan (274)	Duke University (180)
	University of Pittsburgh (265)	Columbia University (178)
	University of Pennsylvania (264)	
Emory University (260)		
Yale University (246)		
National institute	National Cancer Institute (302)	National Cancer Institute (295)
	National Institute of Health (288)	National Institute of Health (245)
		National Heart, Lung, and Blood Institute (215)
		National Institute of Mental Health (187)
Company		GlaxoSmithKline (262)
		Merck Sharp & Dohme Corp. (256)
	Merck Sharp & Dohme Corp. (311)	Novartis (216)
	GlaxoSmithKline (289)	AstraZeneca (215)
	Pfizer (288)	Eli Lilly and Company (196)
		Sanofi (187)
Others		Bristol-Myers Squibb (186)
	Massachusetts General Hospital (324)	Massachusetts General Hospital (222)
	Mayo Clinic (305)	Mayo Clinic (194)
	M.D. Anderson Cancer Center (246)	

Note: The numbers in parentheses are degrees, i.e., numbers of collaborators for each agency.

It is apparent that the number of clinical trials and the number of collaborators show different rank statistics. The global companies perform a large number of clinical trials, while the universities have many partnerships. In the clinical trial collaboration network, global companies and universities can be regarded as the main leaders and intermediators, respectively.

Some agencies seem to perform the roles of leader and intermediary at the same time: the University of California, San Francisco among the universities; the National Heart, Lung, and Blood Institute and National Cancer Institute among the national institutes; GlaxoSmithKline, Pfizer, Merck Sharp & Dohme Corp (New Jersey, USA) among the companies; and the Massachusetts General Hospital, Mayo Clinic, and M.D. Anderson Cancer Center among the others.

4.3. Maps of Clinical Trial Networks

The maps of collaboration networks were constructed using the VOSviewer application version 1.6.5, which is a software tool for constructing and visualizing bibliometric networks that was

developed by the Center for Science and Technology Studies (CWTS), Leiden University. Figures 5 and 6 show the networks of the 2000s and 2010s, respectively, with different colors per cluster (a group of agencies). The original numbers of agencies are 1467 and 3087 in the 2000s and the 2010s, respectively. When the collaboration networks are created with the original data, isolated groups with a small number of nodes (agencies) exist. The networks are regenerated with the remaining 1378 and 2857 agencies in the 2000s and 2010s, respectively, after removing the isolated groups to create the maps for the largest components, and the clusters are expressed. There are 32 groups in the 2000s and 30 groups in the 2010s. This means that the average size of the collaborative groups increased from 43 in the 2000s to more than twice as many—95—in the 2010s, which implies that partnerships have increased among agencies participating in clinical trials.

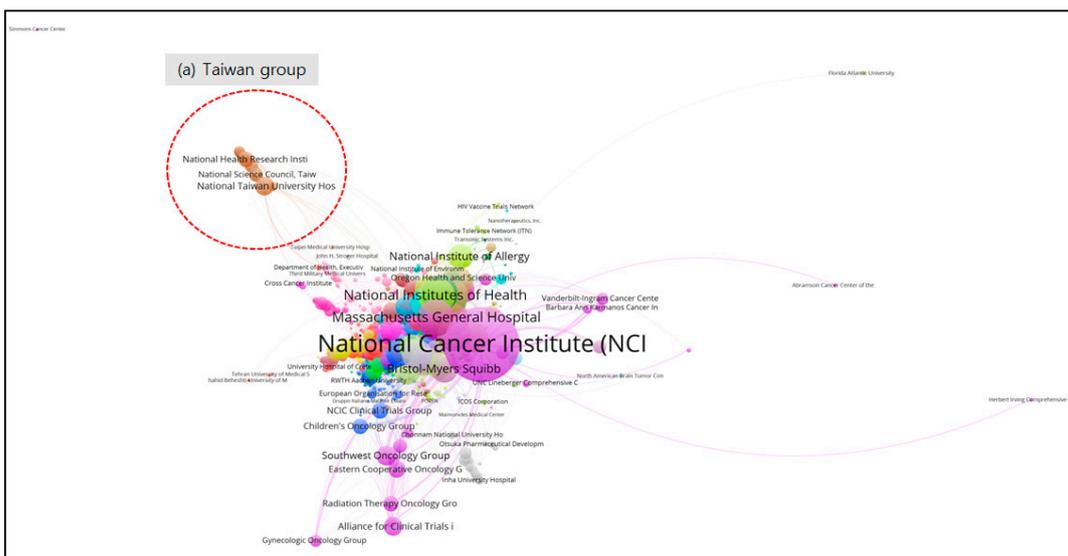


Figure 5. The map of the clinical trial collaboration network in the 2000s (the size of the node corresponds to the degrees of each institution).

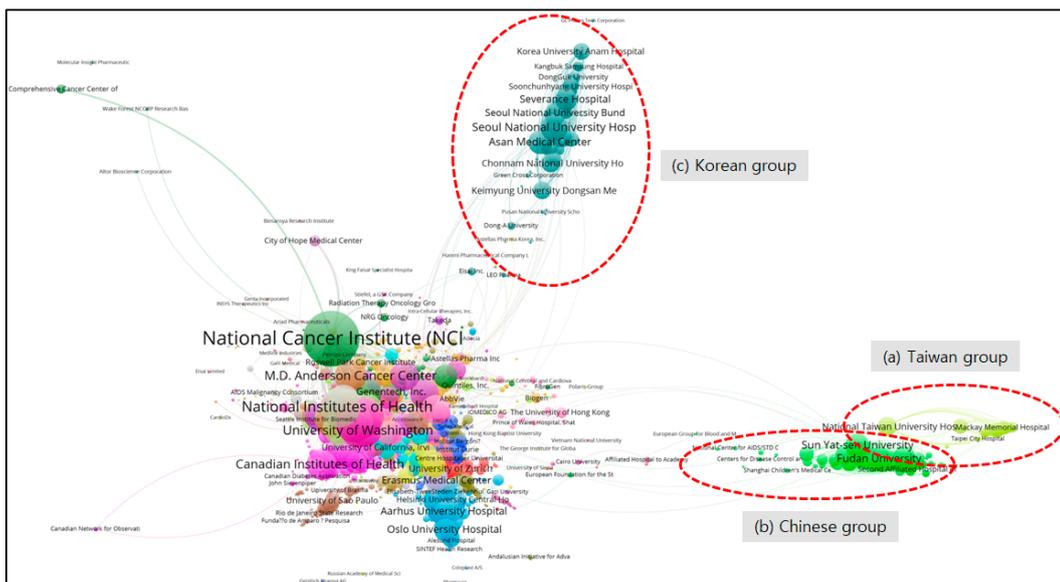


Figure 6. The map of the clinical trial collaboration network in the 2010s (the size of the node corresponds to the degree of each institution).

As shown in Figure 5, there is a large cluster component (or main cluster group), with the National Cancer Institute as the hub in the map of the 2000s, and a separate group of Taiwanese institutions located at the top-left corner of the network. In the 2010s, as shown in Figure 6, two more separate groups, apart from the Taiwan group (a), were formed, each of which comprises South Korean (c) and Chinese agencies (b). The separation from the main cluster group can be understood as a matter of closeness. In other words, the members in a separated group have a closer relationship with members in the same group than with ones in the other groups. Therefore, Taiwanese agencies are closer and have a more active relationship with each other than with foreign agencies. The Chinese and South Korean groups are close to the large cluster components of the network in the 2000s, but they form separate groups in the 2010s, as does the Taiwanese group. This could be an indication of a recent shift from international to national cooperation. It can be interpreted as the efforts of China and Korea to expand their influence on the East Asian pharmaceutical market, and seek a niche market by leading the clinical trials for East Asia region.

Table 6 shows several major agencies in Taiwanese, Chinese and South Korean groups selected by degree: the National Taiwan University Hospital and Chang Gung Memorial Hospital in the Taiwanese group; Fudan University and Sun Yat-sen University in the Chinese group; and the Seoul National University Hospital, Asian Medical Center, Inje University, and Samsung Medical Center in the Korean group.

Table 6. Major agencies in the Taiwanese, Chinese, and Korean groups (the agencies with high degrees).

Group	Major Agencies in Taiwan, China, or Korea (with High Degrees in the 2010s)	Major Agencies in Taiwan, China, or Korea (with High Degrees in the 2000s)
Taiwan	National Taiwan University Hospital Chang Gung Memorial Hospital Taipei Veterans General Hospital, Taiwan Kaohsiung Medical University	National Taiwan University Hospital Chang Gung Memorial Hospital National Health Research Institutes, Taiwan National Cheng-Kung University Hospital
Chinese	Peking Union Medical College Hospital Fudan University Chinese PLA General Hospital Sun Yat-sen University	Fudan University Sun Yat-sen University Shanghai Jiao Tong University School of Medicine Huazhong University of Science and Technology
Korean	Seoul National University Hospital Asan Medical Center Inje University Samsung Medical Center	Asan Medical Center Seoul National University Hospital Samsung Medical Center Inje University

5. Discussion

This study aims to examine whether the clinical trials database, which has not been utilized frequently, could be a probe for understanding the pharmaceutical industry. More specifically, the aim of this study was to investigate whether the clinical trials database is an appropriate information source for exploring partnerships in the pharmaceutical industry. Three explorative studies were performed to obtain a basic understanding of the clinical trials database: studies on the distribution of the number of clinical trials and degrees, the major players of the clinical trial network, and maps of the clinical trial networks. This study contributed to the understanding of a macroscopic collaboration relationship that can be constructed from the database and cover about 40 years of practices. In the first instance, it provided clues to understanding changes in clinical trials collaborative relationships by decade. The characteristics, including the major agencies as the leaders or intermediators of each decade, were identified. The maps of the clinical trials collaboration networks helped understand the overall structure of the global clinical trials and macroscopic changes. This study is different from previous ones on the clinical trials database. Studies such as Glass et al. (2015) [18], Califf et al. (2012) [26], and Thiers et al. (2008) [28], used statistical analyses to understand clinical trials database. Roumiantseva et al. (2013) [27] were interested in the

types of sponsors, rather than their collaborative relationships. The main difference between this study and existing ones is that the collaborative relationships in the pharmaceutical industry were examined using the information on participants in the clinical trials database. The results of this study show that the clinical trials database is meaningful as an information source that shows collaborations for new drug developments, so the clinical trials database could be useful as a probe for exploring open innovation in the pharmaceutical industry.

This study mainly focuses on understanding the long-term clinical trials database, and hence the strategies of specific collaborating groups were not identified, which is a limitation of the study. An overview of the change in the macroscopic strategy according to the long-term clinical collaboration network is the main implication of the study.

An analysis of the strategies of specific agencies or collaborating groups (clusters) is one of the research topics for future studies. The analysis of the moves of specific countries and emerging countries on the collaboration network is another research topic to be conducted for the study on the country-specific strategies. As was found in this study, Taiwan, China, and South Korea formed separate groups in the global collaboration network. Identification of the processes through which these countries formed separate groups as a part of a global network could help understand their strategies. The same idea can be applied to understanding the strategies of the leading global pharmaceutical firms as well. The findings obtained in this study could provide global and comprehensive references to the research topics in the future.

6. Conclusions

More than ever, there is a growing demand for healthcare and new drug development. The amount of R&D expenditure in the pharmaceutical industry is on the rise, while the productivity of new drug development is stagnant. The industry is concerned, as these issues are getting more significant. Especially, the cost and time for the clinical trial phase in the new drug development process are so large that they are becoming social issues. Much of the literature refers to open innovation as a solution to this productivity problem.

This study tests the clinical trial database as a probe for observing open innovation in the pharmaceutical industry. This study analyzed clinical trials data from the ClinicalTrials.gov database in order to observe the characteristics of partnerships, and examined whether the clinical trials data were valuable for understanding the current state of open innovation in the pharmaceutical industry. All of the data from ClinicalTrials.gov (1890–2016) was divided into four time periods—1980s, 1990s, 2000s, and 2010s—and analyzed. This study adopted social network analysis (SNA) to construct collaboration data of the clinical trials for each time period. The analysis was composed of three parts: the distributions of the number of clinical trials and degrees, the major players of the clinical trial network, and maps of the clinical trial networks.

This study of the long-term collaboration network, based on the ClinicalTrials.gov database, provides several understandings for the relationships among pharmaceutical companies, research institutes, and universities, and their mechanisms. First, the number of clinical trials per agency has stagnated since the 2000s, but the number of collaborators continues to grow. This can be interpreted as an increase in external cooperation to look for opportunities for open innovation without expanding the number of clinical trials due to the cost and risk burden of clinical trials.

Second, the main leaders carrying out a large number of clinical trials are different from the intermediators establishing many partnerships on the clinical trial collaboration network. Some of the big agencies, such as the National Cancer Institute (Bethesda, MD, USA) and the National Heart, Lung, and Blood Institute (Bethesda, MD, USA) among national institutes, the University of California, San Francisco among universities, and GlaxoSmithKline (Brentford, UK), Pfizer (New York, NY, USA), and Merck Sharp & Dohme Corp. (Whitehouse Station, NJ, USA) among companies, are performing the roles of leader and intermedicator at the same time. However, many of the top 20 agencies in regard to the number of clinical trials in the 2010s were companies, while those of in the rank of degrees

in the 2010s are universities. The interpretation is that the global big pharmaceutical companies are performing the roles of the main leaders, and the universities are the intermediators in the clinical trial collaboration network.

Third, the maps of the clinical trials collaboration network in the 2000s and in 2010s show a large cluster component and a few separate clusters. The large cluster component consists of institutions mostly, except for some of Taiwanese, Chinese, and Korean institutions. In the 2000s, the Taiwanese group was generated separate in the network. In the 2010s, the Chinese, Korean, and Taiwanese groups were generated separately in the network. This phenomenon can be a clue suggesting the different strategic moves of Taiwan, China, and South Korea from other countries.

The results of this study show that the clinical trials database from ClinicalTrials.gov could be used a probe for understanding and examining the pharmaceutical industry. The results help understand the collaborative relationships in clinical trials in the pharmaceutical industry. It provides some information about how the clinical trial collaboration network has changed, and who the major players are. The study also suggests some specific follow-up research topics, such as the collaborative strategies of clinical trials in each country. Research that analyzes the clinical trial collaboration of companies or countries will help understand their detailed strategies. If follow-up studies would be conducted, the utilization of the clinical trials database will be further expanded, and it is expected to help open innovation in the pharmaceutical industry.

Acknowledgments: The content of this paper was presented in SOITmC 2017 conference, June 2017. This study was supported by the research grant K-17-L02-C02-S03 of Korea Institute of Science and Technology Information.

Author Contributions: Heyoung Yang conceived and designed the experiments; Heyoung Yang and Hyuck Jai Lee collected, cleaned and analyzed the data; Heyoung Yang and Hyuck Jai Lee wrote the paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. DiMasi, J.A.; Grabowski, H.G.; Hansen, R.W. Innovation in the pharmaceutical industry: New estimates of R&D costs. *J. Health Econ.* **2016**, *47*, 20–33. [[PubMed](#)]
2. Roh, C.Y.; Kim, S. Medical innovation and social externality. *J. Open Innov. Technol. Mark. Complex.* **2017**. [[CrossRef](#)]
3. Orloff, J.; Douglas, F.; Pinheiro, J.; Levinson, S.; Branson, M.; Chaturvedi, P.; Ette, E.; Gallo, P.; Hirsch, G.; Mehta, C. The future of drug development: Advancing clinical trial design. *Nat. Rev. Drug Discov.* **2009**, *8*, 949–957. [[CrossRef](#)] [[PubMed](#)]
4. Kaitin, K.I. Deconstructing the drug development process: The new face of innovation. *Clin. Pharmacol. Ther.* **2010**, *87*, 356–361. [[CrossRef](#)] [[PubMed](#)]
5. Pammolli, F.; Magazzini, L.; Riccaboni, M. The productivity crisis in pharmaceutical R&D. *Nat. Rev. Drug Discov.* **2011**, *10*, 428–438. [[PubMed](#)]
6. Munos, B. Lessons from 60 years of pharmaceutical innovation. *Nat. Rev. Drug Discov.* **2009**, *8*, 959–968. [[CrossRef](#)] [[PubMed](#)]
7. Yun, J.J. How do we conquer the growth limits of capitalism? Schumpeterian Dynamics of Open Innovation. *J. Open Innov. Technol. Mark. Complex.* **2015**, *1*, 1–20. [[CrossRef](#)]
8. Witt, U. What kind of innovations do we need to secure our future? *J. Open Innov. Technol. Mark. Complex.* **2016**. [[CrossRef](#)]
9. Anil, G.; Dey, A.; Singh, G. Connecting corporations and communities: Towards a theory of social inclusive open innovation. *J. Open Innov. Technol. Mark. Complex.* **2017**. [[CrossRef](#)]
10. Powell, W.W.; Koput, K.W.; Smith-Doerr, L. Interorganizational collaboration and the locus of innovation: Networks of learning in biotechnology. *Adm. Sci. Q.* **1996**, *41*, 116–145. [[CrossRef](#)]
11. Dougherty, D. Taking advantage of emergence for complex innovation eco-systems. *J. Open Innov. Technol. Mark. Complex.* **2017**. [[CrossRef](#)]
12. Loet, L.; Ivanova, I. “Open innovation” and “triple helix” models of innovation: Can synergy in innovation systems be measured? *J. Open Innov. Technol. Mark. Complex.* **2016**. [[CrossRef](#)]

13. DeMets, D.L.; Califf, R.M. A historical perspective on clinical trials innovation and leadership: Where have the academics gone? *JAMA* **2011**, *305*, 713–714. [[CrossRef](#)] [[PubMed](#)]
14. Dickersin, K.; Rennie, D. Registering clinical trials. *JAMA* **2003**, *290*, 516–523. [[CrossRef](#)] [[PubMed](#)]
15. Simes, R.J. Publication bias: The case for an international registry of clinical trials. *J. Clin. Oncol.* **1986**, *4*, 1529–1541. [[CrossRef](#)] [[PubMed](#)]
16. World Health Organization International Clinical Trials Registry Platform. Organizations with Policies. Available online: http://www.who.int/ictrp/trial_reg/en/index2.html (accessed on 12 December 2017).
17. United States National Institutes of Health. ClinicalTrials.gov: History, Policies, and Laws. Available online: <https://clinicaltrials.gov/ct2/about-site/history> (accessed on 12 December 2017).
18. Glass, H.E.; Glass, L.M.; DiFrancesco, J.J. ClinicalTrials.gov: An underutilized source of research data about the design and conduct of commercial clinical trials. *Ther. Innov. Regul. Sci.* **2015**, *49*, 218–224. [[CrossRef](#)]
19. Knoke, D.; Yang, S. *Social Network Analysis*; SAGE Publications Inc.: Thousand Oaks, CA, USA, 2008; Volume 154.
20. Newman, M.E. Coauthorship networks and patterns of scientific collaboration. *Proc. Natl. Acad. Sci. USA* **2004**, *101* (Suppl. 1), 5200–5205. [[CrossRef](#)] [[PubMed](#)]
21. Kretschmer, H. Coauthorship networks of invisible colleges and institutionalized communities. *Scientometrics* **1994**, *30*, 363–369. [[CrossRef](#)]
22. Wang, C.C.; Huang, M.H.; Chen, D.Z. The evolution of knowledge spillover and company cluster in semiconductor industry. *J. Knowl. Econ.* **2012**, *3*, 109–124. [[CrossRef](#)]
23. Liu, F.; Zhang, N.; Cao, C. An evolutionary process of global nanotechnology collaboration: A social network analysis of patents at USPTO. *Scientometrics* **2017**, *111*, 1–17. [[CrossRef](#)]
24. Almgren, K.; Lee, J. An empirical comparison of influence measurements for social network analysis. *Soc. Netw. Anal. Min.* **2016**. [[CrossRef](#)]
25. Fritsch, M.; Kauffeld-Monz, M. The impact of network structure on knowledge transfer: An application of social network analysis in the context of regional innovation networks. *Ann. Reg. Sci.* **2010**, *44*, 21–38. [[CrossRef](#)]
26. Califf, R.M.; Zarin, D.A.; Kramer, J.M.; Sherman, R.E.; Aberle, L.H.; Tasneem, A. Characteristics of clinical trials registered in ClinicalTrials.gov, 2007–2010. *JAMA* **2012**, *307*, 1838–1847. [[CrossRef](#)] [[PubMed](#)]
27. Roumiantseva, D.; Carini, S.; Sim, I.; Wagner, T.H. Sponsorship and design characteristics of trials registered in ClinicalTrials.gov. *Contemp. Clin. Trials* **2013**, *34*, 348–355. [[CrossRef](#)] [[PubMed](#)]
28. Thiers, F.A.; Sinskey, A.J.; Berndt, E.R. Trends in the globalization of clinical trials. *Nat. Rev. Drug Discov.* **2008**, *7*, 13–14. [[CrossRef](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).