Journal of Evolutionary Economics (2021) 31:1433–1464
https://doi.org/10.1007/s00191-020-00716-1

REGULAR ARTICLE

Exploring network dynamics in science: the formation of ties to knowledge translators in clinical research

Bastian Rake[1] · Pablo D'Este[2] · Maureen McKelvey[3]

Accepted: 26 December 2020 / Published online: 3 February 2021

Abstract

From an evolutionary economics perspective, knowledge networks are self-organizing systems. Therefore, studying changes of these systems requires an understanding of how such changes are influenced by both the behaviors and characteristics of key individual actors and the network structure. We apply this perspective to a network of investigators (i.e. lead scientists) and a sample of 9543 Phase 2 cancer clinical trials during the period 2002–2012, in order to examine the structure and explore the dynamics of the clinical trial network. Using temporal exponential random graph models, we examine whether preferential attachment, multi-connectivity, or homophily drive the formation of new collaborative relations to knowledge translators - i.e. investigators with basic and clinical research knowledge. Our results suggest that despite some increased connectivity over time the network remains fragmented due to the considerably growing number of investigators in the network. This fragmentation limits opportunities for knowledge transfer to advance clinical trials. We find that homophily in research fields and investigators' country of affiliation and heterophily in terms of publication output promote the formation of ties to knowledge translators. We find also that multi-connectivity increases the probability of tie formation with knowledge translators while preferential attachment reduces this probability.

Keywords Network dynamics · Preferential attachment · Homophily · Multi-connectivity · Clinical trials · Knowledge translators

JEL classification O31 · D80

✉ Bastian Rake
bastian.rake@mu.ie

[1] School of Business, Maynooth University, Maynooth, Ireland

[2] INGENIO (CSIC-UPV), Universidad Politécnica de Valencia, Valencia, Spain

[3] Institute for Innovation and Entrepreneurship, Department of Economy and Society, University of Gothenburg, Gothenburg, Sweden

1 Introduction

Studying changes of knowledge network systems such as the scientific research system requires an understanding of how these changes are influenced by the behaviors and characteristics of both key individual actors and the structure of the network. In this Special Issue dedicated to Professor Luigi Orsenigo, we examine a specific question which is related to broad themes in his scientific contributions. The notion of network dynamics is linked closely to Professor Orsenigo's work on the relationship between individual scientists, knowledge evolution, and the structure of the pharmaceutical innovation system over time (McKelvey et al. 2004; McKelvey and Orsenigo 2006). He was very interested also in the evolution of networks per se, and the link between the decomposition of knowledge and technology and the structural evolution of industry networks (Orsenigo et al. 1998; Orsenigo et al. 2001).

This paper is related to Professor Orsenigo's research interests as it studies network dynamics in science. The increasing pace of knowledge generation and scientific advances makes it difficult for individual researchers to keep abreast with frontier knowledge even in their core interest areas (Jones 2009). Consequently, scientific research has come to be characterized as involving the formation of collaborative relations among individual researchers. These interactions have become imperative for knowledge generation and scientific advancement, with collaborative work on research projects becoming the dominant mode of generating new (scientific) knowledge (Adams 2013; Wuchty et al. 2007). At the same time, the focus and objective of research activities in many disciplines have shifted away from generation of new knowledge through basic research towards research which addresses societal challenges, which has fairly clearly defined (commercial) application, or which allows for the development of recommendations for practitioners and policy makers (D'Este et al. 2018; Tijssen 2018). However, translation of the knowledge generated in basic research into products, processes, and services which have successful applications in the marketplace has proved difficult for many industries. The science-based biotechnology and pharmaceutical industries are prominent examples of industries which have seen considerable advances in basic research. At the same time, these industries have struggled to maintain their productivity in relation to translating advances in basic bio-medical research into new successful diagnostics and treatments to patients (Pammolli et al. 2011; President's Council of Advisors on Science and Technology 2012). In the debate on policy initiatives to support the translation of basic bio-medical knowledge into clinical applications which respond to patients' needs, it is important to note that researchers on both sides - basic and clinical- have specialized knowledge and skills but follow different research paradigms (Gittelman 2016). In the context of medical research and the pharmaceutical industry, we propose that both individuals and the network structure matter for understanding further knowledge developments and industry dynamics. This notion accords with Schumpeter's idea of change. We assume that the changes in the network of collaborative relations do not emerge as the result of external forces but rather arise from within the economic system - i.e. the decisions of individuals as well as changes in the network structure (Schumpeter 1934).

In this context, we would suggest that individuals who are active basic and clinical researchers could act as knowledge translators. They are knowledge translators in the sense that they support knowledge exchanges between communities following different

research paradigms, and thereby help to advance bio-medical as well as biotechnology and pharmaceutical R&D projects. Stokes's (1997) term describing what we understand as knowledge translators is "Pasteur-type scientists"; like Louis Pasteur as representative of this group, these scientists are interested in both advancing scientific knowledge including understanding natural phenomena and applying these discoveries to practice. Consequently, we conceptualize knowledge translators as conducting basic research which is inspired by potential uses or existing bio-medical challenges. Hence, their research can be defined as providing an interface between basic research and clinical application (Tijssen 2018). Knowledge translators are important also to the success of pharmaceutical development projects since their involvement increases the probability that a project reaches the next stage in the development process (Haeussler and Assmus 2021).

In addition, in the medical research and in the biotechnology-pharmaceutical industry the network structure matters. We contend that knowledge translators do not work in isolation but instead are part of a research system that is characterized by networks of collaborative relations among different types of actors at different levels of analysis. Since knowledge translators are familiar with basic and clinical research paradigms, they are expected to play an active role in these networks and have a rather high probability of collaborating with researchers focusing on either clinical or basic research. Existing work studies these networks by focusing on their structure and composition in the context of the industry's R&D processes and development of the underlying scientific and technological bases (McKelvey and Orsenigo 2006; Orsenigo et al. 1998; Orsenigo et al. 2001).

This paper contributes to the literature by adopting a different empirical perspective on what drives the formation of collaborative ties in the context of clinical trials in pharmaceutical research. More specifically, we explore whether the mechanisms of preferential attachment, multi-connectivity, or homophily drive the formation of ties to knowledge translators in the self-assembling networks of clinical trial investigators. We consider the positions and individual characteristics of the actors in the network as potential drivers of tie formation.

In clinical trials, lead scientists are responsible and accountable for the conduct of the clinical study (Hoekman et al. 2012). However, the management of geographically dispersed clinical trials requires a considerable level of collaboration and the interdependent nature of trial-related tasks requires knowledge from basic and clinical research (Hoekman et al. 2012; Malterud 2001; Patel et al. 1999). Most investigators are trained in one specific scientific discipline and lack knowledge in other disciplines (FitzGerald 2005); in this case, cross-disciplinary collaboration with experienced investigators and professional networks provide opportunities to access the required skills and to transfer knowledge (O' Connell and Roblin 2006; Reagans and McEvily 2003; Uzzi and Lancaster 2003). We assume that similar to other scientific projects, in clinical trials the investigators self-assemble into teams of collaborating researchers: i.e. they decide autonomously whether and how they collaborate with peers (Contractor 2013) linked through collaborative relations across different teams and research projects (Mathieu et al. 2001). The collaborative relations in science emerge as the result of a self-organizing process involving individuals with common interests (Metcalfe et al. 2005) and may be dissolved or renewed based on individual decisions, interests, and

objectives, and the knowledge and expertise needs of the focal research project (Wagner and Leydesdorff 2005; Wang and Hicks 2015).

Studying the drivers of tie formation within self-assembling networks of clinical trial investigators over time is important for enhancing our understanding of how individuals collaborate in clinical research. The biotechnology and pharmaceutical industry offers much anecdotal evidence on the roles and behaviors of clinical trial investigators but empirical evidence based on large scale quantitative analyses is rather scarce. We need a better understanding of the dynamic processes underlying formation of collaborative relations in networks and their development of complex knowledge in fields such as medicine and bio-pharmaceuticals. Formation of new collaborative ties can be promoted by a range of mechanisms which influence the network structure and in turn affect the ability of individuals and entire system to create new knowledge in the future.

The paper is structured as follows. Section 2 describes the theoretical underpinnings of our study; it adopts an evolutionary economics perspective on social networks and explains the mechanisms driving endogenous network dynamics. Section 3 discusses the study context i.e. collaborative relations among investigators involved in clinical trials. Section 4 presents the data and describes the analytical approach, and Section 5 presents the results of our empirical analyses. Section 6 discusses our findings and concludes the paper.

2 Network dynamics from an evolutionary economics perspective

Consistent with Schumpeter's view and the evolutionary economics notion of change, changes to the composition and the structure of networks are based on dynamics that emerge from the system - i.e. the network itself (Witt 2008). This has implications for the present study because social networks consist of sets of nodes or actors - in our case individual clinical trial investigators - and sets of ties representing relationships between the nodes (Brass et al. 2004). Changes to the composition and structure of the network may occur with the entry or exit of individuals and due to the decisions made by the individual actors to maintain or dissolve existing relationships, or forge new relationships with individuals to whom they previously were not directly connected. The decision to forge or dissolve a tie is based on the individual's aims to benefit from the opportunities for knowledge creation and knowledge recombination offered by the network structure (Zaheer and Soda 2009).

The evolutionary view of network dynamics assumes also that the dynamics within networks are neither random processes nor independent of the initial network structure and the sequence of events that led it to change (Glückler 2007; Martin and Sunley 2006; Nelson 1995). Hence, the processes that lead to these outcomes suggest the presence of path-dependencies. More specifically, the evolutionary view implies that the current network structure depends on past events and structures. The positions of individual actors within the network are the outcomes of their positions in the past and the ongoing formation and dissolution of ties which in turn are influenced by different drivers. Individual actor characteristics play a role in influencing tie formation as such characteristics define similarities and differences among actual and potential collaborators. In addition, ties can form endogenously within the network through self-organizing processes whereby tie formation or termination is shaped by the existence

or absence of ties and the network structure associated with it (Lusher and Robins 2013). Hence, analysis of network dynamics from an evolutionary perspective requires investigation of the mechanisms associated to tie formation in relation to the network's structural properties and the dyadic similarities or dissimilarities among network actors' characteristics. Researchers from various disciplines have studied tie formation in networks, from various perspectives and in different contexts, to try to understand the underling mechanisms.

In this paper, we focus on preferential attachment, multi-connectivity and homophily which have been studied in the wider evolutionary economics literature as alternative influences on the development of collaborative relations within networks (Cantner and Rake 2014; Glückler 2007, 2010; Powell et al. 2005). All three mechanisms involve different motivations to establish new ties. Preferential attachment suggests that tie formation is based predominantly on a rich-get-richer process driven by highly connected individuals (Barabási et al. 2002). Multi-connectivity sees tie formation as based on a preference for diversifying connections (Powell et al. 2005). Homophily considers tie formation to be the result of a preference for similarity, and thus it is based on actor rather than network characteristics (Kossinets and Watts 2009; Powell et al. 2005). Collaborative relationships can be based on multiple motivations, and tie formation within a network of clinical trial investigators can be associated to multiple different mechanisms (Lusher and Robins 2013). In what follows, we discuss these three mechanisms and their influence on the network dynamics.

2.1 Preferential attachment

Preferential attachment suggests that tie formation is influenced by individuals' relative connectedness - i.e. the number of their individual existing ties within the network prior to the formation of new ties (Barabási and Albert 1999). Tie formation based on preferential attachment is based on the notion that entrants and less well-connected individuals try to form ties to well-connected and well-established individuals (Barabási et al. 2002). Tie formation based on preferential attachment shows characteristics of path dependency and the assumption that potential collaborators with higher numbers of existing ties have higher "fitness", and therefore will be better collaboration partners (Rivera et al. 2010). This path-dependency leads to the rich-get-richer phenomenon which is based on cumulative advantage where individual actors with more ties are able to increase their number further while less well-connected individuals and entrants are disadvantaged with respect to tie formation (Rivera et al. 2010). Consequently, the mechanism of preferential attachment suggests that well-connected individuals establish more new connections and more quickly compared to their less well-connected counterparts, and this increases their attraction as collaboration partners even further.

The idea of preferential attachment leading to cumulative advantage based on existing numbers of collaborative relations applies to scientific collaboration networks where a rather small number of high-status researchers can be pivotal for connecting researchers within a discipline. Researchers with numerous co-authors attract the attention of researchers entering the field who seek to collaborate with these highly connected and highly visible researchers (Moody 2004; Newman 2001a). Consistent with the general notion that the number of ties reflects the actors' "fitness", less well connected researchers entering the field will use the number of collaborative ties of

their peers as an indicator of the knowledge and the reputation of these established actors (Wagner and Leydesdorff 2005). Thus, it makes sense for them to connect to those researchers that offer the highest potential gains from collaboration in terms of access to knowledge, learning, reputation, and visibility within the academic community. At the same time, established researchers are willing to extend their collaborative relations since additional contacts allow them to increase their involvement in research projects which increases their productivity and recognition within their discipline.

The theoretical arguments related to preferential attachment as a driver for tie formation in scientific networks are supported by empirical evidence from co-authorship networks across various disciplines. Barabási et al. (2002) suggest that the dynamics of co-authorship networks in mathematics and neuro-science are driven by preferential attachment. Similarly, Jeong et al. (2003) find that the rate of new tie formation depends on the number of existing ties in the context of networks of neuroscientists and physicists. Wagner and Leydesdorff (2005) show that increased international collaboration in science is attributable to the mechanism of preferential attachment. Based on the above, we expect preferential attachment to influence the formation of ties in networks of clinical trial investigators.

2.2 Multi-connectivity

Multi-connectivity assumes that collaborative relations are formed based on a preference for diversifying connections across the network. Individuals may develop a preference for diversifying connections since multiple direct and indirect ties to other actors enhance information flow speed, and facilitate access to knowledge dispersed within the network (Powell et al. 2005). Accordingly, individuals with links to more diverse alters in the network are more likely to form new ties than actor pairs with fewer diverse indirect connections. In line with this argument, multi-connectivity promotes network closure and clustering dynamics based on the formation of additional direct and indirect ties between two individuals (Rivera et al. 2010). Tomasello et al. (2017) show that multi-connectivity promotes ties between different clusters in a sparse network structure, thereby increasing system connectivity and creating multiple opportunities for knowledge exchange. Links among different clusters increase the availability of diverse knowledge and other resources and the attractiveness of the network, and provide incentives for the creation of more ties (Powell et al. 2005).

A dense and cohesive network structure provides opportunities for the formation of additional ties within existing clusters. In cohesive networks actors are linked through multiple direct and indirect connections which improve network resilience (Powell et al. 2005). Put differently, the existence of multiple, independent paths through the network implies that if specific individual actors exit the network, the remaining actors can maintain their relationships through alternative connections (Moody and White 2003). Cohesive network structures with multiple different paths through the network support rapid transmission of knowledge (Powell et al. 2005). Multiple different paths ensure that no single individual controls the flows of information, knowledge, and other resources which further increases the attractiveness of the network (Moody and White 2003). However, this reinforcement of existing relations may be at the expense of connections to more distant actors and bridging ties to other clusters. Hence, multi-connectivity might cause network fragmentation through disconnected clusters, despite

promoting the emergence of a cohesive network structure within particular clusters (Tomasello et al. 2017).

Empirical analysis of multi-connectivity as mechanism of tie formation in social networks is scarce and inconclusive. Some works find that multi-connectivity explains ties among different types of organizations (Glückler 2010; Powell et al. 2005). However, at the individual level, Zappa (2016) and Lomi et al. (2013) find that multi-connectivity is related negatively to the probability of advice seeking ties between two individuals.

2.3 Homophily

Empirical observations of many real-world networks suggest that individual actors are less prone to connect to highly connected individuals than the mechanism of preferential attachment would predict. Homophily thus challenges the notion of preferential attachment or multi-connectivity as drivers of tie formation in that it stresses the importance of other dimensions and mechanisms. Specifically, tie formation based on homophily assumes that ties among individuals within a network are established on the basis of actor similarity.

Actors prefer ties to individuals who are similar to themselves in one or several dimensions (Boschma and Frenken 2010). Similarity-based tie formation may be due to fewer opportunities to connect to dissimilar peers within the network, or to homophilous preferences (McPherson and Smith-Lovin 1987). In the former case, Kleinbaum et al. (2013) emphasize that opportunities to associate differ based on how individuals sort, self-select, or are selected into physical and social locations which result in a tendency for homophilous exchange. Preference for homophilous ties responds to a somewhat complementary explanation which builds on the assumption that tie formation in social networks depends on whether individual actors have the time, energy, or interest in the activity. Before forming new ties, individuals assess whether a specific tie might offer greater benefits than available alternatives, and assume that ties to similar alters will offer greater benefits than ties to dissimilar alters (Kossinets and Watts 2009). These benefits may include a higher likelihood of responses to requests for advice, engagement in trust-based activities, and fewer interpersonal conflicts due to different opinions and beliefs (Lazarsfeld et al. 1954; Nebus 2006; Zucker 1986). Consequently, similarity increases the probability of tie formation (Kossinets and Watts 2009).

Empirical evidence suggests that individuals forge links based on similar socio-demographic and biographical characteristics in private and professional contexts (McPherson et al. 2001; Phillips et al. 2013; Ruef et al. 2003). Building on the findings in the wider literature on the mechanisms that drive tie formation in different social networks, we suggest that homophily in terms of investigators' knowledge base and research environment is a potential driver of tie formation. More specifically, we argue that clinical investigators will be more likely to form ties to potential collaborators working in similar fields or research environments since this perceived or actual similarity encourages knowledge exchange and use of mutual knowledge in collaborative research activities (Jha and Welch 2010). Individual investigators prefer to establish ties to similar alters since communication with similar peers is more effective and is associated to a higher potential for knowledge exchange (Jha and Welch 2010; Rogers and Bhowmik 1970). The empirical evidence supports this and suggests that

researchers with similar knowledge who are active in similar research areas are more likely to form collaborative ties (Sie et al. 2012; Zhang et al. 2018).

3 Collaborative relations of investigators in clinical trials: The role of knowledge translators

Clinical trials are an essential part of R&D processes in the bio-pharmaceutical and medical device industries. In addition to contributing to the state of bio-medical knowledge, the objective of clinical trials is to provide evidence about the safety and efficacy of new drugs, new (diagnostic) procedures, or new devices through a multi-phase process of clinical research involving human volunteers. Each clinical trial is led by either a single (principal) investigator or an investigator team. The lead scientists of the clinical trial are responsible and accountable for the conduct of the entire clinical trial. This includes in particular oversight of trial-related research activities and ensuring that the clinical trial abides by established standards of scientific and ethical integrity (Hoekman et al. 2012). Depending on the role of the trial sponsors (i.e. the organization financing the trial), investigators are responsible for the design of treatment protocols, their implementation, selection, enrollment, and treatment of subjects and collection and verification of data, data analysis, and dissemination of trial outcomes - e.g. through conference contributions and scientific publications (Davidoff et al. 2001; Leong 2007; Rasmussen 2005).

There are three reasons why collaboration among investigators in clinical trials is imperative. The first refers to the nature of the knowledge in clinical trials. In order to conduct a clinical trial, the investigators must apply and interpret bio-medical knowledge of diagnosis, treatment, and patient care that is characterized by substantial tacit knowledge (Malterud 2001; Patel et al. 1999). Despite the need for diverse knowledge, the investigators are trained predominantly in only one specific scientific discipline (FitzGerald 2005). The second reason is that collaboration provides opportunities for experienced investigators to learn from one another. Learning from peers may be required since formal training as a trial investigator is virtually non-existent. Collaboration and professional networks enable investigators to acquire the skills and knowledge needed to successfully manage the clinical trial (O' Connell and Roblin 2006). The third reason for collaboration is that many clinical trials take place at multiple sites located across several countries. This geographical dispersion requires investigators with in-depth knowledge about the differences in the pathophysiology and behaviors of subjects in these specific geographic locations, and familiarity with local regulation (Hoekman et al. 2012).

In our interpretation, the simultaneous involvement of investigators in multiple clinical trials links different collaborating groups of investigators and leads to the emergence of a network of collaborating investigators. Knowledge transfer among different sub-groups in the network offers opportunities for individuals and groups to compensate for lack of skills and knowledge in order to improve translation of bio-medical knowledge into new medications (O' Connell and Roblin 2006). From an evolutionary economics perspective, this knowledge exchange is a precondition for the emergence among researchers of a common understanding which translates into successful therapies for medical needs (Metcalfe et al. 2005). We argue that the

involvement of knowledge translators with knowledge in both basic and clinical research is particularly beneficial for the successful transfer of new drugs or medical devices to the next phase of clinical development. Knowledge translators can integrate the different basic and clinical research paradigms (Gittelman 2016). Knowledge translators are oriented to advancing fundamental understanding of disease and its bio-medical roots which is associated to hypothesis generation and testing in basic science. At the same time, knowledge translators combine their aspirations for scientific discovery with application of the knowledge generated from basic research activities in clinical practice to advance developments of medicines and devices to benefit patients.

These characteristics make knowledge translators similar to Stokes's (1997) Pasteur-type scientists. In line with these arguments, Llopis and D'Este (2016) find that balancing working in a (basic) science logic and clinical activities through contact with patients contributes positively to medical innovation. Also, following Haeussler and Assmus (2021), the involvement of investigators with knowledge in basic and clinical research is particularly beneficial for successful clinical trials of new medications or medical devices. Therefore, collaborative relations with knowledge translators in clinical trials provide opportunities for knowledge transfer and learning which are likely to be valuable to innovation performance. Based on these arguments, knowledge translators are likely to be involved in many collaborative relations based on their recognition as valuable partners for other investigators. The self-assembling nature of investigator networks, which means investigators decide with whom they will collaborate, raises questions about whether the formation of collaborative ties to knowledge translators is driven by the mechanisms described in Section 2.

4 Data and empirical approach

4.1 Data sources

4.1.1 Clinical trial data

To obtain information on the clinical trial investigators and their knowledge domains we use data from different sources. Detailed information on clinical trials conducted in the United States and 179 other countries were obtained from the ClinicalTrials.gov database, a comprehensive registry of clinical trials maintained by the US National Library of Medicine at the National Institutes of Health. Among other information it provides clinical trial details such as sponsors (organizations financing the trial), investigators' names, disease areas addressed, and facilities where the trial was conducted.

In our empirical analysis we concentrate on phase 2 clinical trials addressing the disease area of cancer, irrespective of whether a trial is led by a single investigator or an investigator team. One objective of phase 2 trials is to obtain preliminary data on a drug candidate's effectiveness by testing the drug candidate on humans affected by a specific disease and collect data on its efficacy in a specific disease setting. The collection of data based on clinical trials is an activity associated typically with clinical research. At the same time, phase 2 trials are often used to generate hypotheses about the efficacy of a drug candidate, and test these hypotheses in a clinical context. The generation and

testing of hypotheses are activities more often associated to basic research since it can involve getting information on the bio-medical roots of a specific disease as well as the drug candidate's therapeutic action mechanism (Haeussler and Rake 2017). Consequently, a combination of knowledge and expertise in clinical as well as basic research can be particularly valuable in the context of phase 2 clinical trials which provide an ideal case to study collaborations with knowledge translators.

Our sample includes 9543 phase 2 clinical trials with start dates between January 2002 and December 2012 addressing the disease area of cancer. Clinical trials in this disease area are identified by Medical Subject Headings (MeSH), a controlled vocabulary thesaurus for indexing scientific biomedical research provided by the US National Library of Medicine indicating the disease or condition addressed in each trial registered at ClinicalTrials.gov. In our empirical analysis we focus on the broad disease area of "Neoplasms" indicated by the MeSH tree number "C04".

We use the clinical trial data to build a network of collaborative relations among investigators. We study network dynamics and the formation of ties within the network over discrete points in time, by considering the start year of the clinical trials in our dataset. Start year refers to the year when testing of a new drug candidate in a specific phase 2 clinical trial was initiated, as reported in the ClinicalTrials.gov database. Based on arguments about the importance of collaborations in clinical trials outlined in Section 3, we assume a collaboration among investigators exists in a given year if they are listed as investigators in the same specific clinical trial.

4.1.2 Publication data

Scopus provided information on publication profiles of the 11,678 investigators listed in the clinical trials of our sample. Scopus is a comprehensive interdisciplinary database of academic publications that covers a wide range of scientific work on health and life sciences. We apply exact matching based on investigators' last and first names to match investigators with their publications. We considered only publications in the fields of health and life sciences. Investigators' research output is categorized based on the publishing journal using the CHI journal classification (Hamilton 2003). The CHI journal classification has been used across a wide range of publications, including very recent ones (Anckaert et al. 2020; Veugelers and Wang 2019). The CHI classification categorizes journals into four levels from applied clinical research, level 1 called "clinical observation", to basic bio-medical research, level 4. Level 2 or "clinical mix" journals publish observation and investigative studies while level 3 journals report on "clinical investigation" and focus on research-oriented articles addressing fundamental bio-medical research questions related to clinical practice (Hamilton 2003). We categorize publications level 1 and 2 journals as clinical research, and publications in level 3 and level 4 journals as basic research. This dichotomy was validated by three researchers who were asked to categorize a sample of 10 journals as basic or applied (Haeussler and Assmus 2021).[1] As outlined in Section 3, this

[1] The CHI classification of journals has been criticized due to methodological and other concerns (e.g., Tijssen 2010). Consequently, alternative approaches for classifying publications as basic or clinical research have been developed (e.g., Boyack et al. 2014). However, Anckaert et al. (2020) emphasize that different classification approaches lead to very consistent classification outcomes.

distinction is relevant in the context of our study due to the strong divide between basic and clinical research disciplines and also to address our aim to identify scientists who combine knowledge from both areas of research (FitzGerald 2005).

This classification allows us to distinguish whether a specific investigator is best characterized as a researcher engaged in basic research, clinical research, or both. Investigators are considered basic researchers if over 75% of their publications appeared in basic research journals, and clinical researchers if less than 25% of their scientific articles appeared in basic research journals. Investigators with a share of basic research publications between 25% and 75% are considered knowledge translators: i.e. engaged in both basic and clinical research (Haeussler and Assmus 2021). Investigators who publish between 25% and 75% of their research in basic research journals (and the remainder in clinical research journals) have knowledge and skills from basic and clinical research. In contrast, investigators who publish predominantly or even exclusively in basic or clinical research journals are considered as being specialized in the corresponding area. Based on this classification, 200 of the 11,678 investigators in our dataset are classified as conducting predominantly basic research. This rather low number of basic research investigators can be explained by the applied nature of clinical trials that requires interactions with patients in a clinical context that makes conducting clinical trials difficult for pure basic researchers unless they are part of a larger team. 2653 investigators in our sample are knowledge translators and 5041 are clinical investigators. We were unable to classify the remaining trial investigators either. The vast majority of these unclassified investigators (approx. 99%) could not be classified because we could not link them to individual publications records. These unclassified investigators are likely to be medical practitioners who may have substantial clinical experience but seemingly do not have substantial experience in academic research and publishing.

4.2 Descriptive analysis: Network measures

Analyzing collaborative relations among scientists is well-established in social network research (Barabási et al. 2002; Newman 2001b). In our study, we refer to the individual clinical trial investigators as actors in the network. Investigators are connected to one another through being listed as investigators in the same clinical trials. Accounting for relations between investigators through their involvement in the same clinical trials leads to undirected networks in which the relation from investigator i to j is the same as the relation from investigator j to i. Following the literature discussing alternative measures to assess the structure of social networks, we use binary, unweighted, networks to compute multiple network structural properties (Wasserman and Faust 1994). Accordingly, the degree $d(a_i)$ of an actor i within a network represents its actual direct connections to other actors (Wasserman and Faust 1994). Hence, the degree of an actor denotes the number of connections that are incident to it. Accordingly, average degree is the average number of an actor's connections to other actors in the network:

$$\bar{\mathrm{d}} = \frac{\sum_{i=1}^{n} d(a_i)}{n}$$

Wasserman and Faust (1994) describe a component as a subgraph in which a path that is a direct or indirect connection exists between all pairs of actors in the subgraph, and in which there is no path between an actor in the component and any actor that is not part of the component. Put differently, a component is the maximal connected subgraph of the network. Based on this definition, we calculate the share of investigators who are part of the largest component of the network.

The density of a network graph provides a common measure of connectivity and it is defined as the number of realized linkages among the actors in the network divided by the maximal number of possible linkages:

$$\Delta = \frac{\sum d(a_i)}{n(n-1)}$$

where n is the number of actors (investigators) in the network and $d(a_i)$ is the degree of actor i. Density can vary between 0 and 1, where 0 represents a completely unconnected network and 1 represents a completely connected network.

Average path length is another network structural property which accounts for the efficiency of information transmission within the network. Short path length tends to be associated to rapid diffusion of information through the network and less degeneration of knowledge caused by the transmission (Cowan and Jonard 2004). Average path length is defined as the average number of connections along the shortest paths between all the actors in the network:

$$L = \frac{1}{n*(n-1)} * \sum_{i \neq j} p_{ij}$$

In this equation p_{ij} is the shortest path between the nodes i and j. A path between two disconnected actors by definition is indefinite. Since this has a considerable influence on path length, when calculating the average path length, we consider only the lengths of the existing paths.

4.3 Network regression

Network data have specific properties such as presence of interdependencies between network structure and tie formation. These interdependencies challenge the core assumptions of conventional regression models. Although several attempts have been made to address this issue, empirical analyses of network data using conventional regression do not fully capture the endogenous structural effects of networks. In addition, standard errors and significance levels of conventional regression models can be misleading (Wasserman and Faust 1994). Exponential random graph models (ERGMs) are a class of social network methods that do not require the assumption of independent ties within the network. By accounting for cross-dependencies, emerging network structures, and other factors that cannot be addressed through conventional regression methods, ERGMs avoid the shortcomings of conventional regression models for analysis of tie formation (Kim et al. 2016). Hence, ERGMs allow the mechanisms driving the formation of collaborative relationships among the actors in a

network to be studied by taking account of both the endogenous processes within the network and the actor attributes (Lusher and Robins 2013).

ERGMs consider real-world networks to be the outcome of a stochastic process which led to the emergence of the observed network based on a set of hypothetical networks with similar characteristics - e.g. number of individual actors (Robins et al. 2007). The intuition behind ERGMs is that they assess which factors maximize the probability of emergence of the observed network rather than some other hypothetical network with similar characteristics. ERGMs model the effects of interest in relation to the observed network in order to find a distribution of (network) graphs in which the observed network data are central - not extreme - in the distribution. This implies that the outcome of an ERGM is not one single network but rather a probability distribution of network graphs (Robins and Lusher 2013). In ERGMs, researchers model the combination of structures comprising the observed network which allows inferences about the processes that drive tie formation within the network (Robins and Lusher 2013). Put differently, ERGMs test the statistical significance of the mechanisms that lead to tie formation and the emergence of network structures relative to random tie formation, conditional on other effects in the model (Kim et al. 2016). Thus, ERGMs can take account of both exogenous variables and endogenous dependencies (Cranmer et al. 2017).

We use the temporal exponential random graph model (TERGM) which is an extension of the ERGM that accounts for inter-temporal dependencies in networks that are observed across several time periods (Leifeld et al. 2018). TERGMs follow the basic principles of ERGMS but incorporate parameters that consider how previous realizations of the network influence the current network structure.

The ERGM for a network at time t (N^t) where N^t is an adjacency matrix in which $N_{i,j} = 1$ indicates the presence of a tie between individual i and j, can be expressed as (Leifeld et al. 2018):

$$P\left(N^t | N^{t-K}, \ldots, N^{t-1}, \theta\right) = \frac{exp\left(\theta' h\left(N^t, N^{t-1}, \ldots, N^{t-K}\right)\right)}{c\left(\theta, N^{t-K}, \ldots, N^{t-1}\right)}$$

with θ' being the transposed vector of the model coefficients, h is a vector of statistics computed on N and c is a normalizing constant. $K \in \{0, 1, \ldots, \text{T-1}\}$ accounts for networks observed at previous periods in time. The TERGM can be expressed as the joint probability of observing networks (N) between time periods $K + 1$ and T by taking the product of the probabilities of individual networks conditional on the other networks (Leifeld et al. 2018):

$$P\left(N^{K+1}, \ldots, N^T | N^1, \ldots, N^K, \theta\right) = \prod_{t=K+1}^{T} P\left(N^t | N^{t-K}, \ldots, N^{t-1}, \theta\right)$$

We employ the Markov chain Monte Carlo maximum likelihood estimation (MCMC-MLE) procedure for estimating the model parameters (Hunter et al. 2008). The MCMC-MLE delivers accurate estimates for applications that analyze networks over a rather small number of time steps (Leifeld et al. 2018). With respect to interpretation of the TERGM parameter estimates, the size and direction of an estimate provide an

indication of the frequency of the corresponding configuration - i.e. a structural characteristic indicating the presence of ties (Robins et al. 2007). A positive (negative) estimate indicates that the corresponding configuration, conditional on the other effects in the model, is more likely (less likely) than could be expected to occur by chance (Robins and Lusher 2013). Similarly, if the estimate is large and positive (negative) the corresponding configuration is observed more (less) frequently than if the estimate were zero (Robins et al. 2007). The TERM's goodness of fit is assessed by comparing the structure of the set of hypothetical networks generated by the estimation procedure to the structure of the observed networks using the edgewise shared partner, degree, and geodesic distance distributions (Hunter et al. 2008).

4.4 Network regression analysis: Variables

4.4.1 Dependent variable

Our dependent variable is a binary indicator for whether an investigator in the clinical trial network has established a new collaborative tie to a knowledge translator. More specifically, our dependent variable equals one if a new collaborative tie to a knowledge translator emerged in year *t* and zero otherwise. As we focus on the emergence of new collaborative ties to knowledge translators, our dependent variable is equal to zero in year *t* if investigator *i* has had a collaborative tie to knowledge translator *j* in *t-1*. As outlined above, connections to knowledge translators are valuable because they are expected to contribute to the success of phase 2 trials through their broad knowledge in basic and applied research. Phase 2 trials require knowledge in both areas since they focus on both knowledge generation by developing and testing hypotheses, and evaluation of the efficacy of new drug candidates, new procedures, or new devices in specific disease settings.

4.4.2 Independent and control variables

We build on the social network literature and particularly work on the endogenous dynamics of social networks to study which of the three mechanisms - preferential attachment, multi-connectivity, or homophily - drives the formation of new ties to knowledge translators.

Our measures for preferential attachment build on the assumption of a preference for new tie formation to well-connected actors. Consequently, we follow Glückler (2010) and use absolute differences in actor degree (*Degree Difference*) as an indicator of preferential attachment. This measure accounts for differences in the visibility of the investigators in the network based on their collaborative relations.

We build on the notion of presence of multiple (indirect) paths between actors in the network as an indicator of multi-connectivity (Glückler 2010). A path is a sequence of connections always between different actors which is part of a sequence connecting investigators *i* and *j*. We use the maximum number of *Disjoint Paths* between investigator *i* and investigator *j* to indicate multi-connectivity. *Disjoint Paths* refer to the maximum number of paths within the network between *i* and *j* that do not include common investigators other than the investigator where that path starts and the one

where it ends (White and Harary 2001). In other words, *Disjoint Paths* account for alternative routes through the network connecting investigators *i* and *j*.

We use the absolute difference in investigators' publications in the five years prior to the start year of focal trials that list them as investigators as a measure for homophily (*Publication Count Difference*). More specifically, this measure takes into account whether the publication output of two investigators is comparable or whether two investigators show large differences with respect to their scientific output. Hence, it is an indicator of both the investigators' taste for science and their ability to publish in scientific journals. In addition, we use a binary indicator for whether two investigators are conducting research within the same scientific field i.e. medicine, biology, or chemistry, based on the research areas of their publications. We use this indicator as an additional measure of homophily (*Same Field*). In addition, we use a binary indicator for whether two investigators are affiliated to organizations that are located in the same country (*Same Country*), and an indicator for whether two investigators are affiliated to the same type of organization (*Same Affiliation Type*). Following the classification of clinical trial sponsors in ClinicalTrials.gov we distinguish whether investigators work for companies, the National Institutes of Health, other US government organizations, or universities and other non-profit organizations. *Edges* accounts for the number of ties (i.e. non-zero values in the adjacency matrix) in the network. *Time* controls for linear time trends with respect to tie formation.

5 Empirical results

5.1 Descriptive results: Network structure

We start our empirical analysis with a description of the structure of the network of collaborative relations among lead scientists listed in all clinical trials over time. Fig. 1 shows that the share of team-led phase 2 cancer trials (i.e. phase 2 trials with at least two investigators) has increased considerably over time.[2] While only 8.67% of phase 2 trials starting in 2002 were led by an investigator team, the share of team-led trials increased considerably over time to 67.2% for phase 2 trials starting in 2012.

The increase in team-led trials is accompanied by an increase in the average number of investigators per trial (see Table 1). Phase 2 trials starting in 2002 include approx. 1.2 investigators on average while phase 2 trials starting in 2012 list 3.6 investigators on average, slightly less than the maximum of approx. 3.7 investigators per trail observed for trials starting in 2011. We observe that both the number of investigators and their connectivity have increased over time. Investigators leading phase 2 trials starting in 2002 are connected to approx. 0.82 other investigators. With collaboration becoming the rule rather than the exception, the average number of connections to other investigators increases. For phase 2 trials starting in 2011 the average number of connections individual investigators maintain to their peers reaches a maximum of approx. 6.22 and decreases slightly to 5.62 for phase 2 trials starting in 2012.

While these indicators suggest that clinical trial investigators are increasingly interconnected, it should be emphasized that the network remains rather fragmented. The

[2] As some clinical trials do not report investigator names, the shares presented in Fig. 1 do not add up to 100.

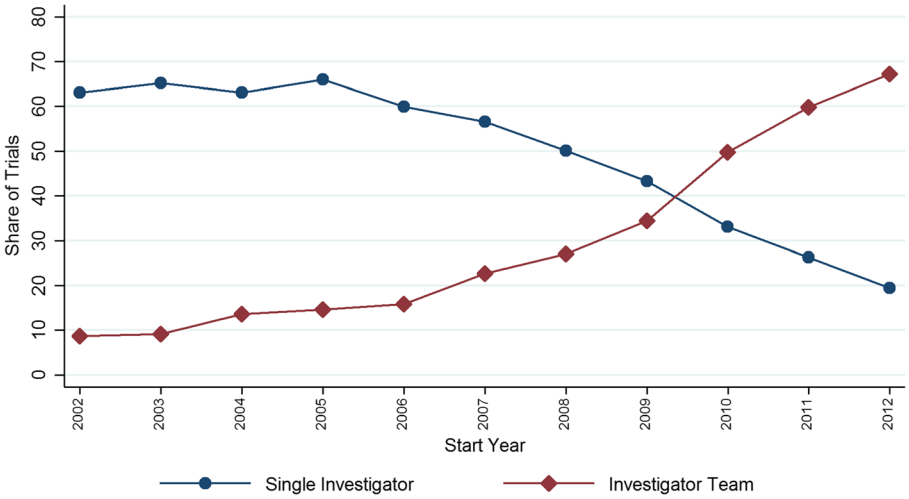


Fig. 1 Share of phase 2 trials led by single investigators and investigator teams

number of components in the network increases considerably over time from 342 in 2002 to 680 in 2007 and decreases in 2012 to 577. Given this seemingly high number of components, the investigator network shows properties different from many other real-world networks which are characterized by a single component in which a considerable share of actors is directly or indirectly linked to one another. In the clinical trial network, the largest component remains rather small. For trials starting in 2002 less than 3% of the investigators form the largest component and this share decreases even further in subsequent years. However, it is increasing from 2009 onwards and in 2012 approx. 5.4% of investigators are part of the largest component.

We illustrate the composition of largest component in Figs. 2, 3, and 4 in the appendix. Similar to the trends indicated by the network measures, these

Table 1 Measures of network structure

Start Year Trial	Number of Investigators	Average Number of Investigators per Trial	Number of Components	Share Largest Component	Density	Mean Degree	Average Path Length
2002	413	1.197	342	0.029	$1.986*10^{-3}$	0.818	1.023
2003	538	1.267	414	0.019	$2.146*10^{-3}$	1.152	1.010
2004	788	1.451	522	0.019	$2.754*10^{-3}$	2.168	1.118
2005	921	1.506	585	0.018	$2.681*10^{-3}$	2.467	1.135
2006	1029	1.511	648	0.017	$2.135*10^{-3}$	2.194	1.103
2007	1352	1.878	680	0.022	$2.634*10^{-3}$	3.559	1.309
2008	1440	2.113	638	0.015	$2.603*10^{-3}$	3.746	1.139
2009	1845	2.659	629	0.024	$2.715*10^{-3}$	5.006	1.363
2010	2309	3.330	594	0.037	$2.473*10^{-3}$	5.708	2.365
2011	2667	3.704	590	0.034	$2.332*10^{-3}$	6.217	1.974
2012	2520	3.597	577	0.054	$2.231*10^{-3}$	5.620	2.726

Figures suggest an increase in the number of investigators who are part of the largest component as well as an increasing number of connections among the members of the largest component. Despite the growth of the largest component, the network remains fragmented into many different sub-clusters (Fig. 5 in the appendix).

The fragmentation of the network has consequences for other network measures. More specifically, we find that the number of investigators and the number of new collaborative relations among these investigators grew at almost the same pace. The rather stable figures for density (number of realized relative to number of potential ties in the network) are empirical evidence of this parallel increase. Average path length is another indicator which suggests that the network remains rather fragmented over time. Average path length shows rather moderate growth from 1.02 for trials starting in 2002 to 2.73 for trials starting in 2012. Short paths length particularly prior to 2010, provides opportunities for rapid transmission of information through the network. However, in clinical trial networks this opportunity for rapid transmission may come at the cost of information being exchanged only within sub-groups or clusters, and not transmitted to the majority of investigators.

We find that the share of knowledge translators among all investigators in phase 2 clinical trials is decreasing over time. More specifically, the results in Table 2 indicate that the share of investigators engaged in both basic and clinical research decreased from approx. 33.8% for phase 2 trials starting in 2002 to approx. 22.9% for phase 2 trials starting in 2012. In contrast, we find that clinical trials are conducted increasingly by specialists in clinical research. The share of investigators involved predominantly in clinical research increases from approx. 36% in for phase 2 trials starting in 2002 to approx. 48.8% for trials starting in 2012. The share of investigators conducting mostly basic research remains quite low over our period of observation and tends to decrease. It should be noted that the absolute number of investigators increases across all three knowledge areas but that the number of clinical investigators increases faster than the number of investigators in the other two categories. In terms of connectivity to other investigators, knowledge translators show similar average degree to investigators with knowledge in only basic or clinical research.[3] However, changes in average degree of investigators classified as basic researchers fluctuate as a consequence of the rather small number of trial investigators in this group.[4] We use the Kruskal Wallis test to explore the differences in the connectivity among the three types of investigators. With an exception for the year 2012, we do not find robust statistical evidence for significant differences among the three investigator types.

[3] It should be noted that compared to clinical and basic research investigators, knowledge translators have a higher median number of publications in the 5 years preceding the start of the trial they are involved in. We see this higher number of publications as an outcome of the collaborative role played by knowledge translators. The development of the unclassified investigators' average degree is similar to the development of the other groups. This finding indicates that unclassified investigators connectivity is - on average – comparable to the connectivity of other groups.

[4] Using the median degree instead of the mean shows a similar trend of increasing connectivity within the network and across the different types of investigators.

Table 2 Investigator types and mean degree

Start Year Trial	Number of Clinical Investigators	Number of Knowledge Translators	Number of Basic Research Investigators	Share Clinical Investigators	Share Knowledge Translators	Share Basic Research Investigators	Average Degree Clinical Investigators	Average Degree Knowledge Translators	Average Degree Basic Research Investigators
2002	149	140	11	35.99	33.82	2.66	1.00	1.78	1.50
2003	208	163	11	38.59	30.24	2.04	1.68	1.22	0.33
2004	310	234	20	39.29	29.66	2.53	2.65	4.09	3.00
2005	383	261	16	41.54	28.31	1.74	2.67	2.46	3.83
2006	449	282	22	43.59	27.38	2.14	2.18	1.68	1.89
2007	520	327	22	38.43	24.17	1.63	4.21	4.06	2.18
2008	480	300	13	33.31	20.82	0.90	4.15	3.89	2.44
2009	859	455	27	46.53	24.65	1.46	5.05	4.48	5.81
2010	1110	501	41	48.05	21.69	1.77	5.62	5.29	6.37
2011	1183	621	44	44.34	23.28	1.65	6.19	5.72	6.10
2012	1231	578	43	48.83	22.93	1.71	5.66	5.04	6.02

Table 3 Results TERGM regression

	(1)	(2)	(3)	(4)
Dependent Variable: New Connections to Knowledge Translators				
Degree Difference	–0.1109 ***			–0.1115 ***
	(0.0036)			(0.0040)
Disjoint Paths		0.9153 ***		0.9074 ***
		(0.0036)		(0.0036)
Publication Count Difference			0.0014 ***	0.0016 ***
			(0.0000)	(0.0001)
Same Field			1.4860 ***	1.4609 ***
			(0.0977)	(0.1268)
Same Country			2.0720 ***	1.7628 ***
			(0.1278)	(0.1646)
Same Affiliation Type			–0.0114	0.1570
			(0.1304)	(0.1620)
Edges	–7.5029 ***	–7.9673 ***	–7.6900 ***	–7.8730 ***
	(0.0298)	(0.0344)	(0.0296)	(0.0346)
Time	–0.0107 *	–0.0436 ***	–0.0245 ***	–0.0371 ***
	(0.0048)	(0.0056)	(0.0048)	(0.0056)
N	27,993,298	27,993,298	27,993,298	27,993,298

Standard errors in parentheses

$p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

5.2 Network regression results

The results of our temporal exponential random graphs models exploring the establishment of new connections to knowledge translators are presented in Table 3. We estimate the TERGMs for a series of investigator networks for phase 2 cancer trials that started between 2005 and 2012. Data prior to 2005 is used to account for differences in the investigators' publication counts, and therefore, not included in the regression analysis. The TERGMs presented in Table 3 include all investigators listed on the clinical trials in the corresponding years. The number of observations in Table 3 refers to the number of dyads among the investigators in the networks in the corresponding years. Hence, they are not directly comparable to the number of observations in most standard regression models. The results indicate that different dimensions of the mechanisms discussed above influence the formation of new ties to knowledge translators. *Degree Difference* as an indicator of preferential attachment has a significantly negative coefficient suggesting that new and less well-connected investigators have a rather low probability to connect to knowledge translators who already have established a large number of collaborative relations. Similarly, well-connected, specialized investigators have a rather low probability of connecting to less-well connected knowledge translators. The coefficients of *Disjoint Paths* as an indicator for multi-connectivity are significantly positive. This finding indicates that investigators with multiple alternative routes to a knowledge translator are more likely to establish a new tie to this knowledge translator.

The difference in the number of scientific articles published by investigators (*Publication Count Difference*) as an indicator of homophily, increases the probability of a new tie to a knowledge translator. This indicates that heterophily rather than homophily with respect to publication output is driving tie formation to knowledge translators. In addition, we find that homophily in terms of working in the same field (*Same Field*) as well as being affiliated to organizations located in the same country (*Same Country*) increases the probability of forming new ties to knowledge translators. In contrast, the coefficient of *Same Affiliation Type* is not statistically significant.

The negative and significant coefficient of *Edges* indicates a low probability that ties to knowledge translators occur at random. The negative coefficient of *Time* suggests that the probability of building a tie to a knowledge translator decreases over time. In terms of goodness-of-fit, the simulated networks fit reasonably well to the observed networks. Appendix Fig. 5 depicts the goodness-of-fit assessment for model 4.

We analyze the robustness of our findings using alternative specifications of the dependent variable. The results of these robustness checks are presented in Table 4 in the Appendix. In model 1, the dependent variable equals one if a new collaborative tie has been formed between investigators - irrespective of whether they are knowledge translators - and zero otherwise. The results are similar to the results of our original analysis. However, the coefficient for *Same Affiliation Type* is statistically significant in the robustness check while it is not statistically significant in our original analysis. In model 2, we define knowledge translators as investigators who have at least one publication in a basic research journal and one publication in a clinical research journal. In model 3, we define knowledge translators as investigators with a share of basic research publications between 33% and 66%. In model 4, our dependent variable equals one if a collaborative tie to a knowledge translator has been newly formed in year t or repeated relative to year $t-1$, and zero otherwise. The results of these robustness checks are similar to the results of our original analysis. Model 5 in Table 4 builds upon our original analysis. In this model, we add a non-linear time trend to the original analysis. The results are qualitatively similar to the original analysis.

6 Discussion

This paper built on an evolutionary economics perspective of the structure and dynamics of social networks. In this view, social networks are self-organizing systems requiring investigation of both individuals' knowledge and the network structure. Changes in the network arise from within the system and are based on the behavior of individual actors who are likely to be influenced by current as well as past structures and events. We applied this perspective to a network of investigators (i.e. lead scientists in phase 2 cancer clinical trials) to enhance our knowledge of clinical trial investigators and their collaborative relations. We explored the structure of the investigator network and the network dynamics by studying the mechanisms driving the formation of collaborative relations to specific individuals who are knowledge translators - i.e. investigators with both basic and clinical research knowledge. We explored whether the endogenous dynamics within the investigator networks are driven by preferential attachment, multi-connectivity, or homophily. Analyzing what drives new connections to these knowledge translators is particularly valuable and relevant since phase 2 trials require knowledge in basic and clinical research.

Our descriptive analysis of phase 2 cancer trials starting between 2002 and 2012 indicates that conducting these trials is increasingly a team activity. Investigators build up a network of collaborative relations to their peers and increase the average number of connections to other investigators over time. However, we found that the network remains rather fragmented. While the number of active investigators grows considerably over time, many investigators are not connected to the largest component of the network but only to smaller sub-components. Although the fragmentation of the network provides many opportunities for tie formation, collaboration, and knowledge exchange, it may have negative consequences in the short-run. More specifically, the fragmentation of the network reduces opportunities for knowledge exchange among investigators belonging to different network sub-clusters. Consequently, we suggest that the process of developing new pharmaceuticals, devices, and procedures may be influenced negatively since fragmentation of the network makes it difficult to distribute (tacit) knowledge about trial management or the success or failure of particular trials among investigators. In extreme cases, lack of opportunities for knowledge exchange can lead to repeated use of less successful practices of trial management in different investigator clusters, or costly repetition of clinical research.

We argue that knowledge translators may be well-equipped to bring together different knowledge paradigms and may facilitate knowledge transfer. Despite the increasing absolute number of knowledge translators, we found that phase 2 clinical trials in the disease area of cancer are more often being led by investigators specialized in clinical research. This casts some doubts about the effectiveness of the efforts to incentivize translational research through dedicated research grants and other means, suggesting that the trend towards specialization may be driving choices against more interdisciplinary profiles of individual researchers. Although interdisciplinary research could increase researchers' visibility it is generally perceived as riskier, and has been shown to be associated to fewer publications and fewer successful grant applications (Bromham et al. 2016; Leahey et al. 2017). Since both factors are related to career advancement in academia, individual researchers may choose to follow a specialized research agenda focused on either clinical or basic research since they perceive this to be a less risky career path. While this choice is reasonable from an individual perspective, we suggest it may have consequences for knowledge generation and the advancement of innovation activities in the bio-medical and bio-pharmaceutical area. Earlier studies indicate that breadth of knowledge across the basic and the clinical research paradigm makes knowledge translators important for trial success (Haeussler and Assmus 2021). The increasing dominance of investigators specialized in clinical research raises questions about whether and to what extent the potential benefits of an interdisciplinary and integrated development process which follows the idea of translational medicine could be realized to successfully address global health challenges.

We took account of the particularities of network data by using temporal exponential random graph models to explore which mechanism - preferential attachment, multi-connectivity, or homophily - drives the endogenous formation of new collaborative links to knowledge translators in the network. Our results suggest that investigators are likely to form ties to knowledge translators with a similar level of connectivity. Large differences in the number of collaborative relations decrease the probability of tie formation. This contrasts with previous findings which suggest that ties among scientists are driven by preferential attachment (Barabási et al. 2002; Jeong et al. 2003;

Wagner and Leydesdorff 2005). We would suggest that investigators are aware of the difficulties involved in maintaining multiple meaningful collaborative links to other investigators. Hence, they may consider it more beneficial to form ties with similar alters as opposed to alters with already high numbers of collaborative links. A large number of already existing ties is likely to reduce the attention paid to new collaborators and in turn, reduce the potential for knowledge exchange. Collaborators with a moderate number of ties who choose to link to knowledge translators with a similar number of connections may be beneficial for both partners, as they result in more intense relationships which lead to successful realization of research ideas (Perry-Smith and Mannucci 2015). In a clinical trial context where an essential part of the process focuses on the application of previously defined research ideas according to pre-specified protocols, and the collection and analysis of data, deeper relationships are particularly beneficial. We found also that the formation of new collaborative ties to knowledge translators is influenced by multi-connectivity. Having multiple independent, indirect connections to a knowledge translator increases the probability of forming a new direct tie to this knowledge translator. Establishing multiple paths through the network ensures access to the knowledge and information flowing through the network even if some direct collaborators abandoned the network. In this case, the importance of links to knowledge translators increases; their broad knowledge base facilitates the flow through the network of basic and clinical knowledge which is required for phase 2 clinical trials.

In the case of scientific publication, we found that the formation of new ties to knowledge translators is driven by heterophily rather than homophily. Our findings indicate that investigators prefer to connect to other investigators with established publication records. We suggest that investigators' publication records may serve as a signal of their scientific and project management abilities and potential to learn from collaboration. In line with the literature, we propose that individuals connect to alters with different but complementary attributes, knowledge, skills, and capabilities (Wagner and Leydesdorff 2005; Xie et al. 2016). This preference for collaboration with dissimilar alters may be driven also by the presence of complementary interests and resources available to researchers of different status. This would suggest that individuals choose to form collaborative ties to other academics either junior or senior to themselves with corresponding lower or higher publication levels (Jha and Welch 2010). Collaborations based on heterophilous preferences in terms of scientific productivity are likely to be beneficial for both more senior and more junior investigators. Junior investigators tend to lack the resources available to more senior colleagues but are more likely to have more time available to work on (additional) research projects. Since senior researchers are likely to have stronger publications records compared to junior investigators, this argument provides support for the idea that collaborative relations among trial investigators may be driven by differences in terms of publication records. In addition, the nature of clinical research and lack of formalized training in conducting clinical trials require that junior researchers need to learn from more senior colleagues about their management. Senior investigators with publication track records should act as mentors to junior investigators inexperienced in the required research and administrative tasks (Bozeman and Corley 2004).

The probability of forming a new tie increases if two investigators publish in the same field or are affiliated to the same national organizations, which points to a preference for homophilous collaborative relations. We interpret this as suggesting that

investigators favor collaborations with knowledge translators with whom they share common language and knowledge about key scientific theories, findings, and methods. Our findings imply that both homophily and heterophily might be at work simultaneously as drivers of tie formation. This is likely to be a consequence of the multifaceted attributes of individuals: when searching for partners actors may seek complementarities in some attributes (leading to heterophily-based tie formation) and similarity in other attributes (homophily-based).

Our study contributes to work on knowledge generation by adding to our understanding of collaborative relations among investigators in clinical research and the mechanisms driving the formation of collaborative ties. More investigation is needed into the role of investigators, their mutual collaborative relations, and the mechanisms that drive the formation of these relations in biomedical research. Although some studies have addressed the importance of trial investigators for successful drug development (Haeussler and Assmus 2021; Flowers and Melmon 1997) and the diffusion of new treatments (Agha and Molitor 2018), or the effects of balancing basic research and clinical activities to medical innovation (Llopis and D'Este 2016), few empirical works examine collaborative networks that include biomedical investigators with both basic and clinical research skills and knowledge (Rake et al. 2017). This absence of empirical evidence provides opportunities for future research into the role of trial investigators and the collaborative structures in which they are embedded.

Future work could extend our analyses by using more detailed measures of investigators' research activities including their cognitive distance to their peers, their involvement in commercialization activities, and differences between star and non-star researchers. Our analysis could also be extended by studying whether collaborative relations differ among different types of investigators and which specific roles they play for making clinical research successful. In addition, future research may study which individual-level and network-level characteristics may be driving differences in collaboration behaviors. Since TERGM estimations require considerable computational power, and particularly if the network includes a large number of actors, future research could extend our analysis by considering other mechanisms which might be related to tie formation in such networks. In addition, future research using TERGMs could focus on repeated interactions among specific types of actors. Another interesting direction for future research would be to examine whether and how the involvement in clinical trials of contract research organizations (CROs) influences the type of investigator selected, and the collaborative relations among trial investigators. Future work could consider the relationship between involvement of CROs and selection of knowledge translators as trial investigators. Finally, it should be noted that clinical trials are a specific context of collaborative relations among individual researchers; therefore, our results may not be applicable to other contexts. This calls for comparable research in different research contexts and organizational settings.

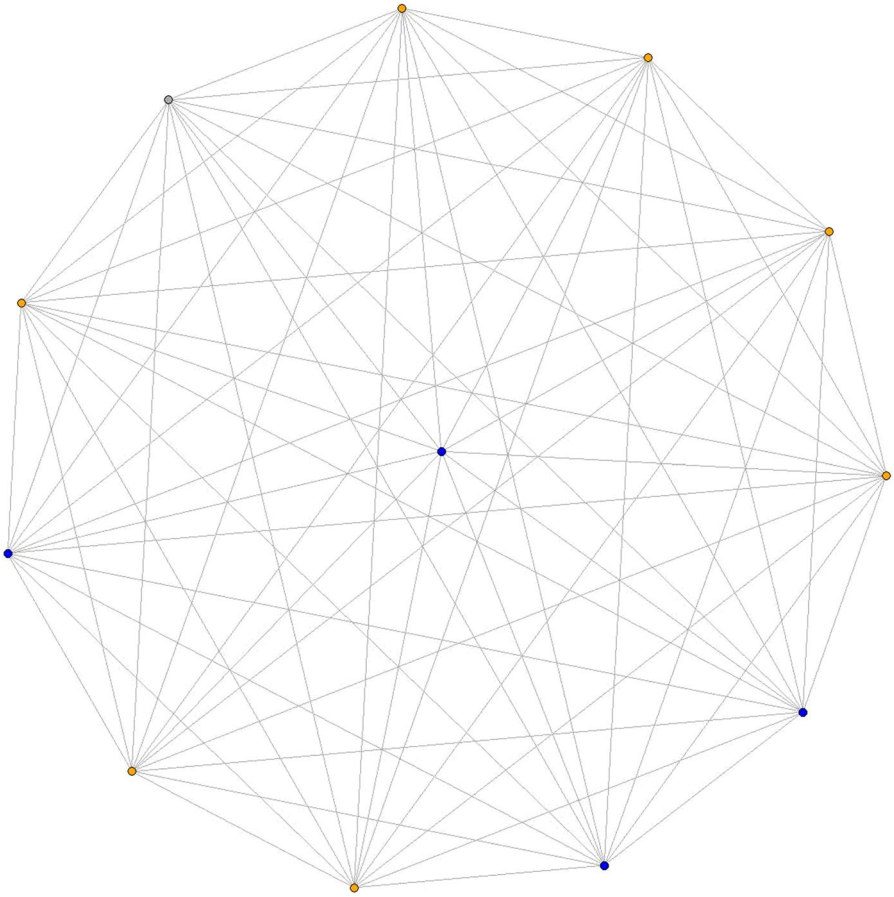
Acknowledgements This work was supported by the Swedish Research Council Distinguished Professor's Programme, Contract number 2017-03360, awarded to Professor McKelvey, for the research program "Entrepreneurial Ecosystems: Transforming society through knowledge, innovation and entrepreneurship". We thank Anne Assmus for sharing data on investigators' publication activities. We thank Mark Bagley and participants in the Workshop on Medical Innovation 2017 held in Reykjavik for discussion of the research idea, and participants in the 17th International Joseph A. Schumpeter Society Conference held in Seoul in 2018, the Workshop in Memory of Luigi Orsenigo held in Milan in December 2018 and the research seminar

held in Spring 2019 at the University of Gothenburg for stimulating discussions and valuable suggestions on earlier versions of the paper.

Compliance with ethical standards This work was supported by the Swedish Research Council Distinguished Professor's Programme, Contract number 2017–03360, awarded to Professor McKelvey, for the research program "Entrepreneurial Ecosystems: Transforming society through knowledge, innovation and entrepreneurship".

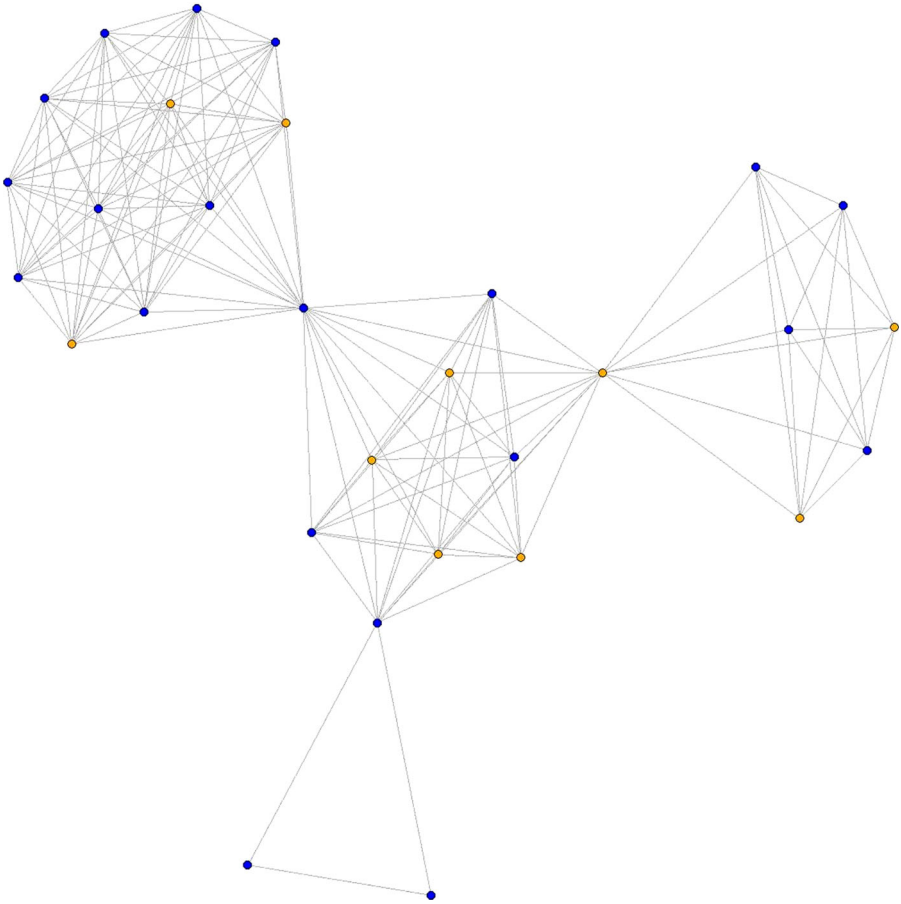
The authors declare that they have no conflicts of interest.

Appendix



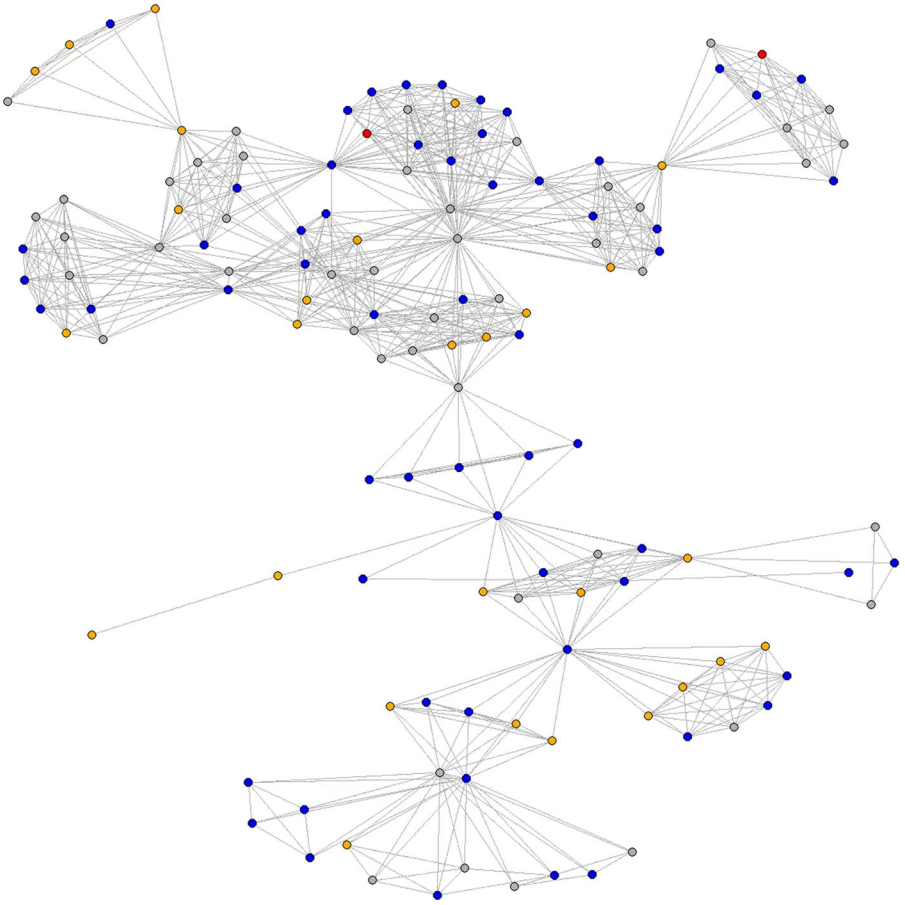
blue: clinical research investigators; orange: knowledge translators; red: basic research investigators; grey: investigators that could not be classified

Fig. 2 Visualization of the investigator network's largest component in 2002



blue: clinical research investigators; orange: knowledge translators; red: basic research investigators; grey: investigators that could not be classified

Fig. 3 Visualization of the investigator network's largest component in 2007



blue: clinical research investigators; orange: knowledge translators; red: basic research investigators; grey: investigators that could not be classified

Fig. 4 Visualization of the investigator network's largest component in 2012

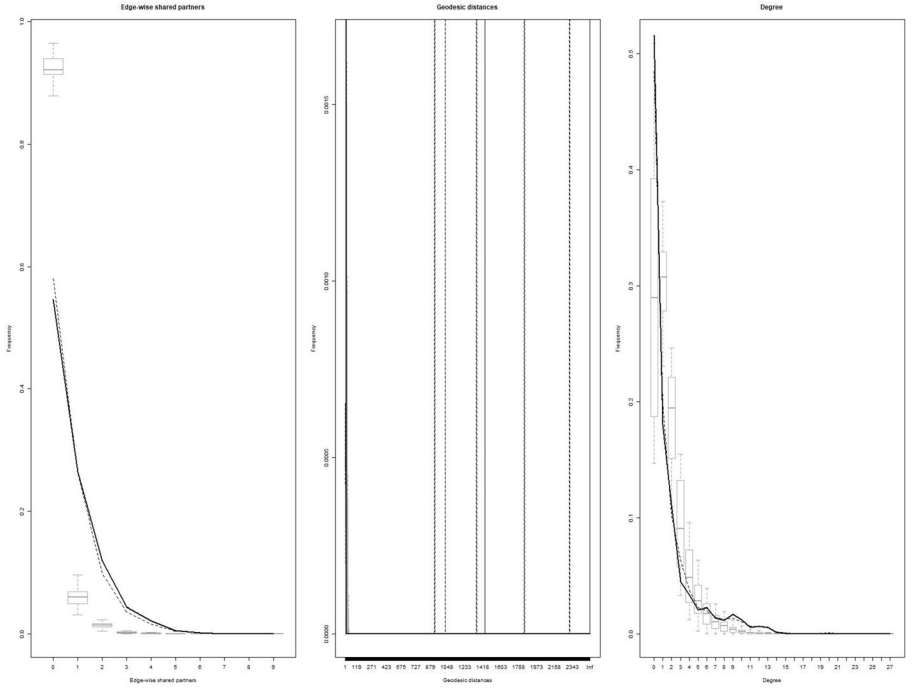


Fig. 5 Goodness-of-fit assessment of model 4 in Table 3

Table 4 TERGM regressions robustness checks

	(1)	(2)	(3)	(4)	(5)
Dependent Variable:	All New Connections	New Connections to Investigators with at least One Basic and One Clinical Publication	New Connections to Investigators Having Between 33% and 6 6% Basic Research Publications	New and Repeated Connections to Knowledge Translators	New Connections to Knowledge Translators
Degree Difference	-0.1417 ***	-0.0955 ***	-0.0943 ***	-0.0974 ***	-0.1122 ***
	(0.0031)	(0.0027)	(0.0045)	(0.0037)	(0.0040)
Disjoint Paths	2.0879 ***	0.8887 ***	0.6632 ***	0.7107 ***	0.9069 ***
	(0.0111)	(0.0028)	(0.0026)	(0.0024)	(0.0036)
Publication Count Difference	0.0014 ***	0.0019 ***	0.0011 ***	0.0016 ***	0.0016 ***
	(0.0000)	(0.0000)	(0.0001)	(0.0001)	(0.0001)
Same Field	1.5139 ***	1.4679 ***	1.8575 ***	1.6811 ***	1.4790 ***
	(0.0894)	(0.0951)	(0.1241)	(0.1113)	(0.1270)
Same Country	1.7120 ***	1.9622 ***	1.9473 ***	1.7406 ***	1.8085 ***
	(0.1133)	(0.1219)	(0.1816)	(0.1495)	(0.1681)
Same Affiliation Type	0.2764 *	0.0602	-0.2257	0.2012	0.2002
	(0.1108)	(0.1228)	(0.1878)	(0.1467)	(0.1655)
Edges	-6.9749 ***	-7.5931 ***	-8.2665 ***	-7.9722 ***	-8.2835 ***
	(0.0232)	(0.0272)	(0.0397)	(0.0334)	(0.0707)
Time	-0.0363 ***	0.0120 **	-0.0377 ***	-0.0153 **	0.1583 ***
	(0.0038)	(0.0043)	(0.0064)	(0.0054)	0.0292
$Time^2$					-0.0194 ***
					0.0028
N	27,993,298	27,993,298	27,993,298	27,993,298	27,993,298

Standard errors in parentheses

$p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

References

Adams J (2013) The fourth age of research. Nat 497:557 EP. https://doi.org/10.1038/497557a

Agha L, Molitor D (2018) The local influence of Pioneer investigators on technology adoption: evidence from new Cancer drugs. Rev Econ Stat 100:29–44. https://doi.org/10.1162/REST_a_00670

Anckaert P-E, Cassiman D, Cassiman B (2020) Fostering practice-oriented and use-inspired science in biomedical research. Res Policy 49:103900. https://doi.org/10.1016/j.respol.2019.103900

Barabási A-L, Albert R (1999) Emergence of scaling in random networks. Science 286:509. https://doi.org/10.1126/science.286.5439.509

Barabási AL, Jeong H, Néda Z, Ravasz E, Schubert A, Vicsek T (2002) Evolution of the social network of scientific collaborations. Physica A 311:590–614. https://doi.org/10.1016/S0378-4371(02)00736-7

Boschma R, Frenken K (2010) The handbook of evolutionary economic geography. In: Boschma R, Martin R (eds) Handbook of evolutionary economic geography. Edward Elgar Publishing, Cheltenham and Northhampton

Boyack KW, Patek M, Ungar LH, Yoon P, Klavans R (2014) Classification of individual articles from all of science by research level. J Informet 8:1–12. https://doi.org/10.1016/j.joi.2013.10.005

Bozeman B, Corley E (2004) Scientists' collaboration strategies: implications for scientific and technical human capital. Res Policy 33:599–616. https://doi.org/10.1016/j.respol.2004.01.008

Brass DJ, Galaskiewicz J, Greve HR, Tsai W (2004) Taking stock of networks and organizations: a multilevel perspective. Acad Manag J 47:795–817. https://doi.org/10.5465/20159624

Bromham L, Dinnage R, Hua X (2016) Interdisciplinary research has consistently lower funding success. Nat 534:684 EP. https://doi.org/10.1038/nature18315

Cantner U, Rake B (2014) International research networks in pharmaceuticals: structure and dynamics. Res Policy 43:333–348. https://doi.org/10.1016/j.respol.2013.10.016

Contractor N (2013) Some assembly required: leveraging web science to understand and enable team assembly. Philos Trans R Soc Lond A Math Phys Eng Sci 371. https://doi.org/10.1098/rsta.2012.0385

Cowan R, Jonard N (2004) Network structure and the diffusion of knowledge. J Econ Dyn Control 28:1557–1575. https://doi.org/10.1016/j.jedc.2003.04.002

Cranmer SJ, Leifeld P, McClurg SD, Rolfe M (2017) Navigating the range of statistical tools for inferential network analysis. Am J Polit Sci 61:237–251. https://doi.org/10.1111/ajps.12263

D'Este P, Ramos-Vielba I, Woolley R, Amara N (2018) How do researchers generate scientific and societal impacts? Toward an analytical and operational framework. Sci Public Policy 45:752–763. https://doi.org/10.1093/scipol/scy023

Davidoff F, DeAngelis CD, Drazen JM, Hoey J, Højgaard L, Horton R, Kotzin S, Nicholls M, Nylenna M, Overbeke JP, Sox HC, van der Weyden MB, Wilkes MS (2001) Sponsorship, authorship, and accountability. JAMA 286:1232–1234. https://doi.org/10.1001/jama.286.10.1232

FitzGerald GA (2005) Anticipating change in drug development: the emerging era of translational medicine and therapeutics. Nat Rev Drug Discov 4:815–818. https://doi.org/10.1038/nrd1849

Flowers CR, Melmon KL (1997) Clinical investigators as critical determinants in pharmaceutical innovation. Nat Med 3:136–143. https://doi.org/10.1038/nm0297-136

Gittelman M (2016) The revolution re-visited: clinical and genetics research paradigms and the productivity paradox in drug discovery. Res Policy 45:1570–1585. https://doi.org/10.1016/j.respol.2016.01.007

Glückler J (2007) Economic geography and the evolution of networks. J Econ Geogr 7:619–634. https://doi.org/10.1093/jeg/lbm023

Glückler J (2010) The evolution of a strategic Alliance network: exploring the case of stock photography. In: Boschma R, Martin R (eds) Handbook of evolutionary economic geography. Edward Elgar Publishing, Cheltenham and Northhampton, pp 298–315

Haeussler C, Assmus, A (2021) Bridging the gap between invention and innovation: Increasing success rates in publicly and industry-funded clinical trials. Res Policy, 50:1041555, https://doi.org/10.1016/j.respol.2020.104155

Haeussler C, Rake B (2017) The changing geography of clinical research: a critical analysis of its drivers. Ind Corp Chang 26:285–310

Hamilton KS (2003) Subfield and level classification of journals. CHI Research Inc CHI No 2012-R

Hoekman J, Frenken K, de Zeeuw D, Lambers Heerspink H (2012) The geographical distribution of leadership in globalized clinical trials. PLoS One 7:e45984. https://doi.org/10.1371/journal.pone.0045984

Hunter DR, Handcock MS, Butts CT, Goodreau SM, Morris M (2008) ergm: ergm: A package to fit, simulate and diagnose exponential-family models for networks. J Stat Softw. 24(3):ihpa54860. https://doi.org/10.18637/jss.v024.i03

Jeong H, Néda Z, Barabási A-L (2003) Measuring preferential attachment in evolving networks. Europhys Lett 61:567
Jha Y, Welch EW (2010) Relational mechanisms governing multifaceted collaborative behavior of academic scientists in six fields of science and engineering. Res Policy 39:1174–1184. https://doi.org/10.1016/j.respol.2010.06.003
Jones BF (2009) The burden of knowledge and the "death of the renaissance man": is innovation getting harder? Rev Econ Stud 76:283–317. https://doi.org/10.1111/j.1467-937X.2008.00531.x
Kim JY, Michael H, Emily CP, Warren B (2016) Understanding network formation in strategy research: exponential random graph models. Strateg Manag J 37:22–44. https://doi.org/10.1002/smj.2454
Kleinbaum AM, Stuart TE, Tushman ML (2013) Discretion within constraint: Homophily and structure in a formal organization. Organ Sci 24:1316–1336. https://doi.org/10.1287/orsc.1120.0804
Kossinets G, Watts DJ (2009) Origins of Homophily in an evolving social network. Am J Sociol 115:405–450. https://doi.org/10.1086/599247
Lazarsfeld PF, Merton RK, others (1954) Friendship as a social process: a substantive and methodological analysis. In: Berger MA, Abel T, Page C (eds) Freedom and control in modern society. Van Nostrand, New York, pp 18–66
Leahey E, Beckman CM, Stanko TL (2017) Prominent but less productive: the impact of Interdisciplinarity on scientists' research. Adm Sci Q 62:105–139. https://doi.org/10.1177/0001839216665364
Leifeld P, Cranmer SJ, Desmarais BA (2018) Temporal exponential random graph models with btergm: estimation and bootstrap confidence intervals. J Stat Softw 83(6). https://doi.org/10.18637/jss.v083.i06
Leong SPL (2007) The role of the principal investigator in Cancer clinical trials. In: Leong SPL (ed) Cancer clinical trials: proactive strategies, vol 132. Springer, New York, pp 157–177
Llopis O, D'Este P (2016) Beneficiary contact and innovation: the relation between contact with patients and medical innovation under different institutional logics. Res Policy 45:1512–1523. https://doi.org/10.1016/j.respol.2016.03.004
Lomi A, Lusher D, Pattison PE, Robins G (2013) The focused Organization of Advice Relations: a study in boundary crossing. Organ Sci 25:438–457. https://doi.org/10.1287/orsc.2013.0850
Lusher D, Robins G (2013) Formation of social network structure. In: Lusher D, Koskinen J, Robins G (eds) Exponential random graph models for social networks: theory, methods, and applications. Cambridge University Press, New York, pp 16–28
Malterud K (2001) The art and science of clinical knowledge: evidence beyond measures and numbers. Lancet 358:397–400. https://doi.org/10.1016/S0140-6736(01)05548-9
Martin R, Sunley P (2006) Path dependence and regional economic evolution. J Econ Geogr 6:395–437. https://doi.org/10.1093/jeg/lbl012
Mathieu J, Marks M, Zaccaro S (2001) Multiteam systems. In: Anderson N, Ones DS, Sinangil HK, Viswesvaran C (eds) Handbook of Industrial, Work & Organizational Psychology: Volume 2: Organizational Psychology. Sage, pp 289–313
McKelvey M, Orsenigo L (2006) Introduction. In: McKelvey M, Orsenigo L (eds) The economics of biotechnology. Edward Elgar, Cheltenham, UK, pp. xi–xiv
McKelvey M, Orsenigo L, Pammolli F (2004) Pharmaceuticals analyzed through the lens of a sectoral innovation syste. In: Malerba F (ed) Sectoral Systems of Innovation Sectoral Systems of innovation: concepts. Cambridge University Press, Issues and Analyses of Six Major Sectors in Europe. Cambridge et al., pp 157–198
McPherson JM, Smith-Lovin L (1987) Homophily in voluntary organizations: status distance and the composition of face-to-face groups. Am Sociol Rev 52:370–379. https://doi.org/10.2307/2095356
McPherson M, Smith-Lovin L, Cook JM (2001) Birds of a feather: Homophily in social networks. Annu Rev Sociol 27:415–444
Metcalfe JS, James A, Mina A (2005) Emergent innovation systems and the delivery of clinical services: the case of intra-ocular lenses. Res Policy 34:1283–1304
Moody J (2004) The structure of a social science collaboration network: disciplinary cohesion from 1963 to 1999. Am Sociol Rev 69:213–238. https://doi.org/10.1177/000312240406900204
Moody J, White DR (2003) Structural cohesion and Embeddedness: a hierarchical concept of social groups. Am Sociol Rev 68:103–127
Nebus J (2006) Building collegial information networks: a theory of advice network generation. Acad Manag Rev 31:615–637. https://doi.org/10.5465/amr.2006.21318921
Nelson RR (1995) Recent evolutionary theorizing about economic change. J Econ Lit 33:48–90
Newman M (2001a) Clustering and preferential attachment in growing networks. Phys Rev E 64:1–4
Newman MEJ (2001b) The structure of scientific collaboration networks. Proc Natl Acad Sci 98:404. https://doi.org/10.1073/pnas.98.2.404

O' Connell D, Roblin D (2006) Translational research in the pharmaceutical industry: from bench to bedside. Drug Discov Today 11:833–838. https://doi.org/10.1016/j.drudis.2006.07.009
Orsenigo L, Pammolli F, Riccaboni M, Bonaccorsi A, Turchetti G (1998) The evolution of knowledge and the dynamics of an industry network. J Manag Gov 1:147–175. https://doi.org/10.1023/A:1009991701667
Orsenigo L, Pammolli F, Riccaboni M (2001) Technological change and network dynamics: lessons from the pharmaceutical industry. Res Policy 30:485–508. https://doi.org/10.1016/S0048-7333(00)00094-9
Pammolli F, Magazzini L, Riccaboni M (2011) The productivity crisis in pharmaceutical R&D. Nat Rev Drug Discov 10:428 EP. https://doi.org/10.1038/nrd3405
Patel V, Arocha JF, Kaufman DR (1999) Expertise and tacit knowledge in medicine. In: Sternberg RJ (ed) Tacit knowledge in professional practice: researcher and practitioner perspectives. Erlbaum, Mahwah, pp 75–99
Perry-Smith JE, Mannucci PV (2015) From creativity to innovation: the social network drivers of the four phases of the idea journey. Acad Manag Rev 42:53–79. https://doi.org/10.5465/amr.2014.0462
Phillips N, Tracey P, Karra N (2013) Building entrepreneurial tie portfolios through strategic homophily: the role of narrative identity work in venture creation and early growth. J Bus Ventur 28:134–150. https://doi.org/10.1016/j.jbusvent.2011.12.002
Powell WW, White DR, Koput KW, Owen-Smith J (2005) Network dynamics and field evolution: the growth of Interorganizational collaboration in the life sciences. Am J Sociol 110:1132–1205
Rake B, D'Este P, McKelvey M, Assmus A (2017) How do networks affect knowledge generation for drug development? Acad Manag Proc 2017:15393. https://doi.org/10.5465/AMBPP.2017.15393abstract
Rasmussen N (2005) The drug industry and clinical research in interwar America: three types of physician collaborator. Bull Hist Med 79:50–80
Reagans R, McEvily B (2003) Network structure and knowledge transfer: the effects of cohesion and range. Adm Sci Q 48:240–267. https://doi.org/10.2307/3556658
Rivera MT, Soderstrom SB, Uzzi B (2010) Dynamics of dyads in social networks: Assortative, relational, and proximity mechanisms. Annu Rev Sociol 36:91–115. https://doi.org/10.1146/annurev.soc.34.040507.134743
Robins G, Lusher D (2013) Simplified account of an exponential random graph model as a statistical model. In: Lusher D, Koskinen J, Robins G (eds) Exponential random graph models for social networks: theory, methods, and applications. Cambridge University Press, New York, pp 29–36
Robins G, Pattison P, Kalish Y, Lusher D (2007) An introduction to exponential random graph (p*) models for social networks. Soc Networks 29:173–191. https://doi.org/10.1016/j.socnet.2006.08.002
Rogers EM, Bhowmik DK (1970) Homophily-Heterophily: relational concepts for communication research. Public Opin Q 34:523–538
Ruef M, Aldrich HE, Carter NM (2003) The structure of founding teams: Homophily, strong ties, and isolation among U.S. entrepreneurs. Am Sociol Rev 68:195–222. https://doi.org/10.2307/1519766
Schumpeter JA (1934) The theory of economic development: an inquiry into profits, capital, credit, interest, and the business cycle. Harvard University Press, Cambridge
Sie RLL, Drachsler H, Bitter-Rijpkema M, Sloep P (2012) To whom and why should I connect? Co-author recommendation based on powerful and similar peers. Int J Tech Enhanc Learn 4:121–137. https://doi.org/10.1504/IJTEL.2012.048314
Stokes DE (1997) Pasteur's quadrant: basic science and technological innovation. Brookings Institute, Washington DC
Tijssen RJW (2010) Discarding the 'basic science/applied science' dichotomy: a knowledge utilization triangle classification system of research journals. J Am Soc Inf Sci 61:1842–1852. https://doi.org/10.1002/asi.21366
Tijssen RJW (2018) Anatomy of use-inspired researchers: from Pasteur's quadrant to Pasteur's cube model. Res Policy 47:1626–1638. https://doi.org/10.1016/j.respol.2018.05.010
Tomasello MV, Napoletano M, Garas A, Schweitzer F (2017) The rise and fall of R&D networks. Ind Corp Chang 26:617–646. https://doi.org/10.1093/icc/dtw041
Uzzi B, Lancaster R (2003) Relational Embeddedness and learning: the case of Bank loan managers and their clients. Manag Sci 49:383–399. https://doi.org/10.1287/mnsc.49.4.383.14427
Veugelers R, Wang J (2019) Scientific novelty and technological impact. Res Policy 48:1362–1372. https://doi.org/10.1016/j.respol.2019.01.019
Wagner CS, Leydesdorff L (2005) Network structure, self-organization, and the growth of international collaboration in science. Res Policy 34:1608–1618
Wang J, Hicks D (2015) Scientific teams: self-assembly, fluidness, and interdependence. J Informet 9:197–207. https://doi.org/10.1016/j.joi.2014.12.006
Wasserman S, Faust K (1994) Social network analysis: methods and applications. Cambridge University Press, Cambridge et al
White DR, Harary F (2001) The cohesiveness of blocks in social networks: node connectivity and conditional density. Sociol Methodol 31:305–359. https://doi.org/10.1111/0081-1750.00098

Witt U (2008) Evolutionary economics. In: Durlauf SN, Blume LE (eds) The New Palgrave Dictionary of Economics. Palgrave Macmillan, Basingstoke

Wuchty S, Jones BF, Uzzi B (2007) The increasing dominance of teams in production of knowledge. Science 316:1036. https://doi.org/10.1126/science.1136099

Xie W-J, Li M-X, Jiang Z-Q, Tan Q-Z, Podobnik B, Zhou W-X, Stanley HE (2016) Skill complementarity enhances heterophily in collaboration networks. Sci Rep 6:18727. https://doi.org/10.1038/srep18727

Zaheer A, Soda G (2009) Network evolution: the origins of structural holes. Adm Sci Q 54:1–31. https://doi.org/10.2189/asqu.2009.54.1.1

Zappa P (2016) Toward a multilevel social exchange theory of advice relations in organizations

Zhang C, Bu Y, Ding Y, Xu J (2018) Understanding scientific collaboration: Homophily, transitivity, and preferential attachment. J Assoc Inf Sci Technol 69:72–86. https://doi.org/10.1002/asi.23916

Zucker LG (1986) Production of trust: institutional sources of economic structure, 1840–1920. Res Organ Behav 8:53–111