

RESEARCH ARTICLE

The evolution and co-evolution of a primary care cancer research network: From academic social connection to research collaboration

Debbie Vermond^{1*}, Esther de Groot¹, Valerie A. Sills², Georgios Lyratzopoulos³, Fiona M. Walter^{2,4}, Niek J. de Wit¹, Greg Rubin⁵

1 Julius Center for Health Sciences and Primary Care, Utrecht University, University Medical Center Utrecht, Utrecht, Netherlands, **2** Department of Public Health and Primary Care, The Primary Care Unit, University of Cambridge, Cambridge, United Kingdom, **3** Department of Behavioural Science and Health, UCL, London, United Kingdom, **4** Wolfson Institute of Population Health, Barts and The London School of Medicine and Dentistry Queen Mary University of London, London, United Kingdom, **5** Population Health Sciences Institute, University of Newcastle, Newcastle, United Kingdom

* D.Vermond@umcutrecht.nl



OPEN ACCESS

Citation: Vermond D, de Groot E, Sills VA, Lyratzopoulos G, Walter FM, de Wit NJ, et al. (2022) The evolution and co-evolution of a primary care cancer research network: From academic social connection to research collaboration. *PLoS ONE* 17(7): e0272255. <https://doi.org/10.1371/journal.pone.0272255>

Editor: George Vousden, Public Library of Science, UNITED KINGDOM

Received: June 22, 2021

Accepted: July 15, 2022

Published: July 29, 2022

Peer Review History: PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: <https://doi.org/10.1371/journal.pone.0272255>

Copyright: © 2022 Vermond et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: Data cannot be shared publicly to protect the privacy of our participants. In smaller scale social network

Abstract

Academic networks are expected to enhance scientific collaboration and thereby increase research outputs. However, little is known about whether and how the initial steps of getting to know other researchers translates into effective collaborations. In this paper, we investigate the evolution and co-evolution of an academic social network and a collaborative research network (using co-authorship as a proxy measure of the latter), and simultaneously examine the effect of individual researcher characteristics (e.g. gender, seniority or work-place) on their evolving relationships. We used longitudinal data from an international network in primary care cancer research: the CanTest Collaborative (CanTest). Surveys were distributed amongst CanTest researchers to map who knows who (the ‘academic social network’). Co-authorship relations were derived from Scopus (the ‘collaborative network’). Stochastic actor-oriented models were employed to investigate the evolution and co-evolution of both networks. Visualizing the development of the CanTest network revealed that researchers within CanTest get to know each other quickly and also start collaborating over time (evolution of the academic social network and collaborative network respectively). Results point to a stable and solid academic social network that is particularly encouraging towards more junior researchers; yet differing for male and female researchers (the effect of individual researcher characteristics). Moreover, although the academic social network and the research collaborations do not grow at the same pace, the benefit of creating academic social relationships to stimulate effective research collaboration is clearly demonstrated (co-evolution of both networks).

Introduction

Until recently, efforts to improve diagnostic accuracy for cancer were based on enlarging capacity in secondary care. This can lead to longer access times, higher costs and greater

analysis studies, who are the respondents, even when using anonymized data, may be obvious to potential readers. Data are available from the data management department of the Julius Center, University Medical Center Utrecht (dm_julius@umcutrecht.nl) for researchers who meet the criteria for access to confidential data.

Funding: This research arises from the CanTest Collaborative, which is funded by Cancer Research UK (<https://www.cancerresearchuk.org/>) [C8640/A23385], of which DV is a PhD student, EdG is a Postdoctoral Researcher, FW is Director (received the award), GR, GL & NdW are Associate Directors and VS is Programme Manager. The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

risks of error and delay [1,2]. Increasingly, primary care is regarded as the optimal setting to initiate health care improvements [3,4]. Timely and adequate diagnosis in primary care is vital for improving diagnostic accuracy in cancer, and therefore more research capacity focused on diagnostic testing in the primary care setting is required [5]. Scientific progress may particularly benefit from multi-disciplinary collaborations between researchers, across different research institutes and countries, as well as across the entire continuum from test development to clinical implementation. Indeed, collaboration between researchers is known to increase scientific productivity and the quality of research compared to individual research efforts [6–9].

Across academic disciplines, networks are developing to connect researchers worldwide and underpin scientific progress [10–12]. In these networks, researchers establish relationships through a variety of social-academic activities and platforms. The CanTest Collaborative (CanTest) is a clear example of such a network and serves as case study for this manuscript (<https://cantest.org>). CanTest was formally constituted in 2017 with funding from Cancer Research UK, building upon several individual collaborations between senior primary care cancer researchers in its participating centres. It comprises nice academic centres in five different countries and across three continents; individual researchers from 10 other academic centres are also involved by invitation [13]. Its main objectives are to increase capacity and sustainability of research into early detection and diagnosis of cancer—recruiting and supporting the development of a new generation of researchers to establish themselves—and to assess and evaluate approaches to improving early detection and diagnosis of cancer in primary care (the work carried out in this study will shed insight in how CanTest has addressed the first objective). By spanning disciplinary, organisational and national boundaries, academic social networks (networks of researchers connected by informal interactions and social relationships) such as CanTest can capture the social substrate of scientific productivity and promote interactions among researchers that facilitate the sharing of meaning and completion of their tasks [14]. Ultimately, through increased social support and better access to critical resources, membership and active participation in such networks is considered to enhance collaboration and scientific outputs (such as joint projects and co-authorship), as well as the individual development of more junior researchers [9,15,16].

However, academic social networks do not necessarily translate into effective collaborative research networks (networks of researchers connected by collaboration); individual characteristics of researchers in networks are considered to play a significant role in their development [17–19]. Two characteristics of researchers that have been thoroughly studied are gender and seniority [20,21]. Females are more often reported to appreciate relationships and the process of collaboration, whereas actual collaborative activity is reported to be higher for males [22–24]. Comparably, where junior researchers are more likely to increase their number of academic social relationships to gain access to new resources for collaboration, senior researchers may be more reserved in creating new relationships because they have many collaborative relationships already [25–27]. Yet, studies that link academic social relationships and collaboration—to explore the effect of gender and seniority in academic social relationships and collaboration in more detail—are lacking.

Another factor that may play a prominent role in explaining the development of networks is individual network positions [28,29]. Along with the individual characteristics of researchers, their positions in relation to each other may steer their relationships. In research on academic collaboration, the logic of network embeddedness (friends of friends tending to become friends) and preferential attachment (individuals seeking out relationships preferentially with others who are popular already) has demonstrated that researchers tend to connect with the connections-of-their-connections and with well-connected researchers [30–33]. Yet, again,

while existing studies looking at the impact of network positions on collaboration have provided a variety of important insights, we still have limited understanding of how individual positions and individual characteristics relate to each other while conditioning academic relationships and collaboration.

Although there is increased financial support for academic networking, there is little published evidence that this type of research collaboration actually accelerates research output [34,35]. Previous research studies have explored separately the development of academic relationships and collaboration (the separate ‘evolution’ of two networks). Yet, although academic interaction is considered an important factor in collaboration, we still have limited understanding of how the expansion of academic social networks translates into effective research collaboration, i.e. to what extent one leads to the other (the ‘co-evolution’ of two networks) [36,37]. Fig 1 visualizes both processes, with the evolution aspect depicted on the vertical axis (solid box) and co-evolution on the horizontal axis (outline box). Evolution is about how a network changes over time, whereas co-evolution describes how changes in one network impact the other network (i.e. how academic social relationships influences collaboration). The objectives of this study are to increase knowledge on (1) the evolution of academic social networks and collaborative networks, (2) the co-evolution of academic social networks and collaborative networks, and (3) the influence of individual researcher characteristics and their network positions on evolution and co-evolution.

Methods

Context

To assess the transition of academic social networks into effective research collaborations, we collected longitudinal network data from an ongoing, international network in primary care cancer research—CanTest. This international research collaborative facilitates international collaboration in primary care cancer research through (1) promoting joint research, (2)

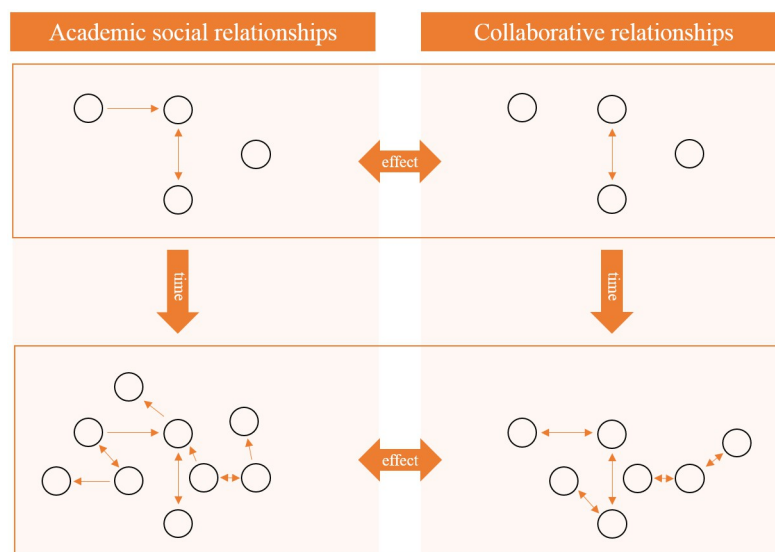


Fig 1. Evolution and co-evolution of the academic social network and collaborative network. The vertical axis shows the evolution of the academic social network and the collaborative network, with two different snapshots in time (top and bottom) for the purpose of illustration. The horizontal axis shows the co-evolution of both networks.

<https://doi.org/10.1371/journal.pone.0272255.g001>

providing various training opportunities and (3) boosting academic exchange. CanTest organizes numerous networking events, the most notable of which is the annual CanTest International School. During this week-long residential activity, researchers in the network are brought together to exchange knowledge and experiences. Two of these International Schools were held in 2018 and 2019. Due to the impact of the coronavirus epidemic, virtual networking events took the place of the 2020 School, and a virtual School was held in 2021. While some network interactions are formally organized, other more spontaneous forms of interaction are also encouraged [38].

Data

Data on the academic social relationships between all 60 researchers in the CanTest network at the time of this study were collected through online surveys at three points in time:

- T0—Point of entry to CanTest (April 2017/2018)—before the first International School
- T1—April 2019—just before the second International School
- T2—June 2019—just after the second International School.

Informed consent was obtained at the start of each survey (S1 Appendix). As is customary in network research, one matrix question, or roster, explored the relationships between researchers in the CanTest network ('academic family') at these three points in time: "Which other members of the CanTest family do you know professionally, and how did/do you connect and interact with them?" [39]. From this roster of all researchers in the network, researchers were asked to identify other researchers with whom they were connected. For each of the 60 researchers in the network, they selected either 'yes' or 'no' for any of four possible ways of being connected: (1) exchanged emails or WhatsApp messages, (2) met at a conference, workshop, meeting or training event, (3) involved in the same individual project, and (4) working at the same institution or department (S1 Appendix). For the analysis, it was necessary to aggregate the results; we did so by recoding any number of ways of being connected (either by 1, 2, 3 or 4 items) as 1 = "yes" and recoding the absence of any connection as 0 = "no". The network was directed so a distinction was made between outgoing relationships (i knows j) and incoming relationships (j knows i). Moreover, since "knowing someone" is assumed to be irreversible, existing relationships could not be terminated but were necessarily maintained. Prior to sending the network survey to the CanTest researchers, the face validity of the survey was assessed by a group of native speakers. They reviewed the survey for ease of use and understanding. All of them deemed the network survey to be acceptable.

We used co-authorship as a proxy measure for research collaboration. Data on co-authorship relations between researchers in the network were derived from Scopus. The time windows searched for the three points in time were 'up to entry to CanTest' (T0), 2018 (T1) and 2019 (T2). Scopus was searched by Author Identifiers and relevant key words (TITLE-ABS-KEY(cancer* OR tumour* OR tumor* OR neoplasm* OR malignan* OR carcinoma* OR sarcoma* OR melanoma* OR lesion* OR leukaemia OR leukemia OR lymphoma* OR myeloma*)) to collect data on the co-authorship relationships between the researchers in the network. Using the Author Identifiers, we corrected for different spelling of researcher names, and merged them when one researcher turned out to have different Author Identifiers. Thereafter, data were entered in a matrix of size 60x60, each row and column representing a researcher. A co-authorship relation between two researchers was coded as 1 = "yes", and absence of a co-authorship relation was coded as 0 = "no". The co-authorship network

Box 1. Definitions

Academic social network	Network of researchers connected by informal interactions and social relationships <i>Who do you know?</i> – self-reported
Collaborative network	Network of researchers connected by co-authorship as a proxy for collaboration <i>Who do you collaborate with?</i> – database-derived

was non-directed (connections between co-authoring researchers are by definition reciprocal) so there is no distinction between incoming and outgoing relationships. [Box 1](#) summarizes how the definitions of the academic social network and the collaborative network were operationalised.

The CanTest member register was consulted to collect individual researcher characteristics. Gender was treated as a constant, categorical actor covariate and was coded as either 0 = “female” or 1 = “male”. Data on researcher seniority was treated as a changing, categorical actor covariate and was coded as 0 = “junior researcher” (early stages of PhD or pre-PhD), 1 = “early-career researcher” (later stages of PhD or early post-doc; also pre-PhD with multiple first author publications in the cancer domain), 2 = “mid-career researcher” (more experienced post-doc; three or more first/last author publications in the cancer domain, supervising more junior researchers/been awarded personal grant(s)), or 3 = “senior researcher” (senior lecturer and above; e.g. been awarded an institutional grant, managing a research group, senior lecturer status). In addition, data on physical workplace (i.e. country and institute) and professional background (i.e. researcher or clinical researcher) were extracted and added to the model to control for sources of scientific embeddedness [8,35,40–44]. Data on the physical workplace covered 18 research institutes across five countries (UK, Denmark, USA, Australia and the Netherlands). The professional background of each researcher was coded as either 0 = “researcher” or 1 = “clinical researcher”. All data were kept in a locked file cabinet and were anonymized prior to analysis to maximize confidentiality.

Model for analysis

We investigated the process of network evolution and co-evolution using stochastic actor-oriented models [45–47]. Statistical analysis of longitudinal network data is not possible with conventional statistical methods assuming independence of observations because, in networks, changing connections are typically interrelated with other simultaneous processes (i.e. changes in other connections in the same network or characteristics of the individual researchers involved). Stochastic actor-based models use a combination of simulation methods with statistical model fitting. For this study, models were estimated with the data-analysis package SIENA in R (Simulation Investigation for Empirical Network Analysis), which is suitable for binary social network data in which a pair of researchers is represented in either state 1 (relationship) or 0 (no relationship) [48]. For all models, t-ratios (indicators of convergence) were obtained of less than 0.1, and overall convergence of less than 0.25, which signals good model convergence [48]. Goodness of fit was assessed with auxiliary statistics (outdegree distribution, indegree distribution and triad census) and was deemed acceptable [49].

To explore the evolution and co-evolution of both networks, we created three models. Model 1 and 2 capture the separate evolution of the academic social network and the collaborative network respectively. Model 3 captures the co-evolution of both networks, exploring the influence of the academic social network on the collaborative network (to measure how both networks are inter-related, considering that (1) some initiatives, although intended to become publications, do not progress, and (2) there are lag phases between informal interaction and publication). All three models contain a combination of effects to control for both individual researchers' positions and an individual researcher's characteristics. Effects express, for example, whether researchers are likely to get to know the connections of their connections (network embeddedness; transitivity), or whether researchers with many relationships are more likely to have additional relationships over time (preferential attachment; indegree popularity). Fig 2 visualizes the effects included in the three models. Detailed explanations of these effects and whether these effects are present in each of the three models are provided in S2 Appendix.

Results

Network development

The development of the academic social network and the collaborative network is depicted in Fig 3, showing a rapidly growing and very dense academic social network, and a collaborative network that is less dense and grows more slowly. Additionally, Fig 4 shows a more detailed view of development of the academic social network, showing the seniority and country of each researcher. A description of the development for both networks is provided in Table 1. The average number of (outgoing) relationships for both networks increased over time, revealing researchers got to know each other as well as starting to collaborate. Furthermore, the academic social network showed a strong tendency toward reciprocity (the co-authorship network is reciprocal by definition). Clusters of researchers were present in both networks.

Next, the association between the networks is given in Table 2. It shows the correlation between the number of (outgoing) relationships for both networks, for each observation moment. These numbers can be regarded as indications of the association between the

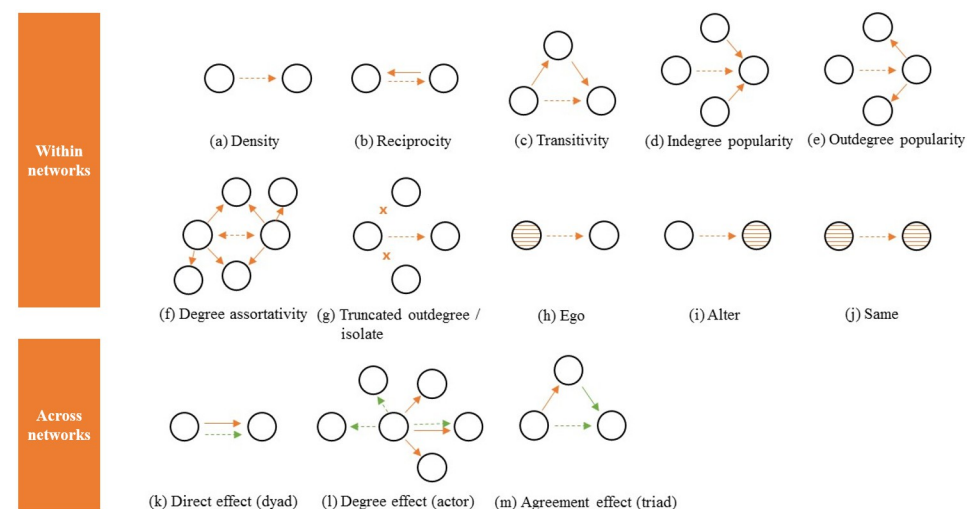


Fig 2. Effects used in the three models. Adapted from Stadtfeld et al [50]. The continuous arrows represent existing relationships at the start of this study; the dashed arrows represent new relationships created over the course of this study. For the cross-network effects, the difference between relationships in the two networks is represented by different colored arrows.

<https://doi.org/10.1371/journal.pone.0272255.g002>

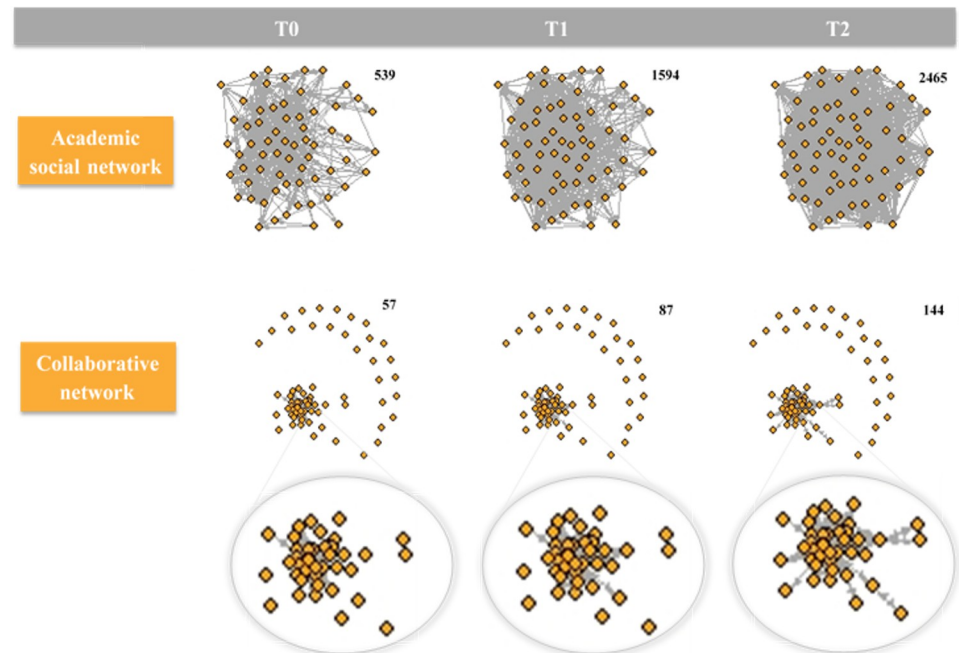


Fig 3. Development of the academic social network and the collaborative network. The top three figures, from left to right, visualize the development of the academic social network over time (i.e. from April 2017/2018 up to June 2019). The middle three figures, from left to right, visualize the development of the collaborative network—zooming in on the core of the network in the bottom three figures. The numbers in the top right corner of each figure represent the number of connections (total outdegree) in each figure.

<https://doi.org/10.1371/journal.pone.0272255.g003>

development of the two networks. The correlations were positive, but decreasing over time, again reflecting a difference in the pace at which both networks developed. In addition, the association at the relationship-level (how many relations between researchers in the academic social network are also present in the collaborative network and vice versa) can be expressed by the Jaccard similarity index. This is a measure of similarity between two sets of data, formally defined as the number of connections in both networks divided by the number in either network, with higher values being indicative for higher similarity. The Jaccard similarity index for each of the three observations was 0.19, 0.13, and 0.11 [47,51]. If independence between

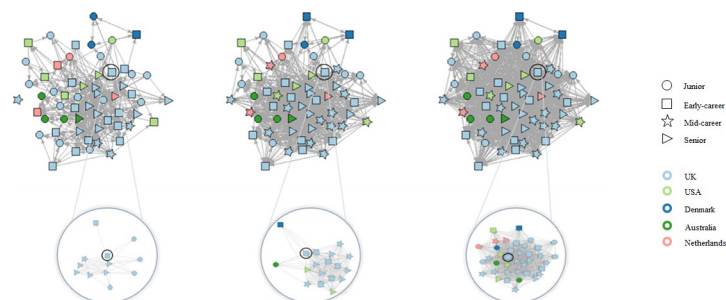


Fig 4. Development of the academic social network showing the seniority and country of each researcher. Again, the top three figures, from left to right, visualize the development of the academic social network over time (i.e. from April 2017/2018 up to June 2019). The relationships of one anonymized, early-career researcher from the UK have been highlighted in the bottom three figures to better illustrate how researchers of different seniority and country get to know each other over time.

<https://doi.org/10.1371/journal.pone.0272255.g004>

Table 1. Descriptives for the evolution of both networks.

	Definitions	Academic social network			Collaborative network		
		T ₀	T ₁	T ₂	T ₀	T ₁	T ₂
Average (out)degree ¹	Average number of (outgoing) relationships amongst researchers	9.0	20.2	39.5	1.9	2.0	3.4
s.d. (out)degree ²	Standard deviation of (out)degrees	6.8	11.4	13.1	3.5	3.8	5.1
Reciprocity ³	If i is connected to j, what is the probability that j is also connected to i?	0.65	0.73	0.79	-	-	-
Transitivity ⁴ (clustering)	If i is connected to j, and j to k, what is the probability that i is also connected to k?	0.51	0.66	0.84	0.53	0.44	0.65

¹ An average outdegree of 9 means that researchers know on average 9 other researchers;

² An s.d. outdegree of 6.8 means that on average the difference between the average outdegree and individual measurements is 6.8;

³ A reciprocity of 0.65 means that 65% of the connections between researchers are mutual;

⁴ A transitivity of 0.51 indicates that 51% of the connections were transitive (referring to network embeddedness/transitivity).

<https://doi.org/10.1371/journal.pone.0272255.t001>

the networks was assumed, the expected Jaccard coefficients would be 0.03, 0.04, and 0.08. The observed values, while not very high, were higher.

Network evolution

The two central columns of Table 3 report the parameter estimates for the separate evolution of both networks with their associated standard errors. Irrelevant effects for the different models are coloured grey. The academic social network (Model 1) exhibited positive *reciprocity* and *transitivity* parameters, indicating that researchers over time tended to get to know (1) those researchers who they were known by, and (2) the ‘friends of their friends’ (i.e. they are likely to cluster in groups). Yet, the negative interaction between reciprocity and transitivity indicates that, compared to one-to-one relationships, in clusters scientists are less likely to reciprocate relationships. The positive *indegree popularity* signifies that well-known researchers are inclined to become even more well-known over time. Conversely, the negative *outdegree popularity* parameter indicates that researchers who know many others are not necessarily well-known to others as well.

With increasing seniority, researchers are less likely to get to know additional researchers within the network (negative *seniority ego*). Similarly, males are less likely to get to know others (negative *gender (male) ego*). The interaction between gender and seniority (*gender (male) x seniority (senior)*) further signifies that males of higher seniority are less likely to get to know additional others. Finally, significant positive effects for *same gender* and *same professional background* assume that researchers are more likely to get to know others of the same gender and of the same professional background.

The negative *degree (density)* parameter for the collaborative network (Model 2) indicates that researchers were selective in their collaborative relationships. Yet, having collaborative relationships increases the initiation of new collaborations (positive *in = outdeg. popularity/activity*)—especially with low-collaborators (negative *degree assortativity*). Although male researchers were less likely to get to know others compared to females, they are more likely to

Table 2. Correlations between the (out)degrees of the two networks, for the three time points separately.

	T ₀	T ₁	T ₂
	Collaborative network	Collaborative network	Collaborative network
Academic social network	0.71	0.39	0.14

<https://doi.org/10.1371/journal.pone.0272255.t002>

Table 3. Evolution and co-evolution of the academic social network and the collaborative network: Parameter estimates and standard errors of SIENA models.

	Evolution				Co-evolution	
	Model 1 Academic social network		Model 2 Collaborative network ¹		Model 3 Academic social network → Collaborative network	
	par.	(s.e.)	par.	(s.e.)	par.	(s.e.)
Within network						
Degree (density)			-3.134***	(0.375)	-5.133***	(1.080)
Reciprocity	5.324***	(1.538)				
Transitivity ²	3.200**	(1.201)	2.232***	(0.362)	2.575***	(0.669)
Indegree popularity ³	0.464*	(0.195)				
Outdegree popularity ³	-0.201**	(0.065)				
In = outdeg. popularity/activity			0.112***	(0.025)	0.152**	(0.055)
Degree assortativity			-0.311***	(0.089)	-0.432**	(0.157)
Outdegree < 10	-12.063***	(1.197)				
Network isolate			-0.749	(0.867)	-1.688	(1.533)
Gender (male) ego	-3.787***	(0.967)				
Gender (male) alter	0.041	(0.079)				
Gender (male) ego+alter			0.254*	(0.119)	0.247	(0.167)
Same gender	0.259***	(0.081)	-0.095	(0.139)	-0.242	(0.192)
Seniority (senior) ego	-1.708***	(0.505)				
Seniority (senior) alter	0.001	(0.031)				
Seniority (senior) alter ³	-0.005	(0.036)				
Seniority (senior) ego+alter			-0.031	(0.044)	-0.016	(0.055)
Seniority (senior) ego+alter ³			0.077	(0.057)	0.086	(0.066)
Seniority similarity ³	-0.008	(0.016)	-0.043	(0.032)	-0.056†	(0.034)
Same professional background	0.106†	(0.062)	0.158	(0.131)	-0.037	(0.154)
Same country	-0.044	(0.110)	0.112	(0.129)	0.026	(0.163)
Same institution	2.395***	(0.347)	0.847***	(0.205)	0.529	(0.236)
Between-network: direct effect						
Academic social network					2.554**	(0.896)
Between-network: degree effects						
Outdegree activity					-0.373	(0.278)
Indegree popularity					0.296	(0.265)
Between-network: agreement						
Academic social network					-0.010	(0.077)
Interactions						
Gender (male) x seniority (senior)	-3.962**	(1.007)	0.151	(0.182)	0.028	(0.266)
Transitivity x reciprocity	-2.149***	(0.782)				

par. = parameter for the effect (estimate); (s.e.) = standard error;

† p < 0.1;

* p < 0.05;

** p < 0.01;

*** p < 0.001.

¹ For the independency assumption of the stochastic actor-oriented model, only papers with a maximum of five co-authors from within CanTest were included.

² gwespFF for Model 1 and gwesp for Model 2/3.

³ Square-root transformed.

<https://doi.org/10.1371/journal.pone.0272255.t003>

collaborate (positive *gender (male) ego+alter*). Finally, collaboration is encouraged by working in the same institution (positive *same institution*).

Network co-evolution

The results for the co-evolution of the academic social network and collaborative network are reported in the right hand column of [Table 3](#). The direct effect of a relationship in the academic social network on the likelihood of a relationship in the collaborative network was positive and significant (*between network: direct effect—academic social network*). When researcher *i* knew researcher *j*, they were likely to start collaboration (i.e. co-authorship) over time. Other cross-network effects were not significant.

Discussion

The CanTest network is clearly successful in connecting researchers with each other. Researchers within the network establish connections promptly and effectively (objective 1: evolution of the academic social network), allowing for considerable exchange of information and ideas to increase capacity and support sustainability of early detection and diagnosis of cancer research. Seniority and gender seem to play a major role in the development of relationships within the network (objective 3: individual researcher characteristics), not affecting who is known or not, but who gets to know others (objective 3: individual network positions). We found that more junior researchers built and expanded their academic network, but that with increasing seniority researchers were less likely to get to know others. This closely aligns with one of the key objectives of CanTest to recruit and support a new generation of researchers to establish themselves and reach early independence [13]. Moreover, compared to male researchers, female researchers seem to expand their academic social networks faster. The interaction between gender and seniority stresses even more how males of increasing seniority are less likely to expand their academic social networks over time.

Yet, in contrast with how more junior researchers expand their academic social networks, they seem to expand their collaborative research networks more slowly (objective 1 and 3: evolution of the collaborative network and the influence of individual researcher characteristics). The positive but decreasing association between the CanTest academic social network and the collaborative network over time further confirms this (objective 2: co-evolution of the academic social network and the collaborative network). The concurrent development of effective research collaboration is indeed a time consuming process—being connected informally, e.g. by being involved in the same project, may lead to collaboration on ideas and study design, application for funding, exchange visits, conduct of research, and only finally co-authorship. Despite this lengthy sequence, our results demonstrate that the transition from social connections into research collaboration does take place: creating an academic social relationship between two researchers significantly increases their chances to collaborate. The existence of the CanTest Collaborative and the structure, events and coordinated communications that go with it, has most likely contributed to the observed results. However, it is not possible to know how the networks studied here would have evolved without the existence of CanTest. The timing of the significant increase in academic social connections since the start of CanTest in April 2017 makes it likely, though, that CanTest has been instrumental in the process of creating and accelerating informal interactions and social relationships and hence collaboration.

Findings from previous studies suggesting higher co-authorship activity for males compared to females are confirmed in the current study [23]. However, we also found evidence to support findings from other studies that females are more appreciative of collaboration and so they expand their academic social networks faster [22]. Frequently assumed tendencies for

same-gender relationships are confirmed as well; however, no evidence was found for same-gender collaborations [52]. Our finding that more senior researchers tend to collaborate more often may have played a role in this. Indeed, based on a large academic bibliographic database research, Combes and Givord (2018) argued that same-gender collaborations occur more commonly at the beginning of a researcher's career and fade with seniority [53]. Furthermore, national background didn't seem to play a role in the development of individual relationships amongst researchers, although the international collaboration does strengthen the network as such. A perhaps surprising finding was that, among participants from the same institution, CanTest seems to have boosted within-own-institution collaborations, beyond boosting between-institution ones. Therefore catalysing new 'internal' network formations seems one of the means by which the network intervention has been effective. This may reflect the fact that most modern universities often encompass multiple campus sites/buildings (the boundaries between them acting as practical barriers) and different departments (the boundaries between them acting as organisational or disciplinary barriers), which mean that the potential for within-institutional collaborations cannot be taken as a given. Therefore, the CanTest network may also have boosted collaborations that—although relating to participants working at the same institution—would not have otherwise happened.

The main strength of the current study is that longitudinal rather than cross-sectional network data of a whole research network was used for understanding network evolution and co-evolution. Specifically, we were able to control for effects from researcher's positions in the network as well as their seniority and gender across three points in time. A limitation of the study was that the study period was relatively short, and co-authorship relationships may take longer to flourish. As a result, the dynamics of collaboration in CanTest may not have been fully captured. Future research will be needed to show whether 'knowing each other' translates into "collaborating with each other" even more when considered over a longer period of time. Moreover, it should be acknowledged that collaborative research activity is not limited to co-author behaviour. Other metrics for collaborative research activity could have been used, e.g. co-funding, but co-authoring activity was preferred as it is the most common metric for collaborative research activity in the literature and its data was readily accessible.

Future research should further address the evolution and co-evolution of relationships and collaboration within (cancer) research networks; in particular between less and more senior researchers as it remains unclear whether it is the bridge function that senior researchers may have towards junior researchers, the density of the CanTest network itself (densely linked networks are more efficient at diffusing information to all their members when compared to sparsely linked groups), or a combination of both that encourages researchers to move forward [25,54]. A combination of a densely linked network and the availability of one or several 'bridging researchers'—often referred to as brokers—might be ideal, pursuing a network that is maximally effective in facilitating collaboration between its members [55]. In addition, more research into the extent to which males and females of different seniority seem to expand their networks is warranted. There is evidence for a 'saturation point' for social connections, arguing that researchers are likely to refrain from initiating new connections if they already have many connections, or for females being more likely to create more diverse social capital [53,56]. Yet, there may be countless other mechanisms that address the interplay between seniority and gender in (academic) social network formation.

Simply establishing the infrastructure for a network of researchers to get to know each other will not necessarily make them collaborate. It may be just a matter of time, but the current study—unique in using longitudinal data to study the co-evolution of social connections and collaboration considering both researcher's characteristics and positions in networks—shows how the seniority and gender of researchers are particularly worth paying attention to

when establishing effective research networks. Increased understanding of how to address and balance researcher's characteristics might help other research initiatives or funding agencies in developing effective research networks to promote research output. This study shows how facilitating and supporting a dense research network, through formal and informal network interactions, positively affects the translation from "getting to know each other" into collaboration—time will tell whether the established social connections will lead to further collaborations in the future.

Supporting information

S1 Appendix. Survey: Matrix question and informed consent procedure.

(DOCX)

S2 Appendix. Effects used in the three models.

(DOCX)

Acknowledgments

We thank members of the CanTest Steering Group for early discussions on this study and the participants for completing the surveys.

Author Contributions

Conceptualization: Debbie Vermond, Esther de Groot, Valerie A. Sills, Georgios Lyratzopoulos, Fiona M. Walter, Niek J. de Wit, Greg Rubin.

Formal analysis: Debbie Vermond.

Methodology: Debbie Vermond, Esther de Groot, Valerie A. Sills, Georgios Lyratzopoulos, Fiona M. Walter, Niek J. de Wit, Greg Rubin.

Writing – original draft: Debbie Vermond, Esther de Groot, Valerie A. Sills.

Writing – review & editing: Georgios Lyratzopoulos, Fiona M. Walter, Niek J. de Wit, Greg Rubin.

References

1. Allgar VL, Neal RD. Delays in the diagnosis of six cancers: analysis of data from the National Survey of NHS Patients: Cancer. *Br J Cancer*. 2005; 92(11):1959–70. Epub 2005/05/05. <https://doi.org/10.1038/sj.bjc.6602587> PMID: 15870714.
2. Macleod U, Mitchell ED, Burgess C, Macdonald S, Ramirez AJ. Risk factors for delayed presentation and referral of symptomatic cancer: evidence for common cancers. *Br J Cancer*. 2009; 101 Suppl 2: S92–S101. Epub 2009/12/04. <https://doi.org/10.1038/sj.bjc.6605398> PMID: 19956172.
3. van Hoof SJM, Quanjel TCC, Kroese M, Spreeuwenberg MD, Ruwaard D. Substitution of outpatient hospital care with specialist care in the primary care setting: A systematic review on quality of care, health and costs. *Plos One*. 2019; 14(8):e0219957. Epub 2019/08/02. <https://doi.org/10.1371/journal.pone.0219957> PMID: 31369567.
4. Berwick D, Nolan T, Whittington J. The Triple Aim: Care, Health, and Cost. *Health affairs (Project Hope)*. 2008; 27:759–69. <https://doi.org/10.1377/hlthaff.27.3.759> PMID: 18474969
5. Walter FM, Thompson MJ, Wellwood I, Abel GA, Hamilton W, Johnson M, et al. Evaluating diagnostic strategies for early detection of cancer: the CanTest framework. *Bmc Cancer*. 2019; 19. ARTN 586. <https://doi.org/10.1186/s12885-019-5746-6> PMID: 31200676
6. Azoulay P, Zivin JSG, Wang JL. Superstar Extinction. *Q J Econ*. 2010; 125(2):549–89. <https://doi.org/10.1162/qjec.2010.125.2.549>
7. Abramo G, D'Angelo C, Di Costa F. Research Collaboration and Productivity: Is There Correlation? *Higher Education*. 2008; 57:155–71. <https://doi.org/10.1007/s10734-008-9139-z>

8. Khor KA, Yu LG. Influence of international co-authorship on the research citation impact of young universities. *Scientometrics*. 2016; 107:1095–110. Epub 2016/05/31. <https://doi.org/10.1007/s11192-016-1905-6> PMID: 27239078.
9. Kyvik S, Reymert I. Research collaboration in groups and networks: differences across academic fields. *Scientometrics*. 2017; 113(2):951–67. <https://doi.org/10.1007/s11192-017-2497-5> PMID: 29081555
10. Page SE. *The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools, and Societies* (New Edition); Princeton University Press; 2007.
11. Austin JE, Seitanidi MM. Value Creation Through Collaboration. In: Renz D, Herman R, editors. *The Jossey-Bass Handbook of Nonprofit Leadership and Management*. Hoboken, New Jersey: John Wiley & Sons, Inc.; 2016. p. 427–43.
12. Adams J. The rise of research networks. *Nature*. 2012; 490(7420):335–6. <https://doi.org/10.1038/490335a> PMID: 23075965
13. N.A. The CanTest Collaborative. <https://www.cantest.org/>.
14. Sonnenwald DH. Scientific collaboration. *Annual Review of Information Science and Technology*. 2007; 41(1):643–81. <https://doi.org/10.1002/aris.2007.1440410121>
15. Burt RS, Kilduff M, Tasselli S. Social network analysis: foundations and frontiers on advantage. *Annu Rev Psychol*. 2013; 64:527–47. Epub 2013/01/04. <https://doi.org/10.1146/annurev-psych-113011-143828> PMID: 23282056.
16. Martín-Sempere MJ, Rey-Rocha J, Garzón-García B. The effect of team consolidation on research collaboration and performance of scientists. Case study of Spanish university researchers in Geology. *Scientometrics*. 2002; 55(3):377–94. <https://doi.org/10.1023/A:1020462712923>
17. Burk W, Steglich C, Snijders T. Beyond dyadic interdependence: Actor-oriented models for co-evolving social networks and individual behaviors. *International Journal of Behavioral Development*. 2007; 31. <https://doi.org/10.1177/0165025407077762>
18. Lee S, Bozeman B. The Impact of Research Collaboration on Scientific Productivity. *Social Studies of Science*. 2005; 35(5):673–702. <https://doi.org/10.1177/0306312705052359>
19. Carswell P, Manning B, Long J, Braithwaite J. Building clinical networks: a developmental evaluation framework. *Bmj Qual Saf*. 2014; 23(5):422–7. <https://doi.org/10.1136/bmjqs-2013-002405> PMID: 24481646
20. West E, Barron DN, Dowsett J, Newton JN. Hierarchies and cliques in the social networks of health care professionals: implications for the design of dissemination strategies. *Soc Sci Med*. 1999; 48(5):633–46. [https://doi.org/10.1016/s0277-9536\(98\)00361-x](https://doi.org/10.1016/s0277-9536(98)00361-x) PMID: 10080364
21. Reagans R, Zuckerman E, McEvily B. How to make the team: Social networks vs. demography as criteria for designing effective teams. *Admin Sci Quart*. 2004; 49(1):101–33.
22. Carr PL, Pololi L, Knight S, Conrad P. Collaboration in Academic Medicine: Reflections on Gender and Advancement. *Acad Med*. 2009; 84(10):1447–53. <https://doi.org/10.1097/ACM.0b013e3181b6ac27> PMID: 19881441
23. Ozel B, Kretschmer H, Kretschmer T. Co-authorship pair distribution patterns by gender. *Scientometrics*. 2014; 98(1):703–23. <https://doi.org/10.1007/s11192-013-1145-y>
24. Szell M, Thurner S. How women organize social networks different from men. *Scientific Reports*. 2013; 3(1):1214. <https://doi.org/10.1038/srep01214> PMID: 23393616
25. Bojanowski M, Czerniawska D. Reaching for Unique Resources: Structural Holes and Specialization in Scientific Collaboration Networks 2019.
26. Bordons M, Aparicio J, González-Albo B, Díaz-Faes AA. The relationship between the research performance of scientists and their position in co-authorship networks in three fields. *Journal of Informetrics*. 2015; 9(1):135–44. <https://doi.org/10.1016/j.joi.2014.12.001>
27. Dias A, Ruthes S, Lima L, Campira E, Silva M, Bragança de Sousa M, et al. Network centrality analysis in management and accounting sciences. *RAUSP Management Journal*. 2019; 55(2):207–26. <https://doi.org/10.1108/RAUSP-02-2019-0021>
28. Joksimovic S, Manataki A, Gašević D, Dawson S, Kovanovic V, De Kerekif IF, editors. Translating network position into performance: Importance of centrality in different network configurations. *ACM International Conference Proceeding Series*; 2016.
29. Lemes Alarcão AL, Sacomano Neto M. Actor centrality in Network Projects and scientific performance: an exploratory study. *RAI Revista de Administração e Inovação*. 2016; 13(2):78–88. <https://doi.org/10.1016/j.rai.2016.03.002>
30. Gonzalez-Brambila CN, Veloso FM, Krackhardt D. The impact of network embeddedness on research output. *Res Policy*. 2013; 42(9):1555–67. <https://doi.org/10.1016/j.respol.2013.07.008>
31. Merton RK. The Matthew effect in science. The reward and communication systems of science are considered. *Science*. 1968; 159(3810):56–63. Epub 1968/01/05. PMID: 5634379.

32. Price DDS. A general theory of bibliometric and other cumulative advantage processes. *Journal of the American Society for Information Science*. 1976; 27(5):292–306. <https://doi.org/10.1002/asi.4630270505>
33. de Sola Pool I, Kochen M. Contacts and influence. *Social Networks*. 1978; 1(1):5–51. [https://doi.org/10.1016/0378-8733\(78\)90011-4](https://doi.org/10.1016/0378-8733(78)90011-4)
34. Robinson TE, Rankin N, Janssen A, McGregor D, Grieve S, Shaw T. Collaborative research networks in health: a pragmatic scoping study for the development of an imaging network. *Health Res Policy Sy*. 2015; 13. ARTN 76. <https://doi.org/10.1186/s12961-015-0067-y> PMID: 26652643
35. Fang W, Dai S, Tang LL. The Impact of International Research Collaboration Network Evolution on Chinese Business School Research Quality. *Complexity*. 2020; 2020. Artn 7528387. <https://doi.org/10.1155/2020/7528387>
36. Moore J, Prentice D, McQuestion M. Social Interaction and Collaboration among Oncology Nurses. *Nursing Research and Practice*. 2015; 2015:248067. <https://doi.org/10.1155/2015/248067> PMID: 26113993
37. Snijders T, Lomi A, Torlo V. A model for the multiplex dynamics of two-mode and one-mode networks, with an application to employment preference, friendship, and advice. *Social networks*. 2013; 35:265–76. <https://doi.org/10.1016/j.socnet.2012.05.005> PMID: 23690653
38. Camarinha-Matos L, Afsarmanesh H. Collaborative Networks: Value creation in a knowledge society. *International Federation for Information Processing Digital Library; Knowledge Enterprise: Intelligent Strategies in Product Design, Manufacturing, and Management*; 2006;207.
39. Butts C. Social Network Analysis: A Methodological Introduction. *Asian Journal of Social Psychology*. 2008; 11:13–41. <https://doi.org/10.1111/j.1467-839X.2007.00241.x>
40. Cummings J, Kiesler S. Collaborative Research across Disciplinary and Organizational Boundaries. 352008. p. 99–118.
41. Iglić H, Doreian P, Kronegger L, Ferligoj A. With whom do researchers collaborate and why? *Scientometrics*. 2017; 112(1):153–74. <https://doi.org/10.1007/s11192-017-2386-y> PMID: 28725095
42. Amuah J, Mutasingwa D, Amankwah E. Enhancing Collaboration between Clinician-Researchers and Methodologists in Clinical Research. *The Journal of pediatrics*. 2020; 221:S58–S61. <https://doi.org/10.1016/j.jpeds.2020.02.035> PMID: 32482237
43. Muia J, Casari C. Reflections on scientific collaboration between basic researchers and clinicians. *Journal of Thrombosis and Haemostasis*. 2016; 14(10):1905–7. <https://doi.org/10.1111/jth.13447> PMID: 27515008
44. de Groot E, Baggen Y, Moolenaar N, Stevens D, van Tartwijk J, Damoiseaux R, et al. Clinician-Scientists in-and-between Research and Practice: How Social Identity Shapes Brokerage. *Minerva*. 2020. <https://doi.org/10.1007/s11024-020-09420-7> PMID: 33041374
45. Snijders TA. Models for Longitudinal Network Data. In: Carrington P, Scott J, Wasserman S, editors. *Models and methods in social network analysis*. New York: Cambridge University Press.; 2005. p. 215–47.
46. Snijders TAB. The statistical evaluation of social network dynamics. *Sociol Methodol*. 2001; 31:361–95. <https://doi.org/10.1111/0081-1750.00099>
47. Snijders TAB, van de Bunt GG, Steglich CEG. Introduction to stochastic actor-based models for network dynamics. *Social Networks*. 2010; 32(1):44–60. <https://doi.org/10.1016/j.socnet.2009.02.004>
48. Ripley RM, Snijders TAB, Boda Z, Vörös A, Preciado P. Manual for Siena version 4.0. Technical Report. Technical Report. Oxford: University of Oxford, Department of Statistics; Nuffield College, 2020.
49. Lospinoso J, Snijders TA. Goodness of fit for stochastic actor-oriented models. *Methodological Innovations*. 2019; 12(3):2059799119884282. <https://doi.org/10.1177/2059799119884282>
50. Stadtfeld C, Takács K, Vörös A. The Emergence and Stability of Groups in Social Networks. *Social Networks*. 2020; 60:129–45. <https://doi.org/10.1016/j.socnet.2019.10.008>
51. Batagelj V, Bren M. Comparing resemblance measures. *Journal of Classification*. 1995; 12:73–90. <https://doi.org/10.1007/BF01202268>
52. Boschini A, Sjögren A. Is Team Formation Gender Neutral? Evidence from Coauthorship Patterns. *Journal of Labor Economics*. 2007; 25:325–65. <https://doi.org/10.1086/510764>
53. Combes S, Givord P. Selective matching: gender gap and network formation in research. Center for Research in Economics and Statistics, 2018.
54. Long JC, Cunningham FC, Braithwaite J. Bridges, brokers and boundary spanners in collaborative networks: a systematic review. *BMC Health Services Research*. 2013; 13(1):158. <https://doi.org/10.1186/1472-6963-13-158> PMID: 23631517

55. Reinholt M, Pedersen T, Foss NJ. Why a Central Network Position Isn't Enough: The Role of Motivation and Ability for Knowledge Sharing in Employee Networks. *Acad Manage J*. 2011; 54(6):1277–97. <https://doi.org/10.5465/amj.2009.0007>
56. Díaz-Faes AA, Otero-Hermida P, Ozman M, D'Este P. Do women in science form more diverse research networks than men? An analysis of Spanish biomedical scientists. *Plos One*. 2020; 15(8 August). <https://doi.org/10.1371/journal.pone.0238229> PMID: [32853227](https://pubmed.ncbi.nlm.nih.gov/32853227/)