Scientific collaboration and team science: a social network analysis of the centers for population health and health disparities

Janet Okamoto, PhD,¹ The Centers for Population Health and Health Disparities Evaluation Working Group,²

¹Mayo Clinic, 13400 E. Shea Blvd., Scottsdale, AZ 85259, USA

²Division of Cancer Control and Population Sciences, National Cancer Institute, 9609 Medical Center Drive, Rockville, MD 20850, USA

Correspondence to: J Okamoto
okamoto.janet@mayo.edu

Cite this as: *TBM* 2015;5:12–23
doi: 10.1007/s13142-014-0280-1

Abstract

The past decade has seen dramatic shifts in the way that scientific research is conducted as networks, consortia, and large research centers are funded as transdisciplinary, team-based enterprises to tackle complex scientific questions. Key investigators ($N=167$) involved in ten health disparities research centers completed a baseline social network and collaboration readiness survey. Collaborative ties existed primarily between investigators from the same center, with just 7% of ties occurring across different centers. Grants and work groups were the most common types of ties between investigators, with shared presentations the most common tie across different centers. Transdisciplinary research orientation was associated with network position and reciprocity. Center directors/leaders were significantly more likely to form ties with investigators in other roles, such as statisticians and trainees. Understanding research collaboration networks can help to more effectively design and manage future team-based research, as well as pinpoint potential issues and continuous evaluation of existing efforts.

Keywords

Research collaboration, Network analysis, Collaboration networks, Health disparities, Scientific collaboration

INTRODUCTION

Health disparities present a complex research problem, as the social determinants of health are influenced by various factors, many of which do not necessarily appear to be obviously linked to health. Trying to understand and disentangle causes and antecedents in order to be able to systematically address and reduce the disparate burden of disease among minority and underserved populations is one of the biggest challenges in current biomedical and public health research. This complexity has created a demand for an integrated approach that brings together multiple disciplines and requires thinking outside of the traditional disciplinary knowledge silos [17]. Partnerships

Implications

Policy: Institutions and funding agencies should offer guidelines and recommended effective practices for research collaboration and develop support mechanisms tailored toward building and sustaining research networks.

Research: More longitudinal studies and development of additional measures of research impact across scientific disciplines are needed in order to more comprehensively assess the value of team-based research.

Practice: Team-based and collaborative research initiatives should focus on creating, supporting, and sustaining mentoring relationships between senior scientists and more junior investigators to ensure a less centralized and more cohesive and connected research collaboration network.

and collaborations promise an efficient and effective approach to this research going forward.

Scientific collaboration to address complex public health problems

The past decade has seen dramatic shifts in the way that scientific research is conducted. Increasingly, traditional single investigator-initiated research projects have been transformed to include multidisciplinary teams of co-investigators and consultants. More dramatic changes occur as research networks, consortia, and large research centers are explicitly funded as transdisciplinary enterprises to tackle complex scientific questions. This shift is resulting in transdisciplinary approaches to disparities research which may facilitate more rapid translation of scientific discovery from “bench to bedside” as advances and innovations in biomedical and public health research increasingly result from multigroup and cross-institutional

collaboration. This has been shown in the literature, with the “nearly universal rise since 1975 in the frequency of collaborations among authors representing different disciplines and located at different universities” [16], indicating an increased priority for collaboration. There is evidence that this type of collaboration is associated with job satisfaction of academic researchers, on par with more traditional measures of satisfaction such as salary and job security [8].

These collaborative networks, comprising investigators from multiple disciplines and research orientations, address the challenges posed by complex, multilevel health disparities issues, through enhanced access to expertise and to specialized infrastructure, placing a premium on knowledge integration and innovation [6]. While scientific research is shifting toward greater collaboration and transdisciplinarity, we need to improve our understanding of the processes and structures that lead to successful scientific collaboration and high-impact, innovative research outcomes. If the ultimate goal of addressing complex scientific problems, such as health disparities, is to be achieved, it will require documentation of successes and limitations by examining intentionally constructed systems to promote and encourage transdisciplinarity [18].

The Centers for Population Health and Health Disparities

Eight geographically dispersed Centers for Population Health and Health Disparities were initially funded from 2003 to 2008. The Centers for Population Health and Health Disparities (CPHHD) initiative was renewed from 2010 to 2015, providing funding for ten centers including three of the original eight centers and seven new centers. In this second round, five centers were funded by the National Cancer Institute (NCI) to focus on cancer health disparities, and five were funded by the National Heart, Lung, and Blood Institute (NHLBI) to focus on disparities in cardiovascular and metabolic health. The Office of Behavioral and Social Sciences Research (OBSSR) at NIH served as an additional funding partner during both funding periods.

Large-scale team science initiatives have been identified as a promising approach to hasten scientific progress to address complex and multifactorial social problems [7, 9, 10, 29, 30]. The transdisciplinary CPHHD program was expressly designed to address health inequities and health disparities, using multilevel approaches based on rigorous scientific methods, with the assertion that “by combining approaches from various disciplines (e.g., those in the physical, biological, social, and behavioral sciences), the CPHHD Program can advance research on disparate health outcomes and probe more deeply into the causes of these outcomes, develop interventions, and identify practice and policy approaches to addressing them.” The aim of funding large, center-grant initiatives such as the Centers for Population Health and Health Disparities is to build capacity in the scientific community for collaboration both within and between

centers and institutions and to use this capacity to obtain support for more sophisticated cross-cutting research questions. The resulting research focuses variously on gene-environment interactions, psychosocial stress, cultural influences, and other individual, interpersonal, environmental, and societal factors contributing to health disparities. Finally, the Funding Opportunity Announcement (FOA) for the current round of CPHHD funding [23] specifically stated that “...investigators should...participate in cross-cutting ‘research interest groups’ and other collaborative activities as a way of advancing trans disciplinary health disparities and population health as a science” and that “interactions (on a regular basis) must be planned with other centers to share information, participate in NIH-directed evaluation activities, promote and coordinate collaborative efforts, identify opportunities for common measures in the field, and review areas of growth and opportunity in the field of health disparities and health inequities.”

Theoretical frameworks for research collaboration

Few studies have examined research *on* research collaboration, pointing to a lack of information on the actual interpersonal relationships between collaborators. Bozeman et al. [5] outlined an organizing framework for the study of research collaboration that classifies attributes of collaboration into three categories: collaborator attributes, attributes of the collaborative relationship, and organizational attributes. Each of these categories is further divided into various subcategories leading to research outputs. Sargent and Waters [26] outlined a model of collaboration similar to Bozeman et al. which they call an inductive process framework of academic research collaboration. They also include three levels of consideration: the context in which the research occurs, the phase of the collaboration or research, and the interpersonal processes that influence collaborative efforts. This model also includes interpersonal factors as important to successful collaborative efforts, but adds contextual factors, external to both the collaborator and the collaborative research. The two frameworks differ in the importance placed on individual or investigator attributes, but both highlight the need for additional research evaluation studies to further elucidate the multifaceted processes that influence successful collaboration. Our report attempts to fill some of those gaps by examining self-reported relationships between collaborators and how they are related to other collaboration attributes, such as the tendency toward transdisciplinary research and conflict.

The present study

For the present study, we documented the interpersonal relationships among collaborators in a large research initiative using social network analysis (SNA) methods. The primary aim was to examine the nature and extent of the working relationships among CPHHD investigators. To assess this aim, we

examined the associations between network position and characteristics of the investigators and consideration of the different types of collaboration that have occurred among these investigators, as well as how the different collaboration types were related to each other. For example, it is important to assess whether collaboration on a grant is associated with other types of collaboration, such as on publications. The structure of the CPHHD initiative's scientific collaboration networks was also investigated, as well as the relationship between network structure and intrapersonal collaboration variables, such as transdisciplinary research orientation and conflict.

METHODS

Sample and data collection

An online survey on baseline social network and collaboration readiness was completed by key investigators involved in the ten CPHHDs. It was conducted from December 2010 through January 2011. Respondents included all center directors, research core principal investigators (PIs), and individual project PIs and other key research personnel, as recommended by each center director.

Each center director provided a standardized informational piece about the survey to investigators and key personnel at their center and invited them to participate. Personal e-mail reminders with survey links were sent by study staff to nonrespondents after the first month and again 2 weeks prior to the close of the survey. Regular reminders were sent to center directors, who were encouraged to send their own reminders to respondents at their centers. Data were collected online. Only summary and aggregated data are provided, to protect the confidentiality of individual survey respondents.

Measures

Sample characteristics

Basic background and demographic information included gender, academic rank, and role within the center. Also recorded was whether the center had been previously funded as a CPHHD. Five categories of academic rank included the following: professor, associate professor, assistant professor, research associate/fellow, and other. Survey data were supplemented with data from the grant applications and communication with the center directors. Such information included the role of each investigator within their center. There were five roles: center director, project PI, research core director, co-investigator, and other.

Intrapersonal orientation to collaboration—Personal orientation to collaboration was measured with three scales, where investigators reported on their disciplinary research orientation, conflict orientation, and views about interpersonal collaboration within their research center.

Research orientation—The Research Orientation Scale (ROS) is a ten-item scale that assesses the disciplinary and/or cross-disciplinary nature of an investigator's values and attitudes regarding research. The ROS is measured on a five-point Likert scale from "strongly disagree" to "strongly agree". Examples of questions on the ROS scale include this disciplinary question, "There is so much work to be done within my field that I feel it is important to focus my research efforts with others in own discipline" and this inter/transdisciplinary question, "In my collaborations with others I integrate theories and models from different disciplines."

This measure was developed by Hall and colleagues [1] to assess the cross-disciplinary continuum, as defined by Rosenfield (1992). However, factor analysis in the original study did not support the ROS as a measure of a *continuum* [9]. Rather, they found that the scale separated into three *distinct* factors assessing disciplinary, multidisciplinary, and inter/transdisciplinary. Factor analysis of the scale with the present study data indicated the same three factors. The inter-disciplinary and transdisciplinary items developed from Rosenfield's four research orientation types loaded onto one factor. Therefore, three composite measures were created by taking a mean of all items that loaded on each factor. For example, the responses from each respondent on the three items that loaded on the disciplinary orientation factor were averaged to create this variable. The Cronbach's alpha reliability coefficients for the sub-scales were 0.63 for interdisciplinary, 0.56 for multidisciplinary, and 0.86 for inter/transdisciplinary orientation.

Conflict—Task conflict was measured using a six-item modified version of Jehn's [15] original four-item measure. Jehn's items were tailored to reflect the team-based research context, and two items were added to enhance overall comprehensiveness of the items. The items were measured on a five-point Likert scale from "Not at all" to "To a very large extent." An example of these items is "To what extent are there differences of opinion about the work in my team?" The items showed good reliability, with a Cronbach's alpha of 0.82.

Views about interpersonal collaboration—Investigators' views about interpersonal collaboration within their CPHHD were measured with four items [9]. The items were measured on a five-point Likert scale from "strongly disagree" to "strongly agree," and a variable was created by taking the mean of all four items. Investigators were asked to rate their views about "collaboration with respect to your research in your center." The interpersonal processes the scale asked about include conflict resolution, communication, trust, and research productivity. The items showed good reliability, with a Cronbach's alpha of 0.77.

Measurement of directed collaboration ties

Investigators were asked to report who they collaborated with from the CPHHD initiative. The survey

included the full list of key investigators that was collected from the grant applications and center directors and was presented alphabetically and separated by center. Investigators indicated whether they currently collaborate with another investigator and/or whether they collaborated with another prior to the start of the CPHHD initiative. Collaboration can be broadly defined, and the meaning can differ greatly between people; therefore, investigators were asked to indicate the nature of each collaboration they reported. Six different types of collaboration were listed, and respondents could select all that apply for each collaborator they nominated. The six collaboration types included the following: study or grant, co-authored publication, co-authored presentation, mentoring or training, committee/work group, and other. The “other” category was included for quality control purposes in order to get a general measure of overall collaboration, in case investigators collaborated with each other on something that did not fall into the five given collaboration types. This type of question allowed for the examination of two general collaboration networks, current and prior to the CPHHD, but it also allowed for more specific investigation of the six separate collaboration type networks for each period (current and prior).

Two measures of network centrality were examined for each investigator in order to better understand their position in the overall scientific collaboration network, closeness and betweenness. Closeness is simply a measure of how close an investigator is to all other investigators in the network [12, 27]. It assesses how well a person can reach all others in the network, either through a direct connection such as having worked on a publication with another investigator or through indirect connections such as a collaborator of a collaborator. The more direct connections and/or the more connected your collaborators are will result in higher closeness centrality. Betweenness centrality is often described as a measure of “bridging” or “brokerage.” A node with high betweenness serves to connect otherwise unconnected nodes [4, 22]. Finally, reciprocity was examined in the analyses. Reciprocity indicates how likely it is that if node A nominates node B as a collaborator, that node B has also nominated node A as a collaborator [12, 27].

Analysis methods

Three different types of analysis were conducted in order to better assess all aspects of the scientific collaboration network of the CPHHD initiative. The three analyses examined different aspects of the networks, including sample characteristic and structural influences. They also considered effects at both the individual and network levels.

Multilevel linear regression—Multilevel linear regression analysis procedures were used to examine associations between investigator characteristics and collaboration readiness variables and the two measures of network centrality, closeness and betweenness. Multilevel

analysis was needed to control for center-level effects, as there are ten individual centers within the CPHHD initiative and it could be expected that investigators within the same center would be more similar to each other. All regression analyses were done using STATA version 11 [28], with social network metrics calculated using UCINET [3].

Correlations of collaboration type—The relationship between the different collaboration types was assessed using the quadratic assignment procedure (QAP, [13, 19, 20]). This correlational procedure is primarily used to examine multiple relations among the same set of actors or the similarity of networks composed of the same set of actors. The QAP first computes Pearson’s correlation coefficient between corresponding cells of two data matrices. Then, an estimate of the significance of the correlation is computed by permuting the elements of the matrix multiple times and counting the number of correlations. The results are presented as a standard correlation coefficient and indicate the level of similarity or overlap among different relations, such as co-authoring a paper and writing a grant proposal together, from the same set of actors. These test procedures were done using UCINET [3].

Exponential random graph models—Statistical network modeling was conducted using the statnet package [11] for R (R [24]). Exponential random graph models are models of network structure and describe how individual and relational characteristics and network properties can shape the overall structure of a network [14]. They are generative models that use Markov chain Monte Carlo (MCMC) maximum likelihood estimation to estimate model parameters and standard errors, based on a fixed number of nodes. Parameter estimates of zero indicate that the effect being modeled occurs at a rate consistent with chance, a positive parameter suggests that the effect is more prevalent, and a negative parameter suggests that the effect is less prevalent than chance, given the other effects in the model. Exponential random graph models (ERGMs) can be interpreted similarly to logistic regression models, with the outcome being the propensity for a tie to exist given the predictors included in the model. Predictors can be characteristics of the node, for example the gender or academic rank of an investigator. They can also be characteristics of the tie, for example the frequency of interactions among two collaborators. Finally, predictors in an ERGM can be structures or processes of the network itself, for example the reciprocity of ties in the network or the tendency for two investigators who are directly linked to share ties with the same collaborators.

RESULTS

Sample characteristics

Table 1 reports the sample characteristics. The survey was completed by 167 out of 177 key investigators (94 %) involved in the ten CPHHDs at the time of data collection (December 2010 to January 2011). Four of

Table 1 | Sample characteristics

	Frequency, %, or mean (SD)	Range	<i>p</i> value ^a
CPHHD centers (<i>N</i>)	10		
Previously funded	3		
Investigators (<i>n</i>)	177		
Gender			0.48
Male	48.0		
Female	52.0		
Academic rank			0.24
Professor	36.2		
Associate professor	23.7		
Assistant professor	21.5		
Research associate/fellow	4.5		
Other	14.1		
Center role			0.09
Center director (PI)	7.3		
Project PI	21.5		
Research core director	11.9		
Project co-investigator	49.7		
Other	9.6		
Intrapersonal orientation toward collaboration			
Research orientation		1–5	
Disciplinary	2.02 (0.67)		0.87
Multidisciplinary	3.43 (0.94)		0.30
Inter/transdisciplinary	4.09 (0.65)		0.48
Conflict	2.06 (0.84)	1–5	0.40
Views interpersonal collaboration	3.93 (1.37)	1–5	0.76
Collaboration network ^b			
Betweenness	0.01 (0.01)	0–1	0.09
Closeness	0.01 (0.002)	0–1	<0.01
Reciprocity	0.33 (0.26)	0–1	<0.01

^a This value is from ANOVAs run to examine differences between centers

^b These network metrics are only for *current* collaborations

the ten CPHHDs had 100 % response rates—there were none with less than 86 %. There were approximately equal numbers of males (48 %) and females (52 %), and there was also a fairly even distribution of academic rank: assistant professor (21.5 %), associate professor (23.7 %), and professor (36.2 %). There were fewer in the research associate/fellow (4.5 %) and other (14.1 %) categories for academic rank, but that was expected as center directors were asked to name key *investigators* for their centers.

Investigators from the different centers did not differ significantly on gender, academic title/rank, or role within their CPHHD, though there is a trend toward a difference among centers for center role. The main reasons for this are the smaller centers with fewer investigators. Investigators were also coded to their highest role, though many perform several roles within a center. Overall, CPHHD investigators reported high levels of collaboration and low levels of conflict on the collaboration scales. There were no center differences on the intrapersonal measures of collaboration.

There was a trend toward a significant difference between centers for betweenness centrality ($F_{9,167}=1.72$, $p=0.08$) and a significant difference for closeness centrality ($F_{9,167}=10.27$, $p<0.001$). This indicates that

some centers have investigators who are closer or who can more easily reach others in the CPHHD network. Reciprocated ties also differed among the centers, with some having more investigators who were connected and who both confirmed the collaboration. Due to these significant differences, center was controlled for in the random effects model of the multilevel regression and these relationships were further examined in the ERGM analyses.

Individual-level network metrics and collaboration

Multilevel regression models were run to examine associations between investigator characteristics and collaboration variables and network measures of centrality (betweenness and closeness) and reciprocity. Models were run both for collaborations reported as prior to CPHHD and current or ongoing relationships (see Table 2). The collaboration networks, both prior and current, included all types of collaboration and were a measure of *any* collaboration. Center level differences were controlled for in the random effects model.

The greatest differences between the prior and current models were with associations between investigator characteristics and the network outcomes.

Table 2 | Standardized coefficients for intrapersonal orientations toward collaboration, network centrality, and reciprocity

	Betweenness		Closeness		Reciprocity	
	Prior	Current	Prior	Current	Prior	Current
Fixed effects estimates						
Investigator characteristics						
Female	0.18**	0.10	0.21**	0.01	0.19*	0.06
Center role	0.25**	0.11	0.12	−0.02	0.11	0.17*
Academic rank	0.17*	0.08	0.04	0.01	0.06	−0.02
Collaboration						
Disciplinary orientation	−0.08	−0.03	0.03	−0.10 ⁺	0.11	−0.03
Multidisciplinary orientation	0.08	0.05	0.08	0.11 ⁺	0.14 ⁺	0.18**
Inter/transorientation	0.17*	0.20**	0.04	0.14*	0.03	0.05
Conflict	0.002	0.01	−0.07	0.02	−0.10	0.06
Views interpersonal collaboration	0.05	0.13	0.10	−0.03	0.16	−0.17
Random effects estimates						
Level 2 (center)	<0.001	0.23	<0.001	0.70	0.27	0.60

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; ⁺ $p < 0.10$

Investigator characteristics were not significantly associated with measures of betweenness, closeness, or reciprocity for current collaborations. The one exception being an association between the role within the center and reciprocity ($\beta = 0.17$, $p < 0.05$), indicating that an investigator's role within the center (e.g., center director, project PI, co-investigator) was associated with whether or not a tie was reciprocated. This association was not significant for collaborations prior to the CPHHD initiative.

All three of the investigator characteristics included in the models—gender, center role, and academic rank—were significantly associated with betweenness centrality in the prior collaboration network. Females ($\beta = 0.18$, $p < 0.01$), investigators with a more prominent center role (role was coded from 1 to 5 with 5 = center director and 1 = other role with project PI, research core director, and co-investigator making up the other categories; $\beta = 0.25$, $p < 0.01$), and those with a higher academic rank ($\beta = 0.17$, $p < 0.05$) had higher betweenness in collaborations prior to the CPHHD initiative. Betweenness centrality is often used as a measure of network brokerage, with those with high betweenness serving as vital connectors between people who would otherwise not be connected. None of these characteristics were significantly associated with betweenness in the current collaboration network.

For the three network metrics, gender was significant for the prior network only. Females had significantly higher betweenness ($\beta = 0.18$, $p < 0.01$), closeness ($\beta = 0.21$, $p < 0.01$), and reciprocity ($\beta = 0.19$, $p < 0.05$) in the collaboration network prior to the CPHHD, but not in the current network, compared to males.

For the collaboration measures, task conflict and views about interpersonal collaboration were not associated with any of the three network measures. Research orientation, however, was significantly associated with network position and reciprocity, primarily in the current collaboration network. Multidisciplinary orientation was significantly associated with

reciprocity for current collaborations ($\beta = 0.18$, $p < 0.01$), indicating that investigators with a multidisciplinary orientation share reciprocated collaborative ties. For the measures of centrality, an inter/transdisciplinary orientation was significantly associated with betweenness ($\beta = 0.20$, $p < 0.01$) and closeness ($\beta = 0.14$, $p < 0.05$). This indicates that those who are more inter/transdisciplinary in their research orientation are closer to everyone else in the network, and they can more easily reach any other person in the network. This usually results in having a better idea of what is going on in the network because they can more quickly and easily access information. They also tend to serve as vital connectors among people in the CPHHD initiative, acting as “bridges” or “brokers” (see Fig. 1).

Collaboration types

A total of 1,519 current collaborative ties were reported by CPHHD investigators, with 93 % of these between investigators in the same center and 7 % occurring between investigators across different centers. There were 1,135 ties reported prior to the CPHHD initiative, with 83 % within the same center and 17 % between centers. The most common collaborative types were grants (55 % of ties) and work groups (51 % of ties) for the current network. This was similar in the prior network, with publications (45 % of ties) reported at a similar frequency to work groups (49 %). Presentations were the most common collaboration type to be reported by investigators from different centers in both the current and prior networks.

A correlational comparison of the different collaboration types was conducted using the quadratic assignment procedure (QAP). The current collaboration network was significantly correlated with the reported collaboration network prior to the CPHHD initiative. All collaboration types were significantly correlated with each other (see Table 3; $p < 0.001$). The highest

Fig 1 | Current CPHHD collaboration network

correlations reported as prior to CPHHD were grants and presentations ($r=0.619$), grants and publications ($r=0.679$), publications and presentations ($r=0.745$), and presentations and training/mentoring ($r=0.612$). For current collaborations, publications and presentations ($r=0.761$) was the only correlation greater than 0.6. High correlations between current and prior networks included grants ($r=0.605$), presentations ($r=0.668$), and training/mentoring ($r=0.627$). These results suggest that certain types of collaboration (e.g., grants) could lead to others or to a stronger collaborative relationship. These possibilities, however, will need confirmation through future longitudinal analysis.

Network structure and collaboration

Table 4 presents exponential random graph model results for both the prior and current collaboration networks. Three models were examined for each network: the first model included only network structural predictors, the second model added difference effects, and the final complete model included collaboration covariates. The dependent variable in this type of model is the existence of a tie, in this case a collaborative relationship tie. Parameter estimates and standard errors are presented in the table.

Structural effects—Structural effects refer to structural characteristics of the entire network, where a negative parameter estimate indicates that the effect is less prevalent than chance and a positive estimate indicates the effect is more prevalent than chance. The edges term is

negative for all models, which is common when examining real-world social networks, and indicates that ties have a low probability of existing in general. Reciprocity effects were significant for all models, meaning that collaborative ties tended to be reciprocated between investigators. Geometrically weighted edge-wise shared partner (GWESP) effects were significant in the final model only for the prior collaboration network. The positive parameter suggests that triangles (three connected nodes or investigators) tend to cluster together in denser regions of the network and measures how many collaborators each pair of connected investigators shares. The model results for the CPHHD prior collaboration network indicated the tendency for investigators who were directly linked to share multiple collaborators. There were significant negative effects for geometrically weighted dyadwise shared partners (GWDSPs) for both the prior and current networks. This parameter examines the number of partners each pair of investigators in the network shares. For investigators who are not direct collaborators, having a collaborator in common decreases the likelihood that they share additional collaborators. This could indicate that when two investigators share a common collaborator, there was not a need for them to be directly connected. The common collaborator could be a source of particular information and/or expertise that they can get directly from them.

Difference effects—Absolute difference effects included in the models are interpreted slightly differently than the structural and covariate effects. These effects

Table 3 | QAP network correlations of collaboration types

	1	2	3	4	5	6	7	8	9	10	11
1 Current grant	1.0										
2 Current presentation	0.483										
3 Current publication	0.547	0.761									
4 Current training/mentoring	0.463	0.582	0.567								
5 Current work group/committee	0.452	0.38	0.383	0.445							
6 Current other	0.238	0.218	0.205	0.225	0.267						
7 Prior grant	0.605	0.454	0.497	0.417	0.384	0.241					
8 Prior presentation	0.415	0.668	0.547	0.448	0.338	0.239	0.619				
9 Prior publication	0.428	0.531	0.59	0.428	0.343	0.228	0.679	0.745			
10 Prior training/mentoring	0.383	0.459	0.438	0.627	0.348	0.22	0.537	0.612	0.58		
11 Prior work group/committee	0.417	0.387	0.398	0.407	0.56	0.245	0.553	0.496	0.5	0.516	
12 Prior other	0.268	0.26	0.247	0.266	0.299	0.405	0.331	0.334	0.327	0.318	0.354

All correlations are significant at $p < 0.001$

model the absolute difference of a variable between investigators who share a direct tie. A negative estimate indicates a propensity for tied nodes to have less of a difference than expected by chance, or somewhat counter intuitively, to be similar on the variable in question, controlling for the other terms in the model. Conversely, a positive estimate indicates a propensity for tied nodes to be different on the variable in question. In the final model, none of the difference effects were significant for the prior network. That is, controlling for the collaboration covariates and structural predictors, there were no significant difference effects in the collaboration network prior to the CPHHD initiative. There were significant difference effects in the current collaboration model, including one for previously funded centers. This variable was included to examine whether investigators involved with a center that was a part of the first CPHHD initiative (three centers were re-funded) collaborated more with each other than with investigators from the newly funded centers. The results indicated that this was not that case and that investigators from previously funded centers had a propensity to collaborate with others who were different than them on this variable or who were involved with newly funded centers. There was a significant negative difference effect for academic rank, more specifically for assistant professors. Assistant professors tended to collaborate more with other assistant professors. Finally, center and research core directors tended to collaborate more with those in other roles. This is not surprising given the nature of these roles to begin with, as facilitators and overall administrators of the research at each site, and also due simply to the fact that there are a limited number of investigators in these roles.

Collaboration covariates—Parameter estimates for the collaboration covariates examine the direct effects of these variables on the tendency for a collaborative tie to exist. The only significant parameter for these covariates was a negative estimate for disciplinary research orientation. This indicates that those with a more disciplinary research orientation were less likely to report collaborative ties.

DISCUSSION

We examined scientific collaboration among investigators involved in a large, cross-disciplinary research initiative to address health disparities. The current collaboration network was clearly “inward” focused, with more dense ties within each center than between them, and more clustering or clumping together of investigators, likely around specific research projects. Only 7 % of current ties were between investigators from different centers. This is to be expected from centers and projects in the process of planning and implementing the first year of research activities and formalizing internal, site-specific infrastructures and protocols. Therefore, it was not surprising to see an initial decrease in attention to collaborations outside

Table 4 | Structural and investigator predictors of scientific collaboration in the Centers for Population Health and Health Disparities (CPHHD)

	Collaboration prior to CPHHD			Current collaboration		
	(1) <i>B</i> (SE)	(2) <i>B</i> (SE)	(3) <i>B</i> (SE)	(4) <i>B</i> (SE)	(5) <i>B</i> (SE)	(6) <i>B</i> (SE)
Edges	-3.60 (0.76)***	-3.94 (1.32)**	-3.61 (15.28)	-3.53***	-4.21 (1.28)***	-4.59 (4.29)
Structural predictors						
Reciprocity	2.02 (0.07)***	2.16 (0.07)***	2.05 (0.57)***	1.36 (0.86)***	1.34 (0.08)***	1.36 (0.19)***
GWESP (clustering) ^a	1.35 (0.45)**	1.58 (0.79)*	1.62 (0.61)**	1.86 (0.54)***	2.26 (0.80)**	1.75 (1.48)
GWDSP (structural equivalence) ^b	-0.18 (0.02)***	-0.18 (0.05)***	-0.18 (0.01)***	-0.27 (0.01)***	-0.27 (0.07)***	-0.24 (0.02)***
Absolute difference effects						
Female	0.15 (0.04)***		0.14 (0.15)		0.06 (0.04)	0.05 (0.05)
Previously funded center	0.09 (0.04)*		0.05 (0.14)		0.12 (0.04)**	0.14 (0.06)*
Academic rank						
Professor	-0.11 (0.05)*		-0.10 (0.27)		-0.14 (0.05)**	-0.10 (0.08)
Associate professor	-0.11 (0.05)*		0.005 (0.65)		-0.08 (0.06)	-0.04 (0.07)
Assistant professor	-0.15 (0.07)*		-0.13 (0.14)		-0.39 (0.08)***	-0.31 (0.09)***
Research associate/fellow	-0.04 (0.07)		-0.07 (0.54)		-0.18 (0.08)*	-0.12 (0.08)
Center role						
Center director	-0.05 (0.04)		-0.01 (0.13)		0.14 (0.05)**	0.19 (0.08)*
Project PI	-0.22 (0.05)***		-0.15 (0.32)		0.10 (0.06) [†]	0.09 (0.09)
Core director	-0.35 (0.07)***		-0.17 (0.90)		0.22 (0.07)***	0.22 (0.08)**
Co-investigator	-0.02 (0.20)		0.27 (3.03)		0.32 (0.17) [†]	0.41 (0.27)
Collaboration covariates						
Disciplinary research orientation		0.03 (0.31)				-0.08 (0.04)*
Multiresearch orientation		-0.02 (0.13)				←-0.001 (0.03)
Inter/transresearch orientation		0.03 (0.52)				0.02 (0.06)
Task conflict		-0.02 (0.42)				0.06 (0.07)
Interpersonal collaboration		-0.07 (0.70)				0.11 (0.11)

* $p<0.05$; ** $p<0.01$; *** $p<0.001$; [†] $p<0.10$

^a The tendency for nodes who share a tie to have shared collaborators

^b The tendency for any pair of nodes (regardless of whether a tie exists between them) to have shared collaborators

an investigator's own center in the current network. It will be interesting to see how between center collaborations unfold as the CPHHD initiative matures. One of the secondary goals of the initiative is to foster cross-center collaborations to facilitate linkage of projects and data and sharing expertise between the different centers. An indication that there should be a future increase in collaboration between centers is the percentage of cross-institutional ties that were reported prior to the beginning of CPHHD activities (17 % of all ties). It should be assumed that as the initiative matures, the percentage of cross-center ties should increase, and eventually exceed, the pre-funding level of scientific collaboration.

The finding that investigator characteristics (gender, academic rank, and role within the center) are associated with network position and reciprocity in the prior but not the current CPHHD networks is unexpected and will take additional longitudinal analysis to fully understand. However, this may suggest a shift in how disparate investigators and groups are connected when they become a part of a large grant-funded research center. This initiative may somehow serve as an "equalizer" in terms of network position, where expertise and knowledge become more important for collaboration than formal structures and roles. It could also be that the initiative provides different formal collaborative structures not related to these particular characteristics.

The high correlation between collaboration on grants and publications and presentations is intuitive. It also makes sense that prior collaboration is highly correlated with current collaboration. Future longitudinal analysis of CPHHD collaboration should provide more detail about which types of collaboration are likely to lead to other types of partnerships. For example, do collaborators on a grant first collaborate on publications or is it the other way around? The results reported here are cross-sectional and only correlational in nature and cannot define causal direction.

The reasons for why shared partner distribution was significant for the prior, but not the current, collaboration network is noteworthy. It may indicate a more distributed network where collaborators seek to minimize redundancy in the current network while seeking parsimony. It could also be an indication of more formal processes occurring as a result of the initiative. There is also the notion of workload when involved in the beginning of a large initiative with each center having multiple projects that could lead to investigators needing to manage the greater demands of getting a new project up and running by being more selective and strategic with their collaborations early on. If an investigator's collaborator already has a tie to another investigator, it may be more efficient with limited social and collaborative resources to *not* share a tie with that same investigator as they can already be reached indirectly. This may not be the case when investigators are not going through this initial "start-up" phase, when collaborative resources and time are

more plentiful. A direct tie with an investigator to whom a collaborator is connected may be more efficient in this case, cutting down on the need to relay information through another.

The finding that assistant professors tend to collaborate more with other assistant professors could suggest that they are more comfortable collaborating with those in similar positions than with investigators with more seniority. It could also be that they have smaller, less diverse, collaboration networks during this early career stage. As one of the secondary aims of the CPHHD initiative is to foster training and to help develop the next generation of health disparities researchers, this finding could also suggest an area for future intervention. More attention, emphasis, and support could be given to the development of strong mentoring relationships between junior and more senior investigators involved in the CPHHD centers in the coming years of the initiative. Exposure of young investigators to this type of initiative, with a focus on cross-disciplinary collaboration as key to addressing complex health outcomes across multiple levels, could help shape views of the problem, the scope and innovativeness of the questions that are asked, and the methods used in future health disparities research and in the long term be one of the more enduring outcomes of the initiative.

The results from these data have some limitations that must be noted. The investigators included in the analysis were, in part, funded to participate in a large cross-disciplinary initiative and may show some bias toward support for this type of work or organization of research. They may also be predisposed to collaboration and/or a cross-disciplinary outlook. Therefore, generalizability is naturally limited. There was variation among the research orientation scale, a measure that attempts to capture the cross-disciplinary orientation of an investigator. However, the mean scores for the three identified factors at the baseline measurement point showed that CPHHD investigators report high multi- and inter/transdisciplinarity and ceiling effects are possible. To fully understand whether these investigators are predisposed to collaboration and/or cross-disciplinarity, some type of comparison group is needed.

Another potential limitation is that the extent to which the collaborations among CPHHD investigators were "manufactured" or organized through formal structures outside of an investigator's control versus more informal but potentially more meaningful channels is not clear. For example, if some or all of the centers constructed working groups around center activities and "assigned" investigators to these groups, are true and meaningful collaborations still being captured with these data? It also is not known if groups that are "manufactured" or "assigned" can or cannot provide impactful and important scientific collaboration. This is something that was not explicitly measured on the baseline instrument but could be examined in future longitudinal measurement.

A final limitation of note is that due to the baseline nature of the study, there is not yet a

way to link collaboration to scientific impact or outputs. This will be an important task in future work, as it is not enough only to have successful collaboration if the science being produced from such relationships does not produce meaningful impact.

Additional empirical and theoretical work is needed to understand the processes of efficient and effective scientific collaboration and to understand the impact it can have on addressing complex public health problems and moving the science of health disparities forward. Work in the organizational sciences has examined teams and collaboration (for example, see [2, 21, 25]), but additional understanding of how that type of research, traditionally conducted with businesses and corporations and other types of performance teams, translates to scientific teams and team-based research is needed. There are many similarities and general features of teams that are likely to easily translate, but there are other aspects specific to teams entrenched in the scientific enterprise that may require modification of theories, methods, and frameworks before they can more effectively inform the processes and practices of team-based research.

This research provides additional understanding of team-based research using empirical data provided by scientists involved in an ongoing team science initiative. Much of the study of team science to date has had to rely on archival data, such as publication or grant information, or is measured retrospectively after the research has concluded. The present study adds to a growing body of empirical studies of team science by using self-report data collected directly from researchers currently participating in team-based research. Network analytic methodologies also move beyond self-report measures alone to provide a greater depth of understanding of collaborative social structures and patterns, which can serve to inform the design and support of future team science efforts.

Acknowledgments: The CPHHD evaluation working group provided guidance for survey development and implementation and provided feedback and direction on the structure of the manuscript and early drafts. The CPHHD evaluation working group includes the following: Tom Belin (University of California, Los Angeles), Shirley Beresford (University of Washington), Susan Cahn (University of Illinois, Chicago), Jarvis Chen (Harvard University), Luis Falcon (University of Massachusetts Lowell), Darla Fickle (Ohio State University), Dave Flum (University of Washington), Melissa Gorsline (Ohio State University), Tim Johnson (University of Illinois, Chicago), Peter Kaufmann (National Heart, Lung, and Blood Institute), Miyong Kim (Johns Hopkins University), Molly Martin (Rush University), Dorrie Rhoades (University of Colorado, Denver), Shobha Srinivasan (National Cancer Institute), Maihan Vu (University of North Carolina), and Richard Warnecke (University of Illinois, Chicago).

The authors would like to thank the CPHHD Steering Committee for feedback on early drafts. The CPHHD Steering Committee includes the following: Alice Ammerman (University of North Carolina), Dedra Buchwald (University of Washington), Lisa Cooper (Johns Hopkins University), Alex Ortega (University of California, Los Angeles), Electra Paskett (Ohio State University), Lynda Powell (Rush University), Beti Thompson (Fred Hutchinson Cancer Research Center), Katherine Tucker (Northeastern University), Richard Warnecke (University of Illinois, Chicago), and David Williams (Harvard University).

Finally, the authors would also like to thank Shobha Srinivasan and Kara Hall of the National Cancer Institute for guidance on survey development and implementation, analysis, and feedback on early manuscript drafts.

Conflict of interest: Janet Okamoto, the lead author, declares that she has no conflict of interest. The CPHHD evaluation working group is a collection of individuals serving as representatives for each of the ten Centers for Population Health and Health Disparities, NCI, and NHLBI and as a group have no conflicts of interest to report.

Adherence to ethical standards: All procedures were conducted in accordance with the ethical standards of the responsible committee on human experimentation and with the Helsinki Declaration of 1975, as revised in 2000.

- Adler NE, Rehkopf DH. U.S. disparities in health: descriptions, causes, and mechanisms. *Annu Rev Public Health*. 2008; 29: 235-252.
- Aram JD, Morgan CP, Esbeck ES. Relation of collaborative interpersonal relationships to individual satisfaction and organizational performance. *Adm Sci Q*. 1971; 16(3): 289-297.
- Borgatti SP, Everett MG, Freeman LC. *UCINET for Windows: software for social network analysis*. Harvard: Analytic Technologies; 2002.
- Borgatti SP, Mehra A, Brass DJ, et al. Network analysis in the social sciences. *Science*. 2009; 323(5916): 892-895.
- Bozeman B, Fay D, Slade CP. Research collaboration in universities and academic entrepreneurship: the state-of-the-art. *J Technol Transfer*. 2013; 38: 1-67.
- Contractor NS, Monge PR. Managing knowledge networks. *Manag Commun Q*. 2002; 16: 249-258.
- Croyle RT. The National Cancer Institute's transdisciplinary centers initiatives and the need for building a science of team science. *Am J Prev Med*. 2008; 35(2): S90-S93.
- Grimwade A, Park P. How they measure up: scientific institutions. *The Scientist*. 2003; 17(20): 17.
- Hall KL, Stokols D, Moser RP, et al. The collaboration readiness of transdisciplinary research teams and centers: findings from the National Cancer Institute's TREC year-one evaluation study. *Am J Prev Med*. 2008; 35(2): S161-S172.
- Hall KL, Stokols D, Stipelman BA, et al. Assessing the value of team science: a study comparing center- and investigator-initiated grants. *Am J Prev Med*. 2012; 42(2): 157-163.
- Handcock, M.S., Hunter, D.R., Butts, C.T., Goodreau, S.M., Morris, M. *statnet: software tools for the statistical modeling of network data* 2003. URL <http://statnetproject.org>.
- Hawe P, Webster C, Shiell A. A glossary of terms for navigating the field of social network analysis. *J Epidemiol Community Health*. 2004; 58(12): 971-975.
- Hubert LJ, Schulz J. Quadratic assignment as a general data analysis strategy. *Br J Math Stat Psychol*. 1976; 29: 190-241.
- Hunter, D.R., Handcock, M.S., Butts, C.T., Goodreau, S.M., Morris, M. *Ergm: a package to fit, simulate, and diagnose exponential-family models for networks*. *J Stat Soft*. 2008, 24 (3). URL <http://www.jstatsoft.org/v24/i03/paper>.
- Jehn KA. A multimethod examination of the benefits and detriments of intragroup conflict. *Adm Sci Q*. 1995; 40: 256-82.
- Jones BF, Wuchty S, Uzzi B. Multi-university research teams: shifting impact, geography, and stratification in science. *Science*. 2008; 322: 1259-1262.
- Klein JT. *Crossing Boundaries: Knowledge, Disciplinarity, and Interdisciplinarity*. Charlottesville: University Press of Virginia; 1996.
- Klein JT. Evaluating interdisciplinary and transdisciplinary collaborative research: a review of the state of the art. *Am J Prev Med*. 2008; 35(2): S116-123.
- Krackhardt D. QAP partialling as a test of spuriousness. *Soc Networks*. 1987; 9: 171-186.
- Krackhardt D. Predicting networks: nonparametric multiple regression analysis of dyadic data. *Soc Networks*. 1988; 10: 359-381.
- Letsky MP, Warner NW, Fiore SM, eds. *Macro cognition in Teams: Theories and Methodologies*. Farnham: Ashgate; 2008.
- Marsden PV. Egocentric and sociocentric measures of network centrality. *Soc Networks*. 2002; 24(4): 407-422.
- National Cancer Institute Request for Application (RFA). NIH-supported Centers for Population Health and Health Disparities (CPHHD), 2008. <http://grants.nih.gov/grants/guide/rfa-files/RFA-CA-09-001.html>.
- R Development Core Team. *R: a language and environment for statistical computing*. R Foundation for Statistical Computing 2012, Vienna, Austria. URL <http://www.R-project.org>.

25. San Martin-Rodríguez L, Beaulieu M-D, D'Amour D, et al. The determinants of successful collaboration: a review of theoretical and empirical studies. *Journal of Interprofessional Care*. 2005; 19(1): 132-147.
26. Sargent LD, Waters LE. Careers and academic research collaborations: an inductive process framework for understanding successful collaborations. *J Vocat Behav*. 2004; 64: 308-319.
27. Scott, J., & Carrington, P.J. (Eds.). *The SAGE Handbook of Social Network Analysis*. SAGE publications 2011.
28. StataCorp. *Stata statistical software: release 11*. College Station: StataCorp LP; 2009.
29. Stipelman B, Feng A, Hall KA, et al. The relationship between collaborative readiness and scientific productivity in the Transdisciplinary Research on Energetics and Cancer (TREC) centers. *Ann Behav Med*. 2010; 39: S143.
30. Stokols D, Misra S, Moser RP, et al. The ecology of team science—understanding contextual influences on transdisciplinary collaboration. *Am J Prev Med*. 2008; 35(2): S96-115.