Author’s response to reviews

Title: Drawbacks and benefits associated with inter-organizational collaboration along the discovery--development--delivery continuum: A cancer research network case study

Authors:

Jenine K Harris (harrisj@wustl.edu)
Keith G Provan (kprovan@eller.arizona.edu)
Kimberly J Johnson (kijohnson@gwbmail.wustl.edu)
Scott J Leischow (SLeischow@azcc.arizona.edu)

Version: 2 Date: 11 April 2012

Author’s response to reviews: see over
Dear Drs. Mittman & Eccles,

Enclosed is our revised paper, “Drawbacks and benefits associated with inter-organizational collaboration along the discovery—development—delivery continuum: A cancer research network case study.” We appreciate the helpful and detailed reviewer suggestions and made several major changes to the paper based on these suggestions and the guidance from the editor, including:

1. Added operational definitions and exact wording for the questions eliciting the data analyzed in this study;
2. Added an explanation of the network characteristics (non-uniform degree distribution and transitivity) that differentiate observed networks like this one from simple random graphs, and are the reason behind the use of geometrically weighted terms in the statistical modeling of networks;
3. Added Table 2 showing the proportion of time spent on each activity by each organization;
4. Added a sensitivity analysis in the data management section to examine the use of unconfirmed ties;
5. Added additional information clarifying the criteria for model fit and model selection;
6. Added id numbers to the nodes in the network figure and put the benefits and drawbacks in order for easier interpretation;
7. Incorporated additional text explaining ties to the 2010 Provan, Leischow, and colleagues study and making explicit the differences between the 2010 study and this one;
8. Added more context for the findings related to the delivery network and for how the findings from this study contribute to existing literature on collaborative science;
9. Revised the limitations sections in light of the revisions to the methods.

In addition to these major changes, we made numerous minor edits throughout the document. The major changes, including substantial edits and additions of text, are highlighted in the manuscript. Point-by-point responses to reviewer concerns are included below.

Thank you in advance for your time and effort. We look forward to hearing from you.

Sincerely,

Jenine K. Harris, Ph.D.
Assistant Professor
George Warren Brown School of Social Work
Washington University in St. Louis
314-935-3522 (p)
314-935-3756 (f)
Response to Reviewer 1

Major Compulsory Revisions

1. The authors situate their study within the framework of team science. However, the study fails to consider the nature of these collaborations which makes the connection to team science somewhat difficult to determine. For instance, are they interdisciplinary, transdisciplinary, or multidisciplinary collaborations? Provan et al. (2008) for example notes the nature of collaborations as a key determining factor in these collaborative networks, and, at minimum, this would seem to be an important variable in the analysis.

We agree that the nature of the collaborations between any two partners is likely to be important, especially for explaining outcomes. The Provan et al. (2008) paper explored this issue, as have others, but like most team science research, the focus was on collaboration among individual scientists. Also, their study did not distinguish whether the focus of the research being conducted was discovery, development, or delivery. In contrast, our paper addresses this DDD distinction directly, but not other aspects of the nature of the collaboration process, and examines the relationships among a diverse set of organizations involved in the research process. While the absence of further detail on the nature of each collaborative relationship we examined can certainly be considered a limitation of our work, no study is able to adequately capture all aspects of collaboration. Our research complements the team science literature but is not specifically a team science study since we did not examine the actual work relationships among individual scientists. To reflect this fact, we have specified more clearly how our study differs from the team science literature, thus offering an alternative way of thinking about scientific collaboration.

2. Additionally, collaborative readiness for such collaborations and collaborative leadership are key factor suggested to determine their success (Gray, 2008; Hall et al., 2008; Stokols, Misra, Moser, Hall, & Taylor, 2008; Syme, 2008). As such, there appears to be a missing variable in understanding the link between the pattern of interorganizational collaboration and the benefits/drawbacks of such collaboration.

This is also a good point, although we do not have data to examine this possible missing variable. We focus only on the relationship between collaboration and attitudes toward benefits and drawbacks from this collaboration during the three sequential components of cancer research. If top managers in organizations see substantial benefit to collaboration and few drawbacks, they are likely to encourage future collaborative efforts by the researchers who work for their organization. It may well be the case that readiness to collaborate is an important additional variable to consider, and along with other variables, would make for an interesting future research project. We make note of all this in the conclusion section of the paper when we discuss limitations.

3. It is unclear how the particular model was chosen. Goodreau suggests minimally the inclusion of k-triangles, as do several others (Robins, Pattison, Kalish, & Lusher, 2007; Robins, Snijders, Wang, Handcock, & Pattison, 2007; Snijders, Pattison, Robins, & Handcock, 2006). The attributes of the organizations considered are minimal. Also, it is not
clear what measure you are associating with each of the variables, (e.g., attribute associated with edge, homophily)

The geometrically weighted terms proposed by Hunter et al (GWD, GWDSP, and GWESP) and included in the full model (Table 3) are equivalent to Snijders and colleagues’ geometrically weighted degree, alternating k-twopath, and alternating k-triangle terms (respectively), with a few minor differences in the calculation of Snijders and colleagues’ degree term and the Hunter and colleagues version of GWD. Like the k-triangles and other two terms recommended by Snijders, these terms are designed to account for the non-uniform degree distribution and transitivity seen in most networks; here, however, they did not improve model fit and so were not retained. Some additional descriptive measures of the degree distribution and transitivity in the networks may shed some light on why these terms were not useful; this is now included in the paper. Specifically, the transitivity and degree distribution in the DDD networks do not exhibit the extreme deviations from what would be seen in a simple random graph and where they did deviate, the local terms were sufficient to capture this characteristic.

References:


We collected a few additional organizational attributes in the survey (e.g., which of 18 cancers were addressed by each organization, which of 13 activities were emphasized at the organization, how much spending was devoted to cancer activity), however, we did not feel that these were central to the research question at hand. In addition, the number of independent variables in the model was already quite high, and, although there is little (if any) discussion of power related to ERG-p* models, we felt that, without a strong theoretical reason, increasing the number of independent variables was not wise.

All of the terms included are node attributes. We added this parenthetical to the table labeling in Table 3 and clarified interpretations accordingly.

4. Finally, assuming edge was the parameter associated with the drawbacks/benefits, I am not sure I follow the logic. Are you trying to suggest that the perceived benefits/drawbacks of collaboration make it more or less likely for the organization to be in collaborations? If so, how do you interpret the multiple negative weights for benefits? An alternative that would be important to address is that these perceptions are outcomes of collaboration. Organizations that collaborate more tend to have more negative perceptions of collaboration than organizations that have fewer collaboration partners.

The wording of the drawbacks and benefits questions would indicate that these perceptions followed collaboration rather than preceded it. We have clarified the interpretations of the coefficients throughout the manuscript as suggested; thank you for this helpful comment.
Minor Essential Revisions

5. Measures paragraph- Operational definitions for discovery, development, and delivery and the exact operational definition of collaborative involvement should be provided.

The definitions consist primarily of examples; they have been added to the measures section.

6. Measures paragraph-Second to last sentence-Which earlier studies were consulted to develop this list? Are there any assurances of the sufficiency of the list?

As noted in the revised version (last paragraph of methods section), the list of benefits and drawbacks was developed based on earlier work by Provan and colleagues [2003] combined with discussions with Coalition research committee members.

7. Data Management- Paragraph 1- While I agree that the authors should symmetrize the matrix, there are a few minor corrections that need to be made. First, ERGM can handle directed ties. While Goodreau uses an undirected network as well, there are developed measures for arcs, reciprocity, etc. for use with directed networks.

This sentence was meant to justify making the links binary (rather than weighted) and not the decision to make them undirected; we have clarified this section.

8. Second, I would like the authors to consider the disagreements again. Many studies of friendship networks, for example, use the presence of agreement between the respondents as evidence of the reliability of the measure.

We discussed the issue of disagreements among respondents regarding their relationships and believe that the strategy we chose is appropriate. Specifically, unlike a friendship tie, which is a very personal relationship and makes sense to confirm since both members of a friendship are personally involved, a collaborative tie between organizations doesn’t necessarily directly involve both people in the dyad who are acting as representatives of their larger organizations. Because of this, we believe that, if one person knew of a collaboration, it is more likely than not that the collaboration existed. In past research, network scholars have used both confirmed and unconfirmed ties, with no clear agreement as to which approach is best; there is no specific guidance regarding this decision around ERGM (Cranmer, 2011). We now discuss this issue much more fully in the Data Management section. We also include a sensitivity analysis comparing the patterns of benefits and drawbacks across confirmed and unconfirmed versions of the networks.

9. Data Management-Paragraph 2- It is unclear to me why you felt it necessary to dichotomize these values. This is not a limitation associated with ERGM. Was there something about the nature of the distributions that led you to dichotomize?

We were concerned about the number of independent variables in the model; including all three categories for each of the benefits and drawbacks would result in a model with 14 additional independent variables, greatly increasing its complexity.
10. I would like the authors to mention which package they used for analysis. Because of the inclusion of AIC, BIC, I would assume the ERGM package in R. This is not a small matter in this case because ERGM because the packages are new and rely (assuming you are using maximum likelihood estimators) on simulations.

We used R-statnet; we added this information to the methods section.

11. Results- Paragraph 3- because your definition of collaboration appears to include the sharing of resources, presence of joint programs, etc., I would be cautious about interpreting betweenness centrality in terms of information flow. In a purely communication network, this would be a logical interpretation, but it does not necessarily hold for other types of relations.

Given earlier remarks regarding model selection and the structural terms in the model, and concerns by both reviewers about the betweenness measures, we revised the paper to focus more on the degree distribution and degree centrality rather than betweenness. The section in question was deleted as a result.

**Response to Reviewer 2**

Overall: This manuscript is certain to make a meaningful contribution to implementation science (as well as across the research continuum) as well as to the Science of Team Science. The findings support some likely anticipated relationships across the research continuum (e.g., positive association between knowledge sharing and collaboration) and highlight important new relationships, which can serve as the basis for future investigations and indicate potential leverage points for future policies and interventions to enhance scientific collaborations (and thereby leading to more efficient and effective investments and advancements in science).

Thank you for the encouraging comments.

12. I assume given the "open peer review system" of this journal, a key article was mistakenly listed as "author citations". The presumed "author citation" was identified by this reviewer as follows: Provan, Keith G., Leischow, Scott J., Keagy, Judith, and Nodora, Jesse. 2010. Research Collaboration in the Discovery, Development, and Delivery Networks of a Statewide Cancer Coalition. Evaluation and Program Planning, 33: 349-355.). After reading this associated manuscript, I believe the current manuscript may benefit from additional connections to this existing work (which document alternative analyses of the same general dataset used in this current work under review) as well as potentially clarifying decisions in the current manuscript that could be seen as potentially conflicting by readers of both articles. I provide more specific suggestions in the related sections below.

See below for our responses to your specific questions regarding our earlier paper. In addition, we have added text and specific references to the 2010 paper in several sections of our this revised manuscript.
Abstract:
Minor Essential Revisions
13. There seems to be a discrepancy in findings between the abstract, table, and discussion section related to - "In the delivery network, those who collaborated were less likely to see enhanced influence on treatment and policy and greater quality or frequency of publications as benefits of collaboration."[in abstract]. I believe this relationship is about the development network?

Yes, thank you for noticing this discrepancy. We have corrected the abstract.

14. "are" is missing in the second to last sentence of the results section.

We corrected this omission.

Background:
Minor Essential Revisions
15. "s" missing from "conflict" in 3rd sentence of 2nd paragraph

This was corrected.

Methods
Major Compulsory Revisions
16. Please clarify the sample size for the study, including:
   a. Addressing/discussing potential issues of power,

There is no guidance around issues of power for ERGM that we know of.

   b. Identifying the unit of analysis for the ERGM (e.g., explaining and specifying the # of edges or dyads)

We added a sentence to the analysis section explaining that the outcome is the likelihood of a tie and that there were 153 possible ties.

17. Include number for each network (discovery, development, delivery)
   a. Further clarify n=153 for dyads - where did this number come from, what is it based on?

We added the formula for calculating the number of dyads to the data management section. It is n*(n-1)/2, which is 18*17/2=153 in this case.

   b. Discuss issues of dependency in the data for the benefits and drawbacks given the data were collected at the organizational level.

We were not sure what you meant by this comment. Our assumption is that the concern is related to the fact that organization-level data were collected from individuals – our key informants. As we note in our response to point 28, although use of a single key informant is
common in organizational network studies, we recognize the limitations of this approach. Our key informants were asked to respond on behalf of their full organization or the cancer research unit of their organization. We cannot be sure that an individual’s responses to questions about benefits and drawbacks might simply reflect their own personal experiences and not those of others in their organization. We now note this in the conclusion. If our interpretation of your comment here is not correct, we will certainly address your concern.

c. Decision to use unconfirmed vs confirmed ties

In the previous study using this data set (Provan et al., 2010), confirmed ties were used since the focus of that paper was on comparing the structure of each of the three networks; hence, we used a more conservative approach where the existence of a tie had to be confirmed by both participants. This paper is more focused on ties than on whole network structure. We now discuss our rationale explicitly (comparing to the 2010 paper) in the Data Management section.

18. (Discretionary Revisions) Consider including information from Provan et al. (2010) regarding the similarity in patterns of relationships with confirmed and unconfirmed. Also, although I found the argument for using unconfirmed ties reasonable - I felt I wanted perhaps a bit more explanation based on the specifics of this study given the fact that the previous study based its results and made an argument for using the confirmed relationships. Was the decision based on a power issue in order to conduct the analyses in the current study?

We made the decision to include unconfirmed ties for the reasons stated in the manuscript in the Data Management section. There is currently no specific guidance around rules for the coding of ties when using ERGM (see Cranmer, 2011 now cited), so the selection of a strategy is based on what is most appropriate for the study at hand. In this case, a collaborative tie between organizations is less personal than a friendship tie and may not directly involve both people in the dyad. We believe that if one person knew of a collaboration, it is more likely than not that the collaboration existed, especially since the key informant may simply not be aware of the tie because it does not involve him/her directly. Since the focus of the 2010 study was on explaining overall network structure, full agreement about the presence of every tie was an appropriate conservative approach for indicating who was connected to whom. This study’s focus is on the relationship between dyadic ties and attitudes (regarding benefits and drawbacks) and not on whole network structure per se, suggesting the value of retaining all ties.

We conducted a sensitivity analysis examining the similarity or differences in patterns of benefits and drawbacks when using confirmed versus unconfirmed ties. We found the patterns quite similar across the confirmed and unconfirmed tie networks. This is discussed in the Data Management section but in the interest of space (and because we did not want to detract from our main argument), extensive details are not provided. We can, of course, provide this information to the reviewer and the editor if you feel it is important to do so.

Finally, we do not know of any specific guidance around power in ERGM.

Minor Essential Revisions

19. Table 1:
c. Include n’s in table

The frequencies were added to Table 1.

d. Within the benefits and within the drawbacks, please place the items in numerical order to increase the reader’s ability to quickly glean the information presented.

The benefits and drawbacks were put in descending order.

d. Please include exact item related to the respondents’ estimates of the amount an organization spent on discovery, development, and/or delivery.

This was added to the Measures section.

20. Discretionary Revisions

a. Consider adding more specific rationale to help clarify the choice to use betweenness centrality vs. degree centrality (as used in prior study). There is a good explanation of what betweenness centrality is, but given the discrepancies between the two studies (Provan et al. 2010) a brief statement of how the metric is specifically relevant to the research questions in this study and/or more relevant for such questions than degree centrality may be useful.

In revising the descriptive measures of the networks to demonstrate the reasons why the structural terms typically used in statistical network models did not increase model fit in this study, we decided that degree centrality was a more important metric to focus on. We deleted the sections on betweenness and modified the figure to show degree rather than betweenness.

b. Include a table that includes each organization and the percentage of resources developed to each - discovery, development, and delivery.

We added this table as Table 2.

Results:

21. Minor Essential Revisions

c. Figure 1:

i. Please include a more complete key that indicates the metric (betweenness centrality) related to the size of the nodes

We added the node size metric to the legend; see response to #21 above for more detail.

ii. Please label the nodes with either a numerical value or some other identifier (e.g., representation of organization’s name) to enable the reader to see how each organization changes its relationship with others across the different network diagrams. For instance, the large central “discovery” node in the Discovery Collaboration Network is less prominent in the
other diagrams – labeling the node will enable the reader to determine its position in the other diagrams and see potential patterns.

We numbered each node to facilitate comparison across networks; thank you for this suggestion, we believe readers will find this information useful, especially given the addition of Table 2.

d. Figure 2:
   iii. Given this paper is not a paper focused on methodology, I do not think this figure is necessary as Table 3 provide ample and necessary information regarding model fit.

The AIC and BIC are not recommended for more complex ERGM since they were measures developed for models meeting the assumption of independence; although they do have the tendency to be consistent with the simulated measures of fit. We have reduced the complexity of Figure 2, but have retained it.

Discretionary Revisions
22. Can you explain the discrepancy between being involved in either the discovery, development or delivery networks with another organization, but not spending any resources on that particular “D”?

Involvement at a low level in any of the three networks does not necessarily require resource expenditure. An organization/respondent may be involved because of their interest in being “in the loop” regarding that particular D, primarily for information sharing purposes or relationship development that may ultimately help them in other D&D areas.

Discussion and Conclusion:
Minor Essential Revisions
23. Discuss limitations due to sample size and dependencies in data.

We added some additional discussion of dependencies in the data in the section on model estimation; however, there is no guidance that we know of on sample size limitations for ERGM.

24. Greater emphasis or discussion is warranted for the fact that nearly all of the drawbacks were significant in the delivery network whereas only geographical differences were significant for discover and development, perhaps highlighting the intense challenges experienced when collaborating towards a goal that involves stakeholders with different expectations, goals, and incentives.

We added text to the results section making the point that the delivery model was different from the other two models with numerous significant barriers. We also added text to the conclusion highlighting the challenges of collaborating on delivery research that the reviewer identified here and in comment #25.

25. More clearly discuss the positive relationships found for the delivery network, for instance, perhaps insufficient resources to support collaboration is an indication that collaborations
were occurring, but that resources needed to support the collaborations in the delivery context were underestimated when the collaborations were initiated.

We added some text to the discussion about the positive relationships in the delivery network, and specifically about the insufficient resources barrier.

Discretionary Revisions

26. Discuss the finding that more benefits than barriers overall were identified - consider this in contrast with the fact that only one significant relationship was found for the benefits - how might this be related constrained range in response for the benefits verses the greater degree of variability found for the drawbacks?

Good point. This may be due to perceived benefits of collaboration being consistent to what is experienced in an actual collaboration. We added a sentence to the discussion to address this possibility.

27. Please carefully describe the interpretations of the “double negatives” to reduce the cognitive load on the reader (e.g., negative relationship with a drawback - difficulty due to geographical differences is negatively related to collaboration)

We revised the results section for clarity, eliminating double negatives wherever possible.

28. What are the implications/limitations/considerations for the interpretation of items such as “loss of control/autonomy over decisions” that are answered by one representative of the organization?

Although use of a single key informant is common in organizational network studies, we certainly recognize the limitations. Our key informants were asked to respond on behalf of their full organization or the cancer research unit of their organization. We cannot be sure that such issues as loss of control/autonomy might simply reflect the individual’s own personal experiences and not those of others in the organization. We now note this in the conclusion.

29. Was there a specific reasons that the names of the organizations were included in the prior manuscript and not the current one? The inclusion of names or at least numerical labeling may help with interpretation.

We added numerical labeling to help with interpretation.

30. Might there be extraneous factors that position the centrally positioned “development” organization throughout these networks? For instance, is this largest organization with the largest overall budget and greatest number of investigators – and if so, how might these variables influence the network characteristics?

It is certainly true that other, possibly extraneous factors like resources, might determine network position. However, our study was not an attempt to explain why one organization versus another
might be in a central network position. Rather, we simply determined which organizations were involved in the network, with which other organizations, and in what ways (i.e., D, D, or D).

31. Consider potential connections to additional literature regarding impact of distance and factors that influence collaborative success, such as:

Great suggestions. We added references and text to the section of the discussion about the distance barrier.

32. Consider potential connections to literature related to collaboration with respects to translation at the delivery stage of research such as:

We reviewed Stokols work (which was cited in the introduction) and identified an article by Emmons and colleagues that was cited in 2008 Stokols article (from the same special issue). The Emmons article speaks specifically to barriers in delivery work. We have incorporated this into the discussion.

33. Consider potential connections to literature related to outcomes such as new methods as well duration and intensity of collaboration as related to outcomes (including frequency of publications), such as:

We reviewed these articles, and, although we did not reference them in the current draft since we did not examine outcomes, they will be key in our future directions in this area and we thank the reviewer for pointing them out.