3rd International Conference of Science of Science & Innovation, July 1-3, 2024
National Academy of Sciences, Washington, DC, USA

Building and Characterizing a Collaboration Network of the US National Cancer Institute's Extramural Workforce

Margaret Gratian, Nan Ma, Laura Ellis, Holly Wolcott, and Christine Burgess

Keywords: *scientific collaboration, measuring collaboration, National Cancer Institute*
Field: *Social Network Analysis*

Extended Abstract

Scientific collaboration has become increasingly important to accelerate progress in biomedical research[1]. Collaboration can be a key method to address challenges to progress that arise from increased specialization of fields, siloed institutions, fragmented data ecosystems, rapid generation of large and complex datasets, and the growing expenses of conducting research and clinical trials[2]. In the United States, nationwide efforts such as the White House Cancer Moonshot Initiative[3,4] and the National Cancer Plan[5] bring the goal of fostering greater collaboration to the forefront of cancer research.

As the United States' principal agency for cancer research and training and the world's largest funder of cancer research, the National Cancer Institute (NCI) supports a substantial extramural workforce of investigators through grants.[6] Given the importance of collaboration in cancer research, we seek to understand and characterize collaboration among the NCI-supported cancer research community.

In this talk, we will discuss how we conceptualized and developed a collaboration network of NCI's extramural workforce. We will provide insight into the network analysis methodology we applied and share our initial findings. We will also discuss challenges with measuring collaboration and how we sought to mitigate them in our work. Lastly, we will discuss how one can use this network and accompanying analysis methodology to inform strategies and policies for fostering collaboration.

NCI supports its extramural workforce through a variety of grant mechanisms.[7] These mechanisms are divided into groups representing the types of funding NCI provides, such as Research Grants, Career Development Awards, or Research and Training Fellowships. Because we were interested in measuring scientific collaboration, we chose to focus on Principal Investigators (PIs) who were awarded a research-focused grant to build our network. This includes Research Project Grants such as the R01, R37, R56, and U01 and Research Center Grants such as the P50 or U54.

Scientific collaboration may take many forms and some aspects of it may be more tangible than others. Critical to the conceptualization of our network was the concept of *measurable* collaborations. To determine if PIs had collaborated – that is, worked together in an interdependent fashion toward a scientific outcome – we required evidence of the collaboration in the form of co-funding on an NIH base project or co-authorship on a publication. Not all collaborations will result in co-funding or co-authorship. So, while this does not give us a holistic view of collaboration, it does give us a measurable one.

We built our network from a PI-perspective, adding PIs as nodes of the network if they had a qualifying research award in Fiscal Year (FY)* 2017 – FY 2022. Edges were added based on the presence of co-funding or co-authorship between two PIs and therefore represent pairwise collaborative events. To better capture trends in collaboration over time, we included co-funding between PIs that occurred between FY 2012 – FY 2023 and co-authorship between PIs that occurred between 2012 – 2023.

Our resulting network consists of over 10,000 PIs representing the nodes of our network. We identified co-funding between these PIs on over 6000 unique NIH base projects and co-

authorship on over 200,000 publications. These projects and publications culminate in over one million network edges and over 200,000 unique pairwise collaborations between PIs.

Once we had developed our network, we applied a variety of network analysis methodologies to characterize the network. We focused on identifying and applying metrics that could tell us about the *strength* of the network. In other words, could we see evidence of a robust research community among NCI's extramural workforce? For example, node degree indicates the number of unique collaborators per PI, while the number of connected components provides insight into how interconnected the network is overall.

We also conducted various temporal analyses of the data to understand how collaborations shift over time. We explored whether co-authorship is likely to follow co-funding and investigated the role NCI funding plays in sustaining collaborations and promoting new collaborations to form.

We identified several challenges with measuring collaboration that must be considered when interpreting the results of an analysis. First, collaborations can take time to produce measurable events – the absence of co-funding or co-authorship does not imply the absence of a collaboration. Second and related to this point, PIs in the network began collaborating at different points in time and therefore reach collaborative milestones (co-authorship or co-funding) at different points in time as a result. Third, when attempting to measure the impact of an intervention on collaboration, it is important to recognize that most interventions are likely introduced in a setting where collaboration is already occurring.

Despite these challenges, a thoughtfully constructed dataset representing measurable collaborations can still be a useful tool to understand and characterize collaborations. Our talk will conclude with insights into how we have used our network of extramural PIs to evaluate initiatives such as the Cancer Moonshot and its goal of fostering greater collaboration and how we plan to use our network going forward.

*The United States Fiscal Year begins on October 1st and ends on September 30th. For example, Fiscal Year 2017 began on October 1st, 2016 and ended on September 30th, 2017.

References

1. National Research Council. *Enhancing the Effectiveness of Team Science.* Washington, DC: The National Academies Press. 2015. https://doi.org/10.17226/19007.
2. Blue Ribbon Panel. *Cancer Moonshot Blue Ribbon Panel Report.* October 17, 2016. 2016. Accessed March 22, 2024. https://www.cancer.gov/research/key-initiatives/moonshot-cancer-initiative/history/blue-ribbon-panel-report-2016.pdf
3. National Cancer Institute. *Cancer Moonshot℠ Progress.* Updated February 1, 2023. Accessed March 22, 2024. https://www.cancer.gov/research/key-initiatives/moonshot-cancer-initiative/progress
4. Sharpless N, Singer D. *The Cancer Moonshot: A Midpoint Progress Update.* Cancer Currents Blog. January 11, 2021. Accessed March 22, 2024. https://www.cancer.gov/news-events/cancer-currents-blog/2021/cancer-moonshot-midpoint-progress-update
5. National Cancer Institute. National Cancer Plan Goals. Accessed March 22, 2024. https://nationalcancerplan.cancer.gov/goals
6. National Cancer Institute. National Cancer Institute Overview and Mission. Updated April 6, 2018. Accessed March 22, 2024. https://www.cancer.gov/about-nci/overview
7. National Institutes of Health. Types of Grant Programs. Updated April 11, 2023. Accessed March 22, 2024. https://grants.nih.gov/grants/funding/funding_program.htm