A Network Perspective on Interdisciplinary Disease Funding and Research Output

Anbang Du* Michael Head † Markus Brede*

Interdisciplinary health research involves integrating knowledge and methods from various subjects to tackle complex problems and foster innovation. Prominent funders like the National Institute of Health, the U.S. National Science Foundation and the UKRI all have established their own funding projects for interdisciplinary research. There is a significant body of work in the field of scientometrics on funding allocations [1]; however, up to our best knowledge, there is limited research examining interdisciplinarity, particularly in the context of health funding. Closing this gap will contribute to a better understanding of health research funding and efforts to promote interdisciplinarity.

Our research addresses this gap by representing diseases and their connections as a network, which we use to quantify how interdisciplinary funding allocation has changed over time. For this purpose, we treat a set of 31 infectious diseases as interconnected elements and compare the resulting networks constructed for funded and non-funded research outputs. As we will discuss in the talk, this approach to quantify interdisciplinarity in research allows us to assess whether combinations of disease areas receive funding in proportion to their respective correlation strength as anticipated by the research output: i.e. if research into particular combinations of diseases is "under-" or "over-funded".

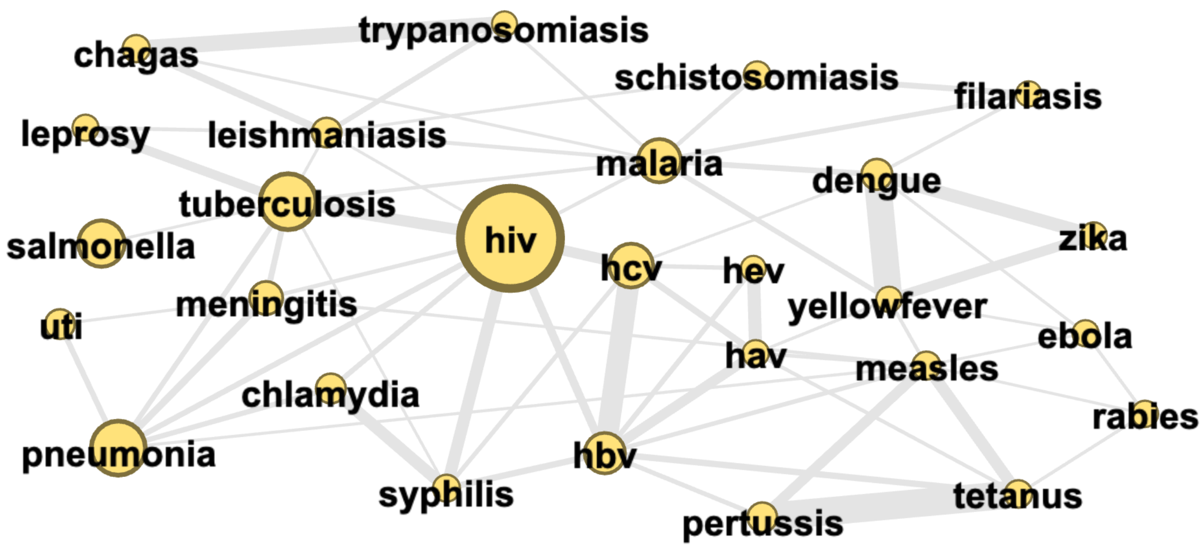
More specifically, we have collected data about publications on these 31 diseases from the Web of Science Core Collection database (WOS) and have also used an global infectious disease funding dataset created at the University of Southampton [2]. Connection strengths between disease areas are assessed by extracting data on the co-occurrence of terms corresponding to infectious diseases in the titles, abstracts, or author keywords of papers contained in WOS. Supposing we consider disease areas i and j and find s_i and s_j occurrences of the disease areas and c_{ij} co-occurrences, connection strengths are quantified by the cosine similarity $c_{ij}/(s_i s_j)$ [3]. Carrying out the above procedure results in a network whose connection strengths represent correlations between disease areas as perceived in all published work on the topics. We argue that this network can be interpreted as representing cross-disciplinary connections, as illustrated in Figure 1A. To track the funders' allocations into interdisciplinarity, we also construct a network for publications in WOS, but restrict this to literature that acknowledges funding.

Networks for connection strengths for all WOS-represented research and research that acknowledges have been obtained covering the period between 1990 and 2023. We then proceed by comparing the evolution of the two correlation networks over time. For example, Figure 1B gives plots that illustrate the evolution of interdisciplinarity for the time periods $1990-2006$, $2007-2015$, and $2016-2022$, each containing a sample of published papers of roughly the same size. To assess differences in correlations between pairs of diseases, we plot correlation strengths measured in publications resulting from funded research vs correlation strengths as measured in all publications excluding funded publications, so that the panels contain data for $31\times 30/2$ pairs of diseases for the respective time periods. Beyond labels for outliers, the panels in Figure 1B also contain regression results and a guiding line $y = x$ for ease of comparison. Presenting data in the above way allows us to easily assess whether connection strengths in funded research differ from connection strengths in unfunded research.

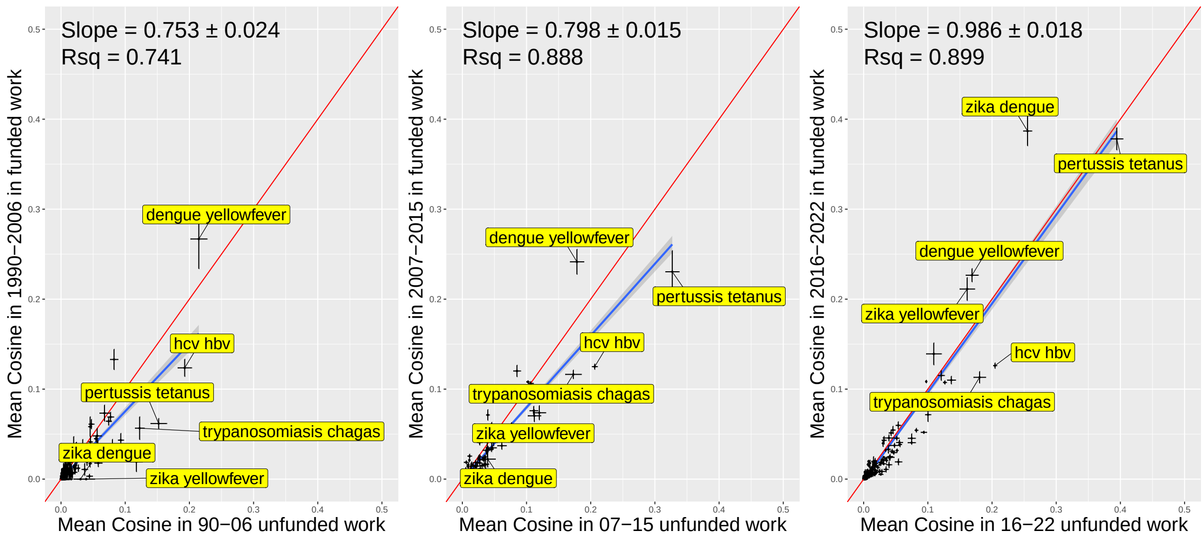
Analysing the evolution of the correlation networks in Figure 1B, we make the following observations: (i) funded research about infectious diseases has generally become more interdisciplinary over time as reflected by the increase in the slope from 0.75±0.02 (1990-2006) to 0.80±0.02 (2007-2015) to 0.99±0.02 (2016-2022); (ii) the dispersion of degrees of interdisciplinarity in infectious disease research has reduced over time as reflected by the increase in the R-squared from 0.74 (1990-2006) to 0.89 (2007-2015) to 0.90 (2016-2022); (iii) changing trends in particular outliers over time. Beyond the analysis discussed above, the talk will also give an overview of an inter-temporal network comparison of the two correlation networks and discuss the practical implications of findings from the network comparisons for funding allocation in infectious disease research. The evolution of the interdisciplinarity of prominent global funders including the NIH and Gates Foundation will also be discussed in the talk.

Our research offers a network-based perspective on the interdisciplinary disease funding landscape, which can serve as a valuable resource for informing global health funders and stakeholders such as the WHO and shaping future resource allocations in this field.

*Agents, Interaction and Complexity, Faculty of Physical Sciences and Engineering, University of Southampton
†Clinical Informatics Research Unit, Faculty of Medicine, University of Southampton



(A) Network of Infectious Diseases in WOS 1990-2022



(B) Evolution of Interdisciplinarity of Infectious Diseases Funding

Figure 1: (A) The network of diseases representing the average cross-disciplinary connection for all literature related to the 31 diseases in WOS 1990-2022. Node size corresponds to the fraction of published work for each disease, and link width corresponds to correlation strengths. HIV is the most interdisciplinary disease with a weighted degree of 0.56, while urinary tract infection is the least with 0.074. (B) The evolution of funders' interdisciplinarity in infectious diseases for the time periods 1990-2006, 2007-2015, and 2018-2022, respectively. The y-axis represents the correlation strength for pairs of diseases in funded research, the x-axis represents the correlation strength for pairs of diseases in unfunded research. The slope of the fitted line, the standard error of the slope, and the R-squared value for each plot is given in the top left corner. The red line in each plot represents the 45 degree line. Yellow labels represent the top six outliers noted in the 2016-2022 period, and their corresponding positions are shown in the 1990-2006 and 2007-2015 plots.

References

1. Kalachikhin, P. A. Scientometric Instruments of Research Funding. *Scientific and Technical Information Processing* **45,** 28–34 (2018).
2. Head, M. G. *et al.* The allocation of US$105 billion in global funding from G20 countries for infectious disease research between 2000 and 2017: a content analysis of investments. *The Lancet Global Health* **8,** e1295–e1304 (2020).
3. Eck, N. J. V. & Waltman, L. How to normalize cooccurrence data? An analysis of some well-known similarity measures. *Journal of the American Society for Information Science and Technology* **60,** 1635–1651 (2009).