2023

Virus Outbreaks and International Scientific Collaboration: A Quantitative Analysis of NIH GenBank Metadata (1992-2018)

Will Devitt
Syracuse University

Sarah Bratt
University of Arizona

Jian Qin
Syracuse University

Jeff Hemsley
Syracuse University

Recommended Citation

Dewitt, W., Bratt, S., Qin, J., & Hemsley, J. (2023).Virus outbreaks and international scientific collaboration: A quantitative analysis of NIH GenBank metadata (1992-2018).

Virus Outbreaks and International Scientific Collaboration: A Quantitative Analysis of NIH GenBank Metadata (1992-2018)

Description/Abstract

Countries around the world suffer from outbreaks of viral diseases. These outbreaks are worsened by socio-economic inequality, both nationally and internationally. To combat these outbreaks, scientific research is conducted. Our paper analyzes the NIH GenBank virus research network (a net record of virus research from 1992-2018) to discover viral outbreaks and ask how the network responded in the event of these outbreaks, with interest in the changes in collaboration between research teams, increased activity, and when + where this activity occurred with the objective of examining the equity of the international response to public health events. We conclude that the GenBank virus network during this time period was generally equitable - rich Western nations consistently contributed to research on viruses that did not necessarily effect them. However, this begs the question of the definition of equity - should poorer countries, who have historically suffered at the hands of their Northern counterparts, be funded to conduct their own virus research, or is the position of the United states acceptable in the pursuit of research equality, as long as the US acts selflessly? We further discuss the definition of an equitable network, propose methods for increasing international research equality, as well as examine the global impacts of disease outbreak and discuss the motivation behind increasing equity and the international scientific community's ability to react to crises.

Keywords

Collaboration networks, GenBank metadata analysis, virus outbreaks

Disciplines

Library and Information Science | Scholarly Communication

Creative Commons License

Virus Outbreaks and International Scientific Collaboration: A Quantitative Analysis of NIH *GenBank* Metadata (1992-2018)

Will Devitt[1], Sarah Bratt[2], Jian Qin[3], Jeff Hemsley[4]

336 Hinds Hall Syracuse, NY 13022

Metadata Lab[1,]*

Abstract

Countries around the world suffer from outbreaks of viral diseases. These outbreaks are worsened by socio-economic inequality, both nationally and internationally. To combat these outbreaks, scientific research is conducted. Our paper analyzes the NIH GenBank virus research network (a net record of virus research from 1992-2018) to discover viral outbreaks and ask how the network responded in the event of these outbreaks, with interest in the changes in collaboration between research teams, increased activity, and when + where this activity occurred with the objective of examining the equity of the international response to public health events. We conclude that the GenBank virus network during this time period was generally equitable - rich Western nations consistently contributed to research on viruses that did not necessarily effect them. However, this begs the question of the definition of equity - should poorer countries, who have historically suffered at the hands of their Northern counterparts, be funded to conduct their own virus research, or is the position of the United states acceptable in the pursuit of research equality, as long as the US acts selflessly? We further discuss the definition of an equitable network, propose methods for increasing international research equality, as well as examine the global impacts of disease outbreak and discuss the motivation behind increasing equity and the international scientific community's ability to react to crises.

---

1. Introduction

Science of science research has analyzed how scientific collaboration networks respond to disease outbreaks. In the case of a global public health event, federal funding agencies and private institutes mobilize and organize resources to respond to the crisis. The urgency of

---

[*] Corresponding author
Email address: `jqin@syr.edu` (Metadata Lab)
[1] wsdevitt@syr.edu
[2] sebratt@arizona.edu
[3] jqin@syr.edu
[4] jjhemsle@syr.edu

a global pandemic can often require organizations to redirect resources, leaving some topics and geographic areas neglected in favor of others, including human resources. Scientists may direct their attention away from current research to focus on the outbreak.

Ample studies of scientific collaboration using publication metadata have revealed how scientific collaboration networks respond to public health crises. However, as data intensive research figures more prominently into addressing infectious disease outbreaks such as SARs, MERs, and SARS-CoV-2 (COVID-19), it is crucial to understand how scientists collaborate in producing critical data about the virus, not only publications, in response to such crises.

The study reported in this paper investigates the scientific collaboration networks through mining the metadata from *GenBank*, one of the international open data repositories curating genetic sequences that is maintained by the National Center for Biotechnology Information (NCBI), to understand how scientific collaboration networks react to infectious disease outbreaks. The focus of this study is on how scientific and human capital are (re)directed to address crises. In addition to GenBank metadata, we also extracted funding data from NIH RePORT data portal to integrate it with the GenBank metadata. Using social network methods and natural language processing techniques, we analyzed the collaboration networks related to infectious diseases to identify patterns of funding, network structural properties, and scientific knowledge outputs. We examined the collaboration networks related to several outbreaks, including Dengue, West Nile, Zika, and HIV, and analyzed collaboration network patterns and funding allocation as well as their impact on the wake of global pandemics. The paper concludes with a discussion and future work by addressing the question: how equitable are scientific collaboration networks when a global disease outbreak occurs?

2. Relevant Research

Studies of scientific collaboration networks have generalized that these networks usually are scale-free [1], namely the degree distribution following a power law, and self-organizing with tendencies of assortative mixing (nodes with high degrees are more likely connected with those similar to themselves) [2]. While these network properties have been proved to be common across disciplines, we know very little about how scientific collaboration networks self-organize in response to public health crises such as infectious disease outbreaks, much less about the factors affecting such self-organizing processes. This literature review focuses on the equality of scientific collaboration.

Inequalities in health lead to poorer life outcomes overall. Inequalities have micro and macro scale impacts on individuals and countries. Research on global health inequalities exists shows that these exist in health research both within and between countries[3] - whereas capacities to act on these inequalities are incredibly important and a primary goal of science as a method to improve the world and the lives of the people living within it. That is the larger goal of our research - to understand the actions and tendencies of the global scientific community, and to propose actions on the root causes of these differences.

Research output on health inequalities (papers like ours) has grown exponentially over the last 30 years [4]. This is indicative of the growing interest in examining and solving these problems that impact billions of people daily.

A 2000 World Health Organization bulletin considers this. Composed of a critical reflection on the issue of global health inequality, it proposes direct action and a re-framing of the issue at hand in a way that we, the authors of this paper, absolutely agree with. It proposes two initial steps for action on the issue: the promotion of recognition among scientists of the great importance of implementing just and distributed health policy, regardless of differences, and the reformation of health policy goals to recognize, specifically, the differences in health conditions between poor and rich countries. [5] This is the primary investigative point of our paper: to understand if the work distribution within the GenBank virus research network is equitable. I.e, are rich nations supporting poor nations by contributing significant amounts of resources to combat diseases that primarily affect those poorer nations? Is there a relative lack of research on these diseases at all? Who does the research in these various situations?

3. Methods

3.1. Discovering Top Network Viruses and Outbreaks

For this analysis, we used two sets of virus research activity – submissions to the network and publications. These datasets extend from the years 1992 to 2018, with each row in each file storing a virus and the corresponding number of submissions or publications on that virus for that year. To discover the top overall viruses in the combined (submission and publication) network, we iterated through the entire dataset, counting the total number of occurrences of specific viruses throughout the years to determine the top studied viruses in the network. With this determination, we then trimmed the dataset to include only values for the identified top viruses. We then created a time-lapse visualization [6] enabling us to identify the major increases in virus research and when they occurred, hypothesising that a substantial increase in network activity would correspond to a real-world disease outbreak. With this visualization of the trends, we were able to discover three distinct increases in network activity in certain years corresponding to global health events. These are detailed under Section 3, Findings.

3.2. Analyzing Nationalities of Researchers

3.2.1. Purpose

Global health events are fundamentally unequal due to the vast gaps in wealth and health outcomes between the richest few nations and the rest, especially between the global North and South. [3] In our analysis of the GenBank virus network, we sought to examine the international equity of virus research. In essence, the question is: who conducts research during global disease outbreaks? Do rich Western nations like the United States contribute significantly to research on outbreaks that affect other, poorer countries? Is there a global research response, or is it localized to the affected nations? Poorer nations, especially in the global south, suffer from a relative lack of health research (eg on specific viruses that largely effect, say, tropical regions) compared to rich Western nations. We seek to analyze the international equity of the GenBank virus network via the examination of collaboration metrics - and to determine the potential ways the international virus community could become more capable of swift reaction to outbreaks in disadvantaged countries and regions.

3.3. Extracting Data Sources

The data used in this analysis is part of a broader project on scientific collaboration networks. The data used for this project was originally extracted from the National Center for Biotechnology Information's (NCBI) nucleic acid sequence database, *GenBank*. We extracted the metadata records via FTP download from the earliest entries available to 2018. The data was then parsed into a relational database to enable filtering and data manipulation. The records of interest can be grouped into two types of references: data submissions ("direct submissions") and publications (including white papers, conference proceedings, but the bulk are journal articles). In addition to *GenBank* records, the NCBI Taxonomy database was merged with the GenBank records to add organism classification information. The database of NCBI metadata records contain bibliographic information such as author name, year of publication, journal, organism taxonomic id, among others. A subset of these NCBI metadata were extracted to analyze research equity in public health outbreaks (described in details in the subsections below).

3.3.1. Analyzing Research Output Metrics

We determined that the best way to measure a country's research output is by determining the number of funded direct submissions to the network. The number of submissions, following from the number of funded authors, is viewed as a quantitative measure of the degree of funding a nation is committing to the research effort on a particular virus, and the degree of collaboration within the network. To this end, we analyzed the entire database of direct network submissions from 1992-2018, identifying the nationality of authors and the journals they submitted to within the timelines of each virus outbreak, identified via the methods above. The outbreaks identified by the first methods section were an increase in Zika virus research from 2015-2017, Dengue virus research from 2007-2012, and West Nile virus research from 2008-2012. We then determined the various sources of this research by country, in order to uncover who does research in public health outbreaks. Determining if countries experiencing public health events are left out to dry or assisted by the international research community is one of the primary goals of our research.

3.4. Analyzing Collaboration Network Metrics

We wanted to understand the trends in collaboration between authors belonging to the virus communities, so we collected the data which consisted of information about all the viruses and segregated it based on the tax-id's belonging to a particular virus group. This segregated data was the base for our further analysis. The methodologies for investigating each virus network were quite similar, but the inferences were different for each virus group. The methodology is as follows: With the segregated file we have for a particular virus, our attribute of interest is the 'authors' field. We analysed only the authors who collaborated with at least one other author and build an edge list using all combinations. Once the edge list is ready, we filter out the tuples belonging to a particular year and simplify the network to contain only one edge if there are multiple edges between two nodes. At the conclusion of our data pre-processing, we discover and examine network statistics such as number of nodes, number of edges, density, assortivity, giant component size and more for each year. These yearly statistics were visualized to draw insights and trends.

4. Findings

4.1. Uncovering Outbreaks via Network Activity

From the network analysis described above, we discovered a few notable increases in research at unique time intervals for specific viruses. These network increases suggest significant impact of the studied viruses upon the global population, as the GenBank network is definitively international. The most notable network increases, and the ones we chose to study, are an increase in Zika virus research from 2015-2017, Dengue virus research from 2007-2012, and West Nile virus research from 2008-2012. Further investigation revealed that each of these increases in research correspond to a real-world outbreak or epidemic of the relevant virus, generally affecting multiple countries globally. [7] [8] [9] [10] [11] [12] [13] [14] [15] [16] [17] [18] This is the extent of our scope, as other outbreaks were less notable.

4.2. Zika Outbreak, 2015-2017

After discovering a substantial increases in Zika Virus research from 2015-2017 [6], and verifying that there was a worldwide Zika health crisis in this time-frame, [7] [8] [9], we analyzed the number of authors and submissions to the network per country, and compared this with the nations most affected by the outbreaks. This outbreak primarily affected South and Central American nations as well as the United States and had a further (yet more minor) global impact. Specifically, Brazil as the host of the 2016 FIFA World Cup was particularly affected by Zika.

4.2.1. Zika Equity

The following figure represents the volume of submissions, per country, to the network on the Zika virus during the length of its global outbreak. The United States and United Kingdom are highly represented, contributing 46.4% and 29.3% of Zika submissions respectively, producing most of the research output on Zika during the outbreak. Following are China at 8.1%, Mexico at 6.9%, and Brazil at 5.9%. Notable is the research coming from Brazil and

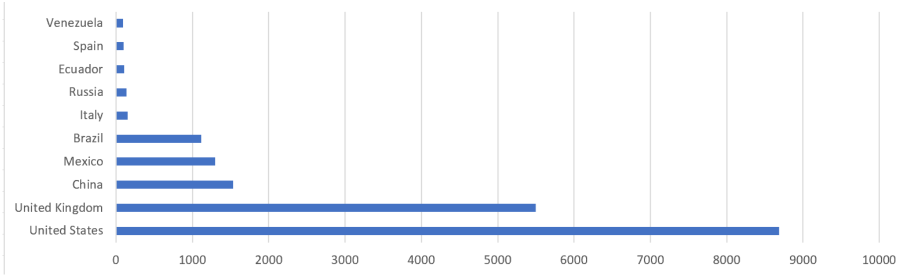

Figure 1: Zika direct submissions per country, 2015-2017

Mexico (and to a lesser extent, Venezuela and Ecuador.) As nations that bore the brunt of the Zika outbreak, their significant research contributions and submissions represent the high level of funding those countries committed to the research effort. This suggests that, for the Zika outbreak, the GenBank virus network has a fair degree of equity.[19] Were the United States and United Kingdom, with their vast resources, not represented in the data and instead there were only large outsized contributions from the affected nations, we would be deeply concerned about the equity of the network. However, from these results it appears that the network is indeed fairly equitable. Additionally, this data suggests that China is also a significant contributor to international virus research efforts. Zika research was conducted on a large scale in nations that were not directly affected, as well as in the nations who suffered to a lesser extent. This level of international collaboration is a good sign of equitable policies. Being that it is the most recent virus of study, this suggests international collaboration within the virus network is increasing. Next, we must analyze the funding received from the NIH by the affected nations and otherwise, and the distribution of those funds.

4.3. Dengue Outbreak, 2007-2012

We discovered a significant jump in Dengue virus research output over the years 2007-2012 [6], and verified the corresponding existence of a worldwide Dengue outbreak within this time period. This outbreak was global, and largely consisted of the Dengue virus 1, 2, and 3 subtypes. It most significantly affected poorer, more tropical nations. Some such nations that experienced substantial Dengue outbreaks were Brazil, India, Pakistan, Vietnam, and Singapore. [10] [11] [12] [13] [14] It is worth noting that in the last 50 years, Dengue virus has had about a 30fold increase in incidence.[20]

4.3.1. Dengue Equity

The included figures represent our findings on the research output per nation by volume of submissions on the three Dengue subtypes during the global outbreak from 2007-2012. As we can see, the United States is represented here to such a degree that the contributions from other nations appear inconsequential, making up a whopping 98.1% of all Dengue

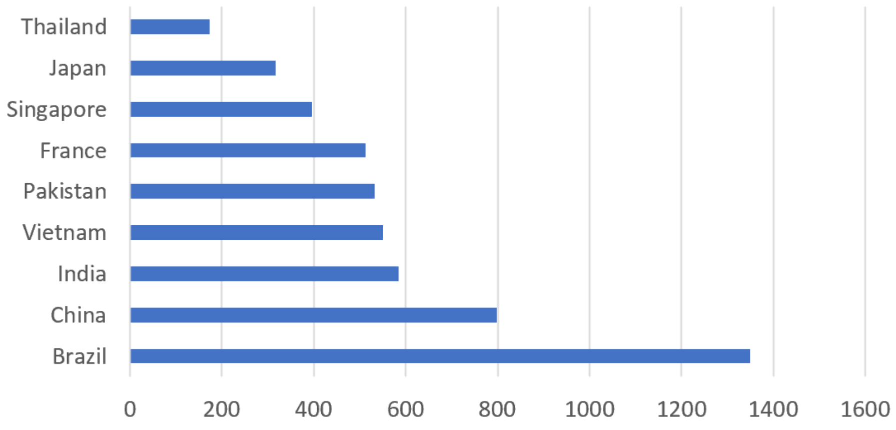

Figure 2: Dengue direct submissions per country, 2007-2012, US excluded

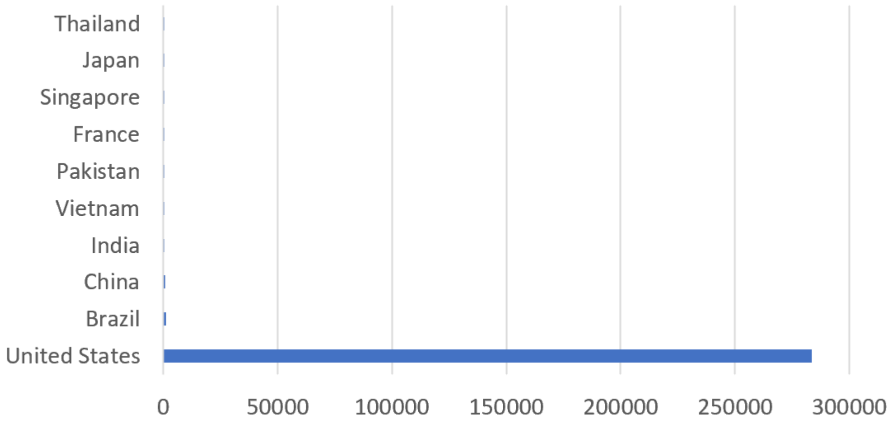

Figure 3: Dengue direct submissions per country, 2007-2012, US included

research during this timeframe, with Brazil, China and India following up at 0.46%, 0.27%, and 0.2% respectively. Thus, we have included two figures as to better understand the research outputs of each involved nation, while acknowledging the massive role the United States played. The massive submission volume of the United States, while not being itself affected to the same degree poorer and tropical nations were [11], again suggests that the network is fairly equitable. Given additional factors that would prompt research interest, like the discovery of the strain Dengue 3 [14] and the 30fold increased incidence of Dengue worldwide over the past 50 years, there would be much motivation for research powerhouses (ie the United States) to conduct research on viruses like Dengue, that do not have much negative impact on the United States itself.

Investigating the figure that excludes the United States also brings insight. Many of the badly affected countries, such as Brazil, India, Pakistan, Vietnam, and Singapore contributed to the research effort as well - quantitatively not to the same degree as the United States, but very significant themselves, continuing the pattern of affected countries contributing to scientific research. The question remains - to where did the NIH distribute funding for Dengue research?

4.4. West Nile Outbreak, 2008-2012

Our data analysis uncovered significant research activity on the West Nile virus from the years 2008-2012 [6] and corroborated this finding with a real-world outbreak, primarily in the United States, with minor outbreaks around the same time in Italy and Spain. [15] [16] [17] [18] This major US outbreak was big news at the time, and resulted in government actions towards mosquito control nationwide, among other public responses.

4.4.1. West Nile Equity

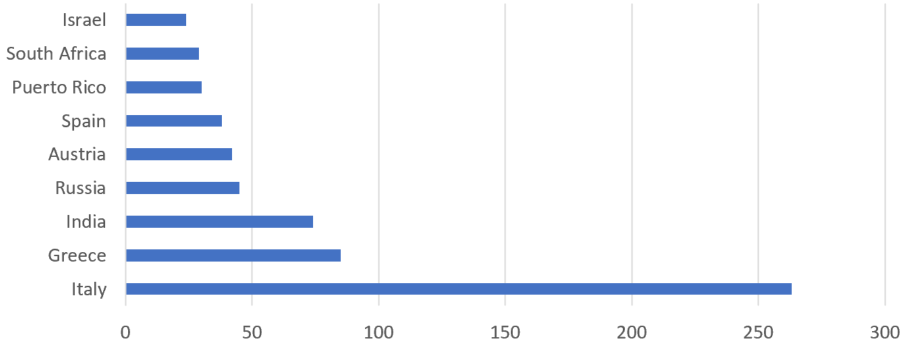

Figure 4: West Nile direct submissions, 2008-2012, US excluded

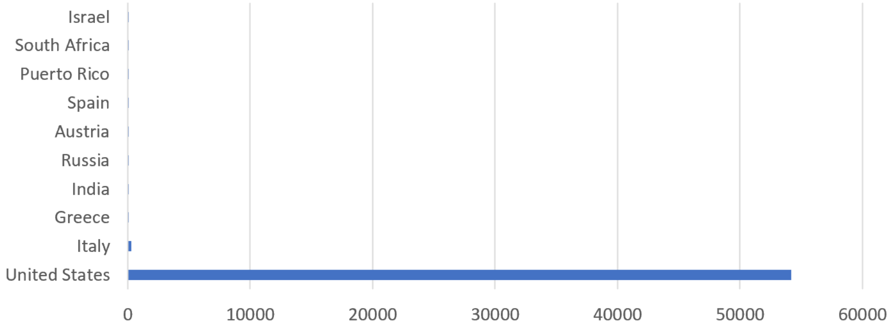

Figure 5: West Nile direct submissions, 2008-2012, US included

The included figures represent the same as the others - the volume of submissions per country on the West Nile virus over the given time period of the outbreak. The United

States again contributed a majority of West Nile research within this timeframe, coming in at 98.8%. Italy is up next at a minuscule 0.47%. Quantitatively, affected nations such as Italy and Spain have made minimal contributions compared to the volume of research coming out of the United States, but we still see a continuation of the pattern of affected nations contributing to research. The West Nile outbreak is a somewhat unique case - in this time period, the United States was far and away the most affected nation. As a wealthy western nation, it is positioned well.

4.5. Collaboration Network Metrics

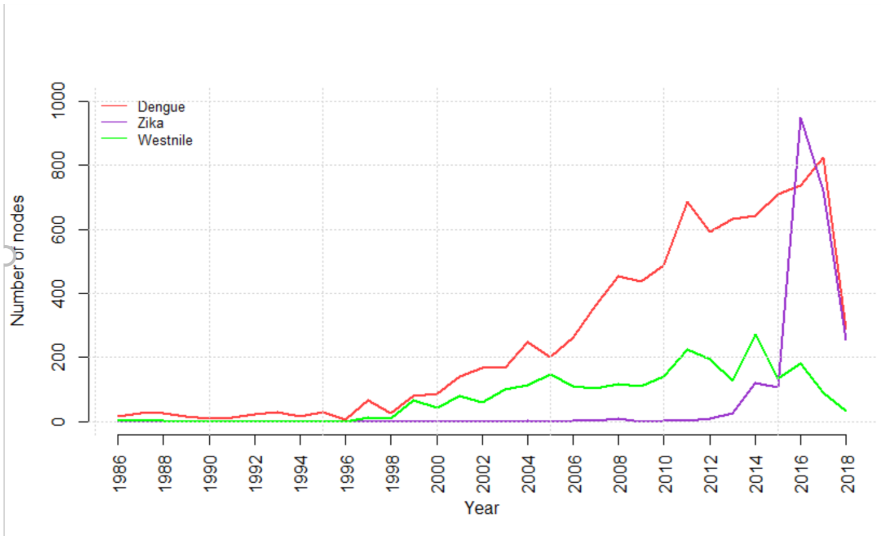

Figure 6: Total virus network node (author/research team) count, 1992-2018

Within the Dengue virus network, we observed that from 1996 to 2017, the number of nodes followed a positive trend - which implies that many authors became interested in Dengue viruses at that time. However, after 2007, the trend saw a very steep positive curve - suggesting even more authors became interested in the research after that year. This provides further verification that our analysis of the Dengue network is correct. As more authors begin to research about a topic, the density of network decreases and so was observed for Dengue - the massive increase in network volume decreased the network density.

Within the Zika virus network, we observed that the number of nodes and edges exploded significantly in 2016, which can be linked with its outbreak during that period. The edges in the network represent relationships between authors. It was only during the outbreak that Zika virus received attention from the researcher community - before 2015, there was almost no research being conducted. The size of the giant component was also noticed to decrease during this period. Also, our inference about density hold true as initially, when there were few nodes (authors), the density was very high. When the number of nodes increased, the

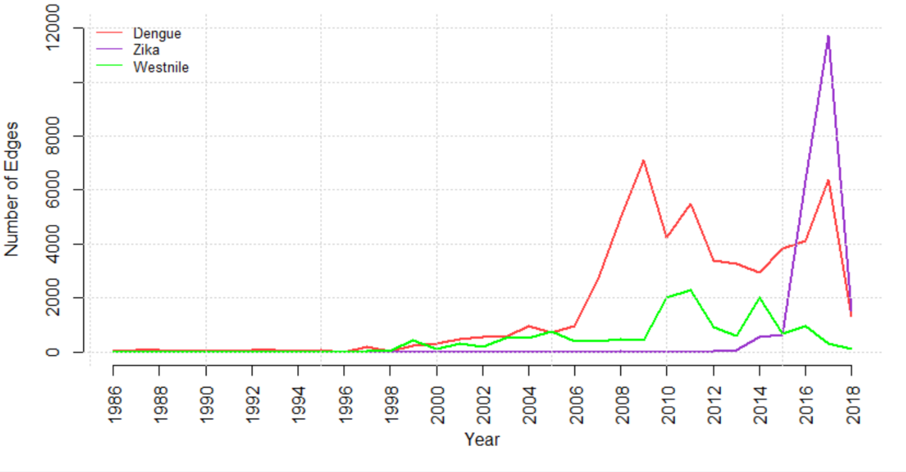

Figure 7: Total Virus Network Edge (collaboration) Count, 1992-2018

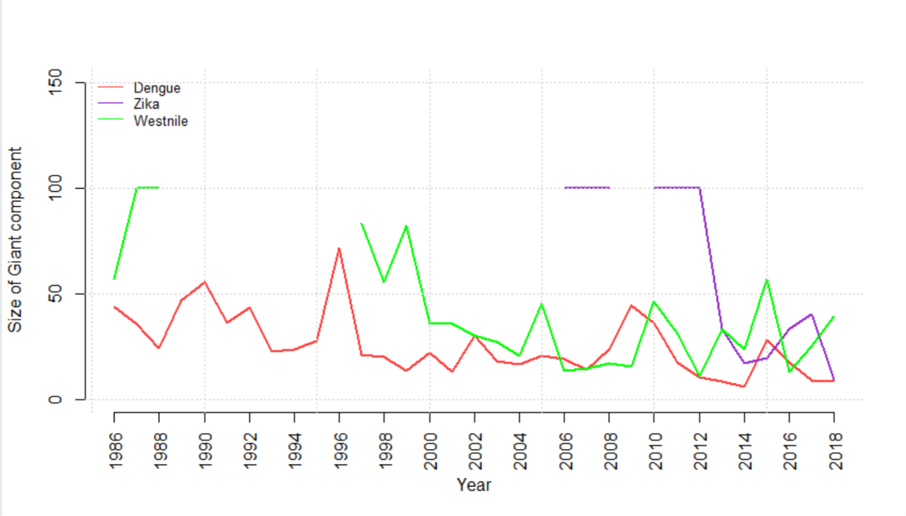

Figure 8: Virus Network Giant Component, 1992-2018

density was detected to decrease. This is due to the fact that the increase in nodes was so great, there was a decrease in the realized connections between researchers.

In the West Nile network, a similar post-outbreak trend was observed as the nodes and edges gradually increased and spiked in 2008 as an effect of outbreak. However, the density of clusters in this network is higher than that of the other two viruses. This suggests that the authors in the West Nile community are densely interlinked and do not usually tend to

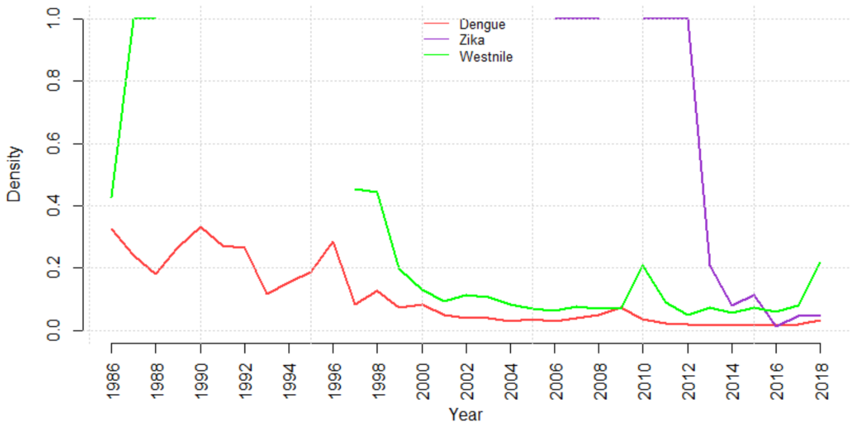

Figure 9: Virus Network Density, 1992-2018

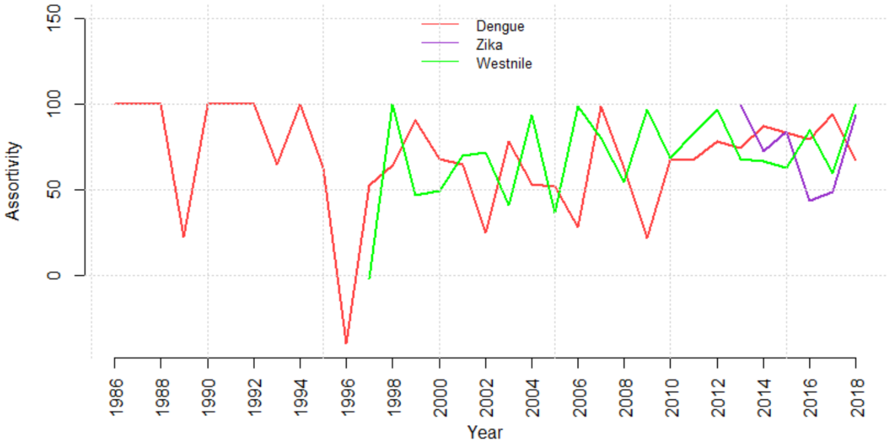

Figure 10: Virus Network Assortitivty, 1992-2018

collaborate with authors outside their network, which was evident from the high assortitivity values. This may relate to the fact that the outbreak largely effected the United States (who also conducted the majority of the research on this virus.)

5. Correlation with the COVID-19 Pandemic Response

The SARS-CoV-2 pandemic and its rapid and devastating effect on almost every nation on the planet is an unprecedented global public health crisis, prompting the response of scientists in every country that can adequately fund research. The focus of our research is understanding how scientific collaboration networks respond, quantitatively and capitalistically, to disease outbreaks - making the COVID-19 pandemic an ideal candidate for analysis and comparison.

Thankfully, the analysis of SARS-CoV-2 scientific output has been done for us, in a wonderful publication by Joanna Nowakowska, Joanna Sobocińska et. al.[21] This paper is a bibliometric study on the first 3 months of research performed on COVID-19 from about January 1 to March 31, and identified an "unprecedented" international response to the pandemic. The authors' methods are similar to ours in measuring the cumulative output of research papers as scientific output. China contributed the largest cumulative amount of COVID-19 research at 37.3%, with the EU coming in second at 22.6% and the United States at 9.4%. The most global disease in the last 100 years produced a powerful response by the global scientific community.

As the authors of the paper conclude, "the present bibliometric study clearly shows the capabilities of modern science to respond to an emerging health crisis, and most likely, to any other future threats." "...The present bibliometric survey underlines that science plays a key role in response to emerging global threats – a notion that has to be considered and acknowledged by political leaders regarding other future risks, be they health-related or not, to the human population."[21]

The scientific community's response to COVID-19 proves that the community is capable of incredible things in the name of the greater good. International public health events severe to the degree of COVID-19 clearly demonstrate this power. Our work in this present study demonstrates that, albeit on a smaller scale, the scientific community tends to respond with strength to public health issues. The COVID-19 pandemic showed us that a high degree of international research equity is possible - and this should be striven for in terms of any disease that makes an international impact. The knowledge that the scientific community is capable of this level of action can and should guide funding policy in the future, and is key to creating a more equitable world.

6. Correlation with International HIV Research

The international research response to HIV (human immunodeficiency virus)/AIDS (acquired immunodeficiency syndrome) are an exemplary case of public health outbreaks and equity. HIV research started in France and quickly became a global health concern [22]. Over the course of public health outbreaks, HIV/AIDS is distinctive, given its wide-reaching and "long-lasting demographic, social, economic and political impacts" [23]. The global response was unprecedented, in both research funding allocated and amount of scientific output the number of countries rising to the task of and relied on sequencing data to a large extent, pooled in large data repositories.

Previous work on international HIV research network analysis found that collaboration among different countries tended to be concentrated in the global north. That is, countries who have been historically well-resources such as the United States and European nations, tend to contribute higher numbers of publication outputs to HIV research [24]. Network analyses of HIV research collaboration networks have also shown that researchers in high-income countries, as might be expected, tend to form the center of the HIV network and rank high in the network centrality measures, such as degree centrality and betweenness centrality (ibid). These suggest disparities between the overall contribution of low-income and low-middle-income countries to the HIV epidemic, though the disease remains a major cause of death and human rights concern primarily in those lower-income bracket countries. The enduring concern is due to the obstacles lower-income countries face in access to treatment and research capacity-building, among other barriers. In this analysis, we found HIV networks to reflect the global nature of the disease's history.

7. Conclusions, Discussion and Future Work

7.1. Conclusions and Discussion

Our examination shows that the United States and other rich, Global North countries conduct the majority of health research on the three outbreaks we identified in the GenBank network. This suggests the network is equitable - e.g, when the global outbreak of Zika virus occurred, the United States was not the most affected - South and Central American nations were, but the United States conducted most of the relevant research. This is an intuitive result, but it further asks the question of our definition of equity. Is equity in this case provided by this redistribution of funds and scientific power from rich nations, or should the affected be enabled further by their richer counterparts to conduct their own research? While the current state of the network shows that the United States is not abandoning its Southern neighbors, to what degree must we accept the role of Western nations as benevolent rulers rather than help the affected help themselves? Wealthy Western nations, by and large, have always benefited from the Global South. From the Transatlantic Slave Trade to the colonization of Africa in the 1800's to the interference and toppling of South and Central American governments by the United States during the Cold War, Western and Global North nations must be willing to rescind their past misdeeds and abuses and comprehensively reinvest in the development of health research in their southern counterparts.

7.2. Proposals

The WHO bulletin mentioned in our introduction serves as a wonderful starting point. We must focus on the wealth disparity between nations if we want to make any substantial progress towards equity in health research. In pursuit of that goal, we must propose funding policy recommendations on the basis of our work. As the COVID-19 pandemic showed both the international scientific community's ability to react to crises and that the economical interdependence of the world's nations increases the fallout from international health events, the NIH and other governmental health research bodies should increase their funding levels to both international institutions and academic bodies within poorer countries. We see an example of this in the NIH's July 2020 initiative to fund data science and health research innovation in Africa[25]. According to the NIH, funding programs like this will help improve health outcomes in Africa over the long term. Outbreaks affect everyone, and funding should be distributed to everyone in order to adequately protect and reinforce the global economy and disadvantaged peoples.

Due to globalization, it would be beneficial to streamline international scientific collaboration in a directed manner, and to promote establishment and betterment of research institutions in under productive nations. This can be most prudently accomplished by large scale sponsorship of scientific activities, increased and equitably distributed funding from the wealthiest and most powerful nations, and centrally-facilitated intergovernmental collaboration to coordinate each of these factors. It has become increasingly clear through our research that the global scientific community must be poised (and financially enabled to be so) to swiftly react to infectious disease outbreaks and other international catastrophes. The COVID-19 pandemic not only shows that this is possible, but that doing so can have incredible results. The politics and economics of doing so are far beyond the scope of this paper, but it is clear to the authors of this paper that the ability of the scientific community to react to crises is enormous (as COVID-19 shows) and can be sufficiently supported to react to crises worldwide promptly and effectively.

7.3. Future Work

Future work must be conducted - not only science-of-science research like ours, but large scale financial and academic efforts by the international community to address the inequalities between nations, especially the rich and the poor. Specifically, the authors of this paper did not consider present international research funding, only the quantitative network characteristics of the NIH GenBank collaboration network. Our results certainly show that most virus research tends to occur in wealthy states (like the US.) But to propose specific funding increases, more research must be done. Also pressing is the coming effects of the global climate crisis. Climate disasters are worsening across the planet and will continue to do so as feedback loops increase in the absence of meaningful state action. [26] Scientists predict that climate change will lead to massive refugee crises due to water, land and food shortages. This will in turn lead to worsening living conditions and increased zoonotic transmission of diseases to human as massive migrations of both humans and animals occurs. [27] Climate change is a looming disaster on the horizon, and the scientific community must be poised to act to its full effectiveness. If it is, millions of lives may be saved.

Acknowledgement

We thank Shripad Laddha for his technical assistance in data analysis and processing.

Funding Information

Research reported in this publication was supported by National Science Foundation Award No. 1561348 and the National Institute of General Medical Sciences of the National Institutes of Health under Award No. R01GM137409. The content is solely the responsibility of the authors and does not necessarily represent the official views of National Science Foundation and the National Institutes of Health.

References

[1] Z. N. E. R. A. S. A.L Barabási, H Jeong, T. Vicsek, Evolution of the social network of scientific collaborations, Physica A: Statistical Mechanics and its Applications 311 (3) (2002) 590–614. `doi:https://doi.org/10.1016/S0378-4371(02)00736-7`.

[2] J. J. Ramasco, S. N. Dorogovtsev, R. Pastor-Satorras, Self-organization of collaboration networks, Phys. Rev. E 70 (2004) 036106. `doi:10.1103/PhysRevE.70.036106`.
URL `https://link.aps.org/doi/10.1103/PhysRevE.70.036106`

[3] J. M. P. J. B. Lucinda Cash-Gibson, Diego F. Rojas-Gualdrón, Inequalities in global health inequalities research: A 50-year bibliometric analysis (1966-2015), PLoS One`doi:10.1371/journal.pone.0191901`.

[4] e. a. Louise Bouchard, Research on health inequalities: A bibliometric analysis (1966–2014), Elsevier`doi:10.1016/j.socscimed.2015.07.022`.

[5] D. Gwatkin, Health inequalities and the health of the poor: What do we know? what can we do?, Bulletin of the World Health Organization.

[6] W. Devitt, Total network activity for top 10 viruses, 1992-2018, `https://public.flourish.studio/visualisation/3992014/` (2020).

[7] V. Sikka, V. K. Chattu, The emergence of zika virus as a global health security threat: A review and a consensus statement of the indusem joint working group, J Glob Infect Dis`doi:10.4103/0974-777X.176140`.

[8] M.-e. a. Juniorcaius Ikejezie, Craig N. Shapiro, Zika virus transmission — region of the americas, may 15, 2015–december 15, 2016, MMWR Morb Mortal Wkly Rep.`doi:10.15585/mmwr.mm6612a4`.

[9] A. A.-M. E. G. Christopher Chang, Kristina Ortiz, The zika outbreak of the 21st century, Journal of Autoimmunity`doi:https://doi.org/10.1016/j.jaut.2016.02.006`.

[10] e. a. Congcong Guo, Global epidemiology of dengue outbreaks in 1990–2015: A systematic review and meta-analysis, Front Cell Infect Microbiol.`doi:10.3389/fcimb.2017.00317`.

[11] R. H. M. e. a. Olivia Brathwaite Dick, José L. San Martín, The history of dengue outbreaks in the americas, Am J Trop Med Hygdoi:10.4269/ajtmh.2012.11-0770.

[12] S. L. R. Karen A Mangold, A review of dengue fever: a resurging tropical disease, Pediatr Emerg Caredoi:10.1097/PEC.0b013e31828ed30e.

[13] A. Schipani, Dengue fever outbreak in bolivia (2009).

[14] L. N. E.E. Ooi, A. Wilder-Smith, D. Gubler, The 2007 dengue outbreak in singapore, Epidemiol. Infect.doi:10.1017/S0950268810000026.

[15] E. V. Caren Chancey, Andriyan Grinev, M. Rios, The global ecology and epidemiology of west nile virus, Biomed Res Intdoi:10.1155/2015/376230.

[16] S. D. V. C. T. R. A. K. e. a. Hadfield J, Brito AF, Twenty years of west nile virus spread and evolution in the americas visualized by nextstrain, PLOSdoi:https://doi.org/10.1371/journal.ppat.1008042.

[17] e. a. Luisa Barzon, Monia Pacenti, Large human outbreak of west nile virus infection in north-eastern italy in 2012, Virusesdoi:10.3390/v5112825.

[18] Epidemiological update: 2012 west nile virus outbreak in the usa.

[19] W. Devitt, Syracuse-metadata-research, https://github.com/williamdevitt7/Syracuse-Metadata-Research (2021).

[20] M. L. Phillips, Dengue reborn: Widespread resurgence of a resilient vector, Environ Health Perspectdoi:10.1289/ehp.116-a382.

[21] M. L. Z. L. P. R. Joanna Nowakowska, Joanna Sobocinska, When science goes viral: The research response during three months of the covid-19 outbreak, Biomedicine & Pharmacotherapydoi:10.1016/j.biopha.2020.110451.

[22] R. C. Gallo, A reflection on hiv/aids research after 25 years, Retrovirology 3.1doi:doi.org/10.1186/1742-4690-3-72.

[23] A. W. Smith, Julia H., The history of aids exceptionalism, Journal of the International AIDs Societydoi:doi.org/10.1186/1758-2652-13-47.

[24] R. S.-G. R. R. G. S. M. Z. G. A. M. F. Tazio Vanni, Marco Mesa-Frias, International scientific collaboration in hiv and hpv: A network analysis, PloS one, 9(3), e93376doi:doi.org/10.1371/journal.pone.0093376.

[25] NIH, Nih to invest 58m to catalyze data science and health research innovation in africa, NIH News Releasesdoi:https://www.nih.gov/news-events/news-releases/nih-invest-58m-catalyze-data-science-health-research-innovation-africa.

[26] e. a. Eric W. Wolff, Feedbacks on climate in the earth system: introduction, Philos Trans A Math Phys Eng Sci`doi:10.1098/rsta.2014.0428`.

[27] N. J. C. Konstans Wells, Host specificity in variable environments, Trends in Parasitology`doi:10.1016/j.pt.2019.04.001`.