landscape view of the use of MIDAS resources as well as the extent or reach of MIDAS dissemination efforts.

B. Results

MIDAS Models and Resources

1. MIDAS Resources and Accessibility

STPI identified 40 MIDAS-funded resources from various public websites and repositories, publications, and from investigator interviews. About half of these resources (n=23) were available for download. The degree of accessibility, however, varied greatly. Some resources were hosted on their own websites and user interface while others were still in their native programming language and were found through supplements of their publication articles. A list of the resources identified is provided in Appendix G.

2. MIDAS Resource Foundational Publications

Of the 40 MIDAS resources identified in Appendix G, STPI collected foundational publications for 27 resources, published from 2007-2015. These publications were identified either through direct citation or searches through the Web of Science database, and then verified through the publication lists of MIDAS resource developers and investigators. All but two of the twenty-seven publications were published as peer-reviewed journal articles.[11] The study team gathered bibliometrics as a measure of the use of MIDAS resources as well as the effectiveness of MIDAS sharing and dissemination practices. The journal impact factor, total citation counts as indicated by Web of Science and Google Scholar, and citation count normalized to publication date were determined for each seminal publication. These data are presented in Appendix G.

Ten of the 27 seminal publications identified averaged more than ten citations per year since publication. Resources with citation per year counts higher than 30 include Repast Simphony, Project Tycho, FluTe, and the spatial simulator developed by Steven Riley. A full listing of these publications and their average citation per year is listed in Table E-1 of Appendix G. Resources with the highest overall citation counts include Repast Simphony, Project Tycho, FluTe, Malaria tools, EpiFast, EpiSimdemics, the Global Epidemic Model, and the spatial simulator developed by Steven Riley.

Data Sharing

[11] The foundational publication for Agent Zero is a book, and the publication for IRED was found on the website for John Hopkins' National Center for the Study of Preparedness and Catastrophic Event Response (PACER).

3. Data Sharing Policy

A key objective of the MIDAS program was to create an open data sharing policy, in which data would be shared freely among MIDAS investigators and models, tools, and other resources would be made publicly available to the greater scientific community. Public availability of MIDAS infrastructure varies. While many tools are open-access and available on host websites or other sources, such as GitHub, some tools are still inaccessible due to log-in requirements. Additionally, while certain tools are available to all MIDAS researchers, they may be inaccessible to the greater scientific community.

The original MIDAS RFA[12] described a "Data Release and Sharing of Results and Resources" policy, which specified roles for the MIDAS Steering Committee, the NIGMS, and the MIDAS awardees. The MIDAS Steering Committee would be responsible for establishing the final data release policy; NIGMS would enforce a policy for timely release of data, to be implemented in consultation with the MIDAS Steering Committee; and grantees would share their findings with other members of the network, make results and models available through the MIDAS database, and include a description of their proposed guidelines for data release.

The subsequent 2005 RFA[13] described in more detail the role of the MIDAS Steering Committee with regard to data sharing and added specifications to the requirement that investigators submit plans for sharing research data and research resources. Subsequent RFA notices included greater specificity for investigator data sharing requirements with respect to sharing model organisms, genome-wide association studies, software development, and communications with U24-award centralized informatics resource. Between 2005 and 2009, there was a shift from emphasizing public access to MIDAS tools and resources toward prioritizing intra-MIDAS sharing of these resources.

In 2009, MIDAS published the "Data Sharing, Intellectual Property, and Publication Policy" which described the guidelines and policies for MIDAS investigators to follow, including policies on data sharing, model sharing and intellectual property. The policy required the investigators to deposit documented results (data, tools, models, analyses, or other intellectual property) into the MIDAS database at least once every 90 days. This database is accessible for MIDAS researchers, but requires permission from the MIDAS informatics services group for anyone outside of MIDAS.

Centralized Informatics Resources

[12] RFA-GM-03-008.

[13] RFA-GM-05-011.

4. Information Resource

The MIDAS technology resource was originally developed with an objective to "design, implement, and manage systems to collect, store, and disseminate data to modelers, both within the MIDAS Network as well as those in the wider research community," amongst other responsibilities. At the inception in 2004, MIDAS funded the Information Technology Resource, managed by Research Triangle International (RTI). The ITR has since changed management from RTI to the University of Pittsburgh and is now called the Informatics Services Group (ISG). The technology resource was intended to support MIDAS investigators in collecting, organizing, and curating data, providing computational tools and simulations, creating models and tools to further research, and making all data and tools publicly available. The 2003 MIDAS RFA called for an informatics group that would perform a variety of functions, including:

1. Creating a centralized database to store and display information for the MIDAS Network
2. Monitor the quality and operation of the database
3. Facilitate sharing of the database across Federal agencies and within the scientific community
4. Provide documentation on all models and tools
5. Curate data according to MIDAS Steering Committee guidelines
6. Propose a data release policy addressing intellectual property rights
7. Generate sophisticated simulated data to test models and collect data to validate models.
8. Provide monthly progress reports to members of the MIDAS network

5. Information Technology Resource

From 2004-2013, the Information Technology Resource (ITR) was managed by Research Triangle International. According to interviews with MIDAS investigators and program staff, in 2005 the ITR had proposed a set of data-centric activities to support MIDAS research investigators, but were asked by program staff to pivot these activities early in the award to focus on building a supercomputer infrastructure and capabilities. In response, the ITR collaborated with the Virginia Bioinformatics Institute and obtained access to an NSF-funded national system of supercomputers called TeraGrid. TeraGrid's supercomputing system provided the capability to perform complex simulations and models to support MIDAS research. However, according to MIDAS investigators, the resource was underutilized by MIDAS researchers, whom often preferred to have programmers and resources in-house, rather than communicate long-distance with RTI to perform their simulations. Investigators who were well versed in supercomputing used

their own university resources to develop and run their models as it was cheaper or free as compared to using the TeraGrid resource. Additionally, investigators preferred to use their own models because they were well-understood, and did not have time to learn to use the new resource. Finally, investigators that did not know how to fully utilize supercomputing resources did not approach the ITR. As the focus for the ITR was to support investigators through supercomputing, and since investigators created models and tools independently, the ITR did not develop models and tools during its first five-year award.

During the second round of funding from 2008-2013, the ITR shifted back to data development and collection as a central focus. Investigators requested large and varied types of data, and ITR began by obtaining particular types of data on which the models were dependent. The ITR focused much of their data development on developing synthetic populations, expanding upon what was first introduced by investigators at Virginia Tech University. The ITR created a now-defunct program called Synthia, which developed US-based synthetic populations and incorporated other kinds of population characteristics such as poverty, and was one of the major accomplishments of the second round of the award.

Another major activity for the ITR was to collect and catalog all the MIDAS tools and resources in a central location. The MIDAS Model Repository was created to provide the models, descriptions of the models, and uses in recent publications. This effort aimed to alleviate the burden of making models share-able by documenting their prescribed usage.

In 2010, the ITR created the MIDAS Software Sharing and Information Outreach Network (MISSION), a group of modelers and programmers that focused on issues related to development of tools and code. This forum allowed MIDAS modelers to discuss best practices and ideas on software development. The MISSION group created a software catalog by surveying investigators to identify what tools and resources had been developed and were currently under development. This effort differed from the MIDAS Model Repository in that the latter was aimed towards documenting existing models, documenting their intended usage, and how they had been used in peer-reviewed publications. The software catalog was intended to capture all the software, tools, and code that had been created and which were under development.

6. Informatics Services Group

In 2014, the technical resource was awarded to the University of Pittsburgh's Informatics Services Group. This group has gone further in its efforts to categorize MIDAS resources and provide technical services to the MIDAS Network. The ISG transitioned the previous MIDAS website created by ITR, Epimodels.org, to a modern content management system using Drupal. They maintained the websites' private back-end which allows investigators to log-in and share information and documents such as meeting notes.

The ISG changed the way publications and resources were being retrieved. Previously, the system employed basic bibliography retrieval but is now using an ontology-based semantic web technology, accessible at OBC.ide. This web resource collects and organizes the disparate MIDAS resources, articles, models, and other artifacts.

The ISG is also focused on creating standards for the models and tools. They have implemented the web-based Apollo system and are creating an infrastructure for standardizing model development and model interactions for disease outbreaks. The Apollo standard offers standard representation used to integrate disease models.

The ISG has continued to provide MIDAS researchers agent-based synthetic populations. While the previous ITR utilized a program called Synthia, the ISG produced an alternative capability to provide customizable populations based on US Census data variables. Providing resources such as synthetic populations, including synthetic ecosystems, is one of their major activities.

C. Outcomes

1. Decentralization of MIDAS Resources

Assessing the efficacy of MIDAS infrastructure is a challenge. Initially, the aim was to house all models, tools, and resources in a centralized location, but the program-level organization of MIDAS infrastructure has shifted to a more de-centralized system. Models and tools are housed on the computer systems of individual MIDAS researchers and research groups, making many tools difficult to access. Collecting and identifying MIDAS infrastructure for this evaluation necessitated a search across many sources[14] and websites, some of which required permissions to access or are now defunct. Additionally, the MIDAS computing infrastructure has moved to distributed computing, instead of a centralized supercomputing system as in the first iteration of the ITR. Software and tools reside on various systems belonging to investigators, and the interface among those systems is now provided by the ISG. In order to more effectively share and disseminate the MIDAS infrastructure, the technical resource would need to collect MIDAS resources and data repositories from previous resource surveys (i.e., MISSION's software catalog) and data collection efforts by the previous ITR as well as other MIDAS researchers.

2. Data Sharing

Despite agreement among MIDAS investigators on data sharing requirements associated with the MIDAS award, compliance with and perceived utility of data sharing

[14] Sources include MIDAS Portal (Epimodels.org), Centers of Excellence websites, MIDAS Network Software Repository (MISSION), Web of Science, and MIDAS investigator interviews

agreements varies greatly. While researchers believe that sharing of datasets would be very useful, that is not the case with MIDAS models and tools. Many researchers develop models tailored to particular data sets and parameters to answer specific research questions and these models may not be useful to other researchers who have different questions to answer. In that case, researchers tend to build and use their own models and do not focus on sharing. In addition, the time and resources needed for documentation to make models and tools shareable is burdensome to investigators. However, the technology resource has made efforts to take on these tasks to alleviate that burden (e.g., the MIDAS Models Repository done by the ITR).

3. Sharing, Tracking, and Dissemination

The complexity and diversity of MIDAS infrastructure presents challenges in meeting the goals of sharing and dissemination of MIDAS infrastructure. MIDAS researchers showed differing levels of awareness for how their tools and resources were being used by the public outside of MIDAS, and many indicated that they did not have robust mechanisms for tracking dissemination and usage. Some researchers post their tools and resources on GitHub, but are unaware of how GitHub may track usage of the resources. Others use Google Analytics to track usage, including total number of page views, number of software downloads, queries made, etc. A few researchers responded that they did not have any formalized mechanism of tracking usage of their tools when hosted on their webpages or on other sites, but that they know informally and anecdotally who uses their products, because users typically contact them with questions on their models.

A challenge in assessing the sharing and dissemination practices of the MIDAS infrastructure is in developing sharing metrics that are appropriate for the specific type of resource. Complex mathematical models that can help determine the rate of disease spread may have different purposes and audiences than systems that employ graphical user interfaces to allow non-technical audiences to visualize and analyze epidemiological data. It is important to be able to define and identify the types of resources in order to develop appropriate metrics for sharing and dissemination.

4. MIDAS Website

While the MIDAS website has been changed and improved to include more modern systems, certain aspects of the MIDAS website are unused. In recent years, the private back-end of the MIDAS website has not been used by investigators in the network, and many of the documents and meeting notes date prior to 2014. Additionally, many links are non-functional or loop back to the home page, and certain historical documents such as the Historical Data and Documents Catalog (HDDC), developed by the ITR, are no longer available and must be recreated. However, indexes for diseases have been subsumed by the OBC.ide system.

5. MIDAS Technology Resource

MIDAS investigators have differing opinions on the efficacy of the MIDAS technology resource. Early in the award, the ITR was heavily underutilized because investigators preferred to create and develop their own tools, while others did not understand the role of the ITR or how to collaborate with the group. In recent years, the resource group has provided more support by providing data sets and producing synthetic populations and synthetic ecosystems, but it remains unclear if the ISG can best support MIDAS researchers, and how to encourage researchers to collaborate with the ISG.

6. MIDAS Organization and Structure

To assess the organization and structure of the MIDAS, STPI examined the relationships, activities, and management structures that encompass the program from its inception to the present. STPI interviews with MIDAS investigators and program staff revealed that the culture of MIDAS as an additive and collaborative network of U awards was strongly integrated with the program's organization and structure.

In 2004, MIDAS began by awarding four separate U01 awards to investigators located at four institutions across the country, one of which was awarded to Research Triangle International to establish the ITR. Four additional U01 investigators were granted in 2005, and three more were added in 2007, to expand the scope of infectious disease modeling research conducted by the network of investigators. For more information, see Appendix A.

Beginning in 2009, two of the original 2004 U01 awardees competed for and were awarded U54-based centers of excellence that conduct policy research, outreach, and dissemination activities in addition to modeling research. The additional activities, responsibilities, and research personnel now working at these centers created smaller networks within the overall MIDAS network. MIDAS investigators described that centers have collaborative research benefits not present with a set of R01 awards, where ideas are able to develop more effectively with fewer barriers between them. Additionally, investigators noted that the pace of research is different at centers, as there are complementary skill sets in the center that researchers can draw from. Finally, centers allow for the support of many young researchers such as students and post-docs, which investigators say, enables greater focus on research than some awards without a cadre of young trainees.

The network saw a significant expansion of investigators between 2009 and 2014, more than doubling the number of new investigators in the network. Three new U01 investigators were added in both 2009 and 2011. Five U01 investigators were added in 2014, along with a new host for the U24 award (University of Pittsburgh's ISG), and a third U54-based center of excellence was awarded to a previous U01 awardee. Subsequently, investigators cited a growing sense of competition within the MIDAS network, and especially among the three U54 centers, as the network has expanded. Currently, the primary programmatic mechanism for intra-MIDAS collaboration is the annual scientific meeting, where all awardee laboratories, their collaborators, and other interested public policy officials gather for a two-day program of scientific talks and

discussion. As a collaboration promoting activity, the awardee annual meeting is commonly shared mechanism amongst other NIH programs with U54 awards. Yet, there are additional collaboration-promoting mechanisms that are employed by other NIH programs that are not part of MIDAS's structure. These include development of and specialized funding for cross-network collaboration, active staff coordination for out-of-network collaboration, and active project monitoring by staff to avoid duplication of effort. In addition, those programs entrusts the program steering committee with reviewing the progress of individual awards, promoting collaboration, encouraging community outreach, and allowing the scientific director to have a voting seat on the committee. See Appendix H for additional examples of collaboration promoting activities of selected NIH/NCI programs.

From its inception until her retirement in the fall of 2014, Dr. Irene Eckstrand served as the MIDAS scientific director. Investigators credit Dr. Eckstrand with establishing connections between the network and policy officials, guiding the scientific direction of the network, lending the credibility to the network, and expanding the program's scientific diversity. Retirement of Dr. Eckstrand was seen by many investigators as a significant loss of structure and leadership to the network.

Concurrent with Dr. Eckstrand's retirement was realization that the program's oversight structures, the steering and executive committees needed to be restructured and repopulated. From 2009 to 2014, the steering committee, which made recommendations to NIGMS regarding scientific direction of the program, was composed of user groups such as public health officials, representatives of the research community, and NIH scientist administrators with relevant expertise. The executive committee, which coordinated and managed the MIDAS Network, was MIDAS principal investigators.

Starting in 2015, the steering committee was repopulated with only MIDAS investigators and program staff, who would now guide the scientific direction of the program. A new external advisory committee was also created, composed of public health officials and representatives of the outside research community, to provide guidance on relevant issues, but not oversight. No additional information on these two organizational committees is available, as they were in process of being created at the time of this study.

7. Summary of Findings

This chapter summarizes the findings detailed in Chapters 2-6.

A. Bibliometric Analyses of MIDAS Scientific Productivity and Impact

Award- and publication-level analyses were performed on MIDAS-funded research publications in comparison to other NIH-funded publications on infectious disease modeling research. Productivity was assessed using overall publication count and publication count normalized by year and funding dollars (publications/per $100,000 of direct support).

Bibliometric analyses indicate MIDAS U54 and U54-co-funded publications tend to have higher impact and are published in higher ranking journals than publications from the comparison groups, and MIDAS U01 awards. U01-R01-co-funded publications typically have higher impact than R01 comparison group publications and are published in higher ranking journals than MIDAS U01 and R01 comparison group publications alone. R01 comparison group and U54-R01-co-funded publications have marginally higher citation counts than MIDAS U54 publications with no other significant citation count differences between groups.

Bibliometric analysis indicates that MIDAS U54 awards outperform their sister U01 awards and other National Institutes of Health (NIH) awards in the productivity and impact of infectious disease modeling research. However, MIDAS cooperative U01 awards are not necessarily more productive nor more impactful than a comparison group of similar R01 research awards.

B. Topic Modeling Analysis of Novel Research Areas

To assess the unique and unanticipated nature of MIDAS-funded research, STPI topic modeled the titles and abstracts of 1,229 journal articles attributed to MIDAS and non-MIDAS grants.

Using this analysis, the results indicate that the MIDAS program supports research in certain areas of infectious disease modeling to a greater degree than the comparison awards. These areas include the modeling of transmission dynamics, modeling of drug and antibiotic-resistant pathogens, and epidemic forecasting from non-traditional sources. MIDAS also supports researchers that publish in areas of research that are not unique to MIDAS, but which are still highly relevant to the field of infectious disease research, such as the modeling of economics, cost, and outcomes of public health interventions and research targeting specific infectious diseases.

While the topic modeling analysis indicates that MIDAS supports valuable research, including areas that are unsupported by other NIH funding, an external panel of experts assessed only two areas as being especially novel. Panelists noted that these research areas presented numerous novel results that enhanced understanding of the interplay between the immune system and disease serotypes, and contained innovative applications of mathematical ideas and methods to disease modeling. Panelists also regarded epidemic forecasting research using non-traditional sources as unique to a program like MIDAS.

C. Descriptive Analysis of MIDAS Education and Training Activities

STPI tracked the careers of MIDAS-supported students and post-doctoral researchers (postdoc) to understand MIDAS's impact on the education and training of future researchers in the field of infectious disease modeling or related areas. STPI identified 151 students and postdoctoral trainees supported by MIDAS at some point during the period of 2004-2014, and successfully traced 143 of them. From those 143 students and postdoctoral trainees, STPI found that approximately half of them are still students or postdoctoral trainees.

Given that MIDAS is still a relatively young program, it seems very logical that most of these students and trainees are still in the early stages of their careers. However, 36% of the traceable students and trainees currently hold academic faculty positions, and an overwhelming majority of them (92%) hold these positions in health-related fields. Despite the limited sample size of former students and trainees who now hold faculty positions, these results provide some evidence to suggest that MIDAS is a significant contributor to capacity building efforts within the field of infectious disease modeling. Moreover, expert panelists found the program's training activities as "impressive", based upon the number of student and postdocs trained and other educational activities.

D. Assessment of MIDAS Outreach Activities

All mentions of outreach activities within MIDAS annual reports, both explicitly and indirectly mentioned, were tabulated with the year in which they were reported and the award number. Surveying across the activities mentioned in the annual reports, STPI found a wide range of activities that MIDAS investigators considered to be "outreach." Nearly half of these mentioned instances of outreach were presentations at conferences. MIDAS investigators often created informational material or supported the booth at the both academic and public health conferences aimed at both researchers and students. Many of the annual reports did not report any outreach activities, but it is unclear whether this is simply reporting error or an indication of no outreach activity. Given that many reports, especially in earlier years, limited the progress section to two pages, it is possible that researchers did not prioritize listing this in the given space.

Those investigators who did perform outreach activities noted in their interviews that cuts in funding would affect outreach activities first. Many stated that they had proposed outreach programs, especially short courses and conferences that were not possible due to limited funds. Targeted advertising of these resources to members beyond the infectious disease modeling community would be an easy and worthwhile outreach effort.

E. Assessment of MIDAS Influence on Infectious Disease Preparedness and Response Policy

STPI assessed the efficacy of MIDAS interactions with policy makers, collaborative activities, and the extent to which MIDAS played a role in public health or public policy decision-making. Sixty instances of policy-related activities were categorized into four major categories—11 networked interactions, 29 research collaborations, 13 epidemic response support, and 7 public health tools.

In general, there was consensus among the interviewees, both from the research and policy sides, that MIDAS played a role in connecting researchers with public officials. Most researchers felt that it was a useful and productive relationship. The three centers of excellence (U54 awards) had particularly close ties with policy makers at the national level and acknowledge MIDAS as being instrumental in creating and fostering those connections. Policy makers, alike, expressed interest in getting to know and work with modelers on a more consistent basis. Furthermore, most researcher investigators responded positively about working with public health departments and even indicated that being a MIDAS-funded researcher served as a stamp of validation when interacting with local officials.

Overall, policy makers (and researchers alike) indicated that there is likely great potential in modeling for helping public health but that the infrastructure and dynamics of the actual interactions must be fine-tuned for greater effect. MIDAS has certainly made progress in bringing together a research community with an interest in policy and enabled collaborations that have had some impact on policy. Policy officials suggested workshops focused on public-facing modeling tools and training public health officials in using them. While it was hard for policy makers to directly attribute policy outcomes to modeling, done through MIDAS, most felt that the funds were important in sustaining research that was flexible and available for collaboration.

F. Assessment of MIDAS Modeling Infrastructure

From the inception of the program, one of MIDAS's major objectives has been to develop and refine models, tools, and resources for infectious disease modeling. STPI identified MIDAS-developed infrastructure – or models, tools, datasets, and other

Award-Level Productivity Metrics

Publications per year and publications per $100,000 total costs were compared at the award level. To present a picture of productivity by year since funding start, the average annual distribution of publications for each award is also presented.

The following four comparative bibliometric analyses were conducted at the award-level including an intra-MIDAS comparison.

1. MIDAS U01 versus R01 Comparison Group
2. MIDAS U54 versus Center Comparison Group
3. MIDAS U54 versus R01 Comparison Group
4. MIDAS U01 versus MIDAS U54

Pairwise comparisons were conducted between groups using the Wilcoxon rank-sum test. The Wilcoxon rank-sum test is a non-parametric alternative to the *t*-test to test whether two samples come from the same population.[19] The test returns a *p*-value that is reported in addition to the median and interquartile range.[20] Effect size, *r*, represents the magnitude of the difference between groups. An *r* between 0.10-0.30 is considered a small effect, 0.30-0.50 an intermediate effect, and 0.50 and higher is a strong effect.[21] Since the MIDAS U54 (n = 3) and center comparison (n = 2) groups have small sample sizes, Wilcoxon rank-sum tests cannot be done for comparison numbers 2-4. In these cases, the mean and standard deviation are reported as descriptive statistics.

Results and Findings

Below are the results and findings from each of the four comparisons for the award-level productivity metrics.

MIDAS U01 versus R01 Comparison Group

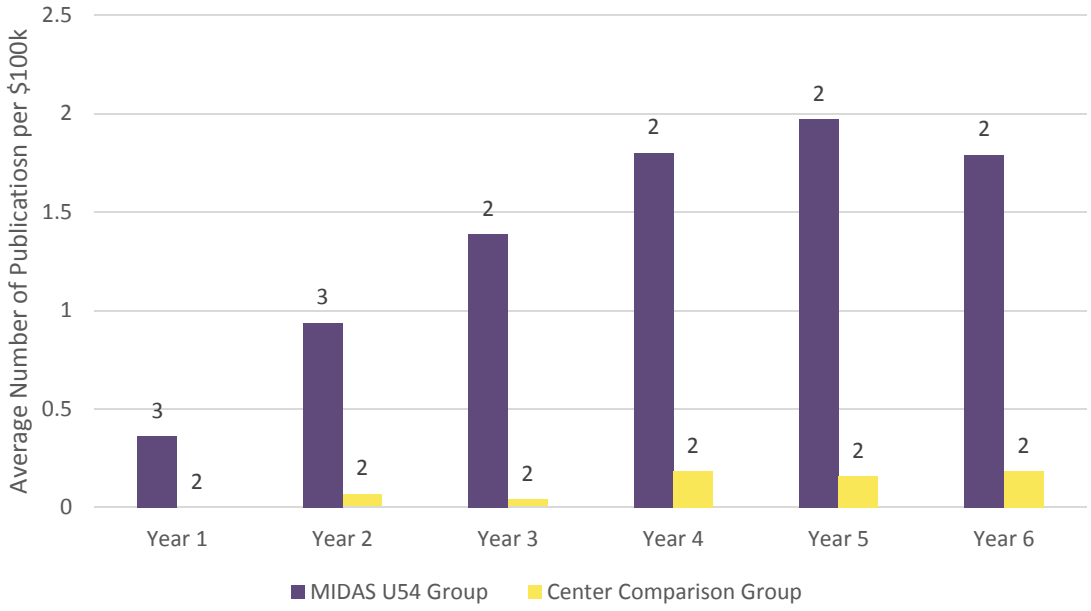
MIDAS U01 publications per year had a higher median than the R01 comparison group (see Table B-6 for descriptive statistics for each group). When normalized for cost, publications per $100,000 did not significantly differ. Figure B-1 shows the annual average distribution of publications per $100,000 total costs.

---

[19] Henry B. Mann and Donald R. Whitney, "On a Test of Whether One of two Random Variables is Stochastically Larger than the Other," *The Annals of Mathematical Statistics* (1947): 50–60.

[20] The median is the number separating the higher half of a data sample with the lower half. Interquartile Range (IQR) is a measure of variability based on dividing a data set into four equal parts. The values that divide each part are Q1, Q2, Q3 and the IQR is Q3 – Q1.

[21] Jacob Cohen, *Statistical Power Analysis for the Behavioral Sciences.* 2nd ed. (Hillsdale: Lawrence Erlbaum Associates 1988).



Note: Because one of the MIDAS U54 awards is recent, data labels refer to the number of MIDAS U54 or center comparison group projects included in the average

Figure B-2. Average Publications per $100k Total Costs over 6 Years

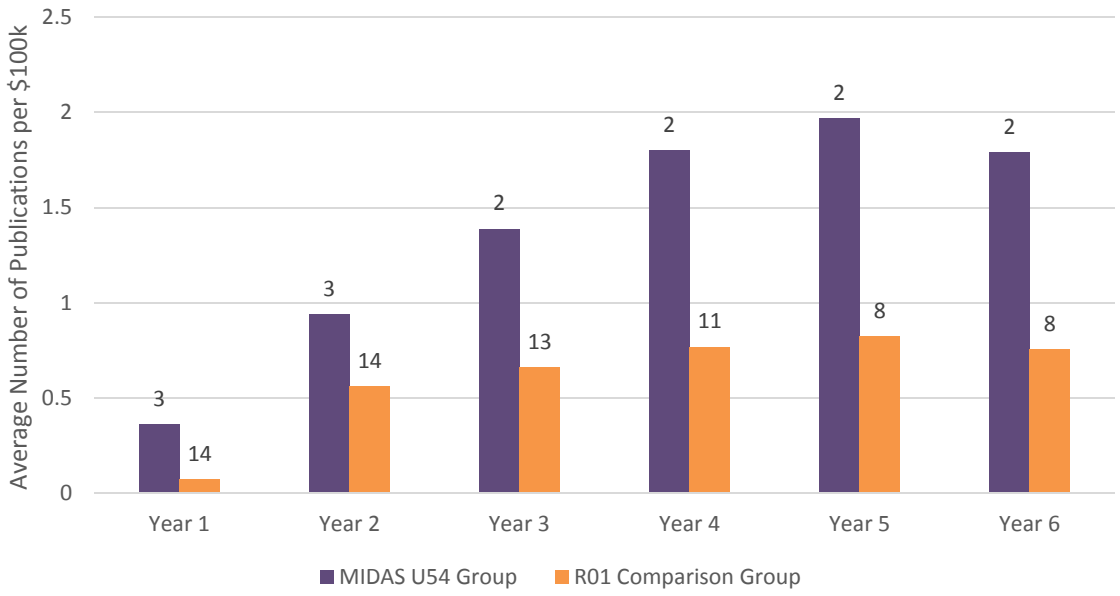
MIDAS U54 awards are as or more productive both per year and per $100,000 total costs than the center comparison group awards.

MIDAS U54 versus R01 Comparison Group

MIDAS U54 mean publications per year and per $100,000 total costs is higher than the R01 comparison group but this cannot be concluded with any statistical significance due to small sample size (see Table B-8 for descriptive statistics). Figure B-3 shows the annual average distribution of publications per $100,000 total costs.

Table B-8. Award-Level Productivity Metrics

Metric	MIDAS U54 Mean (SD)	R01 Comparison Group Mean (SD)
Publications per Year	37.78 (10.19)	3.66 (3.13)
Publications per $100k Total Costs	1.46 (0.30)	0.83 (0.63)



Note: Data labels refer to the number of U54 or R01 projects included in the average

Figure B-3. Average Publications per $100k Total Costs over 6 Years

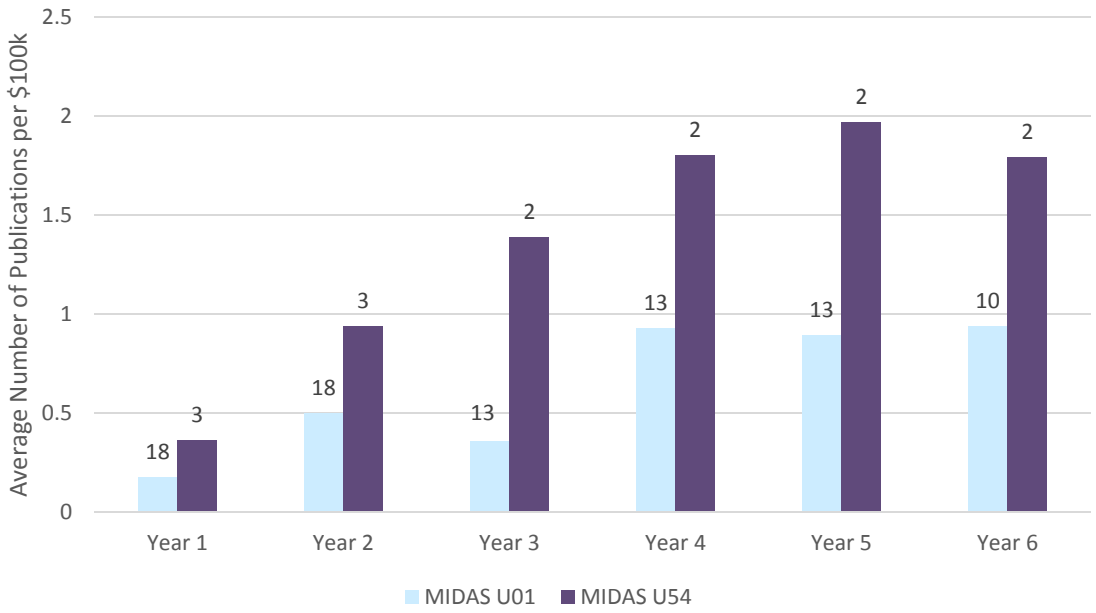
MIDAS U54 awards are as or more productive than R01 comparison group awards both per year and per $100,000 total costs.

MIDAS U01 versus MIDAS U54

MIDAS U54 mean publications per year and per $100,000 total costs is higher than the MIDAS U01 awards but this cannot be concluded with any statistical significance due to small sample size (see Table B-9 for descriptive statistics).

Table B-9. Award-Level Productivity Metrics

Metric	MIDAS U01 Mean (SD)	MIDAS U54 Mean (SD)
Publications per Year	5.88 (3.00)	37.78 (10.19)
Publications per $100,000 Total Costs	1.09 (0.73)	1.47 (0.30)



Note: Data labels refer to the number of U01 or R01 projects included in the average

Figure B-4. Average Publications per $100k Total Costs over 6 Years

MIDAS U54 awards are as or more productive than MIDAS U01 awards both per year and per $100,000 total costs.

Summary of Findings

MIDAS U01 awards are more productive per year than the R01 comparison group, but there is no significant difference between MIDAS U01 and R01 comparison group productivity per award dollar. MIDAS U54 awards are as or more productive per year and per dollar than every comparison group, including MIDAS U01 awards, as shown by award-level metrics.

Publication-level Impact Metrics

Citation count, Impact per Publication (IPP), SCImago Journal Rank (SJR), and Source Normalized Impact per Paper (SNIP) as defined in the introduction to this chapter were compared at the publication-level. Since the landscape of NIH-funded infectious disease modeling research is fairly small, there were co-funded publications among all the comparison groups. In order to disaggregate overlap, when analyzing publication-level metrics, publications funded by both groups in the comparison were categorized into a 'co-funded' publication group.

The following four comparative bibliometric analyses were conducted at the publication-level including an intra-MIDAS comparison.

1. Single Investigator Comparison (MIDAS U01 vs. R01s vs. Co-funded)
    - MIDAS U01: 390 publications
    - R01 comparison group: 289 publications
    - Co-funded: 18 publications
2. Center Comparison (MIDAS U54 vs. Centers vs. Co-funded)
    - MIDAS U54: 534 publications
    - Center comparison group: 64 publications
    - Co-funded: 4 publications
3. U54 vs. R01 (MIDAS U54 vs. R01s vs. Co-funded)
    - MIDAS U54: 527 publications
    - R01 comparison group: 296 publications
    - Co-funded: 11 publications
4. Intra-MIDAS Comparison (MIDAS U01 vs. MIDAS U54 vs. Co-funded)
    - MIDAS U01: 344 publications
    - MIDAS U54: 474 publications
    - Co-funded: 64 publications

Three-way comparisons were conducted for each comparison listed above using the Kruskal-Wallis test by ranks. The Kruskal-Wallis test by ranks is a non-parametric method for testing whether three samples come from populations with the same distribution—it evaluates differences in medians among groups[22]. If the Kruskal-Wallis test returns a significant *p*-value, this indicates that at least two groups in the analysis originate from different distributions or that there is a significant difference in the medians. When the overall Kruskal-Wallis test is significant, the next step is to conduct pairwise comparisons using the Wilcoxon rank-sum test, a non-parametric alternative to the *t*-test, to determine where the difference among the groups lies.

Results and Findings

Below are the results and findings from each of the four comparisons for the publication-level impact metrics. Boxplots of the results can be found below.

[22] William H. Kruskal and W. Allen Wallis, "Use of Ranks in One-Criterion Variance Analysis." *Journal of the American Statistical Association* 47 (260, 1952): 583–621.

Single Investigator Comparison (MIDAS U01 vs. R01s vs. Co-funded)

The Kruskal-Wallis test conducted to evaluate differences among the three Single Investigator Comparison Groups (MIDAS U01, R01, and Co-funded publications) on median citation count, IPP, SJR, and SNIP was significant for IPP and SJR (Table B-10).

Table B-10. Single Investigator Three-way Comparison Test Results

Metric	*p*-Value
Citations	0.37
IPP	**0.04**
SJR	**0.05**
SNIP	0.20

Follow-up tests were conducted to evaluate pairwise differences among for only those metrics which resulted in significant p-value in the table above. The results of these tests indicate a significant difference in SJR between the MIDAS U01 publications and the co-funded publications (Table B-11). The median SJR for the co-funded publication group is higher than the MIDAS U01 publication group. The results also indicate a significant difference in IPP and SJR between the R01 comparison group publications and the co-funded publications. The median IPP and SJR for the co-funded publication group is higher than the R01 comparison group. There were no statistically significant differences between the MIDAS U01 and R01 comparison group publications for IPP or SJR.

Table B-11. Single Investigator Pairwise Comparison Test Results

Group 1	Group 2	Metric	Group 1 Median (IQR)	Group 2 Median (IQR)	*p*-Value	Effect size, *r*
U01	**R01**	**IPP**	3.27 (2.23)	2.95 (2.36)	0.09	0.07
		SJR	1.64 (1.60)	1.49 (1.72)	0.42	0.03
U01	**Co-funded**	**IPP**	3.27 (2.23)	3.85 (5.14)	0.10	0.08
		SJR	1.64 (1.60)	2.44 (3.25)	**0.03**	**0.11**
R01	**Co-funded**	**IPP**	2.95 (2.36)	3.85 (5.14)	**0.03**	**0.12**
		SJR	1.49 (1.72)	2.44 (3.25)	**0.02**	**0.14**

U01-R01-co-funded publications typically have a higher IPP than the R01 comparison group publications as well as a typically higher SJR than the MIDAS U01 publications or R01 comparison group publications. It is unclear why the co-funded publications have a higher impact by the IPP and SJR metrics, but there were no significant differences between the three groups for citation count, nor SNIP.

Center Comparison (MIDAS U54 vs. Centers vs. Co-funded)

The Kruskal-Wallis test conducted to evaluate differences among the three Center Comparison Groups (MIDAS U54, Center, and Co-funded publications) on median citation count, IPP, SJR, and SNIP was significant for IPP and SJR (Table B-12).

Table B-12. Center Comparison Three-way Comparison Test Results

Metric	*p*-Value
Citations	0.19
IPP	**≤0.01**
SJR	**≤0.01**
SNIP	0.09

Follow-up tests were conducted to evaluate pairwise differences among the three groups in the above significant metrics (see Table B-13 for descriptive statistics for each group). The results of these tests indicate a significant difference in IPP and SJR between the MIDAS U54 publications and the center comparison group publications. The median IPP and SJR for the MIDAS U54 publications is higher than the center comparison group publications. The results of the pairwise comparisons also indicate a significant difference in IPP and SJR between the center comparison group publications and the co-funded publications. The median IPP and SJR for the co-funded publications is higher than the center comparison group publications. There were no statistically significant differences between the MIDAS U54 publications and the co-funded publications for either IPP or SJR.

Table B-13. Center Comparison Pairwise Comparison Test Results

Group 1	Group 2	Metric	Group 1 Median (IQR)	Group 2 Median (IQR)	*p*-Value	Effect size, *r*
U54	**Center**	**IPP**	3.53 (2.69)	1.62 (2.69)	**≤0.01**	**0.22**
		SJR	1.96 (1.82)	0.81 (1.26)	**≤0.01**	**0.23**
U54	**Co-funded**	**IPP**	3.53 (2.69)	13.64 (22.54)	0.64	0.02
		SJR	1.96 (1.82)	10.36 (18.40)	0.51	0.03
Center	**Co-funded**	**IPP**	1.62 (2.69)	13.64 (22.54)	**0.04**	**0.26**
		SJR	0.81 (1.26)	10.36 (18.40)	**0.04**	**0.26**

MIDAS U54 and U54-Center-co-funded publications typically have a higher IPP and SJR than the center comparison group publications. There were no significant differences between the three groups for citation count or SNIP.

U54 vs. R01 Comparison (MIDAS U54 vs. R01s vs. Co-funded)

The Kruskal-Wallis test conducted to evaluate differences among the three groups in the U54 vs. R01 Comparison (MIDAS U54, R01, and Co-funded publications) on median citation count, IPP, SJR, and SNIP was significant for all four metrics (Table B-14).

Table B-14. U54 vs. R01 Three-way Comparison Test Results

Metric	*p*-Value
Citations	**0.01**
IPP	**≤0.01**
SJR	**≤0.01**
SNIP	**0.02**

Follow-up tests were conducted to evaluate pairwise differences among the three groups in the above significant metrics (see Table B-15 for descriptive statistics for each group). The results of these tests indicate a significant difference in citation counts, IPP, and SJR between the MIDAS U54 publications and the R01 comparison group publications. MIDAS U54 publications have a higher median IPP and SJR than R01 comparison publications but a lower median citation count. Pairwise comparison results also indicate a significant difference in citation count, SJR, and SNIP between MIDAS U54 publications and co-funded publications. Co-funded publications have a higher median citation count, SJR, and SNIP than the MIDAS U54 publications. The test results also indicate a significant difference in IPP, SJR, and SNIP between the R01 comparison group and the co-funded publications. The co-funded publications have a higher median IPP, SJR, and SNIP than the R01 comparison group publications.

Table B-15. U54 vs. R01 Pairwise Comparison Test Results

Group 1	Group 2	Metric	Group 1 Median (IQR)	Group 2 Median (IQR)	*p*-Value	Effect size, *r*
U54	**R01**	**Citations**	7.00 (15.00)	8.00 (20.25)	**0.05**	**0.07**
		IPP	3.53 (2.70)	3.00 (2.34)	**≤0.01**	**0.16**
		SJR	1.80 (1.70)	1.48 (1.69)	**≤0.01**	**0.12**
		SNIP	1.36 (0.77)	1.29 (0.74)	0.07	0.06
U54	**Co-funded**	**Citations**	7.00 (15.00)	17.00 (58.50)	**0.02**	**0.10**
		IPP	3.53 (2.70)	5.25 (10.41)	0.19	0.06
		SJR	1.80 (1.70)	3.19 (5.41)	**0.04**	**0.09**
		SNIP	1.36 (0.77)	2.23 (3.70)	**0.02**	**0.13**
R01	**Co-funded**	**Citations**	8.00 (20.25)	17.00 (58.50)	0.07	0.10
		IPP	3.00 (2.34)	5.25 (10.41)	**0.04**	**0.12**
		SJR	1.48 (1.69)	3.19 (5.41)	**0.01**	**0.15**
		SNIP	1.29 (0.74)	2.23 (3.70)	**0.02**	**0.13**

The R01 comparison group publications and U54-R01-co-funded publications typically have higher citation counts than the MIDAS U54 publications. The MIDAS U54 and U54-R01-co-funded publications typically have a higher IPP and SJR than the R01 comparison group publications. The U54-R01-co-funded publications typically have a higher SJR than the MIDAS U54 publications, and they have a higher SNIP than both the MIDAS U54 publications and the R01 comparison group publications. The higher impact metrics from the co-funded may be the result of one productive U54-R01 collaborative relationship between Dr. Susan Huang and other MIDAS researchers. Dr. Huang was a co-investigator on a MIDAS U01 award while at Harvard University, but has continued to collaborate with various MIDAS investigators on her own R01 awards, since leaving Harvard and the U01 award and moving on to University of California, Irvine.

Intra-MIDAS Comparison (MIDAS U01 vs. MIDAS U54 vs. Co-funded)

The Kruskal-Wallis test conducted to evaluate differences among the three groups in the Intra-MIDAS Comparison (MIDAS U01, MIDAS U54, and Co-funded publications) on median citation count, IPP, SJR, and SNIP was significant for IPP, SJR, and SNIP (Table B-17).

Table B-17. Intra-MIDAS Three Way Comparison Test Results

Metric	*p*-Value
Citations	0.23
IPP	**≤0.01**
SJR	**≤0.01**
SNIP	**≤0.01**

Follow-up tests were conducted to evaluate pairwise differences among the three groups in the above significant metrics (see Table B-18 for descriptive statistics for each group). The results of these tests indicate a significant difference in IPP, SJR, and SNIP between MIDAS U01 publications and MIDAS U54 publications as well as the co-funded publications. Both the MIDAS U54 publications and the co-funded publications have a higher median IPP, SJR, and SNIP than MIDAS U01 publications. There were no significant difference between MIDAS U54 publications and co-funded publications.

Table B-18. Intra-MIDAS Pairwise Comparison Test Results

Group 1	Group 2	Metric	Group 1 Median (IQR)	Group 2 Median (IQR)	*p*-Value	Effect size, *r*
U01	**U54**	**IPP**	3.27 (2.30)	3.53 (2.73)	**≤0.01**	**0.13**
		SJR	1.37 (1.56)	1.87 (1.86)	**≤0.01**	**0.10**
		SNIP	1.21 (0.78)	1.35 (0.77)	**≤0.01**	**0.09**
U01	**Co-funded**	**IPP**	3.27 (2.30)	3.31 (2.40)	**0.02**	**0.12**
		SJR	1.37 (1.56)	1.80 (1.77)	**0.03**	**0.11**
		SNIP	1.21 (0.78)	1.46 (1.00)	**0.05**	**0.10**
U54	**Co-funded**	**IPP**	3.53 (2.73)	3.31 (2.40)	0.86	0.01
		SJR	1.87 (1.86)	1.80 (1.77)	0.56	0.03
		SNIP	1.35 (0.77)	1.46 (1.00)	0.65	0.02

MIDAS U54 and U54-U01-co-funded publications typically have higher IPP, SJR, and SNIP than MIDAS U01 publications alone. There were no significant differences in citation count.

Appendix C.
Comparative Topic Modeling Analysis

NIH funds infectious disease modeling research through a number of programs, not limited to MIDAS. As part of its evaluation of the MIDAS program, STPI was asked to compare the research outcomes of MIDAS with other NIH-funded research in infectious disease modeling. To address this question, STPI topic modeled the titles and abstracts of journal articles in order to characterize and compare the research outcomes of MIDAS-funded researchers against those of non-MIDAS researchers within the domain of infectious disease modeling.

Approach

STPI topic modeled the titles and abstracts of 1,229 journal articles attributed to MIDAS and non-MIDAS grants, with some dually-attributed articles. STPI directed the model to ignore disease-specific terms in addition to standard stop words when analyzing the content of each article, in order to reflect the disease-agnostic nature of MIDAS research and encourage modeling algorithms to identify similarities between articles beyond the examination of specific diseases. These stop words are:

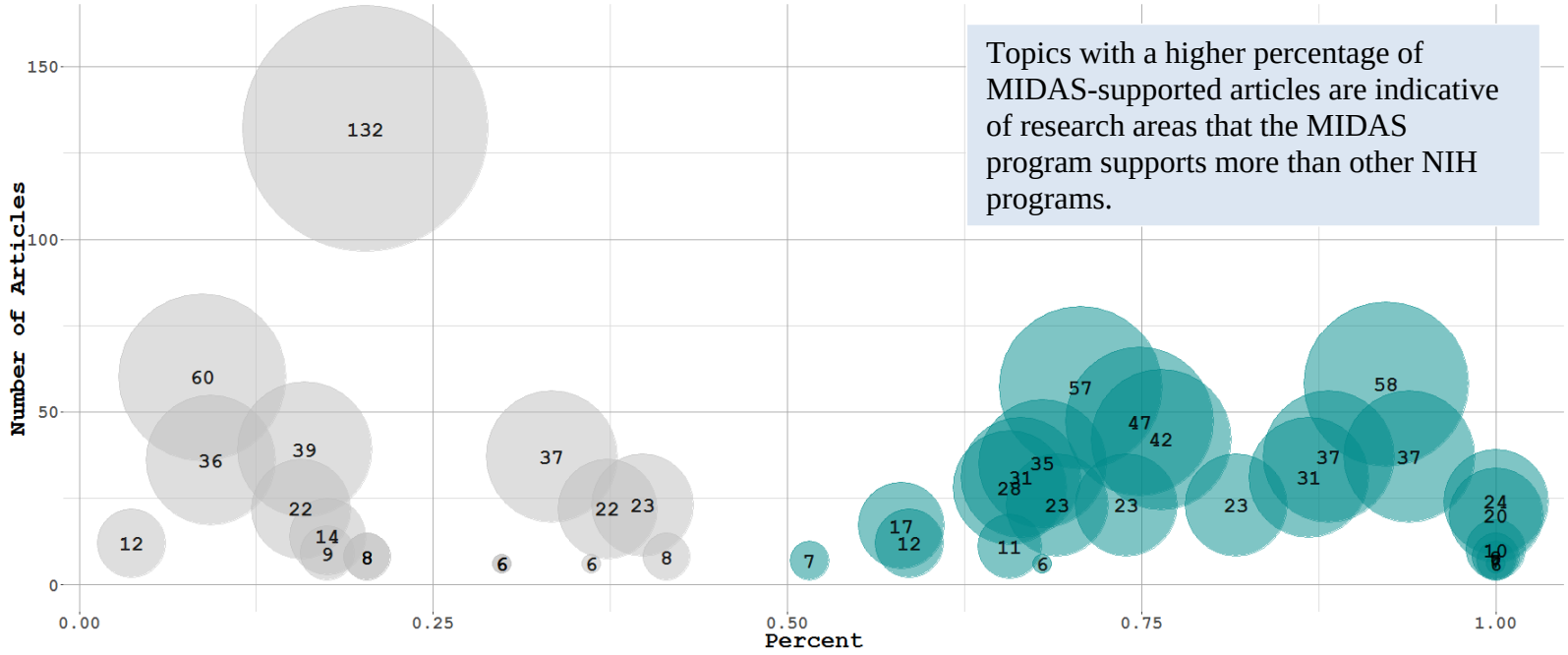
aedes	falciparum	mouth	shigella
aegypti	feral	mrsa	smallpox
aids	fever	mycobacterium	soybean
aureus	flu	nile	spp
avian	foot	norovirus	swine
brucellosis	hcv	pandemic	syncytial
ca	hiv	papillomavirus	tb
caries	hpv	pdm	tuberculosis
chikungunya	human	pertussis	typhoid
cholera	influenza	plasmodium	valley
cholerae	malaria	pneumonia	vally
cov	mdr	polio	vibrio
dengue	mdrtb	poultry	virus
denv	measles	rabies	west
dhf	meningitis	rift	wnv
disease	meningococcal	rust	
ebola	mers	rvf	

STPI's model produced 155 topics. Of these, STPI reviewed the content of 47 to characterize the discrete research areas represented in the set of journal articles. STPI selected these 47 topics based on the criteria that only topics that scored higher than 0.20 coherence and included 6 or more journal articles would be reviewed. These criteria helped ensure that the topics reviewed would mostly represent meaningful and representative research areas, and were reached based on an initial assessment of the topic model outputs.

Of the 47 topics, 41 could be identified to meaningful research areas. To differentiate MIDAS and comparison group research outcomes, STPI calculated the percentage of articles within each of these topics that were attributed to MIDAS grants. Percentages were weighted in proportion to the number of MIDAS and non-MIDAS articles in the original set of articles, and articles attributed to both MIDAS and comparison group grants were counted towards both MIDAS and comparison group percentage. Research areas represented by topics with a higher proportion of MIDAS-attributed articles were identified as the research areas which differentiated MIDAS from other infectious disease modeling research, and are discussed below.

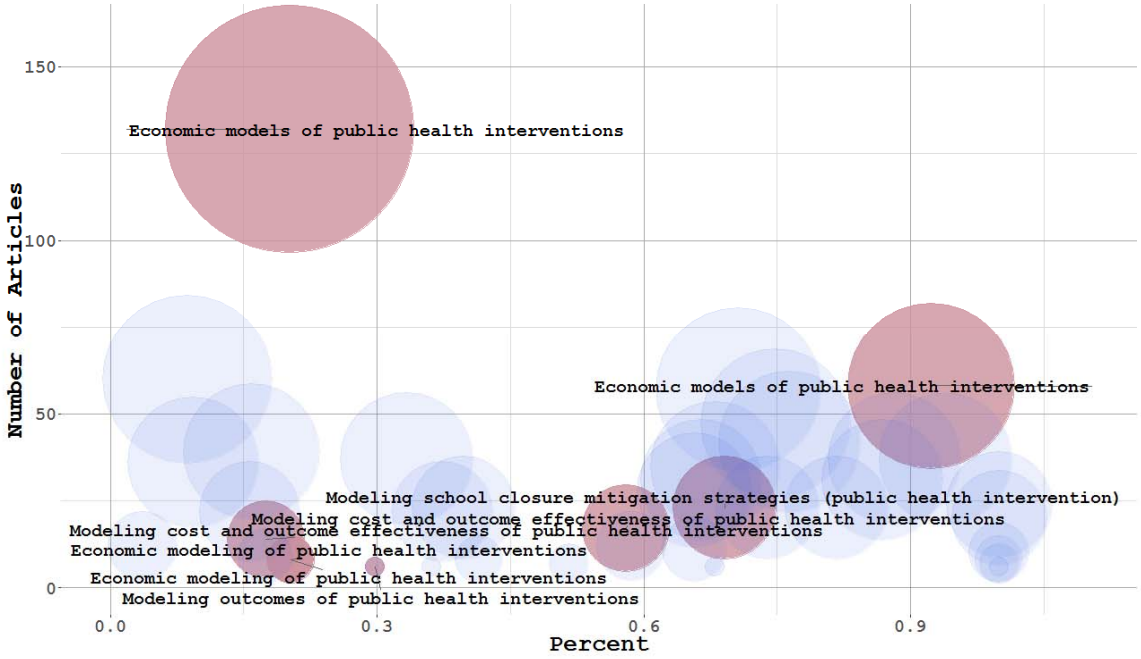
Results and Findings

Table 5 in the body of the paper presents the labeled topics and the number of articles within each topic, ranked by the weighted percentage of documents within each topic attributed to MIDAS. STPI identified 24 topics as "majority" MIDAS topics (higher than 50% weighted percent) and 17 topics as non-majority. This information is also displayed graphically in Figure C-1.



Note: Each circle represents a labeled topic, where size of circle corresponds to the number of articles within that topic. Topics vary widely in the percentage of articles that are MIDAS-supported. Topics are distributed horizontally by the percent of that topic's articles which are attributed to MIDAS grants and vertically by number of articles. Topics that are scored as greater than 50% MIDAS are colored teal, rather than gray.

Figure C-1. Topics Distributed by Percent MIDAS Articles and Number of Articles



Note: MIDAS researchers conduct research related to the modeling of economics, cost, and outcomes of public health interventions, but so do researchers not affiliated with MIDAS.

Figure C-2. Topics on Modeling Economics, Cost, and Outcomes of Public Health Interventions

In addition to identifying the areas of research that distinguish MIDAS, STPI notes two additional findings:

- Some research areas are represented among both majority MIDAS topics and majority non-MIDAS topics. For instance, 8 topics ranging from 18-92% MIDAS were identified as related to models of the economics, cost, and outcomes of public health interventions (represented in Figure C-2).
- STPI primed the topic model to group articles around similarities beyond the specific diseases modeled—dengue, HIV, influenza, etc.—by directing the model to ignore certain disease-specific terms. Some topics produced by the model were still characterized by the discussion of specific diseases rather than the modeling techniques or methodologies that are central to MIDAS research. These are presented in Table C-1.

Table C-1. Topics Corresponding to Research Surrounding Specific Diseases

% MIDAS	Articles	Topic Label
100%	24	Modeling of influenza outbreaks
82%	23	Modeling of Ebola outbreaks
68%	35	Avian influenza surveillance and modeling
36%	6	Modeling influenza
16%	22	Modeling and surveillance of Vibrio Cholera
9%	36	Modeling cost and outcome effectiveness of HIV interventions

These findings are noteworthy because they highlight the limitations of topic modeling. The appearance of identical research areas across multiple topic areas (Figure C-2) attests to the notion that language can reflect but never monopolize ideas, since different language can be used to describe the same concepts. Similarly, the emergence of disease-centric topics even with disease-related stop words removed showcases that deeper language similarities exist between papers discussing a disease. Not all the topics formed by topic models may reflect real or meaningful differences in subject matter, but often they do.

Appendix D.
Scientific Expert Panel Assessment

STPI created an initial pool of panelist candidates from NIGMS staff recommendations of people familiar with the MIDAS program, including grant reviewers, as well as identifying plenary speakers and conference organizers from conferences where MIDAS researchers have presented or published. To narrow down the pool of candidates, only those whom have not received MIDAS funding, and hold an academic appointment of Assistant Professor or above were considered. STPI made final panel selections in consultation with NIGMS staff.

Table D-1. Scientific Expert Panel Members.

	Name	Title	Department	Institution
1	John Brownstein	Assistant Professor	Biomedical Informatics	Harvard University
2	David Gurarie	Professor	Mathematics	Case Western University
3	Helene Carabin	Presidential Professor	Biostatistics and Epidemiology	University of Oklahoma Health Sciences
4	Abba Gumel	Professor	School of Mathematical and Statistical Sciences	Arizona State University
5	Joseph P. Messina	Professor	Geography	Michigan State University
6	Robert Spear	Professor	Environmental Health Sciences	University of California, Berkeley
7	Andrew Dobson	Professor	Ecology and Evolutionary Biology	Princeton University

Expert Panel Assessment Protocol

The expert review questionnaire reproduced herein included a short introduction, instructions, a packet of information to review, and questions at the end. To elicit the panel's views of the MIDAS program's education and training mission, STPI provided the

panelists with the following introduction, followed by the Education and Training Outcomes Section of this report.

> "The MIDAS network has a threefold mission; research; education, training, and outreach; and policy support. As part of fulfilling its education and training mission, MIDAS awardees are required to, "provide national leadership in training a new generation of infectious disease modelers." Please review the following information about the career outcomes for MIDAS-supported students and postdocs, and provide your feedback in the space provided at the end."
>
> [Education and Training Outcomes Section]
>
> "Given the data/information above, and your own experience with academic training, do you think MIDAS has fulfilled its mission of building capacity in the field of infectious disease modeling research? Please comment below:"

To add context to the results of the topic modeling analysis, which identified 11 research topics that differentiate MIDAS from non-MIDAS research (Table 6), the expert panel was asked to assess the novelty of those 11 topics. STPI provided the panelists with the following introduction, followed by the five-earliest published abstracts from each of the 11 topics.

> "STPI's analysis of MIDAS-funded research revealed a number of general research areas where MIDAS is particularly active, especially in comparison to other NIH-funded infectious disease modeling research. Please review the attached document which contains five MIDAS-supported research article titles and abstracts for 11 separate research areas (55 abstracts in total). At the bottom of each set of 5 abstracts, provide your opinion as to whether the research was novel to the field at the time of publication. You can assess the novelty of the research using the 1-5 numerical rating system at the end of each set of abstracts (1 = research was **not novel** at the time of publication; 5 = research was **very novel** at the time of publication). Use N/A if you are not familiar with the subject matter. Please also provide written feedback in the text fields at the end of each set of abstracts." [*Note: Author names and publication titles have been redacted.*]

Results: Education and Training Outcomes Expert Panel Assessment

Panelist 1:

> It is encouraging to see that MIDAS successfully trained 150 individuals in the field of disease modeling, a field where there still clearly lacks expertise. The split among undergraduate, graduate and post-doctoral students is expected for this field given the need for a very strong mathematical background to develop unique models. Hence, I feel that MIDAS has fulfilled its mission in terms of the number of trainees. However, it would have been

helpful to see the data in terms of the number of trainees per PI or per investigator involved in the funded grants and maybe collect some data on the satisfaction of the trainees with their training and current research. In evaluating where the trainees are now, the distribution of the trainees is also to be expected. It would have been helpful however to know how long each training lasted to evaluate if more trainees may have been trained with the available resources. Nonetheless, it is encouraging to see that most post-doctoral trainees now hold an academic position which will undoubtedly contribute to continuing to grow the building of capacity in this field. It would be important to if those now in academic positions are teaching and mentoring students of their own. It is also encouraging to see that 6 of the trainees are in public health positions. Such positions are likely to have more of an impact on using modeling to influence policy. To conclude, I do believe that MIDAS has fulfilled its mission in building capacity and that what has been built will lead to more students being trained by the trainees of these grants.

Panelist 2:

It is quite impressive that MIDAS produced 51 tenure-track position holders over the 10-year period (in addition to 6 graduates working in government agencies and about 50% of the 143 traced supported students are still in graduate schools or pursuing postdoctoral studies). I think MIDAS has fulfilled its capacity building objective over this time period.

Panelist 3:

Infectious disease modeling is an emerging interdisciplinary field. The research and teaching faculty is typically spread among other (traditional) programs, like Mathematics/Statistics, Biology, Computer science, Public health, Medicine et al, which often makes training and capacity building a challenging task MIDAS excellence centers and individual projects plays instrumental role in organizing, directing education and training activities on different levels (from undergraduate through post-doc). Among other activities I mention regular seminar series and summer graduate Schools and programs. Some of my students have attended such programs (CICID), and found them very beneficial for their study and research. Another useful contribution to research and education are modeling and dataset tools (e.g. Tycho, GLEAM et al)

Panelist 4:

To the basic question of whether the program has successfully fulfilled a training mission, I offer a qualified "Yes." This decision is based on two specific considerations. First, without other details regarding the MIDAS program, particularly performance relative to similarly funded NIH programs it is impossible to assess relative program performance. Second, the trend in

funding over time would clearly influence the trends in placement. For example, if the program increased in funding over time the numbers of students funded would be expected to increase resulting in low numbers of currently tenure system employed former students relative to the population as a whole. Assuming that the program has been relatively stable the population performance looks acceptable. The definitions of types of positions are important, though. Post-docs and non-tenure track faculty mean different things at different institutions, and at many places mean exactly the same thing. Also, this may be factor in the higher number of post-docs funded than graduate students. In general, these positions are much closer in structure than would be grouping non-tenure track and tenure track faculty. Given this, the number of tenure line appointments is a bit disappointing. This may be, at least, in part due to the timing of the MIDAS program when virtually all the universities paused in hiring. This unfortunate timing may have permanently removed some very qualified people from the traditional tenure system and also delayed any successes of MIDAS generally. I was pleasantly surprised by the field data. Of the charts, this looks to be the best indicator of training pathway success. Overall, the MIDAS program appears to have meaningfully contributed to the pool of scientists prepared to deal with infectious disease threats.

Panelist 5:

I do not recall how long the MIDAS program has been in operation or have a measure of its overall scale, but I find the productivity numbers impressive. In particular, to have 36% of the trainees now in faculty positions, tenure-track or otherwise, is a large number particularly in view of half of the total trainees still being in post doc or graduate student positions. The distribution of trainees into various employment sectors is about what I would expect, but the number in government public health departments is disappointing. This suggests the diffusion of modeling skills into the practitioner community is still modest at best.

Panelist 6:

The figures provided suggest that MIDAS did an excellent job in its educative mission of capacity building in the area of mathematical models of infectious disease. My biggest concern is the lack of detail about the more specific areas in which the students and post-Docs described were trained. If it is in the labs whose work is described [in the Scientific Outcomes section of the expert panel assessment], then I worry we have a group of people who have been trained in a somewhat superficial and ineffective way and at a level that is not inspiring to aspire to.

Panelist 7:

I am somewhat connected to MIDAS and have seen people supported under MIDAS that are now independent faculty. Therefore, I absolutely [believe that] MIDAS itself has trained the next generation of disease modelers with a unique set of skills that are beyond their advisors. Because of the new availabilities of computation and new types of data, absolutely you can credit MIDAS with a lot. In terms of bang for your buck, can't make that comment without benchmarking. But MIDAS did meet the basic bar of building capacity. Looks like the number of people trained is within reason for funding.

I am not aware of people that have gone on to public health departments, most of the ones I have known have gone on to independent academic careers and starting to be successful on their own. From a public health standpoint, it is great news to hear that that is one path that the training program has led toward. If think about the conventional jobs for disease modelers, working at public health department or CDC is not standard. So for any number of people to have roles in these areas is interesting because they are not standard epidemiologists and public health professionals. I was expecting that number to be zero, so six is pretty good.

Results: Scientific Outcomes Expert Panel Assessment

Research Area/Comments	Panelist 1	Panelist 2	Panelist 3	Panelist 4	Panelist 5	Panelist 6	Panelist 7
Rating for Area 1:	**4**	**4**	**5**	**NA**	**5**	**2.25**	**NA**
Comments	The research focusing on ADE is of particular interest and relevance in an area where vector borne viruses with multiple strains are re-emerging. A better understanding of what makes strains interaction become pathogenic is innovative. The research on Strept is also of interest, although its novelty may not be quite as marked.	The phenomenon of antibody-dependent enhancement (or immune enhancement hypothesis) in diseases with multiple heterologous serotypes poses significant public health challenge vis a vis the population-level control of such diseases. Modeling, and providing insight into, such complex dynamics is challenging. The series of abstracts under this Research Area present numerous novel results which enhance our understanding of interplay between the immune system and the various disease serotypes. In particular, the first paper under this category gave one of the earlier modeling results that showed, using data for dengue serotypes, and that outbreaks of the four dengue serotypes occur asynchronously. The authors further showed that for sufficiently small antibody-dependent enhancement, the number of individuals infected with each serotype synchronizes (where the outbreaks occur in phase). Chaotic dynamics were observed when the antibody enhancement exceeds a certain threshold (and, in this case, the total number of individuals infected with each serotype desynchronize	Papers 1-2 give novel application of the dynamical system theory (chaos) to explore regular and asynchronous outbreak patterns for multi-serotype infections (dengue).Papers 3-4 adopt methods of population biology to immune regulation of multiple-strain pathogens, to explain strain diversity. The last one attempt to unravel the structure of nontypeable invasive diseases. All offer innovative applications of mathematical ideas and methods to disease modeling		My rating here is based on the first three cited papers since last two do not appear to involve models. Characterizing the behavior of the ADE effects in the context of transition to chaotic behavior strikes me as very interesting and potentially very important.	There's some interesting theoretical stuff here, but it declines into anecdote and wishful hand-waving when tested against data.	

Research Area/Comments	Panelist 1	Panelist 2	Panelist 3	Panelist 4	Panelist 5	Panelist 6	Panelist 7
Rating for Area 2:	**4**	**3**	**3**	**2**	**2**	**1**	**3**
Comments	How different informational sources influence health behaviors, including hygiene, social avoidance and vaccination is key to controlling outbreaks and diseases. The research on behaviors associated to influenza and across communities are novel in identifying which source of information may impact behavior the most. The ABM is particularly novel and could really help public health professionals in planning information campaigns. The research on factors associated with influenza uptake is not as novel and has been explored in different context including that of people in high risk groups such as people living with HIV AIDS. The research on HO-MRSA is not particularly novel, just very descriptive.	The series of papers under this category present results that were fairly known during the periods they were published. The one that stands out (and seems new to me) is that pertaining to the knowledge of cause of influenza as well as perceived PHP (personal hygiene practices) and intention alone is an unreliable predictor of future vaccine uptake.	Papers 1-3 apply behavioral (decision making) models and statistical data analysis to study the effect of information on behavioral patterns and vaccination during H1N1 outbreak in Hong-Kong. #4 is primarily statistical analysis of nosocomial MRSA. #5 applies agent-based methodology (information networks) to health related behavioral patterns in different communities. I can't judge novelty of 1-3, #4 seems fairly routine, #5 more original. My overall assessment is tentative.	This is a difficult grouping to assess as a single set. However, none of the papers was particularly novel. These are all mostly traditional social science. The first two should really be one paper. Other NIH programs cover these topics.	This RA seems a fairly standard application of modeling to attempt to understand behavioral factors and, as such, is hypothesis generating at best. The use of agent-based models in this context may be innovative.	This is classic social science survey work – vacuously building castles of policy on sands of weak correlation. Illustrates all the costs and none of the rarely realized benefits of involving social scientists in epidemiology.	
Rating for Area 3:	**3**	**4**	**3**	**2**	**3**	**2**	**2**
Comments	The geographical data to develop ABM are by far the most novel of this section. There is an increasing number of data sources to better understand movements of people which in turn influence contacts that are the base of transmission of infections. The use of such complex data in more powerful computers, which could not be done in the past is novel. The other topics are not as exciting and details on the use and effectiveness of the MREP would have been most welcome to really evaluate how novel this actually is.	The design of models repository (MREP) database represents a really laudable contribution to the disease modeling and public health community. It allows modelers and public health practitioners to run various disease scenarios using relatively easy-to-use and realistic models. The design of GIS-based and geo-spatial databases for agent-based models is quite interesting. The other results under this category seem fairly standard.	The main thrust are computer tools and support systems for organizing and storing geospatial human population and mobility data collected from different sources, that could be used for agent based modeling and prediction. While useful development, the innovation component seems limited	This topic is an important topic and certainly deserves to be part of MIDAS. None of the papers selected as representative are particularly novel. These are clearly incremental papers and datasets. I am aware of the synthesized population data set, but few if anyone outside of RTI seems to have access to it. This kind of product is valuable, but without public access, of essentially no use.	The linkage of GIS-based data with agent-based models is an obvious, but important, idea. The work reported above contributes to the development of the often complicated infrastructure necessary to utilize the improvements that should result from marrying the two concepts.	There's some worthy stuff here, but it strikes as more book-keeping than model development. I don't feel at all inclined to rush out and try and locate any of the papers described and see no use for any of them in classes that I teach in this area. Assembling a large agent based model for the entire US based on movement data from 2002 is a useful academic start, but I would be very circumspect about any projections developed from epidemiological models that use these data. Here I note in passing that previous detailed agent-based models that are similar to those described produced exciting movies of influenza spread in the US, but these bore	

Research Area/Comments	Panelist 1	Panelist 2	Panelist 3	Panelist 4	Panelist 5	Panelist 6	Panelist 7
						essentially no resemblance to the patterns of spread observed when the influenza strain arrived from Mexico in 2012. Similarly assembling a library of models strikes me as a somewhat vacuous exercise; new models are appearing on almost a daily basis and there are enough good textbooks that most disease outbreaks can be modelled quite quickly from the basic tools already available, rather than retro-fitting older models that are often over-specified.	

Research Area/Comments	Panelist 1	Panelist 2	Panelist 3	Panelist 4	Panelist 5	Panelist 6	Panelist 7
Rating for Area 4:	**4**	**3**	**3**	**4**	**2**	**2.5**	**3**
Comments	My novelty score is mostly based on the last 3 abstracts. Understanding the role of HCW in the transmission of influenza is absolutely essential and had not been assessed in this type of context. The last abstract is extremely novel and could have very important impacts on attributing the source of infections, and responsibilities. The 4th abstract could lead to better and faster control of hospital-acquired infection. The first abstract was in the previous section. The second could be of interest but no results are presented.	One of the key new results under this category is setting up alternative vaccine sites does not necessarily increase vaccine coverage (as the sites may not cover some of the other key segments of the society). The result on the importance of incubation period as tool for identifying source of infection isn't new.	Statistical tools and data analysis to assess vaccine efficacy and efficient implementation, the role of health care workers in spreading influenza; detection of hospital outbreaks, and diagnostics of respiratory infections. Sound work, but doesn't look particularly innovative in terms of models or methodology	A mixed group of papers with varying impact. The topic is clearly and important piece in MIDAS. This ought to be a core topic. I have up-scored this because the theme is important. Only evaluating the papers, my score would drop to 3.	Most of the foregoing papers do not fall into my classification as applications of mathematical modeling.	There's some interesting work here, but while I would be delighted if it were produced by undergraduates working on their Senior research thesis, I'm not sure it qualifies as cutting edge work in the development of the next generation of epidemic models.	
Rating for Area 5:	**5**	**4**	**4**	**4**	**3**	**3.75**	**4**
Comments	The research on the impact of travel and travel restrictions is novel and, to my knowledge, had not been explored at the time of publication. The analysis of personal behavior using 1918 influenza outbreak are also novel. Altogether, this section was very novel at the time of their publication.	The first paper addresses an important problem of finding the best control strategy in an epidemic setting with multiple co-circulating diseases. The abstract did not, however, state what the ``best control strategy'' is. It is well-known that the timely implementation of control measures (pharmaceutical or non-pharmaceutical) is effective in combating disease outbreaks. Further, public awareness (or public health education campaigns) geared towards encouraging members of the populace to minimize engaging in practices that increases their risk of acquiring infection is also a well-known measure for effective control of disease outbreaks. One notable novel result is that restricting international travels, if combined with the implementation of other control measures, can cause important delays that reduces the risk of increased outbreaks.	Analysis of flu intervention strategies, ranging from theoretical (simplified) SIR model of #1, to more advanced distributed metapopulation systems and air-travel connectivity ##2-4); reconstruction of incidence time series of 1918 pandemic. Some of these works used innovative approaches and modeling.	Solid mathematical modeling papers and topics. Prior to my first service on the MIDAS panel these topics are what I imagined would be funded. These kinds of papers are products only of a program like MIDAS and NIGMS.	The first 2007 paper might be quite interesting, but there is no description of the methodology used. The third 2007 was certainly a timely and important subject as is the 2008 paper on the same topic using a modeling approach. The 2009 paper is apparently a novel application of a mathematical method used in optics, but not sure it can be called modeling in the usual sense.	There are some very interesting things being done here that are both useful and insightful. The retrospective work on the 1918 Influenza epidemic is insightful on several levels, not least because it illustrates how human behavioral modification is still likely to be more influential in determining the course of an outbreak than many types of public health intervention! The examination of the level of detail needed to model influenza distribution by global airline routes is also important and provides some important insights.	

Research Area/Comments	Panelist 1	Panelist 2	Panelist 3	Panelist 4	Panelist 5	Panelist 6	Panelist 7
Rating for Area 6:	**4**	**3**	**5**	**NA**	**4**	**2**	**5**
Comments	The research on influenza was really novel at the time of publication given the emergence of H5N1. These models helped design response. Exploring state response plans in conjunction to these models would have been even more novel. The research on oncho is less novel.	These studies essentially emphasize the known fact that a hybrid approach (of combining multiple non-pharmaceutical and pharmaceutical intervention strategies) is more effective in combating diseases like strategies than implementing a singular control strategy. The first paper shows that a hybrid strategy (involving the use of antivirals, pre-vaccination and quarantine) could lead to effective control of pandemic influenza strain even with a basic reproduction number (R_0) of a value as high as 2.4. This is an interesting result, and agrees with many of the studies published on related epidemics (including the 2003 SARS outbreaks).	Papers 1-2 apply large scale stochastic simulation model of flu pandemic and preventive strategies for realistically structured host populations (Thailand, US); #3 is the review of the current containment plans; #4 applies similar methods to #1-2 for Onchocerciasis treated with ivermectin; #5 explores surveillance methodologies for close contact groups. The works (particularly ##1-2) develop many novel methods and techniques	I marked this N/A because I do not understand the grouping. One is a duplicate from an earlier group and I don't see how it fits the others. I read one of these papers previously. I like the topic but these papers as a group were not particularly novel.	This area is the closest to my original interpretation of the scope and objectives of the MIDAS program and results it would produce. The scale of the stochastic model reported in the first 2006 abstract certainly has novel elements and the onchocerciasis paper is unusual if not novel in addressing the heterogeneity exposure to environmentally-mediated infectious agents. The latter also provides another example of the effectiveness of targeting treatments addressed years ago by Woolhouse et al.	There is some interesting stuff here, but nothing remarkably new, nor novel. The large scale stochastic simulations for influenza spread always look very impressive, but as mentioned above they bore very little resemblance to the spread of the swine flu influenza that appeared from Mexico in 2011/12. We've known for ages that targeted intervention will be more efficacious for treatment of Onchocerciasis and other nematode infections – all very worthy, but not very useful until we know what determines increased susceptibility and have mechanisms to detect hosts with high levels of infection.	
Rating for Area 7:	**3**	**4**	**4**	**4**	**3**	**2**	**3**
Comments	I cannot comment on the mosquitoes and lemur research. However, the influenza and zoonosis research are somewhat novel. If the workshop on identifying zoonotic emerging infections leads to new approaches, then novelty will be insured. Impact of climate on bird migration is not all that novel, nor is exploring assortativeness in influenza transmission.	The paper on the effect of climate change on the ecology of avian influenza is very interesting. While climate change alters bird migration and influence viral transmission cycle and survival (outside the host), the effect of climate change (environmental factors) on the transmission dynamics of the highly pathogenic avian influenza in domestic poultry. The paper on modeling the host-seeking behavior of mosquitoes is equally interesting, and the results presented, particularly the attack abatement effect of host aggregation, seem new (as of the time of publication). One other notable contribution, under this Research Area, is the laying of a conceptual framework for achieving	Papers #1 reviews data on Madagaskar lemur; #2 gives a broad outline of environmental, climate inputs on avian flu. There is no modeling component in ##1-2. #3 is a theoretical work on assortative mixing and epidemics extinction for flu-like pathogens. #4 is broad outline of surveillance strategies on emerging viral pathogens. #5 develops a realistic model of blood seeking behavior by mosquitos, which could have may potential applications. The novelty rating is an average assessment with a higher rate for #5	This is an important topic and suitable for MIDAS. These papers may have had more impact than any of the other groups of papers.	Very modest use of models. Score based on second 2012 abstract which strikes me as interesting and novel.	This seems to be a mixture of "fishing exercises" – the virus study of Malagasy lemurs looks like a boondoggle to me! It involves no use of models and would produce at best a weak paper in Journal of Wildlife Diseases. The influenza work breaks very little new ground. The work on climate change and avian influenza is essentially vacuous and the workshop on emerging zoonosis is almost a monthly event globally. Most of these have produced nothing substantial that develops the ideas outlined in Lloyd-Smith et al 2012. The spatial models for mosquito behavior are interesting, but I do not see how they will be implemented in any way that is more insightful than Hasibeder and Dye (1992ish).	Wide variety of papers in this research area.

Research Area/Comments	Panelist 1	Panelist 2	Panelist 3	Panelist 4	Panelist 5	Panelist 6	Panelist 7
Rating for Area 8:	**NA**	**NA**	**NA**	**4**	**NA**	**2.5**	**4**
Comments	I do not feel competent to review the novelty of this section.	The abstracts under this Research Area fall outside my general area of modeling expertise.	Papers vary from theoretical study (statistical analysis of person-to-person transmission #1) to syndromatic surveillance (#3), geospatial variability for WNV (#4); SatScan monitoring technologies #2, #5. I can attest novelty of some work #1, #3, but less certain about others.	The development and application of advanced geospatial statistical methods is critical to advance MIDAS. The challenge is moving beyond incremental. These papers are novel but incremental. This is in part the product of the funding model and the tendency to reward past success.	This set of abstracts seems to describe an area with little mathematical modeling but substantial methodological content in statistics.	There are some useful statistical method developed here that may have potential for application to a range of other pathogens.	
Rating for Area 9:	**5**	**4**	**5**	**3**	**NA**	**1.5**	**4**
Comments	Antiviral resistance to influenza treatment and the impact of counselling on improving adherence to HAART are topics that had not been well explored before MIDAS. In fact, oseltamivir was a relatively new drug at the time of the H1N1 and H5N1 outbreaks. So these studies really helped exploring the impact that resistance could have	It is true that the issue of resistance development must be taken into account in times of pandemic outbreaks of (treatable) diseases. It is interesting to know that public health counseling (to improve adherence to the highly-active antiretroviral therapy against HIV infection) ``provides only modest benefit as a tool for HIV prevention but can provide significant benefit for individual patients at an affordable cost".	These papers deal with important evolutionary questions in the spread and control of viral infections, among them host heterogeneity and pathogen adaptation, effect of prophylactic/symptomatic drug treatment and antiviral resistance, the role of stochasticity. I find them highly innovative.	With the exception of the HIV and neuraminidase papers, I've read these for my own work. The first two papers listed are particularly important.	The 2006 abstract sounds interesting but the description is inadequate to judge its novelty. The remaining papers focus largely on the use of models to generate hypotheses related to resistance related factors likely to influence epidemic development. Novelty here requires a detailed knowledge of the resistance literature.	There are again some interesting results here, although all seem to have been produced using the barest bow to epidemiological data. It reads to me as if the author is fascinated by drug resistance, has a set of mathematical tools to model this and will therefore use them to examine vaccine or drug resistance in influenza. An interesting exercise, but a slim long-shot.	
Rating for Area 10:	**5**	**3**	**NA**	**4**	**NA**	**2.5**	**3**
Comments	The GBD initiative of the University of Washington in Seattle has emphasized the need for better global surveillance systems which would generate more valid estimates of prevalence and incident cases of infections, with or without resistance to section. I find this section innovative in its use of modeling to combining sources of data and in trying to obtain more valid estimates of prevalence, especially in resource-poor countries.	The results under this Research Area are fairly standard. For instance, Mycobacterium tuberculosis acquiring mutations during latency is well-known (at the time of publication). It is also known that the use of informatics and genomics can make microbial surveillance more efficient and effective at preventing infections and improving their outcomes. The study on resistant and wild TB says ``Interventions should aim to reduce the infectious duration for those with drug-resistant disease and improve infection control", which is fairly obvious.	#1 testing assessment strategies for TB spread based on public records #2 statistical analysis of spatio-temporal patterns of TB #3 review of surveillance (TB, HIV, malaria) is resource poor settings #4 analysis of multilevel surveillance systems #5 genome sequencing of MTB I can assign novelty rating 4 for #1, but leave N/A for others	These are very good and novel papers. The fit within MIDAS is basically scale, but they are important. I am skeptical that scale should factor prominently in a MIDAS award, but it seems to have percolated throughout the program, not just this theme.	Only the 2010 abstract appears to involve modeling and the remainder various aspects of TB surveillance strategies or pathobiology. The single modeling paper addresses and interesting issue in a straightforward way and would rate a novelty score of 3.	This work really gets to grips with key problems in the control of tuberculosis in India and South America. This ground has been worked before in other areas, but the studies described here take the work further and bring in new approaches that seem to provide new insights.	

Research Area/Comments	Panelist 1	Panelist 2	Panelist 3	Panelist 4	Panelist 5	Panelist 6	Panelist 7
Rating for Area 11:	**2**	**3**	**NA**	**1**	**NA**	**1.5**	**2**
Comments	This section is not particularly novel with the use of survey data to obtain descriptive statistics on some outcomes.	One of the interesting findings under this category is the importance of understanding (and incorporating) cultural/behavioral patterns that are specific to high risk (ethnic) subgroups in the design of effective preventive strategies against the spread of a disease in the (overall) population. It is also interesting to see that lack of ``trust'' (to public health information disseminated to various communities) can lead to different dynamic patterns of belief polarization between racial communities.	#5 develops agent-based modeling for information networks in different communities. Such methodology could be useful for public health assessment and policy making (novelty 4). Other studies are mostly data analysis (N/A)	In the context of MIDAS, these papers have little impact. They could easily, maybe more easily, have been funded and managed from within another program.	These abstracts report rather standard epidemiological analyses of various aspects of diverse diseases using mainly national survey data. The exception is the 2012 agent-based network paper that reports the development of an innovative dynamic model which would rate a novelty score of 4 on its own.	This work seems to be based on a lot of correlation leading to some rather weak causative inference. As far as I can see there has been almost no use of mathematical models. It's therefore not clear to me how significant the reported effects will be in determining observed patterns of disease prevalence and spread.	

Appendix E.
Outreach Activities

Award Number	Year	Reported Outreach Activities
U01GM070698	2013	Develop 10th Anniversary celebratory video.
		Conducted focus groups with educators regarding utility of models in high school and college undergraduate curriculum.
		Supported MIDAS booth at SACNAS, ABRCMS national student conferences.
		Supported MIDAS booth at APHA and NACCHO public health conferences
		Collaborated with NIH to develop hand-washing activity at national Science and Engineering Expo.
U24GM087704	2010	Attended SACNAS with MIDAS booth and materials from research groups.
		Attended ABRCMS with MIDAS materials from research groups and developed an infectious bug interactive game to show dissemination among universities and colleges of the attendees.
		Attended and displayed the MIDAS booth at ABRCMS and SACNAS student conferences.
		Developed a MIDAS outreach newsletter.
	2011	N/A
	2012	Displayed MIDAS booth at SACNAS and ABRCMS and provided MIDAS pop-up tabletop "booth" at Tropical Medicine Annual Meeting and Epidemics.
		Developed a panel application with other MIDAS Researchers for the fall American Public Health Association meetings.
		Attended SACNAS with MIDAS booth-and materials from research groups.
		Attended ABRCMS with MIDAS book and materials from research groups.
		Supported the MIDAS booth and support for workshops and sessions at six conferences and meetings including the American Public Health Association and the National Association of City and County Health Officials Public Health Preparedness Summit. These conferences and meetings have enabled MIDAS to disseminate resources more widely and inform stakeholders about MIDAS. The MIDAS booth presents materials from all of the research groups and occasionally features specific tools and educational opportunities.
	2014	N/A
U01GM087728	2012	N/A
	2013	Promoted MIDAS at the Advancement of Chicanos and Native Americans in Science National Conference
	2014	N/A
U01GM076426	2006-09	N/A
U54GM088491	2010	N/A
	2012	N/A
	2013	Engaged in a K-12 Outreach Partnership with the Univ. of Pittsburgh Health Careers Scholars Academy for a 4-week program for high school students from across

		Pennsylvania interested in public health, research and healthcare careers. Researchers affiliated with MIDAS introduced students to research relating to computational modeling and simulation of infectious disease within the public health concentration course.
		Hosted the 2012 Undergraduate Data Research Palooza, a national undergraduate student competition which engaged students in computational studies in public health and applied student creativity to explore innovative ideas for public health.
		Hosted the Epistemology Think Tank Workshop (May 8-9,2012) - Pittsburgh, PA, a 2-day gathering of 26 experts various aspects of modeling, in policy applications, and in epistemology and philosophy of science, under the auspices of MIDAS and in cooperation with the Center for Philosophy of Science at the University of Pittsburgh.
		Partnerships established with both a leading historically black college & Hispanic serving institution after site visits during the summer of 2012 at University of Texas at San Antonio Honors College (San Antonio, TX) and North Carolina A&T University Honors College (Greensboro, NC).
	2015	N/A
U01GM076499	2007	N/A
	2009	Helped the New Mexico Supercomputing Challenge program (http://www.challenge.nm.org/) and the Santa Fe Institute's GUTS (Growing Up Thinking Scientifically) project (http://www.santafe.edu/education/k12-project-guts.php) establish epidemiological modeling programs for middle and high school students across New Mexico. We trained program staff, teachers and students. We now have hundreds of kids collecting H1N1 data In their schools, downloading data from the web, and making agent based models to evaluate intervention strategies.
U01GM097658	2012	N/A
	2013	N/A
	2014	N/A
	2015	Participated in the following conferences and workshops: o Joint Annual Meeting of the Japanese Society for Mathematical Biology and the Society for Mathematical Biology, Global Disease Forecasting with Wikipedia (July/August 2014). o Mathematical Association of America MathFest Conference, Epidemic Forecasting and Monitoring using Modern Data Assimilation Methods (August 2014). o Approximation, Integration, and Optimization Workshop (September/October 2014). o Society for the Advancement of Chicanos and Native Americans in Science, Alternative Careers in STEM (October 2014). o The Institute for Operations Research and the Management Sciences (INFORMS) Annual Meeting, Global Disease Forecasting using Wikipedia (November 2014). o Blackwell-Tapia Conference, Global Disease Forecasting using Wikipedia (November 2014).
U01GM070694	2007	N/A
	2008	N/A
	2012	Developing the DSI:DC game and have supported a high school intern who received 3rd place in a regional Intel science fair
		Redesigned the Virus Tracker game/simulation previously used in USA Science and Engineering fairs for use at the 2013 Boy Scout

		Jamboree and created a Virus Tracker-in-a-box that we will loan to 8-12th grade teachers.
	2013	N/A
U54GM111274	2015	Creating the NIGMS-supported Summer Institute in Statistics and Modeling in Infectious Diseases (SISMID) in Seattle at U Washington (R25; Director: Halloran). Held for the seventh time, this is a 4 week course with 9 faculty and 15 courses which usually attracts about 170 participants and provides 10 scholarships for students.
		Hosted an International Clinics on Infectious Disease Dynamics and Data that builds capacity in epidemiological modeling in the US and Africa. The fifth annual Clinic on Meaningful Modeling of Epidemiological Data (MMED 2014) was held at the African Institute for Mathematical Sciences (AIMS) in Muizenberg, South Africa on June 2 - 13, 2014. Five group projects from MMED 2014 were considered for follow-up funding from CIDID to support a paper-writing workshop.
		Providing weekly seminars to the HSPH community pertaining to infectious disease epidemiology. Speakers included Dr. Niel Hans, Interuniversity Institute for Biostatistics and Statistical Bioinformatics; Dr. David Smith, Emerging Pathogens Institute at the University of Florida; Dr. Aaron King, University of Michigan; Dr. Rustom Antia, Emory University; Dr. Todd Allen, Massachusetts General Hospital; Dr. Andrea L. Graham, Princeton University; Dr. Stephen Bentley, Welcome Trust Sanger Institute; Trevor Bedford, University of Michigan; and Andrew Rambaut, Institute of Evolutionary Biology, University of Edinburgh.
U54GM088558	2010	Completing an annual symposium on "Surveillance for Decision Making in Emerging Diseases: Lessons from the 2009 H1N1 Pandemic Influenza, which was held June 14-June 15, 2010. The symposium actively included 100 representatives of academia and public health and public health practices and interfaced with outreach efforts to underrepresented groups.
		Enhanced and expanded the Summer Program in Epidemiology for students from under-represented groups. In 2012, seven students attended the program from June 4 - June 30, 2013. 10 students attended intro epidemiology and biostatistics courses, STATA training, and worked on research projects with faculty mentors. Four of the students worked with Dr. Lipsitch on a project titled "Assessing the population level impact of cholera vaccination campaigns"; the remaining students worked on a project titled "Validation study of body size recall in men and women."
	2011	Exhibiting at various conferences which have included, the Society for Advancement of Chicanos and Native Americans in Science (SACNAS), the Annual Biomedical Research Conference for Minority Students (ABRCMS) annual conference, Idealist.org fairs in San Francisco and Los Angeles California, The Morehouse Innovation Expo, the 2011 Emerging Researchers National (ERN) Conference in STEM, and the NIH Graduate and Professional School Fair. Dr. Lipsitch spoke at a special session for minority scientists at the Society for Industrial and Applied Mathematics in Pittsburgh July 2010.
		Creating and coordinating four short courses. The first short course titled, A Practical Short Course on Mathematical Modeling of Infectious Diseases: Preparedness and Response for the 2009 Influenza Pandemic, was held September 8-9, 2010 in Hong Kong, China; a second short course was held June 6-7, 2011 at the Centers for Disease Control, titled A Practical Short Course on Infectious Disease Modeling; the third short course, sponsored by WHO with CCDD collaboration, was held June 14-16, 2011 was titled Short Course on Mathematical Modeling and Influenza Vaccination Strategies, and also held in Hong Kong, China.

		Partnering with the department of biostatistics on their HSPH Summer Program in Quantitative Sciences. This is a 4 week intensive program for 8-10 undergraduate students from underrepresented groups. In an effort to increase our presence within the program and to make students aware of our opportunities, Dr. Lipsitch and a postdoctoral fellow are mentoring three students.
	2012	N/A
	2013	Provided school visits for targeted programs interested in increasing diversity in science at the graduate level.
		Organized an outreach event that took place as part of the April 2015 National MIDAS Network Meeting in Atlanta, Georgia. Twenty students from underrepresented minority (URM) groups joined the MIDAS Network meeting, attending journal clubs, general session talks on HIV modeling and then joined the MIDAS community for networking at the meeting's poster session.
		Exhibited at various conferences including: SACNAS (Society for the Advancement of Chicanos and Native Americans in Science) conference, and the ABCRMS (Annual Biomedical Research Conference for Minority Students) conference. Graduate students from underrepresented groups and educate them about funded opportunities to use their quantitative skills in infectious disease modeling and public health: Equal Opportunity Employer's STEM Diversity Career Expo for the Disabled (Sept,2014); Women in Math in New England (WIMIN) (Sept,2014); Scientist Graduate School & Career Fair (Sept,2014); Grace Hopper Celebration of Women in Computing (Oct,2014); American Indian Science and Engineering Society (AISES) (Nov,2014); Math Alliance's Field of Dreams Conference (Nov,2014). Ms. Larsen also presented at the following conferences on the topics of best practices for applying to summer research programs and the importance of diversity recruitment in STEM fields: Dartmouth SIAM Chapter (Oct,2014); Undergraduate Mathematics Research Conference (March) University of Puerto Rico (Feb,2015); Postdoc Dr. Amy Wesolowski (advised by CCDD's Dr. Buckee; funded by James. S. McDonnell Foundation) presented to the students as an invited speaker on an introduction to mathematical modeling of infectious diseases (Feb, 2015); Infinite Possibilities Conference (for women in underrepresented groups in mathematics) (March, 2015); National Student Leadership Conference (partnership with Harvard Medical School Office of Diversity Inclusion).
		Participated on an alumni panel at Brandeis University for low income/ first generation undergraduate students about how his background has influenced his professional journey as a "citizen advocating for equality for those in need around the world".
		Visited and Spoke at Various Institutions: Visited the Mathematical Theoretical Biology Institute. In addition, Felisa Nobles and Dr. Molly Franke visited the City University of New York, John Jay College in Sept 2012. They met with Dr. Anthony Carpi's group, PRISM (Program for Research Initiatives for Science Majors) as an introduction to a professional career series where 50 students were in attendance. Ms. Nobles, and two students (master's and doctoral level) visited the Health Professionals Mentoring Program at University of California at San Diego in November 2012.
	2015	Hosted the fourth outreach conference, hosting 68 students from underrepresented groups at the Harvard TH Chan School of Public Health for a 1.5 day conference. The conference programming was expanded this year. In addition to offering scientific talks on different topics in infectious disease modeling and professional development sessions, we were also able to run

		an Introduction to Modeling workshop for students. Two URM doctoral students, Corey Peak and Patrick Mitchell (who started in the Harvard Chan Epidemiology program via the CCDD funded Master of Science program) led the workshop. The workshop included an introductory lecture, paper practicum and discussion section, and a computer practicum exercise using an online tool from the Salathé Group. We also hosted Dr. Li Feng, a professor in the Department of Mathematics and Computer Science of Albany State University, a historically black college in Albany Georgia. Dr. Feng joined us to learn more about infectious disease modeling so he can incorporate the subject into the mathematics and computer science curricula at Albany State.
		Partnered with the Harvard Chan Department of Epidemiology to host a cohort of nine undergraduate students from underrepresented groups in our 2015 Summer Program. Two of these students are also alumnae of our annual Outreach Conference. The summer program curriculum consists of three parts: introductory coursework in epidemiology and biostatistics; formal lectures, which are provided by faculty members with different foci in epidemiology; and a group research project where students will investigate a question of public health relevance that interests them. Students also get to take professional development sessions, including GRE prep work, along with other networking opportunities.
		Created the first Annual MIDAS conference for undergraduates promoting diversity in mathematical modeling. As a Center, collaborating with other MIDAS groups on the topic of outreach was a goal and this conference provided a worthwhile opportunity. The conference, held in May 2012, had approximately 80 attendees from institutions across the country. The conference showcased the work of various investigators within the MIDAS network, and showed students the versatility of mathematical modeling in public health. The day consisted of 35 minute presentations from MIDAS investigators and collaborators, a networking lunch with the presenters, and an hour long journal club where students were divided into smaller groups to discuss the paper, "Social Contacts and Mixing Patterns Relevant to the Spread of Infectious Diseases," by Mossong et al.
		Held the third annual symposium on the topic of Antimicrobial Resistance: Biology, Population Dynamics and Policy Options. The symposium was held over two days and included topics on Tuberculosis, MRSA, Pneumococcus Nosocomial Infections, panels and presentations from industry representatives and a poster session. More than 130 people registered for the course.
		Coordinated the fourth short course on modeling; this year's course is scheduled for two sessions from June 17 - June 21, 2012 and will have approximately 100 attendees. The course, titled A Practical Short Course on Mathematical Modeling will be held in Bangkok, Thailand.
		Spoke at conferences geared towards engaging minority students in STEM fields. Murray spoke to an over 250 students at the Annual Biomedical Research Conference for Minority Students in St. Louis, November 10, 2011. Her talk "Understanding the transmission dynamics of drug-resistant Tuberculosis: A multidisciplinary approach," targeted students from a range of backgrounds as she integrated mathematical modeling, genomics and molecular biology into her discussion. She also engaged students at the HSPH booth in the exhibit hall, visited posters, etc.
U01GM087719	2010-5	N/A
U01GM110712	2015	N/A

U01GM110721	2015	N/A
U01GM110744	2015	N/A
U01GM070749	2008-13	N/A
U01GM087729	2010-2	N/A
U01GM076497	2007-10	N/A
U01GM097661	2011-15	N/A
U01GM070698	2007-8	N/A
U24GM110707	2014-16	N/A
U01GM110748	2015	N/A
U01GM076672	2006-09	N/A
U01GM070708	2007-8	N/A

Appendix F.
Examples of MIDAS-supported Policy-Related Activities

Level of Engagement	Researcher	Researcher Affiliation	Public Health Official or Partner	Outcome	Publications and Products
State	Don Burke	University of Pittsburgh		**Public Health Tool:** A modified version of FRED was made into an iPhone app. When released on Twitter, senators in CA used the app to affect legislation on the restriction of vaccine exemptions in CA.	http://fred.publichealth.pitt.edu/measles/
National	Don Burke	University of Pittsburgh	Tim Lant, Nicole Lurie, ASPR/BARDA	**Epidemic Support:** Used modeling to determine patterns of disease spread during the Supported Weekly Interagency Conference Calls for Influenza crisis.	Sourced from interview data
National	Shawn Brown	Pittsburgh Supercomputing Center	PCAST	**Epidemic Support:** Computational modeling at PSC helped policymakers locally and nationally evaluate strategies for responding to the 2009 H1N1 flu epidemic	http://psc.edu/science/2010/h1n1/

National	Shawn Brown	Pittsburgh Supercomputing Center		**Networked Interactions/consultations**: Report to the President on the US Preparations for the 2009 H1N1 Influenza	PCAST Report: Report to the President on U.S. Preparations for 2009 H1N1 Influenza
National	Willen van Panhuis	University of Pittsburgh	CDC, NIH	**Public Health Tool**: Project Tycho - Catalogued contagious diseases in the US from 1888 to the present	https://www.tycho.pitt.edu/
National	Elizabeth Halloran	Fred Hutchinson Cancer Research Center	DOD, HHS, ASPR BARDA	**Epidemic Support**: Supported Weekly Interagency Conference Calls for Ebola	Sources from interview data
National	Ira Longini	University of Florida	Brandon Deen, Los Angeles County Dept of Public Health	**Epidemic Support and Research Collaboration:** Helped Los Angeles County officials develop a measured and appropriate response to the unfolding pandemic and establish reasonable goals for mitigation of pandemic H1N1.	Planning for the Control of Pandemic Influenza A (H1N1) in Los Angeles County and the United States
National	Ira Longini	University of Florida	Juliet R. C. Pulliam, Fogarty International Center, National Institutes of Health	**Research Collaboration:** School-located influenza vaccination programs in public schools showed decreased community risk	School-Located Influenza Vaccination Reduces Community Risk for Influenza and Influenza-Like Illness Emergency Care Visits
National	Ira Longini	University of Florida	BARDA	**Research Collaboration**: Modeled targeted layered containment of influenza pandemic based on scenarios provided by public health officers	Modeling targeted layered containment of an influenza pandemic in the United States

National	Jim Koopman	University of Michigan	CDC, World Health Organization, Gates Foundation	**Networked interactions/consultations:** Workshop on Analyzing the Polio Eradication Endgame. Raised awareness and serious issues regarding global polio eradication	http://psc.edu/science/2010/h1n1/
National	Pejman Rohani	University of Michigan	NIH, Thailand, France	**Research Collaboration**	Deciphering the impacts of vaccination and immunity on pertussis epidemiology in Thailand
National	Mark Lipsitch	Harvard University	Milwaukee Health Dept., NY, CDC	**Research Collaboration:** Modeled the use of Tamiflu to reduce transmission	Oseltamivir for treatment and prevention of pandemic influenza A/H1N1 virus infection in households, Milwaukee, 2009
National	Caroline Buckee	Harvard University		**Research Collaboration:** National Malarial Control and elimination programs	Human movement data for malaria control and elimination strategic planning
National	William Hanage	Harvard University	Dept of Vaccines, National Public Health Institute (KTL), Helsinki, Finland	**Research Collaboration**	Ability of Pneumococcal Serotypes and Clones To Cause Acute Otitis Media: Implications for the Prevention of Otitis Media by Conjugate Vaccines
National	Jeff Shaman	Columbia University	Lynn Finelli, CDC	**Public Health Tools:** For the CDC influenza forecasting challenge -- models were created and improved to help CDC predict and prepare for Influenza	Sourced from interview data

National	Steven Eubank	Virginia Tech	Martin Cetron, M.D., Director, Global Migration and Quarantine, CDC	**Networked interactions/consultations:** Presentation to West Virginia Dept of Health and Human Resources	Community-Wide Mitigation Strategies in Pandemic Planning
Local	Steven Eubank	Virginia Tech	District Epidemiologists	**Public Health Tools:** Worked with the district epidemiologist to identify modeling resources, created a social network scanning tool for foodborne illness and the pertusis outbreak.	http://medicalxpress.com/news/2014-08-social-media-surveillance-tool-vital.html
National	Neil Ferguson	Imperial College	David Swerdlow, CDC and NCIRD	**Epidemic Support:** Working with data from NNDSS, modeled incidence of Pertussis	A Change in Vaccine Efficacy and Duration of Protection Explains Recent Rises in Pertussis Incidence in the United States
National	Neil Ferguson	Imperial College	Thomas W Scott, Fogarty International Center, NIH	**Research Collaboration**: Dengue Control - A project between multiple nations to advise their health boards	Assessing the epidemiological effect of wolbachia for dengue control
National	William Hanage	Harvard University	Cheryl Tar, CDC	**Research Collaboration:** Cholera, *E. coli* projects	Evolutionary Dynamics of Vibrio cholerae O1 following a Single-Source Introduction to Haiti
National	Sara Del Valle	Los Alamos National Laboratory	Matt Biggerstaff, CDC	**Public health tools:** Modeling various scenarios using CDC provided data on community mitigations	http://www.lanl.gov/discover/news-release-archive/2015/December/12.15-flu-season.php
State	Travis Porco	University of California at San Francisco	State of California	**Research collaboration:** Developed a Measles Contact investigation model	The role of vaccination coverage, individual behaviors, and the public health response in

					the control of measles epidemics: an agent-based simulation for California.
Local	Margaret Potter	University of Pittsburgh	Ronald Vorhees, Alleghany Dept. of Health and PA Dept. of Health	**Research collaboration:** Used modeling to decide whether to close schools during Supported Weekly Interagency Conference Calls for Influenza outbreak	Would School Closure for the 2009 H1N1 influenza epidemic have been worth the cost?: a computational simulation of Pennsylvania
Local	Parker Small	University of Florida		**Research Collaboration:** Collaborative studies on local spread of Influenza	Unadjusted and bias-corrected effectiveness estimates of the Alachua County school-located influenza vaccination program for 2011/12 epidemic period.

Appendix G.
MIDAS Infrastructure Data

Table E-1. MIDAS Resources Identified in Literature Review and Interviews, by Institution

Institution	Resource	Developer	Publicly Available?
Argonne National Laboratory	Repast Simphony	Michael North	Y
Columbia University	Columbia Prediction of Infectious Diseases	Jeff Shaman	Y
Imperial College London	Malaria Tools	Jamie Griffin	Y
Johns Hopkins University	LSAM	Jon Parker	N
Los Alamos National Laboratory	PYDA	Kyle Hickman	Y
	QUAC	Reid Priedhorsky	Y
MIT (U54GM088558)	VisuaLyzer	David Reshef	N
Northeastern University	GLEAMviz	Alessandro Vespignani	Y
RTI International	ABM++	Doug Roberts	Y
University of Chicago/Argonne National Lab	TBD	Diane Lauderdale/Charles Macal	N
University of Chicago/Argonne National Laboratory	MRSA ABM Model	Diane Lauderdale/Charles Macal & Michael North	Y
University of Hong Kong	Steven Riley's Spatial Simulator	Steven Riley	N
University of Pennsylvania	PEK	Sky Pelletier	N
	SMK	Sky Pelletier	N
University of Pittsburgh	GAIA	Shawn T. Brown	N
	VELMA	Andrew Stanley Walsh	Y
	Project Tycho	Wilbert van Panhuis	Y
	HERMES	Shawn T. Brown	N
	ISAAC	Russ Schuh	N
	LENA	Margaret Potter	Y

	FRED	Donald Burke	Y
	Global Epidemic Model	Joshua Epstein	Y
	CLARA	Unknown	N
	Agent Zero	Joshua Epstein	Y
	GSAM	Joshua Epstein	N
	Health Care Facility Network Models	Bruce Lee	N
	Pittsburgh/RTI Model	Shawn T. Brown	N
	SEEDY	Colin Worby	Y
	IRED	Jon Parker	Y
	The Ebola Epidemic Chronology	Michael Wagner	Y
	Apollo Location Service	Michael Wagner	Y
University of Texas at Austin	EpiFire	Lauren Meyers	Y
	DiCon	Lauren Meyers	Y
	Texas Pandemic Flu Toolkit	Lauren Meyers	Y
University of Washington	FluTE	Ira Longini	Y
	TranStat	Elizabeth Halloran	Y
Virginia Tech	EpiFast	Keith Bisset	N
	Didactic	Stephen Eubank	N
	EpiSimdemics	Keith Bisset	N
	Indemics	Stephen Eubank	N