DIMACS Center
Rutgers University

Special Focus on Computational and Mathematical Epidemiology

Annual Report

June 2007

Participants who spent 160 hours or more:

Fred S. Roberts, Rutgers University, Principal Investigator; involved in all the working groups and all aspects of the project.

Martin Farach-Colton, Rutgers University, Member of Organizing Committee; has been playing a role in the connections of the epidemiology work to issues of bioinformatics/genetics.

David Madigan, Rutgers University, Member of Organizing Committee; Co-Organizer of Working Group on Adverse Event/Disease Reporting, Surveillance and Analysis; Co-Organizer of Working Group on BioSurveillance Data Monitoring and Information Exchange.

S. Muthukrishnan, Rutgers University, Member of Organizing Committee

Donald Hoover, Rutgers University, Member of Organizing Committee; Co-Organizer of Working Group on Adverse Event/Disease Reporting, Surveillance and Analysis.

Nina Fefferman, DIMACS, engaging in the research, running our "episeminar", mentoring students and postdocs, organizing workshops

James Abello, DIMACS, engaging in the research, running our "episeminar," mentoring graduate and undergraduate students, developing software, organizing workshops.

Other Participants:

Sunetra Gupta, Oxford University, Member of Organizing Committee

David Krakauer, Santa Fe Institute, Member of Organizing Committee

Simon Levin, Princeton University, Member of Organizing Committee; Advisor to PI on all parts of the project.

Marc Lipsitch, Harvard School of Public Health; Member of Organizing Committee

Ilya Muchnik, developing a cancer registry modeling project with James Abello and graduate student David Millman.

Megan Murray, Harvard School of Public Health; Member of Organizing Committee

David Ozonoff, Boston University; Member of Organizing Committee

Burton Singer, Princeton University, Member of Organizing Committee

Daniel Wartenberg, University of Medicine and Dentistry of New Jersey, Member of Organizing Committee

Workshop: Facing the Challenge of Infectious Diseases in Africa: The Role of Mathematical Modeling (funded by a different grant, but clearly influenced by and synergistic with this one) September 25 - 27, 2006

Organizers:
Dominic Clemence, North Carolina AT&T State University
Wayne Getz, University of California at Berkeley
Abba Gumel, University of Manitoba
John Hargrove, South African Centre for Epidemiological Modeling and Analysis (SACEMA)
Edward Lungu, University of Botswana
Fred Roberts, DIMACS, Rutgers University

Workshop: Models of Co-Evolution of Hosts and Pathogens
October 9 - 11, 2006

Organizers:
Viggo Andreasen, Roskilde
Andrea Pugliese, Trento

Workshop: Immuno-epidemiology
December 11 - 13, 2006

Organizers:
Hans Heesterbeek, Universiteit Utrecht
Rob de Boer, Universiteit Utrecht

Working Group Meeting: Spatio-Temporal and Network Modeling of Diseases (Second Meeting)
May 14 - 18, 2007

Organizers:
Valerie Isham, University College London
Alun Lloyd, North Carolina AT&T State University
Denis Mollison, Heriot-Watt University

DIMACS/DyDAn Research Project on Potential Uses of Entropy in Biosurveillance
(funded by the Department of Homeland Security, but collaborating with this project)
May 29 - August 20, 2007

Organizer:
Nina Fefferman, DIMACS

Participants:
Howard University Team:
Abdul-Aziz Yakubu
Devroy McFarlane (graduate student researcher)
Ashley Crump (undergraduate student researcher)

Morgan State University Team:
Asamoah Nkwanta
Anthony Ogbuka (graduate student researcher)
Nakeya Williams (undergraduate student researcher)

US-Africa Advanced Study Institute on Mathematical Modeling of Infectious Diseases in Africa
(funded by a different grant, but clearly influenced by and synergistic with this one)
June 11 - 22, 2007

Organizers:
Brenda Latka, (Program Chair), DIMACS
Wayne Getz, UC Berkeley
Abba Gumel, University of Manitoba
Fritz Hahne, AIMS
John Hargrove, SACEMA
Simon Levin, Princeton University
Edward Lungu, University of Botswana
Fred Roberts, DIMACS
Alex Welte, University of the Witwatersrand

Lecturers:
Jonathan Dushoff, Princeton University
Wayne Getz, University of California, Berkeley
Abba Gumel, University of Manitoba
Suzanne Lenhart, University of Tennessee
James Lloyd-Smith, Pennsylvania State University
Edward Lungu, University of Botswana
Martin Meltzer, CDC

Special Lecturers:
Martin Meltzer, CDC
Fred Roberts, DIMACS

Workshop: Mathematical Modeling of Infectious Diseases in Africa
(funded by a different grant, but clearly influenced by and synergistic with this one)
June 25 - 27, 2007

Organizers:
Brenda Latka, (Program Chair), DIMACS
Wayne Getz, UC Berkeley
Abba Gumel, University of Manitoba
Fritz Hahne, AIMS
John Hargrove, SACEMA
Simon Levin, Princeton University
Edward Lungu, University of Botswana
Fred Roberts, DIMACS
Alex Welte, University of Witwatersrand

Workshop: Systems Biology of Infectious Diseases
August 13 - September 2, 2007

Organizers:
Charles DeLisi, Boston University
Simon Levin, Princeton University

Visitors:

Roman Dementiev, Institute for Theoretical Computer Science, University of Karlsruhe, 3/5-3/22/07
Benoit Gaudin, Computer Science and Informatics Center, University College Ireland, 2/18-3/4/07
Vladimir Gurvich, Russian Academy of Science, 7/1-8/31/06
Jonathan Read, Mathematics Institute, University of Warwick, 3/22-3/30/07
Hans-Joerg Schulz, Computer Science, University of Rostock, 2/18-3/21/07
Christian Tominski, Computer Science, University of Rostock, 3/19-3/29/06

Graduate students:

Qinhe Cheng, Statistics, Rutgers University, winter 06/07
Srinivas Maloor, Statistics, Rutgers, 9/1/05-8/31/06

Other Collaborators

Graham Cormode, Bell Laboratories and AT&T Laboratories, Co-Organizer of Tutorial on Data Mining and Epidemiology

Larry Cox, CDC, Co-Organizer, Working Group: Data De-Identification, Combinatorial Optimization, Graph Theory, and the Stat/OR Interface

Troy Day, Queens University, Co-Organizer, Workshop: Evolutionary Considerations in Vaccine Use

Ding-zhu Du, University of Texas at Dallas, Co-Organizer of Workshop: Combinatorial Group Testing

Alison Galvani, Yale University., Co-Organizer, Workshop: Evolutionary Considerations in Vaccine Use

Abba Gumel, University. of Manitoba, Co-Organizer, Workshop: Evolutionary Considerations in Vaccine Use

Frank Hwang, Chiatong University, Co-Organizer of Workshop: Combinatorial Group Testing

Ramanan Laxminarayan, Resources for the Future, Co-Organizer, Workshop: Economic Epidemiology

Randy Linder, University of Texas, Co-Organizer, Workshop and Working Group: Reticulated Evolution

Catherine Macken, Los Alamos National Labs, Co-Organizer, Workshop: The Epidemiology and Evolution of Influenza

Colleen Martin, CDC, Co-Organizer of Working Group on BioSurveillance Data Monitoring and Information Exchange

Bernard Moret, University of New Mexico, Co-Organizer, Workshop and Working Group: Reticulated Evolution

Alan Perelson, Los Alamos National Labs, Co-Organizer, Workshop: The Epidemiology and Evolution

of Influenza

Allen Rodrigo, University Of Auckland, Co-Organizer, Workshop: Phylogenetic Trees and Rapidly Evolving Pathogens, Co-Organizer of Working Group Phylogenetic Trees and Rapidly Evolving Pathogens

Henry Rolka, CDC, Co-Organizer of Working Group on Adverse Event/Disease Reporting, Surveillance and Analysis; Co-Organizer of Working Group on BioSurveillance Data Monitoring and Information Exchange

Dave Smith, National Institute of Health (NIH), Co-Organizer, Workshop: Economic Epidemiology

Mike Steel, University of Canterbury, Co-Organizer, Workshop: Phylogenetic Trees and Rapidly Evolving Diseases, Co-Organizer, Working Group Meeting: Phylogenetic Trees and Rapidly Evolving Diseases

Katherine St. John, The City University of New York, Lehman College, Organizer of Tutorial: Phylogenetic Trees and Rapidly Evolving Pathogens

Claudio Struchiner, Oswaldo Cruz Foundation, Co-Organizer, Workshop: Evolutionary Considerations in Vaccine Use

Partner Organizations:

Princeton University: Collaborative Research; Personnel Exchanges. Partner organization of DIMACS. Individuals from the organization participated in the program planning, organization, and the research.

AT&T Labs - Research: Collaborative Research, Partner organization of DIMACS. Individuals from the organization participated in the program planning.

Telcordia Technologies: Collaborative Research; Personnel Exchanges. Partner organization of DIMACS. Individuals from the organization participated in the program planning.

NEC Laboratories America: Collaborative Research; Personnel Exchanges. Partner organization of DIMACS. Individuals from the organization participated in the program planning.

Lucent Technologies, Bell Labs: Collaborative Research; Personnel Exchanges. Partner organization of DIMACS. Individuals from the organization participated in the program planning, organization and research.

Office of Naval Research: Financial Support

Alfred P. Sloan Foundation: Financial Support

Burroughs-Wellcome Fund: Financial Support

New Jersey Commission on Science and Technology: Financial Support

IBM Research: Collaborative Research; Personnel Exchanges. Partner organization of DIMACS. Individuals from the organization participated in the program planning.

Microsoft Research: Collaborative Research. Partner organization of DIMACS. Individuals from the organization participated in the program planning.

Avaya Labs: Collaborative Research. Partner organization of DIMACS. Individuals from the organization participated in the program planning.

American Statistical Association: Financial Support

Hewlett-Packard Labs: Collaborative Research. Partner organization of DIMACS. Individuals from the organization participated in the program planning.

Centers for Disease Control: Collaborative Research; Personnel Exchanges. Individuals from the organization participated in the program planning and working group/workshop organization.

Activities

This is a five-year special focus, which started in summer of 2002, following the design of our Center's pioneering special focus on Mathematical Support for Molecular Biology. In this special focus, the Center seeks to:

- Develop and strengthen collaborations and partnerships between mathematical scientists (mathematicians, computer scientists, operations researchers, statisticians) and biological scientists (biologists, epidemiologists, clinicians).

- Identify and explore issues in mathematics and computer science that need to be resolved to make progress on important problems in epidemiology.

- Identify and explore methods of mathematical science not yet widely used in studying problems of epidemiology and introduce epidemiologists to them - with an emphasis on methods of discrete mathematics (including discrete probability) and the algorithms, models, and concepts developed in the field of theoretical computer science.

- Introduce outstanding young people from both the mathematical/computer science and biological communities to the issues and problems and challenges of computational and mathematical epidemiology.

- Involve biological and mathematical scientists together to define the agenda and develop the tools of computational and mathematical epidemiology.

The special focus consists of a research program featuring “working groups” concentrating on specific research topics and a program integrating research and education through a series of workshops and tutorials. The 1994-2000 DIMACS Special Focus on Mathematical Support for Molecular Biology played a central role in laying the groundwork for the field of computational molecular biology, led many fledgling concepts and methods grounded in the mathematical sciences to become standard tools in the biological sciences, produced lasting partnerships between biological and mathematical scientists, and introduced many of today's leaders in computational biology to the field and to each other. This special focus has been doing the same.

Epidemic models of infectious diseases go back to Daniel Bernoulli's mathematical analysis of smallpox in 1760 and have been developed extensively since the early 1900s. Hundreds of mathematical models

have been published since, exploring the effects of bacterial, parasitic, and viral pathogens on human populations. The results have highlighted and formalized such concepts as the notion of a core population in sexually transmitted diseases and made explicit other concepts such as herd immunity for vaccination policies. Relating to persistent infections, key pathogens that have been studied are: Malaria, *Neisseria gonorrheae*, *M. tuberculosis*, *HIV*, and *T. palladum*. Important issues such as drug-resistance, rate of spread of infection, epidemic trends, and the effects of treatment and vaccination all have been addressed through mathematical modeling approaches, which with the help of computational tools have provided new insights. Yet, for many infectious diseases, we are far from understanding the mechanisms of disease dynamics. The strength of the modeling process is that it can lend insight and clarification to existing data and theories. Mathematical models provide a unique approach to representing and studying the integrated behavior of complex biological systems and enable us to compare and contrast existing theories of the dynamic interactions in a complex system. The size of modern epidemiological problems and the large data sets that arise call out for the use of powerful computational methods for studying these large models. As pointed out by Levin, Grenfell, Hastings, and Perelson in a 1997 article in *Science*, "imaginative and efficient computational approaches are essential in dealing with the overwhelming complexity of [such] biological systems." New computational methods are needed to deal with the dynamics of multiple interacting strains of viruses though the construction and simulation of dynamic models, the problems of spatial spread of disease through pattern analysis and simulation, and the optimization of drug design through hierarchical and other search methods on adaptive landscapes.

Statistical methods have long been used in mainstream epidemiology largely for the purpose of evaluating the role of chance and confounding associations. Considerable effort is expended by epidemiologists to ferret out sources of systematic error ("bias and confounding") in the observations and to evaluate the role of uncontrollable error (using statistical methods) in producing the results. Interpretation of the results usually depends upon correlative information from the medical and biological sciences. The role of statistical methods in epidemiology is changing due to the large data sets that are arising and this calls for new methods and new approaches, making use of modern information technology for dealing with huge data sets of information on disease patterns.

A smaller but venerable tradition within epidemiology has considered the spread of infectious disease as a dynamical system and applied difference equations and differential equations to that end. But little systematic effort has been made to apply today's powerful computational methods to these dynamical systems models and few computer scientists have been involved in the process. We hope to change this situation. Probabilistic methods, in particular stochastic processes, have also played an important role. However, here again, few computer scientists have been involved in efforts to bring the power of modern computational methods to bear.

A variety of other potentially useful approaches to epidemiological issues have not yet attracted the attention of many in the computer science community nor are the methods made widely available to biological scientists. For example, many fields of science, and in particular molecular biology, have made extensive use of the methods and techniques of discrete mathematics (broadly defined), especially those that exploit the power of modern computational tools. These are guided by the algorithmic and modeling methods of theoretical computer science that make these tools more available than they have been in the past. Yet, these methods remain largely unused in epidemiology. One major development in epidemiology that makes the tools of discrete mathematics and theoretical computer science especially relevant is the use of Geographic Information Systems (GIS). These systems allow analytic approaches to spatial information not used previously. Another development is the availability of large and disparate computerized databases on subjects containing information on many attributes that might be related to disease status.

The role of discrete mathematics and theoretical computer science has also become important with the

increasing emphasis in epidemiology of an evolutionary point of view. To fully understand issues such as immune responses of hosts; co-evolution of hosts, parasites, and vectors; drug response; and antibiotic resistance; among others, biologists are increasingly taking approaches that model the impact of mutation, selection, population structure, selective breeding, and genetic drift on the evolution of infectious organisms and their various hosts. Epidemiologists are only beginning to become aware of some of the computer science tools available to analyze these complex problems, such as methods of classification and phylogenetic tree reconstruction grounded in concepts and algorithms of discrete mathematics and theoretical computer science and developed in connection with the explosion in "computational biology," a field in which DIMACS has been a pioneer. Many of the recent methods of phylogenetic tree reconstruction resulted from the DIMACS Special Focus on Mathematical Support for Molecular Biology are described in the DIMACS website in the reports on the accomplishments of the earlier Special Focus. Yet, a great deal more needs to be done.

One important modern topic in theoretical computer science that arose in epidemiology is the theory of group testing, which arose in connection with testing millions of World War II military draftees for syphilis. The idea is to avoid testing each individual and instead to divide them into groups and determine if some individual in the group is positive for the disease, updating the process with groups that test positive. The modern theory of group testing is heavily influenced by combinatorial methods, in particular by the methods of combinatorial designs and coding theory, and many modern algorithmic methods, developed by theoretical computer scientists, are not yet widely known or used in epidemiology.

Mathematical methods of formal logic and ordered algebraic systems have been used to develop the foundations for a theory of measurement with important uses in the physical sciences and, more recently, in the social and biological. While this kind of measurement theory has been applied to data analysis in the social and biological sciences, it is virtually unknown in the epidemiology community (where the term "measurement theory" has other connotations), except to the extent that epidemiological studies use principles, grounded in but sometimes challenged by measurement theory, such as that certain kinds of statistical tests are inappropriate for ordinal data.

New interdisciplinary approaches, involving partnerships among mathematical scientists and biological scientists, epidemiologists, and clinicians, offer the promise for making progress on modern epidemiological problems and should take both fields of epidemiology and mathematics/computer science in new and fruitful directions. Mathematical and computational methods seem especially relevant in light of recent modeling approaches to emerging infectious diseases such as the vector-borne diseases from West Nile virus, Eastern equine encephalitis virus and Borrelia burgdorfei (Lyme disease); the spread of "mad cow" disease (transmissible spongiform encephalopathy; and HIV/AIDS. Control measures for these diseases often have counter-intuitive consequences only revealed after sophisticated mathematical analysis. Similar advances as a result of applications of mathematical and computational modeling have not been as evident in the area of chronic disease epidemiology, although work of considerable promise is being done, for example on modeling of the progression of cancer. In this special focus, we are considering both infectious and non-infectious diseases, and we will explore mathematical and computational approaches to animal and plant diseases as well as to human diseases.

The tutorials, workshops, and working group meetings that were held during this year are as follows:

Workshop: Facing the Challenge of Infectious Diseases in Africa: The Role of Mathematical Modeling
(funded by a different grant, but clearly influenced by and synergistic with this one)

Dates: September 25 - 27, 2006
Location: University of the Witswatersrand (Wits), Johannesburg, South Africa
Organizers: Dominic Clemence, North Carolina AT&T State University; Wayne Getz, UC

Berkeley; Abba Gumel, University of Manitoba; John Hargrove, SACEMA Director; Edward Lungu, University of Botswana; Fred Roberts, DIMACS
Attendance: 58

DIMACS in collaboration with SACEMA (the South African Centre of Excellence for Epidemiological Modelling and Analysis) held a 3-day workshop on mathematical modeling and infectious diseases in Africa. The workshop held at the School of Computational and Applied Mathematics at the University of the Witswatersrand (Wits), Johannesburg, South Africa, on September 26-28, 2006, brought together scientists and junior researchers from the US and African countries. Endemic and emerging diseases in Africa provide new and complex challenges for mathematical modeling. The workshop provided an agenda for future collaborations between US and African scientists. It exposed junior US scientists and students to the special challenges of modeling the spread of disease in Africa and the opportunities to collaborate with Africans in developing and applying the tools of mathematical modeling to the tremendous health problems caused by such diseases as HIV/AIDS, tuberculosis and malaria, as well as the possible interventions in the case of major new health threats such as pandemic influenza in a developing region of the world.

Workshop: Models of Co-Evolution of Hosts and Pathogens
Dates: October 9 - 11, 2006
Location: DIMACS Center, CoRE Building, Rutgers University
Organizers: Viggo Andreasen, Roskilde; Andrea Pugliese, Trento
Attendance: 32

It has long been recognized that hosts and pathogens exert strong selective forces on each other. Thus significant coevolution between host and pathogens is to be expected, and with the short generation time of many pathogens, evolution may occur over observable time scales. In fact coevolution has been demonstrated in many host-pathogen systems. For example in the classic gene-for-gene systems, each new resistance gene that is introduced into a (cereal) crop is matched within a few seasons by a virulence gene allowing the (fungal) pathogen to overcome the resistance. The epidemiology of several human diseases can be understood only in an evolutionary context. For long periods Influenza A persistence relies on so-called drift mutations that changes viral antigen sufficiently to allow for reinfection of the same hosts while the evolutionary changes in HIV are so fast that they are an integral part of the infection process within the individual host. More recently it has been proposed that the strain structure in malaria and RSV, among others, should be understood in an evolutionary framework. This workshop focused on evolutionary and coevolutionary processes at the population level while selection processes within the individual host were discussed in other workshops. The first models of host-pathogen coevolution were applications of very general descriptions of coevolution. However, with the increased interest in disease transmission dynamics the focus has now moved to descriptions that explicitly utilize epidemic models to describe the frequency dependent nature of the interaction. The mathematical methods for describing multiple interacting types of the pathogen or the interaction between disease and host genetics are in the process of being developed, but have not yet reached maturity. The workshop brought together mathematical researchers and quantitatively oriented biologists and epidemiologists in the field to discuss the development of mathematical methods, as well as, to explore evolutionary and coevolutionary aspects of a number of host pathogen systems (malaria, influenza, insect-bacculovirus, RSV).

Workshop: Immuno-epidemiology
Dates: December 11 - 13, 2006
Location: DIMACS Center, CoRE Building, Rutgers University
Organizers: Hans Heesterbeek, Universiteit Utrecht, The Netherlands; Rob de Boer, Universiteit Utrecht, The Netherlands
Attendance: 37

Individual hosts differ considerably in the way in which they respond to the same pathogen. This is not only caused by genetic polymorphism determining immune reaction (e.g. MHC), but also by the infection history of the individual (e.g. influenza, parasites with acquired immunity, dengue). An individual's history is a result of the past pattern of transmission in the population. Population transmission (infection pressure) by itself is the collective output of infectious material by the individuals that constitute the population, which in turn is decided by each individual's reaction to the pathogen. This closes a circle of mutual interaction and influence. This cycle influences the population effects of control measures aimed at individuals, and the evolution of resistance and virulence. In order to understand these processes we need a fuller understanding of the immunity-transmission cycle.

Presently immunological theory and epidemiological theory restrict themselves to one part of the cycle, both making “black box,” or rudimentary and idealized, assumptions about the other half (when the influence of this half is considered at all). Almost never is there full feedback between the within- and between-host processes. For the understanding of the evolution of resistance and virulence, however, one has to close the loop. Variants of the infectious agent arise within individuals, but will only be relevant at the population level if they also spread between individuals. One currently lacks the tools to make even qualitative predictions, for example for the population consequences of vaccination when there is great polymorphism in an individual's immune reaction (through genetic polymorphism and/or due to previous memory/exposure). Obvious examples are malaria and dengue.

This 2-day workshop brought together experts on several parts of the immunity-transmission cycle where we want to take, in a structured way, a necessary step towards integration. The central theme was to develop sensible and simple within-host immunological models that can be merged in sensible and simple between-host epidemiological models, and to explore the full cycle and its effects on pathogen evolution, spread and control. Issues addressed included cellular versus humoral immunity, short-lived versus long-lived memory, acquired immunity, polymorphism, evolution of resistance and virulence.

Working Group Meeting: Spatio-Temporal and Network Modeling of Diseases (Second Meeting)
Dates: May 14 - 18, 2007
Location: The International Centre for Mathematical Sciences,(ICMS), Edinburgh, UK
Organizers: Valerie Isham, University College London; Alun Lloyd, North Carolina AT&T State University; Denis Mollison, Heriot-Watt University
Attendance: 43

The focus of this workshop was the transmission dynamics of the spatial spread of infections, primarily on structured populations and networks. The aim was to get together a small group (of around 30) experts with a wide range of approaches to spatio-temporal and network modeling, coming from both biological and mathematical areas, including a) those who can contribute directly to the topics chosen for the sessions, and b) some individuals whose own work may be led directly in these areas, but who will be good value for their general expertise, open-mindedness and readiness to contribute actively to discussions.

Topics for discussion included models for the spread of infection on structured populations and networks, the temporal evolution of networks, modeling of evolution on different spatial and temporal scales, “piecing together” models on different scales, implications for control, model fitting and statistical inference.

The workshop consisted of a mixture of overview papers establishing the “state of the art,” together with short contributed papers and workshop/discussion sessions, where the emphasis should not be on accounts of completed research but on work in progress, unsolved problems and opportunities for

collaboration.

DIMACS/DyDAn Research Project on Potential Uses of Entropy in Biosurveillance
(funded by the Department of Homeland Security, but collaborating with this project)
Dates: May 29 - August 20, 2007
Location: DIMACS Center, CoRE Building, Rutgers University
Organizer: Nina Fefferman, DIMACS
Attendance: 7

Failure of disease surveillance systems to detect and alert authorities early into the onset of a disease outbreak can lead to avoidable increases in the incidence in and mortality from a disease threat, whether that disease results from environmental contamination, natural exposure from zoonotic infections, or purposeful acts of bioterrorism. Many different mathematical and statistical techniques have therefore been proposed to analyze incoming disease incidence reporting and to try and detect a differentiable signal from the unavoidable noise as early as possible after the onset of any event. The possibilities however, are far from exhausted and we have yet to develop a demonstrably reliable early detection method.

The information theoretic measure of entropy in a system fundamentally quantifies the amount of information conveyed within a particular signal. There are many ways of characterizing this measure. Organizer Nina Fefferman has hypothesized that the signal to noise ratio in reported disease incidence data may increase during the occurrence of disease outbreaks and that these increases may be observable when analyzed using entropy measures. If the hypothesized increases are statistically quantifiable, and if they are found to occur early enough in the progression of disease throughout the affected population, this may prove an invaluable tool for early-warning biosurveillance.

This project involves teams of faculty and students from two minority-serving institutions, Howard University and Morgan State University, working at DIMACS for the summer. They are looking at a variety of approaches to address Dr. Fefferman's hypothesis.

US-Africa Advanced Study Institute on Mathematical Modeling of Infectious Diseases in Africa
(funded by a different grant, but clearly influenced by and synergistic with this one)
Dates: June 11 - 22, 2007
Location: AIMS, Cape Town, South Africa
Organizers: Brenda Latka, (Program Chair), DIMACS; Wayne Getz, UC Berkeley; Abba Gumel, University of Manitoba; Fritz Hahne, AIMS; John Hargrove, SACEMA; Simon Levin, Princeton University; Edward Lungu, University of Botswana; Fred Roberts, DIMACS; Alex Welte, Wits University
Attendance: 36

Mathematical modeling of the spread of infectious disease has a long history going back to Bernoulli's modeling of smallpox in 1760. In recent years, models have been vitally important in the development of approaches to such critical diseases as HIV/AIDS, which is of such importance to Africa. Modelers in collaboration with public health officials also played an important role during the 2003 SARS outbreaks and are already working to determine ways to contain the spread of a pending influenza pandemic. The DIMACS/SACEMA/AIMS Advanced Study Institute provided a select group of students the opportunity for exposure to a field where there is a critical shortage of people with the necessary high-level skills and which has many exciting opportunities for research and practical application.

The institute, housed at AIMS, the African Institute for Mathematical Sciences, held a series of lectures and tutorials on the design and analysis of models for the spread of emerging and re-emerging diseases.

The first week provided a basic introduction to mathematical modeling in epidemiology at a fast pace. The introductory week was designed to allow students who have never taken an epidemiological modeling course to acquire the necessary preparatory background they need for the second week. The second week covered more advanced material. Various modeling paradigms were discussed, as well as introductory lectures on related topics. There were a number of hands-on and computer exercises together with group projects to reinforce and extend the various concepts covered.

The main instructors were:

- Jonathan Dushoff, Princeton University
- Wayne Getz, University of California Berkeley
- Abba Gumel, University of Manitoba
- Suzanne Lenhart, Univerisy of Tennessee
- James Lloyd-Smith, Pennsylvania State University
- Edward Lungu, University of Botswana

Special lecturers were:

- Martin Meltzer, Center for Disease Control
- Fred Roberts, DIMACS, Rutgers University

Workshop: Mathematical Modeling of Infectious Diseases in Africa
(funded by a different grant, but clearly influenced by and synergistic with this one)

Dates: June 25 - 27, 2007
Location: Stellenbosch, South Africa
Organizers: Brenda Latka, (Program Chair), DIMACS; Wayne Getz, UC Berkeley; Abba Gumel, University of Manitoba; Fritz Hahne, AIMS; John Hargrove, SACEMA; Simon Levin, Princeton University; Edward Lungu, University of Botswana; Fred Roberts, DIMACS; Alex Welte, Wits University
Attendance: 74

DIMACS, SACEMA, and AIMS held a 3-day workshop on mathematical modeling and infectious diseases in Africa. It served as a capstone to participants in the previous Advanced Study Institute held in June 11 - 22. Mathematical modeling has provided new insights on important issues such as drug-resistance, rate of spread of infection, epidemic trends, and effects of treatment and vaccination. Yet, for many infectious diseases, and in particular many diseases affecting Africa, we are far from understanding the mechanisms of disease dynamics. The modeling process can lend insight and clarification to data and theories. To get the maximum benefit out of mathematical models, however, one needs to specialize them, test assumptions in specific contexts and populations, gather local data to help define key parameters, etc. The workshop's goal was to develop collaborations and communications among US and African senior and junior researchers on these issues, to benefit both sides in their research and the important public health applications of that research.

Workshop: Systems Biology of Infectious Diseases

Dates: August 13 - September 2, 2007
Location: Aspen Center for Physics, Aspen, Colorado
Organizers: Charles DeLisi, Boston University; Simon Levin, Princeton University
Attendance: 24 (registration is still open for this workshop)

The mathematical sciences have long spurred fundamental and applied advances in the life sciences. In the modern era a number of major transformations are associated with specific individuals: the profound

physiological analyses of Helmholtz, and the Crick theory of X-ray scattering from double helices, are only two among many. More recently a major cultural shift toward mathematics and computation has been forced by the Human Genome Project which would not have been possible, absent the methods of mathematics and advanced computation.

The subject on which this workshop focuses, infectious disease, is among the areas in which applied and fundamental quantitative science has played a major role for decades. Contributions range from differential equations models of disease dynamics in human populations--which provides, among other things, the basis for policy--to the quantitative tools of the civil engineer, which have increased life span by decades. More recently, advances in cell biology have transformed our understanding of disease related processes, including molecular changes that occur when pathogens infect cells; physical processes that underlie infection; and systemic changes in the host. We are now poised to understand the biological, chemical and physical determinants of host-to host transmission and changes in host range, and to begin integrating such understanding with changes in demographics, climate, globalization and so on. A central goal of infectious disease research is, therefore, integration across scales. This workshop offers a unique opportunity for physicists to exchange ideas with colleagues in the biomedical sciences on a range of topics relevant to emerging and remerging infectious disease, which cut across scales. These include the physical properties of proteins and nucleic acids and their interactions; the mechanisms governing the behavior of molecular motors; the physics of transcriptional and translational control; learning, memory and adaptation in cells, organs and individuals; and the dynamics of host-host and host parasite interactions.

DIMACS Seminar in Quantitative Biology and Epidemiology

Special Focus visitors Nina Fefferman and James Abello organized the DIMACS Computational and Mathematical Epidemiology seminar during the 2006-2007 academic year. In total, there were almost twenty talks given. The audience came from a wide variety of disciplines, including Biology, Mathematics, Medicine, Physics, Chemistry, Computer Science and more, including regular participants from Rutgers, The University of Medicine and Dentistry of New Jersey (UMDNJ), Robert-Wood Johnson Medical School (RWJMS), Princeton and beyond. The program of this series, including titles and speakers, is given here. Abstract information can be found at http://dimacs.rutgers.edu/SpecialYears/2002_Epid/seminars06-07.html.

An Info-gap Approach to Bio- and Homeland Security
Thursday, September 7, 2006
Speaker: David Fox, University of Melbourne

Surveillance Methods for Drug Safety
Monday, September 11, 2006
Speaker: David Madigan, Rutgers University

A Novel Method for Characterizing and Classifying Dissipative Structures in the Transcriptomics of Budding Yeast: Potential Applications to Genomics, Proteomics, Metabonomics, and 'Cytomics'
Monday, September 25, 2006
Speaker: Sungchul Ji, Rutgers University

Determining Optimal Vaccination Strategies in Dynamic Social Networks
Monday, October 9, 2006
Speaker: Nina Fefferman, DIMACS and Tufts University

Higher Cognition Imaging with Functional MRI

Monday, October 23, 2006
Speaker: Larry Shepp, Rutgers University

Spatial Dynamics of Influenza Hospitalizations in the US Elderly
Monday, November 20, 2006
Speaker: Elena Naumova, Tufts University School of Medicine

Spreading Dynamics on Small-World Networks with a Power Law Degree Distribution
Monday, December 4, 2006
Speaker: Alexei Vazquez, Princeton University

A Tick-Borne Disease Model - Spread and Control of Ehrlichiosis
Monday, February 5, 2007
Speaker: Holly Gaff, University of Maryland

Data Mining for Drug Safety: Statistical Analyses of Spontaneous Reports and Clinical Safety Data
Monday, February 12, 2007
Speaker: William DuMouchel, Lincoln Technologies Division of Phase Forward, Inc.

Link Mining: Current State of the Art
Monday, February 19, 2007
Seminar joint with DyDAn
Speaker: Ronen Feldman, Bar-Ilan University

The Effects of Immune Selection on the Population Dynamics of Pathogens
Monday, March 5, 2007
Speaker: Sunetra Gupta, University of Oxford

Brief Encounters: Social Contacts and the Transmission of Disease
Monday, March 26, 2007
Speaker: Jonathan Read, University of Warwick

On Complexity of Algorithms for Modeling Disease Transmission and Vaccination Strategies
Monday, April 9, 2007
Seminar joint with DyDAn
Speaker: Vladimir Gurvich, Rutgers University

Crime's Ecosystem
Monday, May 21, 2007
Seminar joint with DyDAn
Speaker: Marcus Felson, Rutgers University

Findings

The role of individual choice in the evolution of social complexity

Constant re-evaluation of social affiliations and shifting social network structures can profoundly affect the adaptive fitness of individuals within a population, as well as yielding super-additive effects felt by the population as a whole. To evaluate the impact of different social affiliation choices, and the relative ability of individuals to correctly assess the success of other individuals, Nina Fefferman (DIMACS) and DIMACS postdoctoral visitor Kah Loon Ng (National University of Singapore) have created a set of

mathematical models based on network centrality measures. They choose the hypothetical measures of "popularity", "closeness" and "betweenness" to examine the resulting self-organizations of social groups. Their findings suggest that some different types of social behaviors can lead to the same levels of stability and organizational success, suggesting the possibility that complex organizations could have evolved from simpler ones without any change in the selective pressures acting on the population. A paper describing the results is cited below.

How disease models in static networks can fail to approximate disease in dynamic networks

In the modeling of infectious disease spread within explicit social contact networks, previous studies have predominantly assumed that the effects of shifting social associations within groups are small. These models have utilized static approximations of contact networks. Nina Fefferman (DIMACS) and postdoc Kah Loon Ng (National University of Singapore) examined this assumption by modeling disease spread within dynamic networks where associations shift according to individual preference based on three different measures of network centrality. The results of their investigations clearly show that this assumption may not hold in many cases. They demonstrated these differences in association dynamics do yield significantly different disease outcomes both from each other and also from models using graph-theoretically accurate static network approximations. Further work is therefore needed to explore under which circumstances static models accurately reflect constantly shifting natural populations. A paper describing the results is cited below.

Species-specific behavior affects disease spread throughout and ecosystem

In modeling the dynamics of inter-species disease spread, the focus has up until now been on the contact rates among viable hosts and on the relative physiological susceptibility and infectivity of the various host species. Recent studies of network-based epidemiology have shown that social behavior can greatly impact the within-population dynamics of a disease outbreak; however the majority of these studies have assumed static contact networks. In order to examine the potential influence of social organization in the contributing host species to an ecosystem's susceptibility to disease outbreaks, Nina Fefferman (DIMACS) and postdoc Kah Loon Ng (National University of Singapore) used a set of dynamic social network models with three different populations (representing different species, each with their own social behavior). Into this model they introduced an infectious disease, transmitted via social contact. They then compared both disease incidence in the three populations, and the relative numbers of inter-population transmissions to determine whether different species would experience different disease loads and perform different roles in the propagation of disease throughout the ecosystem, even if all individual physiological characteristics were uniform across species. They found that, not only did populations suffer drastically different disease loads from one another based only on their behavioral organization, but, their behavior also led the populations to act as either 'sources of' or 'buffers to' disease spread among the populations. They therefore concluded that the behavioral ecology of different co-existing species can play a crucial role in the emergence of zoonotic infections and that studies of wildlife disease should begin to incorporate an understanding of social behavior as a population-level etiological phenomenon. A paper describing the results is cited below.

Assessing seasonal variation in multisource surveillance data

A significant proportion of human diseases, spanning the gamut from viral respiratory disease to arthropod-borne macroparasitic infections of the blood, exhibit distinct and stable seasonal patterns of incidence. Traditional statistical methods for the evaluation of seasonal time-series data emphasize the removal of these seasonal variations to be able to examine non-periodic, and therefore unexpected, or 'excess', incidence. Here, Nina Fefferman (DIMACS), Elena Naumova (Tufts), Meena Doshi (Tufts), and Fefferman's student Eric Lofgren presented an alternate methodology emphasizing the retention and

quantification of exactly these seasonal fluctuations, explicitly examining the changes in severity and timing of the expected seasonal outbreaks over several years. Using a PCR confirmed Influenza time series as a case study, Fefferman, et al. provided an example of this type of analysis and discuss the potential uses of this method, including the comparison of differing sources of surveillance data. The requirements for statistical and practical validity, and considerations of data collection, reporting and analysis involved in the appropriate applications of the methods proposed were also investigated. A paper describing the results is cited below.

Name that cluster: text vs graphics

James Abello (DIMACS) and his visitors Benoit Gaudin (from University College, Ireland) and Hans-Joerg Schulz and Christian Tominski (from University of Karlsruhe) worked on clustering analysis problems of the type often arising from epidemiological data. Given a user query, search engines generally return a very sizeable collection of possible answers. Clustering has been proposed as a tool to partition the possible answer set into more manageable subsets of related results. There is no current agreement on the preferred mode of presentation of these clusters. Currently, most search engines display the set of results in an almost pure textual form. This study is a first step to elucidate when and why text appears to outperform graphics for certain fundamental clustering related tasks. To this end, Abello and his colleagues designed three interfaces to display flat clusters of user queries. The interfaces are enhanced with mechanisms by which users provide feedback about the relevance of a cluster for a pre-specified input query. Subsequently, users are asked to provide a name for a given cluster that best describes the cluster contents. They conducted a small number of informal user experiments and derived several hypotheses from their results. To validate the hypotheses, they extended their experiment set up to a web-based platform. This enabled them to reach out to a larger and broader set of participants, which allowed them to apply classical statistical analysis to this "vox populi." Such a statistical approach applied to results obtained from a large number of participants could make their results robust to "unusual behaviors" of just a few participants. A paper describing the results is cited below.

Discrete methods related to disease spread on networks of different characteristics

Studies of the spread and containment of diseases rely on a variety of mathematical and computational techniques. James Abello (DIMACS) and visitor Michael Capalbo (DIMACS) have been concentrating on the use of discrete methods to provide answers to questions related to disease spread on networks of different characteristics. These include sparse power law networks with "high" clustering coefficients. In a paper titled "Max cliques in sparse power law graphs with large clustering coefficients," cited below, they prove that the max clique problem is NP-hard even when restricted to sparse power law graphs with low diameter and "high" clustering coefficients. Motivated by a question formulated by Fred Roberts they introduced the notion of blocking sequences for seed set S of vertices in a connected graph G. Blocking sequences are a formalization of a sequence of sets that if vaccinated will contain the propagation of an infection that starts on the seed set. Since several variations of this problem are NP-hard, Abello and Capalbo developed an approximation algorithm for this problem. They also developed another application of blocking sequences. They showed that for d dimensional integer infinite grids, there exist seed sets S and functions $f(t)$ for which there are no blocking sequences whose sizes are upper bounded by f(t). A paper on blocking sequences is cited below.

Epidemic enhancement in partially immune populations

Jonathan Dushoff (Princeton) and his coauthors observed that a pathogen introduced into a population containing individuals with acquired immunity can result in an epidemic longer in duration and/or larger in size than if the pathogen were introduced into a naive population. They call this phenomenon "epidemic enhancement," and use simple dynamical models to show that it is a realistic scenario within

the parameter ranges of many common infectious diseases. This finding implies that repeated pathogen introduction or intermediate levels of vaccine coverage can lead to pathogen persistence in populations where extinction would otherwise be expected. A paper describing the results is cited below.

Synchronous cycles of domestic dog rabies in sub-Saharan Africa and the impact of control efforts.

Jonathan Dushoff (Princeton) and his coauthors studied synchronous cycles of domestic dog rabies in sub-Saharan Africa and the impact of control efforts. Rabies is a fatal neurological pathogen that is a persistent problem throughout the developing world where it is spread primarily by domestic dogs. Although the disease has been extensively studied in wildlife populations in Europe and North America, the dynamics of rabies in domestic dog populations has been almost entirely neglected. They demonstrated that rabies epidemics in southern and eastern Africa cycle with a period of 3-6 years and show significant synchrony across the region. The observed period is shorter than predictions based on epidemiological parameters for rabies in domestic dogs. They found evidence that rabies prevention measures, including vaccination, are affected by disease prevalence and show that a simple model with intervention responses can capture observed disease periodicity and host dynamics. They suggest that movement of infectious or latent animals combined with coordinated control responses may be important in coupling populations and generating synchrony at the continental scale. These findings have important implications for rabies prediction and control: large-scale synchrony and the importance of intervention responses suggest that control of canine rabies in Africa will require sustained efforts coordinated across political boundaries. A paper describing the results is cited below.

Network metrics reveal differences in social organization between two fission-fusion species, Grevy's zebra and onager

Jonathan Dushoff (Princeton) and his coauthors used network metrics to reveal differences in social organization between two fission-fusion species, Grevy's zebra and onager. For species in which group membership frequently changes, it has been a challenge to characterize variation in individual interactions and social structure. Quantifying this variation is necessary to test hypotheses about ecological determinants of social patterns and to make predictions about how group dynamics affect the development of cooperative relationships and transmission processes. Network models have recently become popular for analyzing individual contacts within a population context. They used network metrics to compare populations of Grevy's zebra (Equus grevyi) and onagers (Equus hemionus khur). These closely related equids, previously described as having the same social system, inhabit environments differing in the distribution of food, water, and predators. Grevy's zebra and onagers are one example of many sets of coarsely similar fission-fusion species and populations, observed elsewhere in other ungulates, primates, and cetaceans. Their analysis of the population association networks reveals contrasts consistent with their distinctive environments. Grevy's zebra individuals are more selective in their association choices. Grevy's zebra form stable cliques, while onager associations are more fluid. They find evidence that females associate assortatively by reproductive state in Grevy's zebra but not in onagers. The current approach demonstrates the utility of network metrics for identifying fine-grained variation among individuals and populations in association patterns. From their analysis, we can make testable predictions about behavioral mechanisms underlying social structure and its effects on transmission processes. A paper describing the results is cited below.

Vaccinating to protect a vulnerable subpopulation

Jonathan Dushoff (Princeton) and his coauthors studied vaccination strategies to protect vulnerable subpopulations. Epidemic influenza causes serious mortality and morbidity in temperate countries each winter. Research suggests that school children are critical in the spread of influenza virus, while the elderly and the very young are most vulnerable to the disease. Under these conditions, it is unclear how

best to focus prevention efforts in order to protect the population. They investigated the question of how to protect a population against a disease when one group is particularly effective at spreading disease and another group is more vulnerable to the effects of the disease. They developed a simple mathematical model of an epidemic that includes assortative mixing between groups of hosts. They evaluated the impact of different vaccine allocation strategies across a wide range of parameter values. With this model they demonstrated that the optimal vaccination strategy is extremely sensitive to the assortativity of population mixing, as well as to the reproductive number of the disease in each group. Small differences in parameter values can change the best vaccination strategy from one focused on the most vulnerable individuals to one focused on the most transmissive individuals. Given the limited amount of information about relevant parameters, they found that changes in vaccination strategy, while potentially promising, should be approached with caution. In particular, they found that, while switching vaccine to more active groups may protect vulnerable groups in many cases, switching too much vaccine, or switching vaccine under slightly different conditions, may lead to large increases in disease in the vulnerable group. This outcome is more likely when vaccine limitation is stringent, when mixing is highly structured, or when transmission levels are high. A paper describing the results is cited below.

Irreversible k-threshold processes: Graph-theoretical threshold models of the spread of disease and of opinion

Paul Dreyer (RAND Corporation) and Fred Roberts (DIMACS) studied models of the spread of disease or opinion through social networks, represented as graphs. In these simple models, vertices are in one of two states, 1 ("infected") or 0 ("uninfected") and change of state takes place at discrete times. They concentrated on the model, called an irreversible *k*-threshold process, where a vertex enters state 1 if at least *k* of its neighbors are in state 1, and where a vertex never leaves state 1 once it is in it. They looked for sets of vertices with the property that, if they are in state 1 at the beginning, then eventually all vertices end up in state 1. Such vertex sets, called irreversible *k*-conversion sets, correspond to vertices that can be infected with a disease or opinion so as to guarantee saturation of the population with the disease or opinion. They found bounds and exact values for the sizes of minimum irreversible k-conversion sets for grids, trees, and regular graphs, and showed that in general the problem of finding this minimum size is NP-complete. They studied ways to "defend" against such saturating sets, for example by "vaccination" or designing network topologies. They also related the problem to the widely-studied firefighting problem, which in the language of epidemiology, allows vaccination to take place every time period.

Malaria model on the impact of drug administration protocols in Mali

Malaria is one of the most life threatening tropical diseases for which no successful vaccine has been developed according to the 2006 UNICEF report on malaria. An effective drug for a person infected with malaria is sulfadoxinepyrimethane (SP). There have recently been trials designed to explore how effective SP is as a temporary vaccine. In 2002, Coulibaly *et al* conducted a study of two groups of people at a site in Bandiagara (Mali), a region of endemic seasonal malaria. The first group was administered SP at the beginning of the wet season, and the second group was given no drug. The two groups were monitored during the entire season for the first malaria episode. It was found that in the first 4 weeks many fewer first episode cases occurred in the first group, but this was slowly reversed between the fourth and the eighth week. During weeks 8-12, the number of first episodes in the first group became larger than in the second group. After 12 weeks, both groups have approximatively the same (but very small) number of episodes. Taken over the entire wet season (May-August in Mali), there was no significant advantage to the group which received SP. Abdul-Aziz Yakubu (DIMACS visitor from Howard University), Avner Friedman (Ohio State University), and Yakubu's student Bassidy Dembele aim to explore whether administering the same amount of SP at a different schedule will be more beneficial. They introduced a mathematical malaria model that allows drug administration under three protocols and two policies. They

describe their results in a paper cited below.

Malaria model with periodic mosquito birth rate

Abdul-Aziz Yakubu (DIMACS visitor from Howard University) and his coauthors introduced a novel malaria model. The new feature was the introduction of a periodic coefficient into the system of ordinary differential equations, which accounted for the seasonal variations (wet and dry seasons) in the mosquito birth rate. They defined a basic reproduction number R_0 which depends on the periodic coefficient and proved that if $R_0 < 1$ then the disease becomes extinct, whereas if $R_0 > 1$ then the disease is endemic and may even be periodic. A paper describing the result is cited below.

The paradoxical nature of locating sensors in paths and cycles

For a graph *G*, a set *D* contained in *V(G)* is called an *r*-identifying-code if for every vertex *x*, the set of vertices at distance at most *r* from *x* is nonempty and for every pair of vertices *x* and *y*, the set of vertices at distance at most *r* from *x* is different from the set of vertices at distance *r* from *y* The various applications of these codes include attack sensor placement in networks and fault detection/localization in multiprocessor or distributed systems. In the extensive literature on this topic, partial results are given about the minimum size of *D* for *r*-identifying codes for paths and cycles and complete closed form solutions are presented for the case $r = 1$. Such problems are important in spread of pathogens through subway tunnels or airport circular tram systems. David Roberts (Georgia Tech) and Fred Roberts (DIMACS) provided complete solutions for the case $r = 2$ as well as developing their own solutions (verifying earlier results) to the $r = 1$ case. These closed form solutions illustrate some surprisingly counterintuitive behavior that arises when the length of the path or cycle or the value of *r* varies.

Higher in utero and perinatal HIV infection risk in girls than boys

Donald Hoover (Rutgers) and his coauthors analyzed mother-to-child HIV transmission rates by sex and exposure time for babies born to HIV-infected untreated African women. Data was analyzed from two independent studies done in Malawi during the 1990s. Infections were established by polymerase chain reaction on blood samples. Odds ratios (ORs) for transmission were examined by period at risk: in utero (infected in umbilical cord blood), perinatal (infected in 1st postnatal blood > or =4 weeks), and postnatal (later postnatal infection). They found that among 1394 singleton births, girls were more likely to become infected than boys. For in utero transmission, the OR was 1.4 (95% CI: 0.9 to 2.2). For transmission during early life (umbilical cord blood not available) the OR was 2.7 (95% CI: 1.5 to 4.9). However, transmission risks in the perinatal and postnatal infection periods did not differ in boys and girls. Among 303 tested twin-birth pairs, girls were at higher risk than boys for in utero (OR: 2.6; 95% CI: 1.2 to 5.8) and perinatal (OR: 1.9; 95% CI: 1.0 to 3.7) infection. Recognized mother-to-child transmission risk factors did not explain the higher risk of infection in girls. They concluded that girls were at higher risk of early (in utero and perinatal) HIV infection than boys. They proposed that minor histocompatibility reactions between maternal lymphocytes and infant Y chromosome-derived antigens reduce the risk of HIV transmission in boys. A paper describing the results is cited below.

Association between the replication capacity and mother-to-child transmission of HIV-1, in antiretroviral drug-naive Malawian women

Donald Hoover (Rutgers) and his coauthors investigated the replication capacity and transmission of human immunodeficiency virus (HIV)-1 in antiretroviral drug-naive Malawian women who had subtype C infection. Infant children of these women received either 1 dose of nevirapine or 1 dose of nevirapine plus 1 week of daily doses of zidovudine. PhenoSense HIV was used to determine replication capacity in 49 women whose infants were infected with HIV-1 and in 47 women whose infants were uninfected by 6-

8 weeks of age. Mean replication capacity was higher in transmitters than in nontransmitters (P=.01). In a multivariate model, higher replication capacity was associated with transmission (odds ratio, 1.45 for each 10% increase in replication capacity [95% confidence interval, 1.11-1.90]; P = .0063), after adjustment for maternal HIV-1 load and other factors. A paper describing the results is cited below.

The impact of breastfeeding on the health of HIV-positive mothers and their children in sub-Saharan Africa

Donald Hoover (Rutgers) and his coauthors assessed the impact of breastfeeding by women infected with human immunodeficiency virus (HIV)-1 on their morbidity and risk of mortality and on the mortality of their children. They analyzed longitudinal data from two previous randomized clinical trials of mother-to-child transmission of HIV conducted between April 2000 and March 2003 in the Republic of Malawi, Africa. Mothers infected with HIV, and their newborns, were enrolled at the time of their child's birth; they then returned for follow-up visits when the child was aged 1 week, 6-8 weeks and then 3, 6, 9, 15, 18, 21 and 24 months. Patterns of breastfeeding (classified as exclusive, mixed or no breastfeeding), maternal morbidity and mortality, and mortality among their children were assessed at each visit. Descriptive and multivariate analyses were performed to determine the association between breastfeeding and maternal and infant outcomes. They found that a total of 2000 women infected with HIV were enrolled in the original studies. During the 2 years after birth, 44 (2.2%) mothers and 310 (15.5%) children died. (Multiple births were excluded.) The median duration of breastfeeding was 18 months (interquartile range (IQR)=9.0-22.5), exclusive breastfeeding 2 months (IQR=2-3) and mixed feeding 12 months (IQR=6-18). Breastfeeding patterns were not significantly associated with maternal mortality or morbidity after adjusting for maternal viral load and other covariates. Breastfeeding was associated with reduced mortality among infants and children: the adjusted hazard ratio for overall breastfeeding was 0.44 (95% confidence interval (CI)=0.28-0.70), for mixed feeding 0.45 (95% CI=0.28-0.71) and for exclusive breastfeeding 0.40 (95% CI=0.22-0.72). These protective effects were seen both in infants who were infected with HIV and those who were not. They concluded that breastfeeding by women infected with HIV was not associated with mortality or morbidity; it was associated with highly significant reductions in mortality among their children. A paper describing the results is cited below.

Outreach Activities

This Special Focus is closely intertwined with our Center's efforts to link mathematics and computer science with biology in the high schools. Project participants have been working with the Summer 2007 DIMACS Bio-Math Connection (BMC), which is aimed at introducing high school math, computer science, and biology teachers to topics at the interface. This project is informing the BMC effort and specific topics from the project are being adapted for use in BMC. The materials developed by BMC participants will consist of modules that can be flexibly adapted for use in a variety of courses at a variety of grade levels in both biology and mathematics. The project is run by DIMACS in collaboration with the Consortium for Mathematics and its Applications (COMAP) and Colorado State University. More information on the Field Testers workshop, that just ended, is available at http://dimacs.rutgers.edu/BMC/FieldTesters/2007/index.html Roughly twenty teachers were trained to use three modules integrating math and biology – Spider Silk, Genetic Inversion, and Biomatrices.

Researchers from Howard University and Morgan State University are spending the summer at DIMACS working on "Potential uses of entropy in biosurveillance" through a Department of Homeland Security-sponsored program aimed at faculty and students from minority-serving institutions. Nina Fefferman (DIMACS) and a team of two professors, two graduate students, and two undergraduate students are looking into coding theory techniques that can be applied to epidemiology. They hypothesize that the signal-to-noise ratio in reported disease incidence data may increase during the occurrence of disease

outbreaks and that these increases may be observable when analyzed using entropy measures. If the hypothesized increases are statistically quantifiable, and if they are found to occur early enough in the progression of disease throughout the affected population, this may prove an invaluable tool for early-warning biosurveillance. This activity has been closely coordinated with and is a clear offshoot of our NSF grant.

Under other funding, but coordinated with and a clear offshoot of this grant, DIMACS and the South African Centre for Epidemiological Modeling and Analysis (SACEMA) held a 3-day workshop on mathematical modeling and infectious diseases in Africa in September 2006. Then in June 2007, DIMACS/SACEMA together with the African Institute for Mathematical Sciences (AIMS), held a two week short course and a subsequent 3-day "capstone" workshop. The short course was aimed at training United States and African graduate students and postdoctoral fellows in mathematical epidemiology and the control of emerging and re-emerging infectious diseases. The capstone workshop served as a culmination to the short course and helped in furthering research and forming collaborations on the modeling of infectious diseases in Africa.

Books

S. Ji, A. Chaovalitwongse, N. Feferman, W. Yoo, and J. E. Perez-Ortin, "Mechanism-based clustering of genome-wide transcript levels: roles of transcription and transcript-degradation rates," *Clustering Problems in Biological Networks*, ed. A. Chaovalitwongse (to appear).

E. Lofgren, N. H. Fefferman, M. Doshi and E.N. Naumova, "Assessing Seasonal Variation in Multisource Surveillance Data: Annual Harmonic Regression," *BioSurveillance 2007*, eds. D. Zeng et al., Lecture Notes in Computer Science, 4506:114-123.

X. Xue and D. R. Hoover, "Statistical methods in cancer epidemiological studies" (to appear).

Papers

J. Abello and M. Capalbo, "An approximation algorithm to the modified quarantine problem on expander graphs" (submitted).

J. Abello and M. Capalbo, "Blocking sequences in infinite grids" (submitted).

J. Abello, and M. Capalbo, "Max cliques in sparse power law graphs with large clustering coefficients," (submitted).

J. Abello, B. Gaudin, H. J. Schulz, C, Tominski, "Name that cluster: text vs graphics" (submitted).

R. J. Biggar, T. E. Taha, D. R. Hoover, F. Yellin, N. Kumwenda, and R. Broadhead, "Higher *in utero* and perinatal HIV-infection risk in girls than boys," *JAIDS*, 41(4):509-13 (2006).

J. D. Church, S. E. Hudelson, L. A. Guay, S. Chen, D. R. Hoover, N. Parkin, S. A. Fiscus, F. Mmiro, P. Musoke, N. Kumwenda, J. B. Jackson, T. E. Taha, and S. H. Eshleman, "HIV-1 variants with nevirapine resistance mutations are rarely detected in antiretroviral drug naïve African women with subtypes A, C, and D," *AIDS Res Human Retroviruses*, 23:764-768 (2007).

J. D. Church, D. Jones, T. Flys, D. R. Hoover, N. Marlowe, S. Chen, C. Shi, J. R. Eshleman, L. A. Guay, J. B. Jackson, N. Kumwenda, T. E. Taha, and S. H. Eshleman, "Sensitivity of the ViroSeq™ HIV-1 Genotyping System for detection of the K103N resistance mutation in HIV-1 subtypes A, C, and D," *J*

Mol Diagnostics, 8:430-432 (2006).

B. Dembele, A. Friedman, and A. -A. Yakubu, "Malaria model with periodic mosquito birth rate," *Mathematical Biosciences* (submitted).

B. Dembele, A. Friedman, and A. -A. Yakubu, "Malaria model on the impact of drug administration protocols" (in preparation).

P.A. Dreyer, Jr., and F.S. Roberts, "Irreversible *k*-threshold processes: Graph-theoretical threshold models of the spread of disease and of opinion" (in preparation).

J. G. Dushoff, J. B. Plotkin, C. Viboud, L. Simonsen, M. Miller, M. Loeb, and D. Earn, "Vaccinating to protect a vulnerable subpopulation," PLoS Med 4(5): e174 doi:10.1371/journal.pmed.0040174.

S. H. Eshleman, J. D. Church, S. Chen, L. A. Guay, A. Mwatha, S. A. Fiscus, F. Mmiro, Musoke, N. Kumwenda, J. B. Jackson, T. E. Taha, and D. R. Hoover, "Comparison of HIV-1 mother-to-child transmission after single dose nevirapine prophylaxis among African women with subtypes A, C and D," *JAIDS* 42(4):518-21 (2006).

S. H. Eshleman, Y. Lie Y, D. R. Hoover, S. Chen, S. E. Hudelson, S. A. Fiscus, "Petropoulos CJ, Kumwenda N, Parkin N, and T. E. Taha, "Association of HIV-1 replication capacity with HIV-1 mother-to-child transmission among antiretroviral drug naïve Malawian women," *J Infect Dis*, 193(11):1512-5 (2006).

N. H. Fefferman and K. L Ng, "The role of individual choice in the evolution of social complexity," *Annales Zoologici Fennici*, 44:58-69.

N. H. Fefferman and K. L. Ng, "How disease models in static networks can fail to approximate disease in dynamic networks," provisionally accepted by *PRE*.

N. H. Fefferman and K. L. Ng, "Species-specific behavior affects disease spread throughout and ecosystem" (in preparation).

T. S. Flys, S. Chen, D. C. Jones, D. R. Hoover, J. D. Church, S. A. Fiscus, A. Mwatha, L. A. Guay, F. Mmiro, P. Musoke, N. Kumwenda, T. E. Taha, J. B. Jackson, and S. H. Eshleman, "Quantitative analysis of HIV-1 variants with the K103N resistance mutation after single-dose Nevirapine in women with HIV-1 subtypes A, C, and D," *JAIDS* 15;42(5):610-3(2006).

K. Hampson, J. G. Dushoff, J. Bingham, G. Brückner, Y. H. Ali, and A. P. Dobson, "Synchronous cycles of domestic dog rabies in sub-Saharan Africa and the impact of control efforts." *Proc Natl Acad Sci U S A.* 2007 May 1;104(18):7717-22.

M. Hauben, D. Madigan, S. Reisinger, A. Hochberg, and D. O'Hara, "Data mining in pharmacovigilence: computational cost as a neglected performance parameter," *Drug Safety* (tentatively accepted).

M. Hauben, L. Reich, C. M. Gerrits, and D. Madigan, "Spontaneous reporting of hyperkalemia and the randomized aldactone evaluation study," *Drug Safety* (to appear).

P.J. Lioy, F.S. Roberts, B. McCluskey, M.J. Lioy, A. Cross, L. Clarke, L.L. Stanton, W. Tepfenhart, M.E. Ferrara, "TOPOFF 3 comments and recommendations by members of New Jersey Universities

Consortium for Homeland Security Research *Journal of Emergency Management*, **4** (2006), 41-51.

J. R. Pulliam, J. G. Dushoff, S. A. Levin, and A. P. Dobson "Epidemic enhancement in partially immune populations," PLoS ONE 2(1): e165. doi:10.1371/journal.pone.0000165.

D. L. Roberts and F. S. Roberts, "Locating sensors in paths and cycles: The case of 2-identifying codes," European J. of Combinatorics, to appear.

J. Shi and D. R. Hoover, "Exact distribution of the smallest multinomial group when each category has equal probability" (in preparation).

S. R. Sundaresan, I. R. Fischhoff, J. G. Dushoff, and D. I. Rubenstein, "Network metrics reveal differences in social organization between two fission-fusion species, Grevy's zebra and onager," *Oecologia* 2007 Feb; 151(1):140-9.

T. E. Taha, N. L. Kumwenda, D. R. Hoover, G. S. Kafulafula , A. Fiscus, C. Nkhoma, S. Chen, and R. L. Broadhead, "The impact of breastfeeding on the health of HIV-1 infected mothers and their children in Sub-Saharan Africa," *Bull World Health Org* 84(7):546-54 (2006).

T. E. Taha, D. R. Hoover, N. I. Kumwenda, S. A. Fiscus, G. Kafulafula, C. Nkhoma, S. Chen, E. Piwowar, R. L. Broadhead, J. B. Jackson, P. G. Miotti, "Late postnatal transmission of HIV-1, breast milk viral load and Nevirapine levels in breast milk and plasma," *J Infect Dis*, 196 (1):10-4 (2007).

T. E. Taha, N. I. Kumwenda, G. Kafulafula , B. Makanani, C. Nkhoma, S. Chen , A. Tsui , D. R. Hoover, "Intermittent intravaginal Metronidazole Gel Treatment of bacterial vaginosis in HIV-uninfected and infected Women: A Randomized Clinical Trial," *PLoS Clinical Trials* 23;2(2): (2007).

Q. Xia and D. R. Hoover, "A procedure for group sequential comparative Poisson trials," *Biopharmaceutical Statistics* (to appear).

Reports

H. Rolka, H. Burkom, G. F. Cooper, M. Kulldorff, D. Madigan, and W-K Wong, "Issues in applied statistics for public health bioterrorism surveillance using multiple data streams: research needs," *Statistics in Medicine*, 2006; **26**:000–000

D. L. Roberts and F. S. Roberts, "The paradoxical nature of locating sensors in paths and cycles: The case of 2-identifying codes)," DIMACS Technical Report.

Talks

N. H. Fefferman, "Does Securing Infrastructure Against Workforce-Depletion Depend on Whether the Risk is Environmental or Infectious?" DIMACS Capstone Workshop on Mathematical Modeling of Infectious Diseases in Africa, June 25 - 27, 2007, Stellenbosch, South Africa.

N. H. Fefferman, "Preparing Societal Infrastructure Against Disease-Related Workforce Depletion." DIMACS Workshop on Facing the Challenge of Infectious Diseases in Africa: The Role of Mathematical Modeling, September 25 - 27, 2006, University of the Witwatersrand, Johannesburg, South Africa.

E. Lofgren, N. H. Fefferman, M. Doshi and E. N. Naumova, "Assessing Seasonal Variation in Multisource Surveillance Data: Annual Harmonic Regression," NSF Biosurveillance Systems & Case

Studies Workshop in May, 2007.

N. H. Fefferman and K. L. Ng, “The impact of shifting social contacts on infectious disease dynamics” invited seminar talk to the Princeton University Dept. of Ecology and Evolutionary Biology and invited seminar talk at the Rutgers Univ. Dept of Ecology and Evolution.

B. Gaudin, “Name that cluster: text vs graphics,” Technology Trends in the United States (InfoVis 2006 Contest), University College Dublin, Ireland.

S. Ji and R. Patel, “Visualizing and classifying dissipative structures (or time series) in two-dimensional principal grids,” DIMACS Workshop on Computational Methods for Dynamic Interaction Networks, Rutgers University, Piscataway, 9/24-25, 2007.

F. S. Roberts, “Meaningless Statements in Epidemiology,” Workshop on Mathematical Modeling of Infectious Diseases in Africa, Stellenbosch, South Africa, June 2007.

F. S. Roberts, “Locating Sensors to Detect Chem, Bio, or Nuclear Threats,” DyDAn program on Mathematics and Homeland Security for High School Teachers, May 2007.

F. S. Roberts, “New Challenges for Modelers of Infectious Diseases of Africa,” Workshop on Facing the Challenge of Infectious Diseases in Africa: The Role of Mathematical Modeling in Johannesburg, South Africa, September 2006.

F. S. Roberts, “Measurement of Pollution,” two hour tutorial, MAA Shortcourse on Environmental Modeling, Knoxville, 8-06.

F. S. Roberts, “Graph-theoretical Problems Arising from Defending Against Bioterrorism and Controlling the Spread of Fires,” DyDAn program on Mathematics and Homeland Security for High School Teachers, May 2007

F. S. Roberts, “Graph-theoretical Problems Arising from Defending Against Bioterrorism and Controlling the Spread of Fires,” DIMACS/AIMS/SACEMA Advanced Study Institute on Mathematical Modeling of Infectious Diseases of Africa, at African Institute for Mathematical Sciences, Muizenberg, South Africa, June 2007.

Q. Xia and D. R. Hoover, “Exact methods for group sequential and stratified comparative Poisson studies,” Pfizer- Rutgers Colloquiumn New Brunswick, NJ, April 27, 2007 (invited).

Main Web Site

http://dimacs.rutgers.edu/SpecialYears/2002_Epid/

Other Specific Products

Web Pages

http://dimacs.rutgers.edu/Workshops/Diseases
Web page for the workshop on “Facing the Challenge of Infectious Diseases in Africa: The Role of Mathematical Modeling.”

http://dimacs.rutgers.edu/Workshops/Pathogens

Web page for the workshop on “Models of Co-Evolution of Hosts and Pathogens.”

http://dimacs.rutgers.edu/Workshops/Immuno
Web page for the workshop on “Immuno-epidemiology.”

http://dimacs.rutgers.edu/Workshops/WGSpatioTemporal2
Web page for the working group meeting on “Spatio-Temporal and Network Modeling of Diseases (Second Meeting)”

http://dimacs.rutgers.edu/Biosurveillance
Web page for the DIMACS/DyDAn research project on “Potential uses of entropy in Biosurveillance.”

http://dimacs.rutgers.edu/BMC/
Web page for the DIMACS BioMath Connection

http://dimacs.rutgers.edu/Workshops/AIMS
Web page for the US-Africa Advanced Study Institute on Mathematical Modeling of Infectious Diseases in Africa

http://dimacs.rutgers.edu/Workshops/AfricaDiseases
Web page for the workshop on “Mathematical Modeling of Infectious Diseases in Africa.”

http://dimacs.rutgers.edu/Workshops/InfectiousDisease
Web page for the workshop on “Systems Biology of Infectious Diseases.”

Software

In 2004 and 2005 DIMACS, through this project, fostered a collaboration between David Ozonoff (Boston University) and Alex Pogel (U. of New Mexico) that has since flourished. During 2006, Pogel's work on Ozonoff's SuperFund grant has seen much progress made toward building and distributing an open source concept lattice analysis tool for epidemiologists, which will generate contingency tables from concept lattices with a few clicks on lattice nodes. This work stemmed originally from this NSF project. Pogel has hired a team of software engineers, written a specification document, and has a well-defined five-year plan for the completion of the project. The first version of this software is expected to appear at SourceForge.net in mid-2008, and the product will be publicized in the 2009-2010 academic year. Also, in Pogel's work at PSL, various features described in the Ozonoff, Pogel, and Hannan contribution to the AMS-DIMACS Special Volume on Computational Epidemiology (“Generalized Contingency Tables”) have been incorporated into Seqer (pronounced “seeker”), the data analysis software developed by Pogel's group over the past three years.

Contributions

Contributions within Discipline

This Special Focus is by nature multi-disciplinary. It involves applying methods of computer science, statistics, and mathematics to the problems in epidemiology. Introducing people to this combination of disciplines has been a key goal and a key accomplishment of this project. Below is a selection of comments from project participants indicating their assessment of the impact of this grant.

Kaus Dietz, Director of the Department of Medical Biometry at the Eberhard-Karls-University of Tübingen in Germany said “In December 2006 I attended the DIMACS-Workshop on Immuno-

epidemiology. I met for the first time Professor Sungchul Ji, Department of Pharmacology and Toxicology, Rutgers University. We started a collaboration on the analysis of genome-wide mRNA levels measured during diauxic shift in yeast cells."

Sungchul Ji from the Department of Pharmacology and Toxicology at Rutgers University said "I want to thank you and your colleagues for organizing the above workshop this week. I learned a lot from the workshop, the first such meeting I have attended in immuno-epidemiology. The meeting opened my eyes to the field, which I may enter in the future as a theoretical cell biologist. One concrete result of my attending the workshop was meeting Klaus Dietz whom Hans Heesterbeek introduced to me. During the meeting Klaus and I began to analyze some of the budding yeast microarray data that I have been working on during the past two years. After taking Kalus to the Newark Airport on the last day, Klaus continued his calculations with his laptop at a bar in the airport and found out (about 30 minutes before his flight time) that my data are lognormally distributed. To me this is a totally unexpected finding and prompted me to write the attached email to ask him to expand his calculations. If Klaus finds lognormal distributions for *certain* variables we will have (I believe) results novel and important enough to construct a short communication to a journal such as the Journal of Theoretical Biology. Again, many thanks for all the good work that you and your colleagues at DIMACS have been doing over the years. I have been benefiting from these workshops enormously, but have been slow in acknowledging my indebtedness to DIMACS until now."

Gavin Welch from the Epidemiology and Biometry Core at Eastern Virginia Medical School said "Thanks for a wonderful series of workshops. As a direct result of the four DIMACS Comp-Epi workshops I attended: (1) One of the participants at the host-pathogen coevolution workshop gave a seminar on his influenza modeling work at my institution. We invited him because I learned of his work at the workshop. (2) I've used several examples in my classes that I've learned about by talking with other participants at DIMACS workshops. Most of the examples came from discussions over lunch."

Abdul-Aziz Yakubu from Howard University said "Participation resulted in collaborations with Avner Friedman and my student, Bassidy Dembele. We're working on Malaria epidemic in Mali. To date, we've submitted a paper for publication. The title of the paper is 'Malaria model with periodic mosquito birth rate.' We're currently working on a second paper on the Impact of Malaria Drug Administration in Mali."

One of the biggest contributions of this project has been the development of a new initiative to bring together scientists from the United States and various African countries to collaborate on mathematical modeling of infectious diseases in Africa. One specific DIMACS workshop was instrumental in leading to the development of the proposed activity in Africa. This was a workshop on "Evolutionary Aspects of Vaccine Use," organized at DIMACS at Rutgers University in June 2005, and stemming from the work of the DIMACS working group on Methodologies for Comparing Vaccination Strategies. The workshop, organized by Troy Day (Queen's University, Ontario), Alison Galvani (Yale), Abba Gumel (U. of Manitoba), and Claudio Struchiner (Oswaldo Cruz Foundation, Brazil), examined general evolution-related questions for diseases for which there is a vaccine or hope for one. There were 30 scientists participating, from diverse backgrounds, including mathematicians, epidemiologists, virologists, immunologists, and vaccine developers. The special problems of vaccination strategies in Africa that arose in this workshop were one of the primary motivations that led Abba Gumel to propose that DIMACS sponsor a short course and workshop that directly focuses on mathematical modeling of infectious diseases in Africa.

Following that workshop, DIMACS and the South African Centre for Epidemiological Modeling and Analysis (SACEMA) held a 3-day Workshop on mathematical modeling and infectious diseases in Africa, at the School of Computational and Applied Mathematics at the University of the Witwatersrand, Johannesburg, South Africa, on September 26-28, 2006. The workshop brought together scientists from

the United States and various African countries, as well as junior researchers and students and provided an agenda for future collaborations. It exposed junior United States scientists and students to the special challenges of modeling the spread of disease in Africa and the opportunities to collaborate with Africans in developing and applying the tools of mathematical modeling to the tremendous health problems caused by such diseases as HIV/AIDS, tuberculosis and malaria, as well as the possible interventions in the case of major new health threats such as pandemic influenza in an underdeveloped region of the world.

Subsequently, DIMACS, SACEMA, and the African Institute for Mathematical Sciences (AIMS), developed plans for a two week short course on mathematical modeling and infectious diseases in Africa. The short course, held at AIMS in Cape Town, South Africa, on June 11-22, 2007, trained junior United States and African graduate students and postdoctoral fellows in mathematical epidemiology and the control of emerging and re-emerging infectious diseases. Those who completed the short course participated in a 3-day "capstone" workshop designed to establish collaborations. The capstone workshop was held at the new conference facilities of the Stellenbosch Institute for Advanced Study (STIAS) on June 25-27.

Long term DIMACS visitor Nina Fefferman said "Throughout the past year, I have continued my research into the varying success of organizational strategies of worker allocation under conditions of disease related workforce depletion. The results of these ongoing endeavors have been presented at two DIMACS workshops. Additionally, work to extend their applicability, incorporating economic modeling, is now underway in collaboration with researcher Dr. Ramanan Laxminarayan (of Resources for the Future), whom I met at a DIMACS workshop in 2005. Further possibilities for extensions of this work to examine alternate scenarios of economic impact are just beginning in collaboration with Dr. Martin Meltzer (of the CDC), whom I met at the recent DIMACS Capstone Workshop in South Africa.

"Due to our mutual involvement with DIMACS, Prof. Dina Fonseca (at the Rutgers Univ. Center for Vector Biology) and I met to discuss our mutual interest in temporal constraints on mosquito reproduction. As an outgrowth of this meeting, we have begun a collaboration to investigate the impact of mosquito habitat selection on invasive spread of a newly introduced mosquito in the state of NJ. Once accomplished, this work will have profound impact on the capability for predictive modeling of mosquito-borne diseases as this species is a known disease vector in its native habitats. This work involved the recruitment of three other researchers, Prof. Rick Lathrop, Prof. Jim Miller, and Prof. Randy Gaugler (of the Rutgers Univ. Depts. of Ecology and Evolution, Oceanography and Marine Science, and the Center for Vector Biology, respectively). Together, the five of us have now submitted a proposal for funding of this work to the NSF Directorate of Biological Sciences.

"As the result of a conversation had during a dinner at the DIMACS workshop on The Epidemiology and Evolution of Influenza, I have organized and am editing the publication of a special issue of the journal *Annales Zoologici Fennici* dedicated to the uses of simulation modeling techniques in the investigation of questions in biology and medicine. This issue is due for publication in October of this year. This issue would never have been organized had it not been for the interest and encouragement expressed at that DIMACS workshop dinner.

"All of these projects and collaborations were made possible by my involvement with DIMACS over the past year. I feel that my perspective, even of my own, individual research has been broadened substantially by this opportunity and I am very grateful. I feel lucky to have been a part of this focus in Computational and Mathematical Epidemiology and am looking forward to continuing with these projects and these people."

Sungchul Ji describes his collaboration with Nina Fefferman as follows: "Dr. Fefferman attended one of my seminars given at the DIMACS Center at Rutgers on April 17, 2006 invited by James Abello. Based

on the content of my seminar entitled 'The Simpson-Elsasser-Wolfram (SEW) Framework for Modeling the Living Cell,' Nina asked me to contribute a paper to a special edition of the journal Ann. Zool. Fennici, for which she is an editor. This led to writing a set of three manuscripts applying the network theory to cell biology. It was through the influence of James Abello, Nina Fefferman and Art Chaovalitwongse whom I met through DIMACS Workshops and Seminars that I was motivated to learn and apply network concepts and theories to molecular and cell biology. In the process, I have been to led to formulate what appears to be a novel theory of biological networks and apply it to modeling the living cell."

Suzanne Lenhart from the University of Tennessee said "I started a new collaboration with Edward Lungu and Abba Gumel and will report back on the progress of that work about HIV epidemic models including the educational information impact." Abba Gumel is from the University of Manitoba in Winnipeg, Canada and Edward Lungu is from the University of Botswana, in Gaborone, Botswana.

Jonathan Read from the Mathematics Institute at the University of Warwick collaborated with Nina Fefferman (DIMACS) on a variety of projects while visiting DIMACS. The primary focus of their work was expanding the research that Fefferman and Kah Loon Ng began to examine under what general rules self-organizing dynamic social networks formed similar convergent structures using centrality measures as a metric for similarity. This work continued at the DIMACS funded workshop in Edinburgh. Second, Read collaborated with Fefferman to design a novel presentational encoding strategy for social network graphs which will allow a rapid characterization of the potential for the spread of disease and/or information within the network depending on the individual through which occurs the primary introduction. Third, Read and Fefferman collaborated on models examining the difference in selective pressures acting on pathogen evolution caused by physiological constraints from host-immune evasion. Lastly, Read presented his own research into the social network structure of academia and during his visit collaborated with James Abello to create a clearer visual representation of his results. Together they were able to construct an image which clearly communicated greater levels of complexity in the network than had been previously possible.

Regarding Jonathan Read, Nina Fefferman said: "During the past year, I met researcher Jonathan Read at the DIMACS workshop on Models of Co-Evolution of Hosts and Pathogens. Together, we are working to develop models of the impact of intra-host immunology and physiology on the selective pressures driving pathogen virulence. This work is ongoing. We also collaborated to expand some of the work done by Fefferman & Ng (see above), attempting to determine under what general rules self-organizing dynamic social networks formed similar convergent structures using centrality measures as a metric for similarity. This work is also ongoing. In addition to those models, I have continued collaborating with Read to design a novel presentational encoding strategy for social network graphs which will allow a rapid characterization of the potential for the spread of disease and/or information within the network depending on the individual through which occurs the primary introduction. This work is ongoing and was continued in person after our initial meeting at the October 2006 workshop during Read's brief visit to DIMACS in March of 2007, and subsequently also at the DIMACS workshop in May 2007 on Spatio-temporal and Network Modelling of Diseases. Each of these DIMACS funded events provided valuable opportunity for furthering our collaboration that would otherwise have been impossible."

Contributions to Other Disciplines

Since the "discipline" is inherently multidisciplinary, there is no separate entry in this section.

Contributions Beyond Science And Engineering

Several of the outcomes of the special focus have immediate applicability to real world problems.

Donald Hoover from the Statistics Department at Rutgers University reports that his theoretical work focuses on statistical methods for testing interventions (such as vaccines) to prevent rare outcomes such as the types that could be of concern for bioterrorism. He has continued and expanded work begun during the earlier years of this proposal on understanding development of maternal and infant HIV resistance to antriretroviral medications that are given to prevent maternal to infant transmission of HIV in Africa. His findings, for example, conclude that girls were at higher risk of early (in utero and perinatal) HIV infection than boys and that breastfeeding by women infected with HIV was not associated with mortality or morbidity; it was associated with highly significant reductions in mortality among their children.

William DuMouchel from Lincoln Technologies spoke on "Data Mining for Drug Safety: Statistical Analyses of Spontaneous Reports and Clinical Safety Data" in the DIMACS seminar. Postmarket data on drug safety is difficult to analyze, particularly because it has no "denominator," that is, there is no record of how many times a drug was used without side effects. Data quality is also an issue, since the same event can be reported multiple times by the doctor, medical facility and the drug company itself. Successful techniques that are robust in the face of these problems need to be developed to identify potential problems with drugs already on the market.

David Madigan from the Statistics Department at Rutgers and his coauthors wrote a report on "Issues in applied statistics for public health bioterrorism surveillance using multiple data streams: research needs." The objective of this report is to provide a basis to inform decisions about priorities for developing statistical research initiatives in the field of public health surveillance for emerging threats. It was written to help public health surveillance authorities to make better judgments on building practical systems because rapid information system advances have created a vast amount of health data pouring in from various reporting sources. They describe the analytic and statistical methodologies that synchronize best with information technologies used to gather the data. Among other things, they describe a space–time statistic that has successfully been used to detect and track public health events of interest.

Several participants in this project, working through the New Jersey Universities Consortium for Homeland Security Research, were observers for the TOPOFF 3 homeland security exercise in New Jersey in April 2005. The exercise involved release of pneumonic plague with symptomatic individuals sent to hospitals and asymptomatic individuals to Points of Dispensing (PODs) to receive antibiotics. They worked with both the NJ Office of Homeland Security and Preparedness and FEMA to provide input to their After Action Reports. They wrote a paper for the Journal of Emergency Management that summarizes their observations about the exercise, with emphasis on the POD strategy.

Through this project, we have made several connections with the Centers for Disease Control and Preparedness. CDC folks have helped organize project workshops and advised us on project design. In turn, the P.I. was invited to help CDC design a mathematical modeling program. That program is now being implemented.

Contributions To Human Resources Development

Several graduate students, undergraduates, and postdoctoral researchers have participated in the program. Local graduate students and many non-local students were also involved as visitors and workshop/working group attendees.

Specifically, the following Rutgers graduate students have undertaken small research projects under support of the special focus.

- Qinhe Cheng worked on a project titled “Causaul Effects in Dropouts from General Drug Trials”
- Srinivas Maloor worked with David Madigan from the Statistics Department.
- Graduate student David Millman has worked on random graph models of how diseases spread with DIMACS visitor Michael Capalbo.

Long-term DIMACS visitor Nina Fefferman has been investigating a variety of entropy metrics using a variety of disease data sources, and temporal and spatial scales. This year she is working with the following professors, graduate students, and undergraduate students:

- Abdul-Aziz Yakubu, Mathematics, Howard University
- Devroy McFarlane, Mathematics, Howard University (graduate student)
- Ashley Crump, Mathematics, Howard University (undergraduate student)
- Asamoah Nkwanta, Mathematics, Morgan State University
- Anthony Ogbuka, Biological Sciences, Morgan State University
- Nakeya Williams, Mathematics, Morgan State University

This project involves teams of faculty and students from two minority-serving institutions, Howard University and Morgan State University, working at DIMACS for the summer. They are funded through a Department of Homeland Security program that aims at getting faculty and students at MSIs to work on homeland security research problems. The project was developed as a result of our NSF project and is closely tied to it.

This project hosted and trained a postdoctoral fellow from National University of Singapore, Kah Loon Ng. Regarding Ng, Nina Fefferman (DIMACS) said “During his post-doctoral work at DIMACS, Kah Loon Ng and I began work together investigating the role of individual-level behavioral choice on the stability of social networks. Since social network contacts provide the means of transmission of infectious disease, we have been examining the divergent disease spread processes that emerge from the variations in stability. Our first paper in this subject was published in January of this year in the biology journal *Annales Zoologici Fennici*. This paper focused on the diversity of stable structures that could influence the evolution of dynamic social networks. Building on that publication, our second paper is already awaiting a final decision after having been provisionally accepted by the physics journal *Physical Reviews E*. In this paper, we demonstrated that disease processes on static networks was insufficient (in many cases) to accurately approximate infection spread on networks with shifting social contacts. As many of the leading studies in this field have relied on the assumption that findings on static networks are easily generalizable to any network sharing similar, average graph theoretic properties, we hope and expect that our work together at DIMACS will have a profound effect on the cutting edge of research in the field of network-based epidemiology. We are also in the process of preparing further papers building on these earlier ideas. Most fundamentally, we are examining whether or not differently behaving, but occasionally interacting networks can yield different patterns of inter-population disease transmission. These studies will be among the first to explicitly examine the network epidemiology of the formation of reservoir populations for zoonotic infections, a subject of great concern, especially considering the diversity of bird behaviors, potentially influencing their social susceptibility to avian influenza. We are also attempting to expand the scope of our investigations to examine the impact of disease avoidance behaviors on the social stability of a network. This work is still in the early stages.

“In addition to our direct research together, Dr. Ng and I have been invited to present some of our findings at various departmental seminars and professional meetings. This past year, we have given invited talks to the meeting of the DIMACS Working Group on “Spatio-temporal and Network Modelling of Diseases.” Based on our presentation at this meeting, we have initiated collaboration with Prof.

Martina Morris, to explore the potential interface between her research and our own. Our work was also presented at an invited seminar talk to the Princeton University Dept. of Ecology and Evolutionary Biology and at an invited seminar talk at the Rutgers Univ. Dept of Ecology and Evolution."

Maite Severins from the Department of Farm Animal Health at Utrecht University in the Netherlands said "I am a Phd student at Utrecht University and attended the DIMACS Workshop on Immuno-epidemiology on December 11 - 13, 2006. This has recently led to a collaboration between myself and Sander van Noort, PhD student at the Theoretical Epidemiology group of the Instituto Gulbenkian de CiEAncia in Portugal, who also attended the DIMACS workshop."

Lastly, the workshops are also helping participants improve their teaching activities. For example, Gavin Welch from the Epidemiology and Biometry Core at Eastern Virginia Medical School wrote that after attending the Comp-Epi workshops he used several examples in his classes that he learned about by talking with other participants at DIMACS workshops.

Donald Hoover from the Statistics Department at Rutgers University has collaborated with Ph.D. statistics students in activities related to this grant. His theoretical work focuses on statistical methods for testing interventions (such as vaccines) to prevent rare outcomes such as the types that could be of concern for bioterrorism. This has lead to a PhD thesis completed by his student Qi Xia. An article from this thesis was accepted for publication in the *Journal for Biopharmaceutical Statistics* and follow up articles form the thesis are planned. He has begun work with a new Ph.D. student in the Department of Statistics, Jing Shi, to expand on these ideas.