Interagency Modeling and Analysis Group

Futures Meeting

IMAG

The Impact of Modeling on Biomedical Research

December 15-16, 2009 ● Bethesda, Maryland

FINAL REPORT

Table of Contents

Executive Summary

The Interagency Modeling and Analysis Group (IMAG) held an IMAG Futures Meeting on December 15-16, 2009 on the National Institutes of Health (NIH) Campus at the Natcher Conference Center. This meeting was open to the public to share in the discussion of a working group who deliberated on the impact of biomedical, biological and behavioral modeling. This meeting was an opportunity to assess the extent to which computational modeling has succeeded to date in making a difference in the broader biomedical research endeavor, and to discuss progress in the field in the context of current challenges and opportunities for biomedical, biological and behavioral modeling.

While presenters and discussants were not charged with reaching consensus or making specific recommendations, a picture emerged from the discussions of accelerating progress and expanding opportunity for computational modeling, especially multi-scale modeling, to become an integral and essential component of training, research and translation in a growing number of biomedical specialities, much as it already has in fields such as physics and engineering. In particular given the availability of new data, tools and methods, the potential for more sophisticated and highly integrated models, that span a wider range of scales of biological organization, and that apply to clinical practice and healthcare, is greater now than at any time in the past.

In some research fields such as computational neuroscience and cardiovascular biophysics, modeling has become increasingly integrated with experimental research. However, challenges remain. In many branches of biomedical science, familiarity with and acceptance of computational modeling remains the exception rather than the norm. There is still a pressing need for more scientists with the interdisciplinary training to carry out this kind of research.

Efforts could be made to achieve closer interactions and better communication between researchers applying modeling to different biomedical disciplines. This could accelerate progress in biomedical modeling, open new fields of biomedicine to modeling and lead to more comprehensive, integrated and reliable models. Application areas with particular promise include drug discovery, predictive analyses of cell signaling pathways and regulatory networks, patient-specific modeling for diagnosis and treatment planning especially for surgery and interventional procedures, and clinical trial design.

Some common conclusions from the discussions about modeling at different scales were:

- Modeling has been successful in almost all biomedical domains and scales;
- Modeling will be increasingly necessary as the complexity of biomedical data and conceptual hypotheses grow;
- Modeling has the potential to accelerate translation from basic science to clinical medicine;
- Modeling can efficiently drive experimentation, data acquisition and technology development; and
- Modeling promotes interdisciplinary research and training.

I. Introduction

In April of 2003 the Interagency Modeling and Analysis Group (IMAG) was formed, starting from a working group, comprised of program staff from nine Institutes of the National Institutes of Health (NIH) and three directorates of the National Science Foundation (NSF). IMAG is now comprised of program directors from nine government agencies in the U.S. and Canada, including the National Institutes of Health (NIH), the National Science Foundation (NSF), Department of Energy (DOE), Department of Defense (DOD), the United States Department of Agriculture (USDA), the Canadian federal research network for Mathematics of Information Technology and Complex Systems (MITACS), the National Aeronautics and Space Administration (NASA), the United States Department of Veterans Administration (USDVA), and the Food and Drug Administration (FDA, joined January 2010).

The purpose of IMAG is to bring together program officers who have a shared interest in supporting modeling and analysis methods in biomedical, biological and behavioral systems. IMAG has promoted and supported a wide variety of modeling over the past six years, notably the multi-scale modeling (MSM) initiative, which was originally a multi-agency FOA and continued on by different IMAG agencies through separate FOAs.

The IMAG/MSM Consortium was started in 2006 with the convening of the initial group of MSM grantees from the Interagency Opportunities in Multi-Scale Modeling in Biomedical, Biological, and Behavioral Systems Solicitation. [1] The Consortium has grown to include principal investigators of other projects funded by other initiatives, but also relevant to the Consortium. All Consortium activities are publicly available on the IMAG Wiki.[2]

As the IMAG/MSM Consortium has matured, IMAG is now also exploring population models that are based primarily on statistical data and methods and the possibilities for leveraging these models with mechanistic models developed by other MSM participants. Furthermore, there is a swelling interest from the international community in coordinating national and regional efforts for multiscale, "Physiome" modeling[3]. In order to sustain and accelerate progress in biomedical modeling, it is therefore important to consider how models can indeed impact broader communities.

On December 15-16, 2009 IMAG held the first IMAG Futures Meeting - *The Impact of Modeling on Biomedical Research*. This meeting was an opportunity to assess to what extent computational modeling has succeeded or failed in making a difference in the broader biomedical research endeavor, and to discuss these issues in the context of current challenges and opportunities for biomedical, biological and behavioral modeling. This was a brainstorming meeting that included government and community leaders, as well as attendees interacting via worldwide videocast.

[1] http://www.nsf.gov/pubs/2004/nsf04607/nsf04607.htm
[2] www.imagwiki.org/mediawiki
[3] Hunter and Borg. Integration from proteins to organs: the Physiome Project. Nat Rev Mol Cell Biol (2003) vol. 4 (3) pp. 237-43

I.1 Contributors

The meeting discussions were organized according to modeling activities related to each of five major levels of organization in the biological hierarchy: population modeling; whole body modeling; cell-tissue-organ modeling; pathways and networks modeling; and atomic and molecular modeling. This report is a compilation of thoughts from participants of the five speaker panels representing these five categories of modeling and discussants who provided comments and feedback in person and via the Internet. The primary authors of this report is Dr. Andrew McCulloch from University of California San Diego, and the other panel chairs (introduced below) – Dr. Sylvia Plevritis, Dr. Donald Bolser, Dr. Joel Stiles, and Dr. Vijay Pande.

Participants were encouraged to reflect on their own fields and expertise and call attention to issues that are unique to each level of modeling (scale) as well as issues that may span across scales. The archived videocast, all meeting presentations, reports and an open post-meeting public commentary period are available on the IMAG wiki (http://www.imagwiki.org/mediawiki/index.php?title=IFM_Announcement).

I.1.1 Population Modeling

- Dr. David Eddy is the Medical Director and Founder of Archimedes, a company that mathematically models medicine. Prior to founding Archimedes, Dr. Eddy was Professor of Engineering and Medicine at Stanford, and the J. Alexander McMahon Professor of Health Policy and Management at Duke University.
- Dr. Bryan Grenfell is Alumni Professor of Biology at Penn State University, where he works in the Center for Infectious Disease Dynamics using theoretical models and empirical data to study the population biology of infectious diseases.
- Dr. Sylvia Plevritis (**Chair**) is Associate Professor of Radiology at Stanford University, where her research program focuses on computational modeling of cancer biology and cancer outcomes.
- Dr. Paolo Vicini is a Research Fellow at Pfizer Global R&D in La Jolla, CA. He has a background in bioengineering and works on translational research in pharmacokinetics, dynamics, and metabolism.

I.1.2 Whole Body Modeling

- Dr. Donald Bolser (**Chair**) is Professor of Physiological Sciences at the University of Florida, where his laboratory studies respiratory physiology and uses computational models to investigate cough mechanisms.
- Dr. Marco Viceconti from Instituto Ortopedico Rizzoli di Bologna develops multi-scale computational models for applications in human orthopedics and trauma. He is actively involved in the leadership of the European Commission-funded Virtual Physiological Human[4] (VPH) Project (see Section VII: International Activities).

[4] The virtual physiological human-A European initiative for *in silico* human modeling. Viceconti M, Clapworthy G, Van Sint Jan S. J Physiol Sci. 2008 Dec:58(7):441-446.

- Dr. Yoram Vodovotz is Director of the Center for Inflammation and Regenerative Modeling at the University of Pittsburgh, where he studies systems modeling of inflammation.

I.1.3 Cell, Tissue and Organ Modeling

- Dr. Andrew McCulloch (**Chair**) is a Professor of Bioengineering at University of California San Diego who works on multi-scale studies of cardiac biology and disease.
- Dr. Jeff Smith is a Senior Investigator at the NIH in the National Institute of Neurological Disorders and Stroke, where he is studying the functional and computational properties of oscillatory motor networks in the mammalian brainstem and spinal cord.
- Dr. Shlomo Ta'asan from Carnegie Mellon University is developing multi-scale models for immunology and infectious diseases including influenza and tuberculosis.
- Dr. Vito Quaranta, a cancer biologist at Vanderbilt University, is applying systems biology approaches to understanding mechanisms of cancer cell invasion and metastasis.

I.1.4 Pathways and Networks Modeling

- Dr. Ronald Germain from the Laboratory of Immunology and the Program in Systems Immunology and Infectious Disease Modeling at the National Institute of Allergy and Infectious Disease, part of the NIH, uses mechanistic cell systems models to investigate questions in immunology.
- Dr. Timothy Kinsella is a radiation oncologist at the Case Western School of Medicine and University Hospitals of Cleveland.
- Dr. Joel Stiles (**Chair**) is Director of the Center for Quantitative Biological Simulation at the Pittsburgh Supercomputing Center with research interests in synaptic and cellular microphysiology.

I.1.5 Atomic and Molecular Modeling

- Dr. Vijay Pande (**Chair**) is Associate Professor of Chemistry, Structural Biology and Computer Science at Stanford University and an investigator of the Physics-based Simulation of Biological Structures (SIMBIOS) National Center for Biomedical Computing (NCBC). His laboratory develops computational models and tools for molecular simulation of protein folding and small molecule drug design.
- Dr. Abby Parrill is Professor and Chair of Chemistry at the University of Memphis, where her research group works on computational structural biology and rational ligand design.
- Dr. Linda Petzold is Professor of Computer Science and Mechanical Engineering at the University of California Santa Barbara. Her research includes the development and analysis of multiscale simulation methods for biochemical reactions and networks.
- Dr. Tamar Schlick is Professor of Chemistry, Mathematics, and Computer Science and a member of the Courant Institute of Mathematical Sciences at New York University. Her group is developing molecular modeling, bioinformatics, and

mathematical methods to study DNA repair and fidelity mechanisms, chromatin folding, and RNA structure and function.

I.1.6 Other Speakers

- Dr. Peter Hunter is a Bioengineer who directs the Auckland Bioengineering Institute at the University of Auckland and leads the international Physiome project under the auspices of the International Union of Physiological Sciences (IUPS).
- Dr. Tina Morrison is a Medical Device Fellow in the Division of Cardiovascular Devices at the Food and Drug Administration (FDA).

I. 2 Meeting Charge

In the context of each of the five scale levels of the biological hierarchy, participants were asked to address the following four major charge topics:

- How modeling has impacted various research fields (success stories and mechanisms)? Is it the onus on modelers to prove that their models are useful to someone else?
- To what extent has the broader research communities accepted modeling as a critical tool for *driving* research or policy (what has worked and what hasn't worked)?
- What are the major challenges to overcome (how do we get there)?
- In what ways can modeling further impact the broader research communities (how far can we go)?

These questions were also to be addressed in the context of the following specific issues:

(1) The current perception of modeling in the biomedical and clinical research community, what needs to change to encourage more acceptance?

(2) Future biomedical and clinical applications for models, based on current success stories, things that couldn't be solved w/o models, time to cure – e.g. comparative effectiveness research.

(3) Future directions for model development – e.g. explicit models for predictions versus implicit or embedded models in technology, enabling technologies and infrastructures for modeling.

(4) Model validation and the availability of appropriate data.

(5) Uncertainty quantification and predictability of outcomes.

(6) The current state of model development – e.g. community-developed versus custom-made models.

(7) The current state of peer review for modeling research – e.g. changes that need to occur in the community.

This report summarizes the presentation and discussion related to modeling at each of the five biological scales under subheadings reflecting the four charge topics: Impact of

Modeling through 2009; Acceptance of Modeling as a Driving Tool; Major Challenges; and Opportunities for Modeling to Further Impact Research and Policy.

Many of the specific questions (1) through (7) as well as other common themes arose in discussion of more than one of the charge topics. We have made an effort to avoid unnecessary duplication while capturing the spirit of the presentations and exchanges.

II. Population Modeling

II.1 Impact of Modeling through 2009

Computational modeling at the population level has had an impact in large part by influencing health care coverage decisions and the development of clinical guidelines.

David Eddy described many successes of the "Archimedes" model. Archimedes is a population-level model built up from underlying physiological pathways. Its uses range from policies (e.g., guidelines, performance measurement, incentives, priority setting, strategic goals, cost and cost-effectiveness analysis, comparative effectiveness analysis) to research planning (e.g. drug portfolio analysis, clinical trial design and production). Currently, Archimedes is being used by pharmaceutical companies, voluntary health organizations (e.g., American Heart Association, American Diabetes Association, American Cancer Society), government agencies (e.g., CDC), policy making organizations (e.g., National Committee for Quality Assurance (NCQA)), insurers, and health plans.

Sylvia Plevritis indicated that statistical models have made a larger impact at the population level than mechanistic (or phenomological models). For instance, statistically-derived tools for risk assessment are routinely used in clinical practice. Examples of such tools for assessing the risk of developing breast cancer include the Gail and Claus models. Similar tools used in genetic counseling include BRCAPro. These models often rely on longitudinal data from large epidemiological studies such as the Nurse's Health Study, Carotene and Retinol Efficacy Trial (CARET), the Women's Health Initiative (WHI), Physician's Study, and extensive family registries such as the Utah Family Registry and the Family Breast Registry. Dr. Plevritis said that mechanistic models at the population level have had less impact than statistical models, but she believes the trend is changing, particularly in areas where traditional statistical approaches are not sufficient in integrating complex datasets. She has applied mechanistic models at the population level to influence cancer screening programs among high risk cohorts who are understudied with more traditional methods. Dr. Plevritis is also part of the NCI Cancer Intervention and Surveillance Network (CISNET), which is a consortium of modelers focused on cancer screening and prevention policies. Collectively, the CISNET models have targeted numerous policy decisions on screening for breast cancer, colon cancer, lung cancer and prostate cancer (cisnet.cancer.gov). One recent example is the updated U.S. Preventive Services Task Force (USPSTF) recommendations on breast cancer screening, which were derived, in part, from the results of CISNET screening models. CISNET has also worked

with the USPSTF regarding recommendations for colon cancer screening, in addition to a variety of other organizations focused on cancer control.

Bryan Grenfell described a substantial body of work demonstrating the value of modeling nonlinear epidemic dynamics for controlling infectious disease. He discussed the role of modeling in the HIV epidemic, Bovine Spongiform Encephalopathy (BSE) / Creutzfeldt-Jakob Disease (CJD), the UK Foot and Mouth epidemic of 2001, antibody resistance, assessing threats and possible control of small pox, and pandemic influenza, including the current H1N1 pandemic. He described the "anatomy" of an epidemic with a model that captures the essence of an epidemiological clock moving populations from susceptible to infected and ultimately recovered states. He gave an example of modeling the occurrence rate of measles in the UK, demonstrating how it is impacted by baby booms and schooling. Dr. Grenfell suggested that the concepts of transmission dynamics such as reproductive ratios are now influential in shaping quite a lot of research, citing the NSF/Fogarty Ecology of Infectious Disease and NIGMS/Midas Programs.

Paolo Vicini discussed the role of modeling in the pharmaceutical industry, with a focus on pharmacokinetic and pharmacodynamic (PK-PD) models. While not without challenges[5], modeling and simulation have had a significant impact on drug development[6], especially in the clinical setting. With the advent of the discipline of pharmacometrics[7], the idea of "model-based drug development[8]" has started to take hold[9]. Pharmacometrics builds on advances made possible by the pharmacokinetic (what the body does to the drug) and pharmacodynamic (what the drug does to the body) frameworks[10] commonly used to map therapeutic dose to effect through individual exposure. Model-based drug development has been facilitated by the ready availability of a variety of computer-intensive modeling tools[11] spanning the development process. Specifically relevant to the population level, nonlinear mixed effects modeling[12] allows

[5] Challenges in the transition to model-based development. Grasela TH, Fiedler-Kelly J, Walawander CA, Owen JS, Cirincione BB, Reitz KE, Ludwig EA, Passarell JA, Dement CW. AAPS J. 2005 Oct 5;7(2):E488-95.

[6] Pharmacokinetics/Pharmacodynamics and the stages of drug development: role of modeling and simulation. Chien JY, Friedrich S, Heathman MA, de Alwis DP, Sinha V. AAPS J. 2005 Oct 7;7(3):E544-59.

[7] Pharmacometrics: a multidisciplinary field to facilitate critical thinking in drug development and translational research settings. Barrett JS, Fossler MJ, Cadieu KD, Gastonguay MR. J Clin Pharmacol. 2008 May;48(5):632-49.

[8] Model-based drug development. Lalonde RL, Kowalski KG, Hutmacher MM, Ewy W, Nichols DJ, Milligan PA, Corrigan BW, Lockwood PA, Marshall SA, Benincosa LJ, Tensfeldt TG, Parivar K, Amantea M, Glue P, Koide H, Miller R. Clin Pharmacol Ther. 2007 Jul;82(1):21-32.

[9] Impact of pharmacometrics on drug approval and labeling decisions: a survey of 42 new drug applications. Bhattaram VA, Booth BP, Ramchandani RP, Beasley BN, Wang Y, Tandon V, Duan JZ, Baweja RK, Marroum PJ, Uppoor RS, Rahman NA, Sahajwalla CG, Powell JR, Mehta MU, Gobburu JV. AAPS J. 2005 Oct 7;7(3):E503-12.

[10] Pharmacokinetic/pharmacodynamic modeling in drug development. Sheiner LB, Steimer JL. Annu Rev Pharmacol Toxicol. 2000;40:67-95.

[11] Simulation of clinical trials. Holford NH, Kimko HC, Monteleone JP, Peck CC. Annu Rev Pharmacol Toxicol. 2000;40:209-34.

[12] Population pharmacokinetics/dynamics. Sheiner LB, Ludden TM. Annu Rev Pharmacol Toxicol. 1992;32:185-209.

pharmacometricians to account explicitly for between-subject variability, both when analyzing data (thus obtaining a rigorous estimate of between-patient unexplained variation) and when simulating potential clinical design scenarios.

II.2 Acceptance of Modeling as a Driving Tool

David Eddy made the point that the acceptance of modeling varies widely. At one end of the spectrum are those who take modeling very seriously, and when the model produces a counterintuitive result, they work hard to learn from the model. At the other end of the spectrum are those who will accept or reject the model based on whether or not its conclusions confirm or conflict with their prior conceptions. He said that, to some extent, the differences are due to differences in the available data. Some disease categories such as heart, diabetes, HIV, and cancer screening have received a great deal of attention in modeling, whereas other disease areas or organs have received very little. Disease categories such as cancer have excellent registries for incidence and staging. Other diseases have very poor information and fewer clinical trials. There are some disease categories, mental health is a good example, where definitions may be still too subjective or variably applied to enable the collection of quantitative data needed to build and validate unambiguous models.

Sylvia Plevritis spoke about the value of a modeling consortium, which is an idea that has been accepted at the NCI with the creation of CISNET, the Cancer Intervention and Surveillance Network[13], mentioned earlier. CISNET is a consortium of investigators from different institutions who develop and apply models to understand the impact of a variety of interventions on population trends in cancer incidence and mortality. In CISNET, multiple research teams work separately, but in parallel, to tackle common questions and compare their model results. If an important finding is robust to a variety of modeling assumptions made by the different modeling teams, then that finding is often accepted as robust and more likely to have an impact.

Bryan Grenfell indicated that modeling still has a long way to go in being accepted as a useful tool in many epidemiological studies. He claimed that the role of modeling in wildlife disease ecology is more developed than human disease epidemiology. He stated that the acceptance of modeling involves a tight integration with data and biological expertise, publication of results (ideally parallel work by several groups) and efforts to engage the people who shape policy.

Paolo Vinci indicated that there is increased interest in using computer-intensive tools in decision making[14] in a variety[15] of settings[16] in the pharmaceutical research community,

[13] http://cancer.cisnet.gov

[14] How modeling and simulation have enhanced decision making in new drug development. Miller R, Ewy W, Corrigan BW, Ouellet D, Hermann D, Kowalski KG, Lockwood P, Koup JR, Donevan S, El-Kattan A, Li CS, Werth JL, Feltner DE, Lalonde RL. J Pharmacokinet Pharmacodyn. 2005 Apr;32(2):185-97.

[15] Utilisation of pharmacokinetic-pharmacodynamic modelling and simulation in regulatory decision-making. Gobburu JV, Marroum PJ. Clin Pharmacokinet. 2001;40(12):883-92.

[16] Pharmacometrics at FDA: evolution and impact on decisions. Powell JR, Gobburu JV. Clin Pharmacol Ther. 2007 Jul;82(1):97-102.

where modeling has manifested as a new discipline of pharmacometrics. Dr. Vicini also pointed out that model validation and uncertainty quantification are fundamental components for model acceptance. Widely appreciated criteria for model testing and ultimately acceptance have been proposed, and while their practical application requires appropriate data, their theoretical and methodological basis has been well worked out in this context. He provided historical reviews that this audience may find useful by way of background[17]. At the population scale, this has facilitated the emergence of approaches to computationally-intensive design of clinical trials[18], with the intention of increasing the probability of successful outcomes.

II.3 Major Challenges

Dr. Eddy emphasized the value added by models in many other fields of human endeavor, such as energy, transportation, construction, space, manufacturing, finance and so forth. He believes that medicine is still lagging far behind. While it is true that biological complexity and variation present special problems, there exist good mathematical and statistical techniques for addressing these. In addition, the problems in healthcare are not necessarily more difficult than, for example, the problems physicists face when they try to measure the speed at which the universe is expanding or the existence of black holes. Dr. Eddy went on to state that to overcome skepticism about modeling in medicine, models need to be well formalized so that an inquirer can understand the model if s/he invests the time and effort to do so. That said, the onus is, and should be, on modelers to prove the value and validity of the model by making it transparent to the user. Toward this goal, methodological advances are needed in the area of assessing model robustness, uncertainty and approaches to model validity, so that even a non-mathematician can trust the model. Ideally, there should be a standard set of criteria for establishing the utility and validity of a model.

Dr. Plevritis argued that the broader research community needs to understand that a model does not need to be right to be useful. A model simply needs to synthesize our current knowledge and make predictions that can be tested. If the predictions are wrong, unfortunately the model is often disregarded even if in such casethe model can be used to trace the deficiencies in our knowledge. This type of analysis can guide investments for future research. Dr. Plevritis also made the point that biomedical model building requires expertise from multiple disciplines, at the very least two: the first relates to the biomedical sciences, and the second relates to technical aspects of model building, parameter estimation and model validation. These disciplines are taught in different schools at most universities, with different cultures. Modelers who work to bridge these domains have to train within both disciplines, albeit to varying degrees. They need to determine what to communicate in each discipline. They need to believe that there will be

[17] Non-linear mixed effects modeling - from methodology and software development to driving implementation in drug development science. Pillai GC, Mentré F, Steimer JL. J Pharmacokinet Pharmacodyn. 2005 Apr;32(2):161-83.

[18] Pharmacometrics and the transition to model-based development. Grasela TH, Dement CW, Kolterman OG, Fineman MS, Grasela DM, Honig P, Antal EJ, Bjornsson TD, Loh E. Clin Pharmacol Ther. 2007 Aug;82(2):137-42.

career opportunities in academia and industry that will recognize value and apply their unique training.

Dr. Vinci argued that the broad acceptance within industry has not yet been matched by an equal level of acceptance in academia. Acceptance may increase if these tools are deployed more broadly within academic research communities and especially if they are used to design experiments. In addition, there is a shortage of well-trained PK-PD and pharmacometric experts. PK-PD modeling (both at the preclinical and clinical scale) requires faculty with a wide set of skills and research areas, including informatics, engineering, statistics and pharmacology. There is some uncertainty as to where this kind of "eclectic professional" will eventually come from (the "pharmacometricians of the future"[19]). Another challenge is certainly linked to model sharing and peer review, but this is not new[20]. Modelers should also do a better job at demystifying their science and make it more accessible to biologists and clinicians, by focusing on the most biologically relevant messages at the expense of technical details.

Brian Grenfell suggested that models need to demonstrate a greater capability to integrate diverse and disparate data. In the case of epidemics, models can do more to integrate the epidemiological and evolutionary dynamics of pathogens. Dr. Grenfell spoke about how models need to capture "several voices." He argued that modeling disciplines in biomedicine are still too siloed, especially across integrative scales. Dr. Grenfell made the point that to bring on more methodological advances, theorists need to be fully integrated into the research, which requires overcoming communication barriers between the modelers and the domain experts. He believes that theorists still are not always integrated closely enough with empirical researchers (as physicists or engineers would be). This can cause obvious limitations in communicating and disseminating results, as well as limiting realism of models and opportunities for model validation if key data sets are neglected. More quantitative training of empirical workers and policy makers would obviously help here, as well as in developing a crucial feel for what's going on "under the hood" of models.

II.4 Opportunities for Modeling to Further Impact Research and Policy

Dr. Eddy believes that for modeling to advance health policy, clinical guidelines and drug development, the models need to capture more fully clinical and biological realism, including the effects of population diversity. In epidemiological studies, models need to capture and utilize more observational data. The data should be person-specific, longitudinal, and contain the important demographic, physiological, intervention, and outcome variables. Paradoxically, some believe that only when models prove to be valuable in integrating data, then investments will be made to collect, archive and maintain such data. However, Dr. Eddy claimed the key is the collection and availability of data. The first step should be to make certain that we have clear reproducible

[19] Where will the pharmacometricians of the future come from? Panel Discussion. American Conference on Pharmacometrics, October 4-7, 2009, The Grand Pequot at Foxwoods Resort, Mashantucket, CT. http://www.go-acop.org/sites/all/assets/5_Monday_PM%20of%20the%20Future.doc

[20] A retrospective on the modeling methodology forum. The review process for modeling papers: a revisionist's experiences and perspectives. DiStefano JJ 3rd. Am J Physiol. 1994 Oct;267(4 Pt 1):E485-8.

definitions of the conditions, any biomarkers, signs and symptoms, and outcomes. The second is to ensure that we have data on the incidence and progression of the conditions as a function of a variety of patient characteristics and risk factors. The third is to ensure that we have clinical trials for the effects of major treatments and that we have clinical data on the accuracies of various tests. With the existence of these data, the fourth concern becomes their availability to researchers. Currently, a great deal of data are collected but kept private (not just by pharmaceutical companies but by academic groups or the trialists themselves). Creating a pool of data (after a suitable time to enable researchers themselves to publish their initial round of reports) would greatly advance the state of modeling.

Sylvia Plevritis stated that modeling has the potential to play a greater role in "Integrative Sciences." Modeling can provide more of an integrated perspective of existing data which may lead to new insights and potentially avoid unnecessary, redundant and sometimes harmful studies and experiments. In addition, models can provide the mechanism to link between a variety of biological and clinical scales. For example, in many population models there are "submodels" (or assumptions) about disease progression. When simulating known trials with underlying disease progression models, inferences can be made about important biological questions such as the rate of cancer progression, the evidence (or lack of evidence) that cancer progresses from specific precursor lesions, and whether or not it does so in clinical stages or cellular grade. Answers to these questions can be linked to the molecular heterogeneity of the disease.

Bryan Grenfell spoke about the necessity to integrate epidemiological and evolutionary dynamics of pathogens (so called 'phylodynamics'). Progress has been made recently (particularly for influenza – and the current pandemic will generate more, hopefully). However, more progress is needed, both in general and for specific infections. Part of the difficulty is because we really need to have the right data and model structures to understand cross-scale dynamics (molecular level to in-host to population) before we can decide which simplifications are reasonable.

Paolo Vinci stated that the possibilities brought forth by computer-based design of experiments in the context of therapeutic development do not need to stop at clinical trials. He believes that there is a chance to impact basic and translational science and academic research, which requires the functional integration of many levels of expertise. In many ways, model-based design and hypothesis testing allow greater generality and flexibility compared with more traditional statistical analysis tools. The onus is on the biologist and clinician to work with a modeling expert to translate their hypothesis in terms that are amenable to computer modeling. This requires the willingness to formulate mechanisms and gather informative experimental data that would shed light on key aspects of pathways, models and especially (again most relevant to the population scale) biological variability.

III. Whole Body Modeling

III.1 Impact of Modeling through 2009

Marco Viceconti gave several examples of the impact of whole body modeling such as prediction of the risk of bone fracture in patients. Whole body models are informed with a combination of patient-specific and population data?. Predictions of this class of models include muscle forces that a patient might exert in various regions of the skeleton during different physiological activities. The goal is first to identify boundary conditions for the organ and then use tissue level models to predict regional stress and strain distributions. This information can then be used to guide rehabilitation efforts.

Another example that Dr. Viceconti illustrated, this time from cardiovascular medicine, was prediction of the risk of rupture of an aneurysm. Again, boundary conditions are predicted and applied to a local vascular model from which risk of aneurism rupture can be estimated.

At Dr. Viceconti's institution, whole body modeling is being used in clinical practice in pediatric oncology. Some osteosarcomas require reconstruction of massive portions of the skeleton; however prosthetics are not possible at this time due to high growth rates of children. Modeling is used to set a rehabilitation strategy and is especially useful when clinical experience is limited due to a low number of patient admissions for this problem.

A recent example of the impact of whole body computational modeling is the approval by the FDA of an *in silico* model of diabetes for use in preclinical testing by Juvenile Diabetes Research Foundation Artificial Pancreas Consortium sites[21]. The goal of the artificial pancreas project is to allow blood glucose levels to be maintained at normal levels without significant patient involvement. Modeling is likely to be an important element in the development and testing of algorithms that control communication between glucose pumps and monitors.

III.2 Acceptance of Modeling as a Driving Tool

Donald Bolser cited experiences from his lab on the use of computational models to study cough mechanisms to illustrate the challenges facing modelers in getting their results accepted by the experimental and clinical communities. He began his talk with a video-fluoroscopic record of disordered swallowing in a patient with amyotropic lateral sclerosis and noted how complex the behavior of swallowing was, even in health. He went on to list several issues with how modeling is perceived by experimentalists and clinicians including: the notion that biological systems are too complex to be modeled; the limited communication between experimentalists and modelers; a limited understanding of what computational modeling can accomplish; and inertia that may impair acceptance of modeling as an important tool for understanding biological systems.

Dr. Bolser presented an example of a complex model of the brainstem neural network that generates breathing and airway protective behaviors, such as cough[22].There are so

[21] www.artificialpancreas.com.

[22] Reconfiguration of the pontomedullary respiratory network: a computational modeling study with coordinated *in vivo* experiments. Rybak IA, O'Connor R, Ross A, Shevtsova NA, Nuding SC, Segers LS,

many elements and interconnections that it is difficult for experimentalists to appreciate the function of this network at the whole body scale. Dr. Bolser went on to suggest that computational modeling and simulation are necessary tools for investigating this system.

Dr. Bolser shared his experiences investigating the complex version of the brainstem network model through simulation. He indicated that his enthusiasm for computational modeling reached a high level when simulations accurately predicted a novel physiological mechanism in the regulation of cough by the nervous system.

He closed with suggestions for enhancing the enthusiasm of experimentalists and clinicians for computational modeling that centered on fostering not just exposure to modeling but hands-on experience. These suggestions included offering frequent workshops and incorporating computational modeling into training environments.

III.3 Major Challenges

Yoram Vodovotz described systems modeling of the whole body inflammatory response in mice. These efforts resulted in an approach that utilized an ensemble of models, each with different parameter values that could be queried.

Dr. Vodovotz indicated that the focus in those models has been to find bottlenecks that are inhibiting translation of basic findings into clinical applications. Challenges include incorporation of mechanisms, methods, and other knowledge that arises from multiple scales, systems biology and computational approaches[23]. Translational bottlenecks can include clinical trials, efforts to improve diagnostics and device design with models. Simulations have to be validated at the clinical level and this goal requires that the models must be informed with data that are obtained from a clinical, as opposed to an experimental, setting. This information can be incorporated through the use of simulated clinical trials and the use of "virtual" patients.

In one study, clinical trials were simulated and compared with actual trials. Simulations predicted that drug treatment was harmful to some patients, resulting in no net benefit. Modeling allowed a new simulated trial to be conducted with a larger numbers of patients in which the influence of age could also be predicted. The results indicated that increasing age results in a poorer prognosis and reduced response to the drug. The goal of this effort was to utilize data that were likely to be available from the clinical setting in order to improve a clinical outcome.

A question was asked regarding conflict of interest and checks and balances that might prevent misuse of a model. Dr. Vodovotz responded that the models related to animal work are all published and that they try to work with agent based models. Making real

---

Shannon R, Dick TE, Dunin-Barkowski WL, Orem JM, Solomon IC, Morris KF, Lindsey BG. J Neurophysiol. 2008 Jul:586(17):4265-4282.

[23] Translational systems approaches to the biology of inflammation and healing. Vodovotz Y, Constantine G, Faeder J, Mi Q, Rubin J, Bartels J, Sarkar J, Squires RH, Okonkwo DO, Gerlach J, Zamora R, Luckhart S, Ermentrout B, An G. Immunopharmacol Immunotoxicol. 2010 Feb 22, PMID: 20170421.

progress in the use of clinical models is currently difficult without the level of resources that industry has available. The proper use of models is important and careful records need to be kept of assumptions that go into a particular model and its effect on quality and efficiency of care.

Another discussant stated that models built for public use can be subject to ongoing modifications as new information is generated and asked if this represented a feedback system. This can be positive as long as this sort of application is supported by the model. In such a situation, commercial interests, and/or academic credit can be a problem for modelers as they may not have the same sort of control as traditional bench or experimental scientists over how their model is applied.

There was discussion of open source codes and how they are viewed. It was generally agreed that there can be significant advantages to open source licenses when they do not include constraints such as those of some "viral" or "copyleft" clauses. The discussion extended to the concept of open standards and formats. This process can be facilitated by setting up consistent models and data, which would assist others in developing their own models.

III.4 Opportunities for Modeling to Further Impact Research and Policy

One opportunity for whole-body modeling to impact research and policy may be in developing clinical work flows and standard operating procedures. Formal work flows are being developed for some conditions, such as osteoporosis, with the goal of being able to deal with many thousands of patients and to balance progressive invasiveness and cost related to risk of intervention.

Several questions dealt with model detail and this led to a discussion of the use of multi-scale models and how they might be linked across scales. An example of this idea is modeling biophysical properties of an aneurism and linking it to cellular models of the underlying pathobiology. Linkage of different models could be accomplished by identification of control points that are common.

Dr. Viceconti responded to a question about hybrid models with a summary of the variety of factors that could contribute to prediction of how tissue responds to loads that included mechanical properties, protein expression, adaptation, and issues regarding how to link different molecular pathways.

One questioner suggested that following individual patients that have been treated based on information provided by a model represents validation. This and another question led to an extensive discussion on validation, prediction, and problem solving.

IV. Cell-Tissue-Organ Modeling

IV.1 Impact of Modeling through 2009

Models that integrate from cell dynamics to three-dimensional heterogeneous organ scale function are now quite well established in several biomedical fields including cardiovascular, pulmonary, neuroscience, immunology, cancer biology, musculoskeletal and orthopedic research. Given that these scales bridge a key gap between bench biology and clinical medicine, we expect to see continued growth in the number and scope of translational applications of the multi-scale models in this category of clinical problems including patient-specific modeling for diagnosis and therapy planning.

Shlomo Ta'asan summarized some lessons from historical examples of mathematical models that now serve as fundamental underpinnings of modern engineering, economics and other fields. However, widespread use of successful models and theories can also lead to dangerous overconfidence and increase the risk of their misuse. The current financial crisis, for example, has been attributed to over-reliance and misuse of financial models. In biomedicine, Dr. Ta'asan pointed to the Hodgkin and Huxley model[24] of the nerve action potential and models of HIV viral replication and growth as examples of very widely used analyses. Jeff Smith noted that the Hodgkin-Huxley theory actually predicted the existence of ion channels that were not known at the time. This has ultimately led to the discovery of channelopathies.

Peter Hunter agreed that biophysically and anatomically based cell-organ level modeling has probably had its greatest impact in orthopedics[25] and cardiovascular applications[26], but he added that it is now involved in nearly all areas of research on human physiology. For example, our understanding of how the transmission of force through the skeleton during locomotion depends on the inhomogeneous and anisotropic material properties of bone is a direct consequence of the application of continuum mechanics using finite element modeling to handle the complex bone anatomy. Similarly, we would have a very poor understanding of pressure and flow in the cardiovascular system without the benefit of solving the underlying physical conservation laws (in this case expressed as the Navier-Stokes equations for Newtonian fluids).

Another example, pointed out by Dr. Hunter, is the analysis of stress and strain distributions in the myocardium throughout the cardiac cycle using large deformation elasticity theory and the inhomogeneous, anisotropic and nonlinear material properties of cardiac tissue. This has yielded a new understanding of how the ejection fraction that characterizes ventricular pump function depends specifically on tissue architecture.

There are many other examples of computational modeling informing our basic understanding of physiological mechanisms and being incorporated into clinical diagnostic or therapeutic applications including: air flow, blood flow and tissue mechanics in the lungs; the electrical wave propagation of muscular excitation in the stomach and intestines; the analysis of normal and abnormal gait with musculo-skeletal

[10] Hodgkin AL, Huxley AF (1952) A quantitative description of membrane current and its application to conduction and excitation in nerve. *J Physiol* 117:500-544.
[25] Blemker et al. Image-based musculoskeletal modeling: applications, advances, and future opportunities. J *Magn Reson Imaging* (2007) vol. 25 (2) pp. 441-51
[26] Smith et al. The Cardiac Physiome: at the heart of coupling models to measurement. *Exp Physiol* (2009) vol. 94 (5) pp. 469-71

models; tumor growth; transport processes in a wide variety of tissues; and adaptive tissue growth and remodeling in vascular aneurysms.

IV.2 Acceptance of Modeling as a Driving Tool

While biomedical research is not typically driven by predictive, mechanistic mathematical models, most hypothesis-driven research is actually guided by conceptual models. As biological data become increasingly complex, high content and high throughput, there is a rapidly growing need for systematic frameworks (models) to interpret them.

Andrew McCulloch observed that, compared with even only five years ago, it is increasingly appreciated amongst experimental biomedical scientists, at least in the fields of cardiovascular biology and neuroscience, that such complexity needs to be handled systematically.

For this reason, Dr. McCulloch and other panelists felt that the increased integration of modeling in biomedical science is inevitable. Increasingly, new structural and functional data are driving new models. There are now good examples of models driving data collection (e.g. the various activities of the Connectomics initiative[27] and the Cell Centered Database[28]). While new data may be driving the development of new models, multi-scale models based on detailed structure and physical principles remain data-limited, especially because of the lack of detailed physical properties to assign to physical structures at the cell, tissue and organ scale, structures that can now be reconstructed with unprecedented resolution.

For the same reasons, modeling is also driving technology development (both software and hardware). Some of the National Centers for Biomedical Computing[29] such as SIMBIOS[30] at Stanford and MagNET[31] at Columbia University as well as several NCRR sponsored Biomedical Technology Research Centers[32] are focused on developing and deploying high quality software for multi-scale modeling at the cell to organ levels. New techniques for imaging, segmenting, reconstructing and annotating three-dimensional anatomical structures are also being driven by the potential for large-scale models. However both software and hardware development are very expensive.

There are good examples of broader fields where physically based multi-scale modeling is driving experimental research: The legacy of Hodgkin and Huxley (1952) has created a strong tradition of ionic modeling that provides a systematic means to interpret measurements in single cells and channels, to control for experimental artifacts (e.g. buffering properties of indicators), to reconstitute data from expression systems into

[27] http://connectomes.org/
[28] http://ccdb.ucsd.edu/
[29] http://www.ncbcs.org/
[30] http://simbios.stanford.edu/
[31] http://magnet.c2b2.columbia.edu/
[32] http://www.ncrr.nih.gov/biomedical%5Ftechnology/biomedical%5Ftechnology%5Fresearch%5Fcenters/informatics%5Fresources/

whole cells *in silico*, and to integrate predictions from single channels or cells to the whole organ. Another example is in experimental biomechanics and mechanobiology where the framework of continuum mechanics has provided a means to relate measurements across scales and to infer cell, tissue and organ mechanical properties. Twenty years ago, the concept of inverse analyses to deduce tissue and organ properties from clinically accessible data in conjunction with models was treated with skepticism; today strain is a common output of clinical cardiac imaging systems, and inverse analysis (e.g. electrocardiographic imaging) is coming into clinical use. There are now excellent examples in orthopedic and vascular surgery of image driven modeling for patient diagnosis and procedure planning (e.g. the OpenSim and SimVascular projects at SIMBIOS), in addition to patient-specific modeling of ablation therapy for atrial fibrillation and cardiac resynchronization therapy for congestive heart failure. During the discussion, it was pointed out that with the rise of patient-specific modeling, it will become important to be aware that model assumptions must reflect the patient population and the differences between distinct and diverse patient populations.

Dr. Hunter pointed out that biophysically based mathematical modeling is accepted as the essential framework for quantitative analysis in all areas of biomedical engineering. Large sections of the biomedical science community, however, are still resistant to modeling. An important lesson from the US higher education system is that training for future biomedical researchers should include mathematical modeling closely integrated with experimental techniques at the undergraduate level (as occurs in the biomedical engineering programs). Discussants noted that there is interest in and enthusiasm for using multi-scale modeling as a driver for interdisciplinary research training (e.g. the HHMI-NIBIB Interfaces Graduate Training Program[33]). Institutes at the NIH are also sponsoring short courses in modeling. While training is effective for the long term, it also requires an interdisciplinary culture and standardized tools with well documented protocols and limitations.

In modern post-genomic science, a key motivation for the broad community of biomedical scientists to accept modeling will be the need for tools that can enhance understanding of biological complexity, that will allow conceptual models to become more quantitative and that will provide a means to integrate large and heterogeneous data sets with legacy data and concepts. This is driving the development of data-driven models, especially in systems biology where comprehensive models of signaling or metabolic networks can become a *defacto* data archive.

Finally, acceptance of models and their accessibility to the research community will require better tools and standards for modeling, standardized protocols and experimental instrumentation and techniques for improved data collection and analysis.

IV.3 Major Challenges

Confidence in model results relies on validation. Since it is not generally, if ever, possible to validate biomedical models in their entirety, validation involves quantitative

[33] http://www.hhmi.org/grants/institutions/nibib.html

experimental confirmation of model results. Dr. Ta'asan pointed out that this can be very costly and time-consuming both because of the large number of possible model experiments that can be conducted, the need for quantitative measurement, and the confounding effects of biological and experimental variability on the interpretation of experimental results. In addition to having confidence in model results, the biomedical community also needs to see the utility of models to experimental science. Beyond the utility of models to reproduce known phenomena, models should also be able to predict new findings or mechanisms and generate new hypotheses. But the effective iteration of theory and experiment usually requires collaboration which in turn is driving the need for interdisciplinary training of a new cadre of scientists educated in both mathematical and computational modeling and experimental biology.

In the discussion that followed Dr. Ta'asan's remarks, the current limits on interdisciplinary collaboration were discussed. Dr. Ta'asan and some others felt that the effective iteration between model and experiment is currently most efficient when multi-disciplinary investigators are working together in the same laboratory or team. This appears to be a function of the biomedical discipline or field. In some better established areas, interdisciplinary collaborations between experimental and computational labs are common, whereas in other fields they remain the exception. Some discussants felt that young people are better equipped today to be trained in both modeling and biomedical science, while others felt that the availability of better mathematical and computational modeling tools for research and education is not yet enough in the biomedical areas to ensure a sound grounding in the fundamentals. Regardless, it will be important for the success of new training programs, that institutions and peer reviewer processes create career paths that allow the graduates of new programs to succeed.

Amongst the key challenges identified by Vito Quaranta was the problem of extending multi-scale models in time to the time courses of development, disease progression and aging, e.g. developing mechanistic models of cancer disease progression. This advance will help to make mathematical modeling more useful for closing the loop between "omics" data and clinical outcomes. For these complex problems, models remain conceptual rather than quantitative, yet conceptual models have been valuable for advancing understanding. If the conceptual models of specific system components can be formalized, then the opportunity arises to integrate these to predict more complex and chronic phenotypes. One class of analysis that may be particularly effective for this kind of modeling is agent-based models[34] which can incorporate empirical rules as well as equations derived from theoretical principles. While model predictions are often wrong, they remain a valuable way for testing the consequences of common assumptions and hypotheses or investigating new hypotheses.

A continuing challenge is the high cost and time required to engineer, develop and maintain high quality software. This is important because models and their results must be reliable, intuitive to investigators in the field and provide convenient ways to make use of diverse biomedical data needed to formulate models at these scales. For example, Dr. Quaranta pointed out that cancer models rely on data from pathologists that is typically

[34] Zhang et al. Multiscale agent-based cancer modeling. *J Math Biol* (2009) vol. 58 (4-5) pp. 545-59

not coded for machines. Existing software environments like Neuron, Genesis, SimVascular and Simmune are helping to encourage the entry of biologists and biomedical scientists to modeling.

There remains a large unmet need for well-curated databases of heterogeneous, multi-scale, structural, functional, cellular, tissue and organ physiological data. Although there are large amounts of biological data at the molecular component level, multi-scale modelers working at the cell to organ scale are relatively starved for data especially at the critical meso-scales between macromolecule and single cell and between single cell and tissues. To help bring the quantitative and predictive power of the mathematical sciences into the biological sciences, and to take advantage of an engineering approach to biological materials and a more systematic approach to the knowledge management of multi-scale physiological data, it is essential that the research community adopt common standards. Some of these are being developed with the VPH-Physiome Project but much more effort is needed on data standards. Discussants also emphasized the importance of collaboration and training in these data warehousing efforts.

In Dr. Hunter's view one of the major challenges is a sociological one: how do we encourage more bioengineering, mathematics and computer science researchers to contribute to international open source projects rather than reinventing the wheel over and over with new software projects? Progress would be much faster with a more coordinated effort and there is certainly no shortage of challenging problems to keep the bioengineers, computer scientists and applied mathematicians busy without developing completely new software tools. The other major challenge mentioned is how to encourage wider uptake of modeling by biomedical scientists and, as suggested above, Dr. Hunter feels the answer to this lies in education.

IV.4 Opportunities for Modeling to Further Impact Research and Policy

Jeff Smith made the case persuasively that increasing impact of modeling at all scales is inevitable. This is largely because biological dynamics are complex and frequently defy intuition and predictability. The goal of multi-scale models of cellular dynamics is to quantify and visualize physiology in four dimensions to aid in understanding, hypothesis testing and experimental study design.

Feedback, feed-forward, stability, and instability are all concepts from control theory that computational modeling has introduced to biology. We expect to see these systems level concepts become more commonplace as biologists have access to models as tools for investigation of complex phenotypes. Indeed the field of computational neuroscience, which is truly multi-scale, has emerged because the brain computes. While this field is still in its formative stage, models are already extending to processes such as neural plasticity, and there is already a high degree of acceptance of computational neuroscience as an integral subdiscipline. A continuing influx of mathematicians, computer scientists,

bioengineers and others trained in modeling into the field of neuroscience has been accepted and widely encouraged.

Interdisciplinary programs and multi-scale modeling are influencing the way that neuroscience is being undertaken. For example, projects like the Biomedical Informatics Research network (BIRN) and the Connectomics Initiative, which aims to map entire brains at synapse resolution, will be generating terabytes of data per day, driven by the capabilities and requirements of computational models. The Blue Brain Project is a similar example for large-scale cortical models.

This is not unique to the nervous system. Spatial organization is critical to all aspects of physiology and enormously complicates systems biology. This is where multi-scale modeling is especially important. The IUPS Physiome Project[35] described by Dr. Hunter illustrates similar approaches especially for the heart, lungs, kidney, and musculoskeletal systems.

Indeed, the extent of the parallel progress within specific biomedical disciplines drew the attention of many discussants to the comparative lack of cross-fertilization between fields such as computational neuroscience, systems biology and physiome research. This probably reflects the reliance of modelers on data specific to their biological system of interest and the priority of modelers to make their research relevant to the larger community of experimental investigators. However, it also suggests an opportunity for the future to expand opportunities for multi-scale modeling groups studying different systems to interact and share experiences and technologies.

Several discussants highlighted the rapid growth and potential for applying cell-organ level multi-scale models to patient-specific diagnosis and planning of procedures, surgeries or medical therapy. Medical device design is another area where there remains very large potential for the application of multi-scale biomedical models. Since the industry is familiar with using models to analyze the performance of the devices, there is less resistance to the concept of modeling the target tissue or organ itself. Rather it is suitable software tools and data sources that are the limitation. During the special session, Dr. Tina Morrison addressed the activities and meetings of the FDA to encourage the development and validation of tools and data resources to support modeling in cardiovascular device design[36].

Dr. McCulloch suggested that one area where multi-scale modeling efforts at the cell-tissue-organ scale have so far failed to have much translational impact is that of pharmaceutical discovery. But these efforts are viewed as premature rather than conceptually faulty. This prompted considerable discussion. Paolo Vicini suggested that the best examples of multi-scale models applied to drug discovery to date are mostly in bridging the gap from preclinical studies to human studies. There are examples where costly failures of drug trials could have been predicted by computational models and others where model results alone could have been sufficient for a drug company to

[35] http://www.physiome.org.nz/
[36] http://www.fda.gov/MedicalDevices/NewsEvents/WorkshopsConferences/ucm135284.htm

modify a clinical trial design. However, this concept is only beginning to gain traction in the industry. Hence, there remains tremendous potential for computational modeling at the cell to organ scales in drug discovery though the challenges are greater than in other applications.

V. Pathways and Networks Modeling

V.1 Impact of Modeling through 2009

By definition, "Pathways and Networks" refer to the interconnectivity of various elements of a system, and these elements can range from individuals in a population, to cells and tissues in the body, to genes and proteins in a cell, or to atoms within a molecule. In addition, network connectivity can span disparate elements that exist at one or more of these spatial scales, so concepts and methods underlying pathways and networks are inherently central to integrative multi-scale modeling.

In perhaps the most general sense, the success of pathways and networks modeling is reflected in the application of graph theory and kinetic analysis to complex network topologies, whether they represent evolutionary dynamics, individuals within a population, physical and regulatory interactions between genes and proteins, or signaling mechanisms and cascades within and between cells. This approach has helped investigators to describe and interpret complex interactions in terms of functional modules and motifs such as positive and negative feedback loops, bistable switches, ultrasensitivity, etc., using much the same methods and terminology developed earlier for various branches of engineering (e.g., electrical circuit design). Particular examples of success include cell cycle dynamics controlled by feedback between cyclin-dependent kinases, engineering of gene regulatory networks into single-celled organisms (synthetic biology), synaptic and neural networks, modeling and simulation of the immune system (e.g., kinetic proofreading of immune signals), and DNA damage and repair (cancer biology). A growing complement of powerful and easily used tools and resources is now available for modeling cell regulatory dynamics.[37]

Ronald Germain illustrated how mechanistic cell systems models have elucidated fundamental scientific questions in immunology such as how the T-cell can specifically respond to such low numbers of foreign ligands when they are presented in an environment full of structurally similar self-peptide/MHC molecule complexes without mass action-driven desensitization or activation mediated by these self-ligands. Models have shown how specific key biochemical feedback loops play a critical role in achieving this discrimination.

V.2 Acceptance of Modeling as a Driving Tool

In light of the growing complexity of networks and pathways, there is no doubt that modeling and simulation are essential to understanding, whether in the context of basic research or clinical applications. As a result, the onus on modelers has shifted at least somewhat, from proof of principle or proof of usefulness, to clarity of presentation and

[37] http://stke.sciencemag.org/cgi/ul/sigtransUl;CAT_7

meaning, so that the experimental community can more easily understand, integrate, and adopt tools and approaches developed by modelers. There are also ongoing sociological issues, but these issues are bidirectional. It remains all too easy for growing modeling communities to segregate themselves from experimental communities, but by the same token, it remains all too easy for experimental communities to remain divorced from modelers. Hypothesis-driven experimental science is in fact driven by models, but often such models are composed of empirical concepts rather than quantitative relationships and parameters. It is imperative that both communities recognize the utility *and limitations* of both conceptual and quantitative models, and that the synthesis of conceptual and quantitative models is the ultimate goal and necessity.

Modeling is now essential to the study of systems with many components and complex connectivity. In certain areas this is not only well accepted, but inescapable – for instance, data analysis of high-throughput experimental investigations, multiple areas of bioinformatics (e.g., multiple sequence alignment and structure prediction), structural biology, genomics and genome-wide association studies, and others. Even outside of high-throughput studies, there is increasing emphasis and use of integrated software environments for multi-scale network modeling and physiological modeling, with considerable effort now aimed at improving user interface design for non-modelers.

The question of "what hasn't worked" still largely reflects outstanding difficulties and complexities of multi-scale biological models. How does one map networks and pathways of interactions into spatially realistic models? How does one decide when a spatially realistic model and stochastic approach are required in place of a single compartment model and the use of continuum methods? How can one efficiently explore the impact of stochastic methods applied to a spatially realistic model? Several key factors underlying all of these questions are the difficulties encountered with multiscale software development and the computational costs incurred with spatial realism and stochastic approaches. Computational cost has also been a critical limiting factor for molecular simulations over many years, with resulting disappointments in so-called "intelligent" drug design. In addition, cross-fertilization of ideas and approaches remains difficult between "data-rich" and (relatively) "data-poor" areas, e.g., systems biology *vs*. classical neuroscience and other subsets of human physiology.

V.3 Major Challenges

Joel Stiles pointed out that one considerable challenge for the future is the computational expense of complex models. This is true for multi-scale models in general, but perhaps especially so when envisioning pathways and networks mapped into spatially realistic models that employ stochastic simulation methods. A very wide range of biological timescales confronts any quantitative model, and it remains quite difficult to push from atomic-molecular-cellular times and events to organ-body-population levels and insights. The time of the computational "free lunch" is now over; single-chip clock speeds used to double every eighteen months or so, with concomitant increases in the speed of simulations, but now semiconductor physics have imposed limits that are difficult to overcome. As a result, multicore technology now accounts for continuing increases in computer speed, but adapting models and simulations to large-scale multicore systems is

often non-trivial. Specialized hardware, such as application-specific integrated circuits, may be increasingly necessary in a variety of forms, but development of any specialized hardware is very expensive and time-consuming.

Even more costly will be the software development required for multi-scale, spatially realistic stochastic models of pathways and networks, designed to address uncertainty quantification, cell-to-cell variability, and many other issues. In addition, such studies will require effective handling of massive data, e.g., immunological and cancer networks, or neural Connectome data at the level of white matter tracts and/or synaptic connectivity in gray matter. From the standpoint of both hardware and software design, there is an ever increasing need for integrated multidisciplinary teams, along with associated plans for training and professional development, people support, and academic rewards.

V.4 Opportunities for Modeling to Further Impact Research and Policy

At the level of cell, tissue, and organ physiology, modeling and simulation are already playing a role in patient-specific testing and intervention. There is no doubt that over time, the same will be true for clinical neuroscience, immunology, cancer biology, and virtually all other areas, as multi-scale networks and pathways are integrated with increasingly realistic and complete models of human systems. For many years, a "realistic" model of a "complete" cell has been viewed as an unreachable grand challenge, both because of the incredible diversity of cellular phenotypes and the concomitant computational challenges. At this point, however, it is time to revisit this challenge and consider exemplars that could become feasible within the next five years or so. For example, a model of a complete neural circuit with explicitly delineated synaptic connectivity, or a model of nuclear structures and gene regulation using hybrid methods spanning atomic, molecular, and subcellular levels of detail could be contemplated.

Dr. Hunter pointed out that this will depend on the development of open model and data standards, model/data repositories and freely available open source software. This should have a big impact on the broader community, provided the user interfaces are designed for biologists or clinicians to use. He said we should be trying to integrate modeling into all areas of biomedical science as mathematics is the language of quantitative science and there is no other way to handle the complexity of biological systems.

Dr. Germain also re-emphasized the need for multidisciplinary groups to achieve success in modeling pathways and networks and discussed the scientific and sociological impediments to success in this arena. He described the new Program in Systems Immunology and Infectious Disease Modeling (PSIIM) at the NIAID, which is structured as a team-based interactive enterprise rather than as the typical independent set of PI-driven laboratories within a department. Each team lead has been recruited for expertise in a key area of systems biology, whether wet lab or computational, including pathway analysis using high throughput RNAi methods, proteomics, genomics/epigenetics, modeling and simulation, and bioinformatics. To facilitate career development in this integrated environment, Dr. Germain worked with the NIH administration to change the tenure rules to recognize the value of each individual's contributions to publications resulting from the coordinated efforts of a team-oriented program of this nature. He also

reviewed how the PSIIM had as major goals the development and broad dissemination of computational modeling and simulation software useful to the average biologist and SOPs for the definition of pathway nodes and edges and quantification of molecular amounts and concentrations, to provide the tools and methods that would allow non-computational investigators to bring modeling into the purview of their research efforts.

Timothy Kinsella described the progress of the NCI's Integrative Cancer Biology Program in improving understanding and management of cancer through the use of predictive *in-silico* models to integrate and explore the large "omic" datasets with clinical information.

VI. Atomic and Molecular Modeling

VI.1 Impact of Modeling through 2009

There have already been many notable success stories from molecular modeling in the field of drug development. A spectacular example of drugs made possible through molecular modeling successes are inhibitors of the two viral enzymes HIV protease and reverse transcriptase (RT) protease inhibitors[38] and RT inhibitors are now regular components in HIV drug cocktail therapies. The SARS virus inhibitor was identified by computer-aided molecular design three years after the global effort coordinated by the World Health Organization (WHO) mapped the SARS genome. Potent thrombin inhibitors for blood clotting diseases were also based on molecular modeling by Merck scientists. Other drugs developed in large part by computational techniques include the glaucoma treatment Dorzolamide, the migraine medication Zolmitriatan, and the well known Sildenafil (initially developed for hypertension and then angina). Notable herbicides and fungicides were also developed by QSAR techniques.

In addition to promoting the successes of modeling, it is important for us to learn from failures, both to see how we can push in certain areas for improvement as well as to educate more broadly the limitations of a given model. Finally, there are cultural barriers and traditional training in many fields is not completely amenable to computational science.

VI.2 Acceptance of Modeling as a Driving Tool

To improve perception and reduce doubts among experimentalists, molecular modelers need to show that computation can be quantitative and useful. One concern is that models of this kind usually have too many parameters to identify with confidence. One can make numerous models that fit data, and look for the most probable models, but we often lack sufficient data to completely specify a model. Thus, we need more experimental data to help further constrain the model details (this could take the form of experimental tests of models or experimental determination of parameters to go into models).

One key is to show the predictive ability of molecular models, as opposed to their ability

[38] Marrone et al. Structure-based drug design: computational advances. *Annu Rev Pharmacol Toxicol* (1997) vol. 37 pp. 71-90

to reproduce the data with which they were constructed. Also, we expect that certain "Killer Apps" will be helpful to drive acceptance. These applications would be classic applications which share challenges with a broad range of other applications and also have a wealth of experimental data in order to test and validate the computation. This would ideally take the form of some success story in which modeling combined with experiment clearly leads to something greater than experiment alone.

The panelists also felt that there is a strong need for mathematics, physics and computation-based biomedical interdisciplinary training, both at the undergraduate and graduate levels as well as continuing professional education of peers and colleagues, for example through symposia, short courses and workshops.

At the molecular scale, there is a stronger culture of community developed codes than community developed models, in part since models generally come from detailed measurements, such as X-ray crystallography. However, there are exceptions, such as models of the nuclear pore complex, which required a collaborative effort, owing to the challenges involved. This suggests that higher order scales would need community development, but not the single-protein scale. Here "model" is typically defined as the atomic coordinates of a structural model of a given protein or protein complex; we stress that the other components of a model, such as force fields have a strong history of community development at the atomic scale.

Peer review for modeling research is a major challenge for the acceptance of modeling in general. There is a Catch-22 that modeling will not be accepted without high profile successes, yet many results will not be published in high profile places without broader acceptance of modeling in general. For example, many biology and chemistry journals return without review modeling manuscripts that do not include experimental validation.

This is a cultural difference from, for example, the physics community, in which theoretical prediction has a long history. Indeed, the best way to make "true predictions" (as opposed to predictions after the fact) is to publish the results prior to experimental validation, but this is currently not supported by many journals. Finally, high profile predictions could encourage experimentalists to test predictions, which would either gain acceptance of these models (if the prediction is successful) or improve those models (if the prediction is not successful). However, such an approach does come with a significant potential pitfall, as a controversial failure of a model could compromise wider acceptance of modeling by the experimental community.

VI.3 Major Challenges

One key area for modeling infrastructure is the ability to share models. There are well known outlets for sharing experimental structural data, e.g. the Protein Data Bank (PDB)[39], but nothing for model sharing. We are really missing this for all levels, even atomic scale models, of structures, which are usually depreciated in the PDB (i.e. not easily accessible or searchable).

[39] http://www.pdb.org/pdb/home/home.do

We need a model sharing resource to share many aspects, including structural data, force fields, and metadata. This also leads to a data provenance issue as well as highlighting the need to have a formal way to facilitate collaboration. There are existing tools for sharing code that could perhaps be used as a model. The PDB is successful because the community knows and accepts that all structures go to the PDB, and the PDB is referenced by journals. One solution would be for journals to encourage some sort of data or code sharing plan, such as the recent addition of a software archive by *PLoS Computational Biology*. Another solution would be for literature search tools like PubMed or the journals themselves to include computational models somehow into their databases, at least as a citation.

VI.4 Opportunities for Modeling to Further Impact Research and Policy

At the atomic scale, calculating the uncertainty of a given model is hard to do quantitatively, so it is often neglected. This suggests the natural question of how can one give an implied uncertainty (e.g. a general sense of the level of statistical uncertainty), at the very least. Unlike in experimental research, our community currently does not have this level of standard. There are examples of uncertainty calculations, especially in methods driven by statistical mechanics, such as Monte Carlo or molecular dynamics calculations, where one can perform error estimates.

This suggests that an important goal would be to raise the bar in the area of statistical uncertainty, especially in order to facilitate acceptance by the medical community. As a community, including the point of view of a scientific journal, we may not want to set 'rules' maybe just recommendations. Finally, we have been discussing statistical uncertainties above, but each type of model also has systematic errors, which are important to consider as well.

The most important aspect of validation is whether a model has predictive value. Clearly, no model is perfect, but the determined value is whether it is predictive and under what regime. Moreover, this means that model invalidation is important, since all models have assumptions and regimes of applicability and we need to know in which regime(s) a given model can be trusted.

VII. International Activities: The European VPH

Dr. Viceconti described an ambitious initiative sponsored by the European Commission known as the *Virtual Physiological Human* (VPH) which aims to create a methodological and technological framework that will enable collaborative investigation of the human body as a single complex system[40]. Thus, the VPH is the framework of technologies and methods that will allow the classical reductionist approach to biomedical research to be complemented by an integrative approach where the detailed understanding of single

[40] http://en.wikipedia.org/wiki/Virtual_Physiological_Human

partitions are composed in order to understand the complex behaviors that emerge by the systemic interaction of the various partitions across dimensional and temporal scales and across different organ and physiological process systems.

As defined by the VPH research roadmap[41], the VPH will make possible the delivery of a radically new approach to healthcare that includes personalized care solutions, more holistic approaches to medicine, and more preventative approaches to treatment.

The VPH vision is tightly linked to the use of computer models, and in particular multi-scale models, as modeling is the only possible way to approach the problem that integrative research poses. It is only when the reductionist knowledge is captured into a predictive model that it is possible to form a systemic understanding.

In Europe, the VPH initiative is driven by a large community of practice, within which researchers, industries and clinical experts collaborate on the development and the refinement of this vision. This community is currently incarnated in the VPH Network of Excellence (VPH NoE)[42]. For more information on the VPH initiative, interested individuals can join the Biomed Town on-line community[43]. Many shared resources are accessible via the VPH NoE website, or through the Physiome Space service[44] and the VPH News feed[45].

There is a long-standing tradition of US-EU cooperation between biomedical researchers, usually by means or bilateral collaborations self-supported or funded by specific US or member state programs. Since the beginning of the VPH initiative in 2005 the community cultivated an international dimension through intense collaborations with non-European colleagues, in particular in New Zealand, Japan and the US. In 2007, a group of large research projects from all over the world signed the World Integrative Research Initiative (WIRI) agreement, which set a common research agenda[46]. In December 2007, representatives of the European Commission and of the U.S. National Institutes of Health (NIH) witnessed the research community signing of the “Osaka Accord on Worldwide Integrative Biomedical Research Cooperation”, which included the WIRI agreement as an annex.

There have been a number of recent and forthcoming VPH conferences. Representatives of the DG Research and DG INFSO of the European Commissions and representatives of IMAG and the National Institute of Biomedical Imaging and Bioengineering (NIBIB) at NIH participated in a panel discussion entitled "Opportunities and challenges for International Cooperation” as part of the ICT-BIO 2008[47] conference in Brussels,

[41] http://www.europhysiome.org/roadmap
[42] http://www.vph-noe.eu/
[43] http://www.biomedtown.org
[44] http://www.physiomespace.com
[45] http://www.biomedtown.org/biomed_town/VPH/VPHnews/RSS
[46] http://ec.europa.eu/information_society/events/ict_bio/2008/index_en.htm
[47]

Belgium. A follow-up event (VPH 2010)[48] is scheduled for September 30th – October 1st, 2010, in Brussels.

On January 2010 a new international action, called ARGOS, was initiated. The European Commission DG RELEX (external relations) funded some policy action under the heading "Transatlantic Methods for Handling Global Challenges in the European Union and the United States", including ARGOS: "Transatlantic Observatory for Meeting Global Health Policy Challenges through ICT-Enabled Solutions". The action explicitly targets electronic health records and the VPH as primary topics. This observatory will run until June 2011, and it is expected to produce a comparative analysis of current US and EU policy approaches.

VIII. Multi-Scale Modeling in Medical Device Design

Tina Morrison described an FDA project led by Dr. Donna Lochner, Deputy Division Director, FDA, on leveraging simulation-based engineering and medical imaging technology for cardiovascular device design. The FDA's goal is to harness computer modeling and medical imaging to enhance the regulatory review process of cardiovascular devices.

Dr. Morrison stated that engineering analysis methods are needed to predict whether a proposed design will function safely. Computer simulation methods have the ability to manage and integrate data from a variety of sources (animal, preclinical and clinical). Together with data mining strategies, computer models can help to determine the most sensitive and critical design areas and improve our understanding of under-represented anatomies (e.g., pediatric patients) and physiologies.

Device manufacturers are familiar with the use of computational engineering tools to evaluate device performance. But one of the key challenges in the arena of implanted medical devices is knowing what the *in-vivo* conditions are and how to translate them into boundary conditions on a computer model of the device itself. The goal of the project is to improve understanding of device performance in all regions of the body, and in under-represented populations.

Through public workshops[49] the FDA has promoted the use of computational modeling in cardiovascular device design and evaluation. During these annual workshops, starting in 2007, clinicians, academicians and medical device manufacturers share their research efforts in advancing computer modeling and gathering medical data for the use of computer modeling. This year, the 2010 workshop will focus specifically on gathering data for computer models and nonclinical models (e.g., bench-testing boundary

[48] http://www.biomedtown.org/biomed_town/VPH/VPHnews/vph2010
[49] FDA/ NHLBI/ NSF Workshop on Computer Methods for Cardiovascular Devices March 18-19, 2008; June 1-2, 2009; June 10-11, 2010

conditions) and demonstrating how these two models work together in the evaluation process of cardiovascular devices.

The Division of Cardiovascular Devices currently has four projects underway to address barriers to implementation of computational modeling in device design and evaluation. These four projects, in particular, will shed light on the biomechanical environment of the aorta, the superficial femoral artery, mitral valve, coronary arteries and atrial septum, which the Division hopes to translate into boundary condition for finite element models or engineering bench tests.

At the FDA, cardiovascular devices undergo a wide range of rigorous engineering tests. This relates closely to validation and the availability of appropriate data. Many of the models submitted to the FDA with regulatory submissions are not validated, and if they are, they are typically only validated for one loading mode or range of deformation. This is a particular concern for the FDA and for the medical device industry.

An important goal shared with the academic community is to develop reference data, guidance documents, white papers or other tools to integrate computational modeling into regulatory evaluation of cardiovascular devices, such as a public database of boundary conditions including imaging data and computer models. The FDA envisions an open-source repository of computer models and image data and representing the diseased vasculature, bench testing parameters and boundary conditions to evaluate fatigue or other properties that relate to the safe and effective design of cardiovascular devices.

Dr. Morrison summarized the discussion by stating that the FDA recognizes that computer modeling can be used to augment bench, animal and clinical testing because no one model can demonstrate safety and effectiveness alone. Through collaborations, the FDA can harness expertise, perspective and information to enhance regulatory evaluation and device development process.

IX. Conclusions

The advantages of modeling are now clear in investigation at all scales of biological organization. At the population level, models are primarily developed to help make decisions that directly impact patients and/or populations. The decision can come in the form of a policy recommendation, a clinical guideline or even the “go-ahead” (or not) to run a new experiment. At all scales, models inform decision by integrating existing clinical or experimental data. In the words of David Eddy, “A model can be viewed as an extension of the mind,” particularly when the data is vast and diverse, and the mind can not fully process its complexity in order to pass judgment toward a decision. The model provides a framework for capturing the complexity and quantifying uncertainty.

In cases where the desired empirical data can not be obtained, because the clinical trial or biological experiment is not feasible, possible or ethical, then models can be used to extrapolate beyond the existing observations. The value of model-based extrapolations is

particularly appreciated in the field of epidemics, since the study of epidemics cannot be conducted via a trial.

Modeling as an integral part of all biological research is now inevitable – it is only a matter of time, but we should work to accelerate the process. For example, the complexity of pathways and networks as we currently appreciate them, already goes well beyond the limits of human intuition. Modeling and simulation of pathways and networks will be essential to new drug design as well as safe and effective gene therapy, especially for the complex multifactorial diseases that we face today. And finally, experience shows that modeling and simulation can be communicated effectively to non-modelers, but it takes substantial time and effort, the proper tools and training, and sustained support. Below we summarize key points that emerged from discussions of each of the four charge topics:

IX.1 Impact of Modeling through 2009

- Statistical models at the population level have had an impact on health care coverage decisions and the development of clinical guidelines, in various settings including infectious disease epidemics.
- Whole body models are increasingly being used in clinical practice, especially in metabolic diseases and orthopedics.
- Models that integrate from cell dynamics to organ scale function are now well established in several biomedical fields including cardiovascular, pulmonary, neuroscience, immunology, cancer biology, and musculoskeletal research.
- Similarly, mechanistic systems models have elucidated fundamental scientific questions in cell signaling and are becoming part of the scientific toolkit for quantitative investigation of signal transduction pathways.
- Modeling and simulation have also significantly impacted drug development, both at the molecular scale in the design of small molecules and at the clinical level through the use of pharmacokinetic and pharcodynamic models in clinical trial design.

IX.2 Acceptance of Modeling as a Driving Tool

- While hypothesis-driven biomedical research is typically guided by conceptual models, the acceptance of quantitative computational models by the biomedical science community varies widely and is not yet widespread, even as modeling has become well integrated in certain fields such as electrophysiology and neuroscience.
- The advent of pharmacometrics has increased interest in using computer-intensive tools in decision making for pharmaceutical research.
- In light of the growing complexity of networks and pathways, the increased integration of modeling in biomedical science is inevitable. Increasingly, new structural and functional data are driving new models.
- Modeling is also driving the development of new software and hardware technologies.
- The onus is on modelers to shift from proof of principle to accessibility and reproducibility so that the research community can more easily understand, validate, integrate, and adopt models and modeling tools.

- In some fields, such as molecular modeling, there is a stronger culture of community developed codes than community developed models.

IX.3 Major Challenges

- The more widespread use of modeling and simulation will depend on the capability to assimilate and integrate in teaching, research and development, diverse and disparate experimental data, biological mechanisms, computational and mathematical methods, and other knowledge that arises from multiple levels of biomedical investigation.
- Modeling will continue to require an understanding of both physics, mathematics and biology, which will continue to drive the need for greater interdisciplinary collaboration and more training programs that bridge different scientific cultures.
- Acceptance of models and their accessibility to the research community will require better tools for modeling, more standardized protocols for assessing model validity, robustness and uncertainty, and improved techniques for data collection and analysis.
- There remains a large unmet need for well-curated databases of heterogeneous, multi-scale, structural, functional cellular, tissue and organ physiological data and infrastructure that facilitates model sharing such as open model and data standards, model/data repositories and freely available open source software.
- The wide range of biological timescales makes it difficult to integrate from atomic-molecular-cellular times and events to organ-body-population levels. Similarly, spatial organization enormously complicates cell systems biology.
- The consequent computational expense of complex biomedical models is a major challenge as is the high cost and time required to engineer, develop and maintain high quality software for implementing complex models with large computational or data handling requirements.
- Modeling investigators in biomedicine are still too isolated from each other, especially across integrative scales and between medical specialties.
- Improved peer review for modeling research is still challenging in part because the means for testing and validation of models and model findings submitted for publication review is usually impractical.

IX.4 Opportunities for Modeling to Further Impact Research and Policy

- Mechanistic models will have an increasing impact at the population level and will become integrated with statistical models.
- As models capture more clinical and biological realism and account for better population diversity, they will continue to have a growing impact on health policy, clinical guidelines and drug development. Data collection and availability will likely determine the rate of this progress.
- Whole-body modeling has an opportunity to influence the development of improved clinical work flows and standard operating procedures.
- Since cell-tissue-organ scales bridge a key gap between bench biology and clinical medicine, there is a substantial opportunity for growth in the number and scope of translational applications of multi-scale models, especially patient-specific modeling for diagnosis and therapy planning.

- The goal of developing a structurally and functionally detailed complete cell model is emerging as a new grand challenge.
- There is substantial untapped opportunity for greater cross-fertilization between biomedical modeling subfields such as computational neuroscience, systems biology and physiome research.
  - An increasing impact of modeling at all scales of research is inevitable. This is largely because biological dynamics are complex and frequently defy intuition and predictability by only human means.

X. APPENDIX A: Public Commentaries

Richard Olshen, Ph.D., Professor of Biostatistics, Department of Health Research and Policy, Stanford University School of Medicine

I have written about simulation and modeling, and about the requirement that a model incorporates all sources of randomness, whether in (frequentistic) conditional distributions of observations given the model and its parameters or in (Bayesian, or random effects) distributions of parameters themselves, given a model. The latter are concerns of geneticists when they study "admixture," indeed of epidemiologists and others who attempt to draw inferences from "non-randomized experiments."

Though I have never drawn formal distinction between them, and the distinction may be too crude at best, it may be worthwhile to distinguish between qualitative and quantitative models, and between purposes for which one constructs models in the first place. In this admittedly crude formulation, it seems that there are at least two very different kinds of qualitative models. One concerns, for example, models of financial outcomes that are at least approximately verifiable from data. Examples have probability generating mechanisms that are one-sided stable laws with reflecting barriers, processes for which sample paths are monotonic except for occasional large downward "jumps." While I am hardly an expert regarding these models, I know that they have found wide application. While one can construct tests of "goodness of fit" of data to such models, and may find confirmatory evidence in data, models cited are not useful for predicting the future of particular sample paths, no matter any ergodic properties. This failure owes to the description of sample paths being at best qualitative. The Markov property is equivalent to future and past being conditionally independent given the present; but given the present, the time until a discontinuity is not only unknown but unknowable from previous data. However, even though such models may not aid prediction – in my view is the principle goal of science – they may be very useful to those who set policy, where goals are not of predicting future properties given the results of individual sample paths thus far.

There are other models that seem to me neither particularly useful for prediction nor helpful aids to setting policy. I have in mind models that entail such as "the path of a coastline is not of bounded variation." Such qualitative conclusions seem to me unverifiable from data. This is not to preclude their mathematical elegance, only to call into question their scientific usefulness. My concern applies even when particular assumptions or findings entail, for example, a precise value of the Hausdorff dimension of some set in question.

Predictive, quantitative models seem to be very different. They may be at least approximately verifiable from data, and also may allow prediction of future from past data. In particular, they may also lend understanding to non-random mechanisms that enabled prediction. One extreme example, were it possible, might be predicting some aspect of folding of a protein from its nucleotide or amino acid

sequence. Even when understanding non-random phenomena is difficult at best, models may be very useful not only for setting policy. Examples include prediction of six month survival for patients who have suffered myocardial infarctions and have survived to leave the hospital.

To summarize, in general models may be qualitative or quantitative. Models may be useful for predicting the future of a particular sample path or for setting policy. These two goals are not the same. At best, models enable quantitative understanding of non-random phenomena that govern prediction and that also are verifiable from data.

Relevant excerpt from the 2009 JASON report on Rare Events (PDF pages 35-44), http://www.fas.org/irp/agency/dod/jason/rare.pdf

"From a programmatic standpoint of funding research, the main problem with standalone research projects that aim to create new (insight) models is that they separate the model's creator from the model's user community, so they tend to face an adoption barrier. Experts are rightly skeptical of new tools developed by non-experts, especially if a model appears complex, mathematical, and highly abstracted rather than hewing closely to real-world data analysis needs. Success of an insight tool should ultimately be judged by how many experts use it and find it indispensable in their work. "Useful to experts" necessarily includes many factors that become just as important as the scientific validity of the model – issues such as software quality and usability, in the case of computer models. Therefore an important part of any research plan to develop new models is the researchers' plan for collaboration and adoption by experts. Will the tool be used and evaluated by real-world analysts? Do they find it useful? Will it spread to other analysts if it is successful?"

James Bower, Ph.D., Professor of Computational Neurobiology University of Texas

"Hypothesis-driven experimental science is in fact driven by models, but often such models are composed of empirical concepts rather than quantitative relationships and parameters."

First, there is no doubt that, by definition, true hypothesis driven science requires models. Assuming that NIH is funding hypothesis driven research, then, in principle, all research proposals should be include or be related to, and the outcomes advance some model. If this is not the case, then the usefulness and even accuracy of the data obtained is suspect.

The important question then becomes how one defines a model. As also stated

above, the vast majority of times the word "model" is used in biology, it is used almost metaphorically, to describe some more or less empirical idea about the relationship between things. Classically, this is done in the form of some kind of block and arrow diagram positing relationships at almost any level from molecular to whole systems. These block and arrow diagrams are now a standard feature at the start of biological papers and presentations, where they give the impression that there is some solid model, or understanding behind the work being presented. However, in my experience, if a model does not have a mathematical structure then the objects being modeled or their relationships remain fundamentally undefined. Rigorous mathematical definitions are important because, otherwise such models can easily be made to account for whatever new data is obtained. Further, without mathematical descriptions it is often not even clear if everyone in a field is talking about the same thing.

Unfortunately, in my view, just because a model has a mathematical basis also does not mean that it is useful as a tool to advance our understanding. The issue of models as tools is, in my view, critical to the eventual success of modeling (and a particular model). I would argue that, at present, the majority of mathematically-based models in biology are not in fact useful in advancing the field, because most are intended to demonstrate principles – rather than as tools for discovery. In other words, most models are built to explain to others (or most often to convince others) how the system works, rather than as a mechanism to discover features of the system not understood previously. Accordingly, I have said many times that if you don't know anything more about the system after you build the model, than you did before, it is of little use. Unfortunately, these are the kinds of models typically built by non-biologists (physicists, engineers, computer scientists) trained in programs to 'bring modeling to biology".

While this might seem at first a subtle distinction, in practice it is not at all. I would say that there is often one clear telltale sign of models built as tools, and that is that they are first and foremost built to be structurally and physiologically realistic without reference to ideas about function. In some large sense, what we are doing in biology is trying to figure out how biological forms reflects function. If the model being constructed abstracts the form at the outset, there is very little chance that one will learn anything one doesn't already know or believe. If, on the other hand, the model is built first and foremost as a realistic representation of the system in question, then through an iterative process involving the interplay between modeling and experimental studies, it is possible to discover relationships you didn't know existed before you started building the model.

What this means is that simply replicating data (often described as a form of model 'prediction), is not a good enough measure of a model, and for certain, overall simplicity is also likely not a virtue. By analogy, Ptolomy's model of the solar system replicated and predicted the position of the planets extremely well. The model was also 'simple, being constructed from well known mathematical objects (circles). But, because it was not in fact anatomically realistic, there was no way to

understand the role of gravitational attraction in governing the structure of the solar system. This required the construction of a "realistic model" (in this case by Newton of the moon's movement around the earth), out of which the inverse square relationship fell out. In my view, biology as a science is in a very analogous circumstance as physics in the 16th century, and will need to travel through a similar developmental path to move forward.

So, in summary and at the most abstract level, I would assert:

1) for progress to be made, all scientific studies need to be formulated and interpreted in the context of models
2) those models must be formulated mathematically so that the components and their interactions are defined.
3) Not all mathematical models are useful – those with the most value need to be constructed based on physiological and anatomical relationships, NOT constructed to demonstrate what one already believes to be the case.
4) The majority of "models" built in biology aren't really models but illustrated stories. The majority of mathematically-based models are built to demonstrate assumed principles rather than discover new principles.
5) This must change if we have any chance of understanding biological systems, including the nervous system
6) Educating future generations of biological scientists will be essential in making this transition.
7) Current leaders in the field – most of whom mostly tell stories, will resist this change as long a possible.

And to conclude, again quoting from the section on network modeling:

"Modeling as an integral part of all biological research is inescapable – it's only a matter of time but we need to accelerate the process."

On this I concur, but how much time is another question.

"And finally, modeling and simulation can be communicated effectively to non-modelers, but it takes time, effort, training, and support"

For the reasons stated above, success for modeling in biology will actually be dependent, in my opinion, on training biologists to build models, not on modelers 'effectively communicating" to non-model building biologists. As in the 16th century in physics, at the start of this effort in biology, the modeler and the experimentalist, must be one in the same. To motivate this change, NIH should require all grant proposals be written with respect to quantitative models, and all published scientific experimental papers should require references to models as well.

XI. APPENDIX B: Meeting Agenda

Tuesday December 15, 2009:

8:00	Welcome - Grace Peng, IMAG Chair
8:20	**Donald Lindberg**, Director, National Library of Medicine, NIH
8:30-10:00	**Population scale:** epidemiology, risk and surveillance models, pharmacokinetic and pharmacodynamic models Participants: David Eddy, Bryan Grenfell, Sylvia Plevritis (Chair), Paolo Vicini Moderators: Patty Mabry (NIH-OBSSR), Timothy Gondre-Lewis (NIAID)
10:00-10:15	Break
10:15	**Richard Nakamura**, Deputy Director, National Institute of Mental Health and Director, Division of Intramural Research Programs NIMH, NIH
10:25-11:55	**Whole-Body scale:** behavior and control systems Participants: Don Bolser (Chair), Marco Viceconti, Yoram Vodovotz Moderators: Peter Lyster (NIH-NIGMS), Nancy Shinowara (NIH-NICHD)
11:55-12:55	Lunch on your own
1:00	**Edward Seidel**, Director, Office of Cyberinfrastructure, NSF (Presented by Abani Patra, IMAG, NSF)
1:10-2:40	**Cell-Tissue-Organ scale:** interactions between all scales Participants: Andrew McCulloch (Chair), Jeff Smith, Shlomo Ta'asan, Vito Quaranta Moderator: Tom Russell (NSF-OIA)
2:40-3:00	Break
3:00	**Jeremy Berg**, Director, National Institute of General Medical Sciences, NIH
3:10-4:40	**Pathways and Networks scale:** molecular interactions, systems biology Participants: Ron Germain, Tim Kinsella, Joel Stiles (Chair) Moderator: Jennie Larkin (NIH-NHLBI)

4:40-5:30	Open Discussion

Wednesday December 16, 2009:

8:00	Welcome Back - Grace Peng, IMAG Chair
8:10	**International Impact of Modeling** - Marco Viceconti, Virtual Physiological Human Network of Excellence
8:20	**David Thomassen**, Chief Scientist, Office of Biological and Environmental Research, Office of Science, DOE
8:30-10:00	**Atomic and Molecular scale:** protein structure interactions Participants: Vijay Pande (Chair), Abby Parrill, Linda Petzold, Tamar Schlick Moderator: Susan Gregurick (DOE-BER)
10:00-10:10	**Leveraging the Simulation-Based Engineering and Medical Imaging Technology Revolutions for Cardiovascular Devices** - Tina Morrison, Food and Drug Administration
10:10-10:30	Break
10:30-11:00	**How Multiscale Modeling can Impact Biomedical and Clinical Research** - Peter Hunter, IUPS Physiome Project
11:00-11:15	Population Scale Summary
11:15-11:30	Whole-Body Scale Summary
11:30-11:45	Cell-Tissue-Organ Scale Summary
11:45-12:00	Pathways and Networks Scale Summary
12:00-12:15	Atomic and Molecular Scale Summary
12:15-12:45	Open Discussion
12:45	Adjourn

XII. APPENDIX C: Working Group Participants

Berg, Jeremy, Ph.D.
National Institute of General Medical Sciences
National Institutes of Health
bergj@mail.nih.gov

Bolser, Donald, Ph.D.
Department of Physiological Sciences
University of Florida College of Veterinary Medicine
bolserd@mail.vetmed.ufl.edu

Eddy, David, M.D., Ph.D.
Archimedes, Inc.
david.eddy@archimedesmodel.com

Germain, Ron, M.D., Ph.D.
National Institute of Allergy and Infectious Diseases
National Institutes of Health
rgermain@mail.nih.gov

Grenfell, Bryan, Ph.D.
Ecology and Evolutionary Biology
Woodrow Wilson School
Princeton University
Grenfell@princeton.edu

Hunter, Peter, Ph. D.
Auckland Bioengineering Institute
University of Auckland, New Zealand
p.hunter@auckland.ac.nz

Kinsella, Timothy, M.D.
Radiation Oncology
Case Western Reserve University
Timothy.kinsella@case.edu

Lindberg, Don, M.D.
National Library of Medicine
National Institutes of Health
lindberg@nlm.nih.gov

McCulloch, Andrew, Ph.D.
Department of Bioengineering
University of California San Diego
amcculloch@ucsd.edu

Nakamura, Richard, Ph.D.
National Institute of Mental Health
National Institutes of Health
rnakamur@mail.nih.gov

Pande, Vijay, Ph.D.
Department of Chemistry
Stanford University
pande@stanford.edu

Parrill Abby, Ph.D.
Department of Chemisty
University of Memphis
aparrill@memphis.edu

Petzold, Linda, Ph.D.
Department of Mechanical Engineering
Department of Computer Science
University of California Santa Barbara
petzold@engineering.ucsb.edu

Plevritis, Sylvia, Ph.D.
Radiological Sciences Laboratory
Stanford University
sylvia.plevritis@stanford.edu

Quaranta, Vito, M.D.
Cancer Systems Biology
Vanderbilt University Medical Center
vito.quaranta@vanderbilt.edu

Schlick, Tamar, Ph.D.
Computational Biology/

Chemistry/Biomathematics
The New York University
schlick@nyu.edu

Seidel, Edward, Ph.D.
Office of Cyberinfrastructure
National Science Foundation
hseidel@nsf.gov

Smith, Jeff, Ph.D.
National Institute of Neurological Disorders and Stroke
National Institutes of Health
js230u@nih.gov

Stiles, Joel, M.D., Ph.D.
Pittsburgh Supercomputing Center
Carnegie Mellon University
stiles@psc.edu

Ta'asan, Shlomo, Ph.D.
Department of Mathematical Sciences
Carnegie Mellon University
shlomo@andrew.cmu.edu

Thomassen, David, Ph.D.
Office of Biological and Environmental Research, Office of Science
Department of Energy
david.thomassen@science.doe.gov

Viceconti, Marco, Ph.D.
Laboratorio di Tecnologia Medica
Instituti Ortopedici Rizzoli
viceconti@tecno.ior.it

Vodovotz, Yoram, Ph.D.
Department of Immunology
University of Pittsburgh
vodovotzy@msx.upmc.edu

XIII. APPENDIX D: All Identified Participants

Last Name	First Name	Email	Organization
Anderson	Vernon	andersonve@nigms.nih.gov	NIGMS/NIH
Bakic	Predrag	Predrag.Bakic@uphs.upenn.edu	University of Pennsylvania
Bassingthwaighte	Jim	jbb@bioeng.washington.edu	University of Washington
Beaumont	Jaques	beaumont@binghamton.edu	Thomas Watson
Berg	Jeremy	bergj@mail.nih.gov	NIGMS/NIH
Bolser	Don	bolser@vetmed.ufl.edu	University of Florida
Brazhnik	Paul	brazhnikp@nigms.nih.gov	IMAG-NIH
Brazhnik	Olga	brazhnik@mail.nih.gov	IMAG-NIH
Bruning	Gert	g.bruning@philips.com	Philips Electronics
Cavelier	German	gcavelier@mail.nih.gov	IMAG-NIH
Ciesielski	Chris	Kcies@math.wvu.edu	West Virginia University
Corn	Milton	corn@nlm.nih.gov	NLM/NIH
Couch	Jennifer	Jennifer.Couch@nih.hhs.gov	IMAG-NIH
Deckhut	Alison	adeckhut@mail.nih.gov	IMAG-NIH
Duerk	Jeffrey	jeffrey.duerk@case.edu	Case Western Reserve University
Eddy	David	david.eddy@archimedesmodel.com	Archimedes Inc.
Germain	Ron	rgermain@mail.nih.gov	NIAID/NIH
Gondre-Lewis	Tim	tglewis@mail.nih.gov	IMAG-NIH
Good	Peter	goodp@mail.nih.gov	IMAG-NIH
Grefenstette	John	gref@pitt.edu	University of Pittsburgh
Gregurick	Susan	Susan.Gregurick@science.doe.gov	IMAG-NIH
Grenfell	Bryan	grenfell@princeton.edu	Princeton University/Penn State University
Herr	Bruce	bh2@chalklabs.net	Chalklabs
Hunter	Peter	p.hunter@auckland.ac.nz	Auckland Bioengineering Institute
Kelley	Chris	kelleyc@mail.nih.gov	NIBIB/NIH
Kinsella	Tim	timothy.kinsella@case.edu	Case Western Reserve University
Kratz	Mary	mkratz@umich.edu	IMAG-DOD
Krotoski	Danuta	krotoskd@mail.nih.gov	NICHD/NIH
Larkin	Jennie	larkinj2@mail.nih.gov	IMAG-NIH
LaRowe	Gavin	gl@chalklabs.net	Chalklabs
Lee	Bruce	byl1@pitt.edu	University of Pittsburgh
Lindberg	Don	lindberg@mail.nih.gov	NLM/NIH
Liu	Aiyi	liua@mail.nih.gov	NIH
Liu	Yuan	liuyuan@mail.nih.gov	IMAG-NIH
Lyster	Peter	lys@mail.nih.gov	IMAG-NIH
Mabry	Patty	mabryp@od.nih.gov	IMAG-NIH
Marcus	Stephen	marcusst@mail.nih.gov	IMAG-NIH
McCulloch	Andrew	amcculloch@ucsd.edu	University of California, San Diego
Nakamura	Richard	rnakamur@exchange.nih.gov	NIMH/NIH
Paluh	Janet	paluhj@rpi.edu	Rensselaer Polytechnic
Pande	Vijay	pande@stanford.edu	Stanford University
Parrill	Abby	aparrill@memphis.edu	University of Memphis
Patra	Abani	apatra@nsf.gov	IMAG-NSF

Peng	Grace	pengg@mail.nih.gov	IMAG-NIH
Petzold	Linda	petzold@cs.ucsb.edu	University of California, Santa Barbara
Plevritis	Sylvia	sylvia.plevritis@stanford.edu	Stanford University
Przekwas	Andrzej	ajp@cfdrc.com	CFD Research Corporation
Quaranta	Vito	vito.quaranta@vanderbilt.edu	Vanderbilt University
Russell	Thomas	trussell@nsf.gov	IMAG-NSF
Sabourin	Stephanie	sabourinsm@mail.nih.gov	IMAG-NIH
Sauro	Herbert	hsauro@u.washington.edu	University of Washington
Schlick	Tamar	ts1@haifa.biomath.nyu.edu	New York University
Schwartz	Suzanne	sus2003@med.cornell.edu	Cornell University
Seidel	Ed		NSF
Shinowara	Nancy	shinowan@csr.nih.gov	IMAG-NIH
Shumei	Sun	ssun@vcu.edu	Virginia Commonwealth University
Sidles	John	sidles@u.washington.edu	University of Washington
Smith	Jeff	jsmith@helix.nih.gov	NINDS/NIH
Spector	Alexander	aspector@jhu.edu	Johns Hopkins University
Stiles	Joel	stiles@psc.edu	Pittsburgh Supercomputing Center
Ta'asan	Shlomo	taasan@gmail.com	Carneige Mellon University
Tao	Feng	ftao_2000@yahoo.com	Omic Biosystems Inc
Thiagarajan	Kristina	thiagarajakd@ornl.gov	Oak Ridge National Laboratory
Thomassen	David	David.Thomassen@science.doe.gov	DOE
Tina	Morrison	tina.morrison@fda.hhs.gov	IMAG-FDA
Tseng	Hung	tsengh@mail.nih.gov	NIAMS/NIH
Udupa	Jayaram	jay@mail.med.upenn.edu	University of Pennsylvania
Viceconti	Marco	viceconti@techno.ior.it	Laboratorio de Tecnologia Medica
Vicini	Paolo	Paolo.Vicini@pfizer.com	Pfizer
Vodovotz	Yoram	vodovotzy@upmc.edu	University of Pittsburgh Medical Center
Ward	Richard	wardrc1@ornl.gov	Oak Ridge National Laboratory
Xia	Ashley	axia@mail.nih.gov	NIAID/NIH
Yang	Liming	lyang@mail.nih.gov	NCRR/NIH
Yao	Alison	yaoal@mail.nih.gov	IMAG-NIH
Ye	Jane	Jane.Ye@nih.hhs.gov	IMAG-NIH

XIV. APPENDIX E: Workshop Planning Committee

Anderson, James, Ph.D.
National Institute of General Medical Sciences
National Institutes of Health
andersj@mail.nih.gov

Brazhnik, Paul, Ph.D.
National Institute of General Medical Sciences
National Institutes of Health
brazhnikp@nigms.nih.gov

Cavelier, German, Ph.D.
National Institute of Mental Health
National Institutes of Health
gcavelier@mail.nih.gov

Couch, Jennifer, Ph.D.
National Cancer Institute
National Institutes of Health
Jennifer.Couch@nih.hhs.gov

Deckhut, Alison, Ph.D.
National Institute of Allergy and Infectious Diseases
National Institutes of Health
adeckhut@mail.nih.gov

Gondre-Lewis, Timothy, Ph. D.
National Institute of Allergy and Infectious Diseases
National Institutes of Health
tglewis@mail.nih.gov

Gregurick, Susan, M.D.
Office of Biological and Environmental Research
Department of Energy
Susan.Gregurick@science.doe.gov

Larkin, Jennie, Ph.D.
National Heart, Lung, and Blood Institute
National Institutes of Health
larkinj2@mail.nih.gov

Liu, Yuan, Ph.D.
National Institute of Neurological Disorders and Stroke
National Institutes of Health
liuyuan@ninds.nih.gov

Lyster, Peter, Ph.D.
National Institute of General Medical Sciences
National Institutes of Health
lys@mail.nih.gov

Mabry, Patty, Ph.D.
Office of Behavioral and Social Sciences Research
National Institutes of Health
mabryp@od.nih.gov

Patra, Abani, Ph.D.
Office of Cyberinfrastucture
National Science Foundation
apatra@nsf.gov

Peng, Grace, Ph.D.
National Institute of Biomedical Imaging and Bioengineering
National Institutes of Health
penggr@mail.nih.gov

Remington, Karin, Ph.D.
National Institute of General Medical Sciences
National Institutes of Health
remingka@nigms.nih.gov

Russell, Thomas, Ph.D.
Office of Integrative Activities
National Science Foundation
trussell@nsf.gov

Sabourin, Stephanie
National Institute of Biomedical
Imaging and Bioengineering
National Institutes of Health
sabourinsm@mail.nih.gov

Shinowara, Nancy, Ph.D.
National Institute of Child Health and
Human Development
National Institutes of Health
shinowan@mal.nih.gov

This meeting was sponsored by NIH-NIBIB, NSF-OCI, NIH-NIGMS, NIH-OBSSR, DOE-BER, and NIH-NIMH.

IMAG is very grateful to Dr. Andrew Mcculloch and to Dr. Sylvia Plevritis, Dr. Donald Bolser, Dr. Joel Stiles, and Dr. Vijay Pande for making this report possible.