Modeling Immune Response to Influenza A Infection by Integrating Quantitative/Computational Technologies

Hulin Wu, Ph.D., Professor

Chief, Division of Biomedical Modeling and Informatics
Director, Center for Biodefense Immune Modeling

Department of Biostatistics and Computational Biology
University of Rochester School of Medicine and Dentistry

hwu@bst.rochester.edu
Outline

• Introduction
• Mathematical Models
• Statistical Methods for Model Identification
• User-Friendly Computer Software Development
• Discussion and Conclusion
Introduction: Modeling Biological Processes

- A multi-disciplinary business
- No model is correct, but some models are useful
- Different purposes need different models
- Different models need different experimental data
- A model that does not fit the data well is not a good model, but a model that fits the data well is not necessarily a correct model
- Develop a useful model for your purpose
Interdisciplinary Biometric Sciences

Integrating quantitative/computational methods and techniques for biomedical research:

- Biomathematics: Mathematical Biology, Theoretical Biology
- Biostatistics
- Biocomputing: Computational Biology, Bioinformatics, Biomedical Informatics, Health Informatics
- Bioengineering or Biophysics
Interdisciplinary Interplay

Biomathematics

Biostatistics

Bioengineering

Biocomputing

Biomedical Research Problems
Division of Biomedical Modeling and Informatics

Founded at the Dept of Biostatistics & Computational Biology, University of Rochester in 2004

- **Biomedical Problems:** HIV/AIDS Treatment, influenza virus infection, and immunology

- **Biomathematics/Bioengineering/Biophysics:** 2 faculty, 1 postdoc

- **Biostatistics:** 4 faculty, 1 postdoc, 4 PhD students

- **Biocomputing/Bioinformatics:** 3 faculty, 6 software developers, 1 postdoc
Funded Projects

- AIDS Clinical Trial Modeling and Simulations (NIH R01 AI 055290, PI: Dr. Hulin Wu)
- Nonparametric Modeling of Long-Term HIV/Cell Dynamics (NIH R01 AI 52765, PI: Dr. Hulin Wu)
- Center for Biodefense Immune Modeling (NIH N01 AI 50020, PI: Dr. Hulin Wu, Bioinformatics Core Director: Dr. Ma, Biocomputing Core Director: Dr. Warnes, Biostatistics Core Director: Dr. Liang)
- The Biomedical Informatics (BI) Key Function, the University of Rochester’s Clinical and Translational Science Institute (CTSI) (NIH UL1 RR024160, Co-PIs: Drs. Dongwen Wang, David Krusch and Hulin Wu)
- Centers of Excellence for Influenza Research (Informatics Core Director: Dr. Jingming Ma; Biostatistics Core Director: Dr. Hulin Wu)
• Center grant: Immune Function and Biodefense in Children, Elderly and Immunocompromised Populations, (NIH N01 AI 50029, Informatics Core Director: Dr. Jingming Ma)

• Analysis of AIDS Data by Using Semiparametrical Models (NIH R01 AI 62247, PI: Dr. Hua Liang)

• Generalized Varying-Coefficient Partially Linear Models (NIH R01 AI 59773, PI: Dr. Hua Liang)

• Other collaboration grants
Center for Biodefense Immune Modeling
University of Rochester

Objectives:

- Develop mathematical/computational models to simulate immune responses to influenza A virus
- Design and conduct experiments to identify, measure and validate the immune response models
- Develop statistical methods for model identifications and predictions
- Develop user-friendly database and software tools for modeling and simulating immune responses to influenza A virus
- Develop an education program to foster the next generation of researchers with multi-disciplinary expertise in mathematical modeling and immunology
- Investigate the feasibility to extend the models for other pathogens
Quantitative Sciences Components

- Mathematical Models
- Statistical Methods for Model Identification
- Computer Software Tools
Forward Problems: Simulation & prediction

Initial models

Biological hypotheses/mechanisms

Feedback information

Software/tools development

Working models

Design new experiments to collect data

Any new/alternative model structures need testing?

Yes

Any model parameters need estimate?

Yes

Update working models

Confirmed new model structure and parameters

Sensitivity analyses & simulation validations

Data analysis, Hypothesis testing & estimation

Database: data management

Deliver to/share with users

Data analysis, Hypothesis testing & estimation

Design new experiments to collect data

Any new/alternative model structures need testing?

Yes

Any model parameters need estimate?

Yes

Update working models

Confirmed new model structure and parameters

Sensitivity analyses & simulation validations

Database: data management

Deliver to/share with users

Forward Problems: Simulation & prediction

Initial models

Biological hypotheses/mechanisms

Feedback information
Mechanism-Based Mathematical Models

- Close interactions between biomathematicians and influenza immunologists/virologists
- What level of details should we model?
  - Purpose of modeling
  - What can we measure in the lab?
  - Complexity and identifiability of the model from experimental data
Two Compartment Flu Model

Airway/Lung

Infected Epithelial cell (E_p*)

Uninfected Epithelial cell (E_p)

Influenza virus (V)

Immature dendritic cell (D)

Virus+ dendritic cell (D*)

Antiviral antibody (A)

Effector CD8 T cell (T_E)

Naive CD8 T cell (T_n)

Mature Virus+ dendritic cell (D_m)

Short-lived Plasma cell (P_s)

Activated B cell (B_a)

Naive B cell (B)

Effector CD4 T cell (H_E)

Naive CD4 T cell (H_n)

Long-lived Plasma cell (P_L)

Spleen/Lymph Node

Airway/Lung

(1) (2) (3) (4) (5) (6) (7) (8) (9) (10) (11) (12) (13) (14) (15)
Mathematical Models: Influenza Virus Infection

Mathematical Models: Airway/Lung Compartment

\[
\begin{align*}
\frac{d}{dt} E_p &= \delta_E (E_0 - E_p) - \beta_E E_p V, \\
\frac{d}{dt} E_p^* &= \beta_E E_p V - k_E E_p^* T_E (t - \tau_T) - \delta_{E^*} E_p^*, \\
\frac{d}{dt} V &= \pi_V E_p^* - c_V V - k_V V A(t), \\
\frac{d}{dt} D &= \delta_D (D_0 - D) - \beta_D D V, \\
\frac{d}{dt} D^* &= \beta_D D V - \delta_{D^*} D^* - \gamma_{D^*} D^*
\end{align*}
\]

\(E_p\): uninfected epithelial cells
\(E_p^*\): infected epithelial cells
\(V\): free influenza virus
\(D\): immature dendritic cells
\(D^*\): virus-loaded dendritic cells
Influenza Virus Infection

Mathematical Models: Spleen/Lymph Node Compartment

\[ \frac{d}{dt} D_M = k_D D^* (t - \tau_D) - \delta_{D_M} D_M, \]
\[ \frac{d}{dt} H_N = \delta_{H_N} (H_{N0} - H_N) - \pi_H(D_M) H_N, \]
\[ \frac{d}{dt} H_E = \pi_H(D_M) H_N + \rho_{H_E}(D_M) H_E - \delta_{H_E}(D_M) H_E, \]
\[ \frac{d}{dt} T_N = \delta_{T_N}(T_{N0} - T_N) - \pi_T(D_M) T_N, \]
\[ \frac{d}{dt} T_E = \pi_T(D_M) T_N + \rho_{T_E}(D_M) T_E - \delta_{T_E}(D_M) T_E, \]
\[ \frac{d}{dt} B = \delta_B(B_0 - B) - \pi_B(D_M) B, \]
\[ \frac{d}{dt} B_A = \pi_B(D_M) B + \rho_{B_A}(D_M + h H_E) B_A - \delta_{B_A} B_A - \pi_s B_A - \pi_l H_E B_A, \]
\[ \frac{d}{dt} P_s = \pi_s B_A - \delta_s P_s, \]
\[ \frac{d}{dt} P_L = \pi_l H_E B_A - \delta_l P_L, \]
\[ \frac{d}{dt} A = \pi_{AS} P_s + \pi_{AL} P_L - \delta_A A. \]
Forward Problem

- X31 infection: virus cleared within 10 days since infection

- CD8+ T cells in airway/lung is approximated as \( \gamma T_E(t - \tau_T) \) where \( T_E(t) \) is the level of effector CD8+ T cells in spleen/lymph node

- respiratory DC migration stops around 2 days [Legge Immunity 2003]
- DC kinetics in lymph node [Belz PNAS 2004] compared with the model solution

[ Data provided by Topham, Belz, Heath, Randall]
Experimental Design and Statistical Methods

- Close interactions among statisticians, biomathematicians and influenza experimentalists
- Mathematical identifiability analysis
- Statistical identifiability analysis
- Statistical methods for parameter estimation and model validation
- Simulation-guided experimental design: what to measure and when to measure?
- Model fitting: experimental data
- Model validation and prediction
CBIM Topham Lab Haemaglutination Assay
X31 Viral Titers in Lung (log10 EID50/mL)
Pointwise Means and Individual Observations*

Data as of 4082008
* = Num of Obs at each titer value
CBIM Zand Lab Elisa Results (0–35 Wks)
Serum X31–Antibody pg/mL
Pointwise Means +/- 2 Std Err

Retrieved from DataTrans on Apr 7 2008
CBIM Zand Lab Elisa Results (0–16 Wks)
Serum X31–Antibody pg/mL
Pointwise Means +/- 2 Std Err

Retrieved from DataTrans on Apr 7 2008
CBIM Zand Lab Elisa Results (0−35 Wks)
X31–Antibody pg/mL
Individual Replicate Observations and Pointwise Means

Retrieved from DataTrans on Apr 7 2008
CBIM Zand Lab Elisa Results (0–16 Wks)
X31–Antibody  pg/mL
Individual Replicate Observations and Pointwise Means

IGG

Retrieved from DataTrans on Apr 7 2008
CBIM Zand Lab Elisa Results (0–35 Wks)
X31–Antibody pg/mL
Individual Replicate Observations and Pointwise Means

IGM

Retrieved from DataTrans on Apr 7 2008
CBIM Zand Lab Elisa Results (0−16 Wks)
X31−Antibody pg/mL
Individual Replicate Observations and Pointwise Means

Retrieved from DataTrans on Apr 7 2008
Inverse Problems for ODE Models

- Mathematical identifiability problem: all parameters theoretically identifiable?

- Statistical identifiability problem: all parameters practically identifiable with presence of measurement errors

- Statistical estimation methods: how to identify (estimate) all identifiable parameters?
Mathematical Identifiability Problem

Under the ideal condition of noise-free observations, can all unknown parameters of the postulated model be uniquely estimated from the experimental data?
Nonlinear ODE models: difficult

- **Power series expansion**
- **Similarity transformation**
- **Implicit function theorem**
- **Differential algebra methods**
Statistical Identifiability Problem

- Monte Carlo simulation approach
- Correlation matrix method
Statistical Estimation Methods for Nonlinear ODE Models

- The nonlinear least squares (NLS) principle:
  - numerically solve the ODE
  - global optimization method: necessary
  - differential evolution algorithm or scatter search method

- Bayesian methods
  - use prior to solve the identifiability problem
  - good for both cross-sectional data and longitudinal data:
    regression model and hierarchical model
  - computation: expensive

- Two-step smoothing approaches
  - avoid numerically solving the ODE
  - easy to implement
  - theoretical properties need to be established
The NLS Method

\[ \frac{d}{dt} X(t) = F[X(t), \theta], \quad X(0) = X_0 \]  \quad (3)

\[ Y(t_i) = H[X(t_i), \beta] + e(t_i), \]  \quad (4)

\[ e(t_i) \sim (0, \sigma^2 I), \quad i = 1, \ldots, n \]

- The NLS method: minimizing

\[ \sum_{i=1}^{n} \{Y(t_i) - H[X(t_i, \theta), \beta]\}^T \{Y(t_i) - H[X(t_i, \theta), \beta]\}, \]

where \( X(t_i) \) evaluated numerically from Eq (3).

How to deal with local minima and non-convergence? Gradient methods vs. global optimization methods.
The NLS Method: How to deal with local minima and non-convergence?

- Gradient methods: not work
Bayesian Methods: Longitudinal Dynamic Model


**Observation Model:**

\[ y_i = V_i(\theta_i) + e_i, \quad [e_i|\sigma^2, \theta_i] \sim \mathcal{N}(0, \sigma^2 I_{m_i}) \]

- \( y_i = (y_{i1}(t_1), \cdots, y_{im_i}(t_{m_i}))^T \): measurements of \( V \)
- \( V_i(\theta_i) \): solutions to the differential equations
- \( \theta_i = (\gamma_i(t), \lambda_i, \rho_i, k_i, \delta_i, N_i, C_i) \): parameters for the \( i \)th subject
- \( e_i = (e_i(t_1), \cdots, e_i(t_{m_i}))^T \): measurement errors

This can be written into a **nonlinear mixed-effects model**.
A Two-Step Smoothing Approach

\[ \frac{d}{dt}X(t) = F[X(t), \theta], \quad X(0) = X_0 \]  
\[ Y(t_i) = X(t_i) + e(t_i), \quad e(t_i) \sim (0, \sigma^2 I), \]  

A Two-Step Estimation Approach:

- **Step 1.** Fit model (7) to obtain the estimates of \( X(t) \) and \( X'(t) = dX(t)/dt \): \( \hat{X}(t) \) and \( \hat{X}'(t) \) using a nonparametric smoothing method.

- **Step 2.** Substitute the estimates \( \hat{X}(t_i) \) and \( \hat{X}'(t_i) \) in the dynamic equation (6) to obtain a regression model:

\[ \hat{X}'(t_i) = F[\hat{X}(t_i), \theta(t_i)] + e_2(t_i). \]  

Then fit the above nonlinear regression model to estimate \( \theta \) from (8).


User-Friendly Software Development

- Friendly to quantitative scientists: mathematicians, statisticians and computer scientists
- Friendly to biological scientists and biomedical investigators
- Two software tools under development
  - Immunological data sharing and data management system: DataTrans
  - Differential equation model simulation and estimation for immunology and infectious disease modeling: DEDiscover
A Comprehensive Data Management System for Immunological Research

An immunological research always involves different types of data and information, like biological samples of participating subjects, experiment raw data from immunological laboratories, processed data by computation tools, and various documents for studies. It also needs an efficient way to handle tremendous phenotype data generated by immunological assays which include flow cytometry, enzyme-linked immunosorbent assay (ELISA), and enzyme-linked immunospot (ELISPOT). In addition, researchers need a standardized tool to record all data and documents based on research workflows, an easy way to save and query data/documents from distributed locations, and an efficient platform for data sharing to the broad research community. To achieve those goals we present a comprehensive Web-based system, DataTrans, for managing data and information in the studies of immunological research.

Partial funding support from

National Institute of Allergy and Infectious Diseases
National Institutes of Health

Contract Number N01-AI-50020

University of Rochester
Center for Biodefense Immune Modeling
School of Medicine and Dentistry
DEDiscover allows the user to enter a differential equation model or to select from a set of pre-defined models. Models may be specified as either ordinary differential equations or delay differential equations.

DEDiscover provides simulation tools and (with version 2.0) parameter estimation tools, which can be easily selected, configured and controlled using simple visual tools. Experimental data, necessary for estimation, can be loaded from standard spreadsheet formats. Simulation results are generated in real time, allowing interactive exploration of the effect of varying model parameters. Parameter estimation can be accomplished using several provided algorithms. With estimation progress displayed during computation, results are displayed in both tabular and graphical formats, and can be exported to standard file formats.

DEDiscover has been designed using a “plug-in” architecture to allow easy addition of new models, model parsers, differential equation solvers, and statistical estimation methods.
Discussion and Conclusion

- Multi-discipline communication barrier
  - Among quantitative scientists: mathematicians, statisticians, physicists, engineers and computer scientists
  - Between quantitative scientists and biological scientists

- Management barrier
  - Lack of leaders who have multi-disciplinary training and knowledge
  - Lack of leaders who have both management skills and research capability with long-term missions

- Education and promotion
  - Special training: conference, workshop, symposium, summer school
  - Regular training: education program and courses
Acknowledgment

- **Statisticians:** Drs. Yangxin Huang, Dachen Liu, Jianwei Chen, Hongqi Xue, Hua Liang
- **Math Modelers/Biophysicists:** Drs. Ha Youn Lee, Alan Perelson
- **Bioengineers:** Dr. Hongyu Miao
- **Biocomputing:** Drs. Jingming Ma, Dongwen Wang, Gregory M Warnes
- **Software/Database Developers:** Jeffrey Williams, Canglin Wu, Alain LeBlanc, Carol Crowley, Tao Wu, Stephen Massaro, Weidong Yin
- **Virologists/Immunologists:** Drs. Lisa Demeter, Xia Jin, Martin Zand, David Topham, Tim Mosmann, Brian Ward
- **Project manager:** Ms. Jeanne Holden-Wiltse
Acknowledgments

- NIAID/NIH grant R01 AI 055290: AIDS Clinical Trial Modeling and Simulations
- NIAID/NIH grant N01 AI50020: Center for Biodefense Immune Modeling