Modeling & Simulation (Computational Immunology)

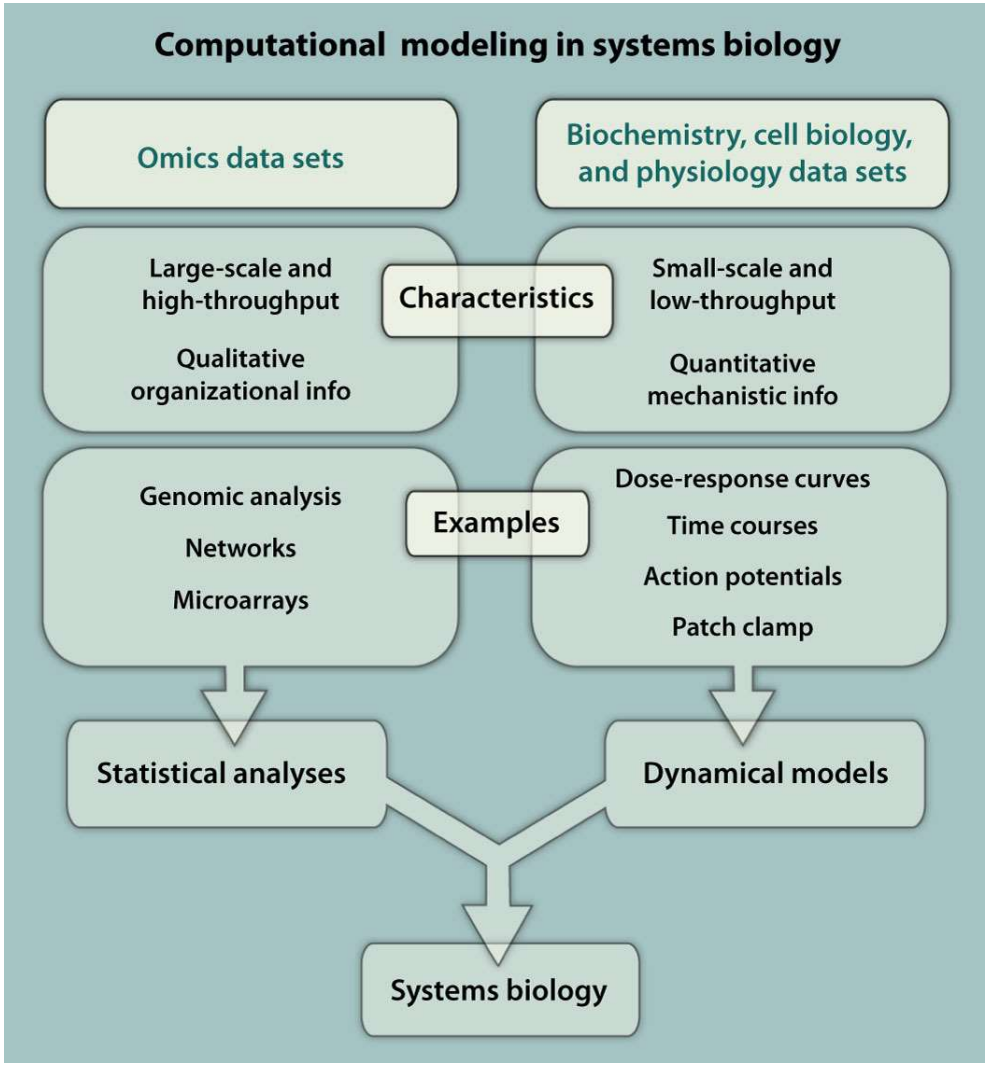
Steven H. Kleinstein

Departments of Pathology and Immunobiology
Yale School of Medicine

steven.kleinstein@yale.edu

March 23, 2020

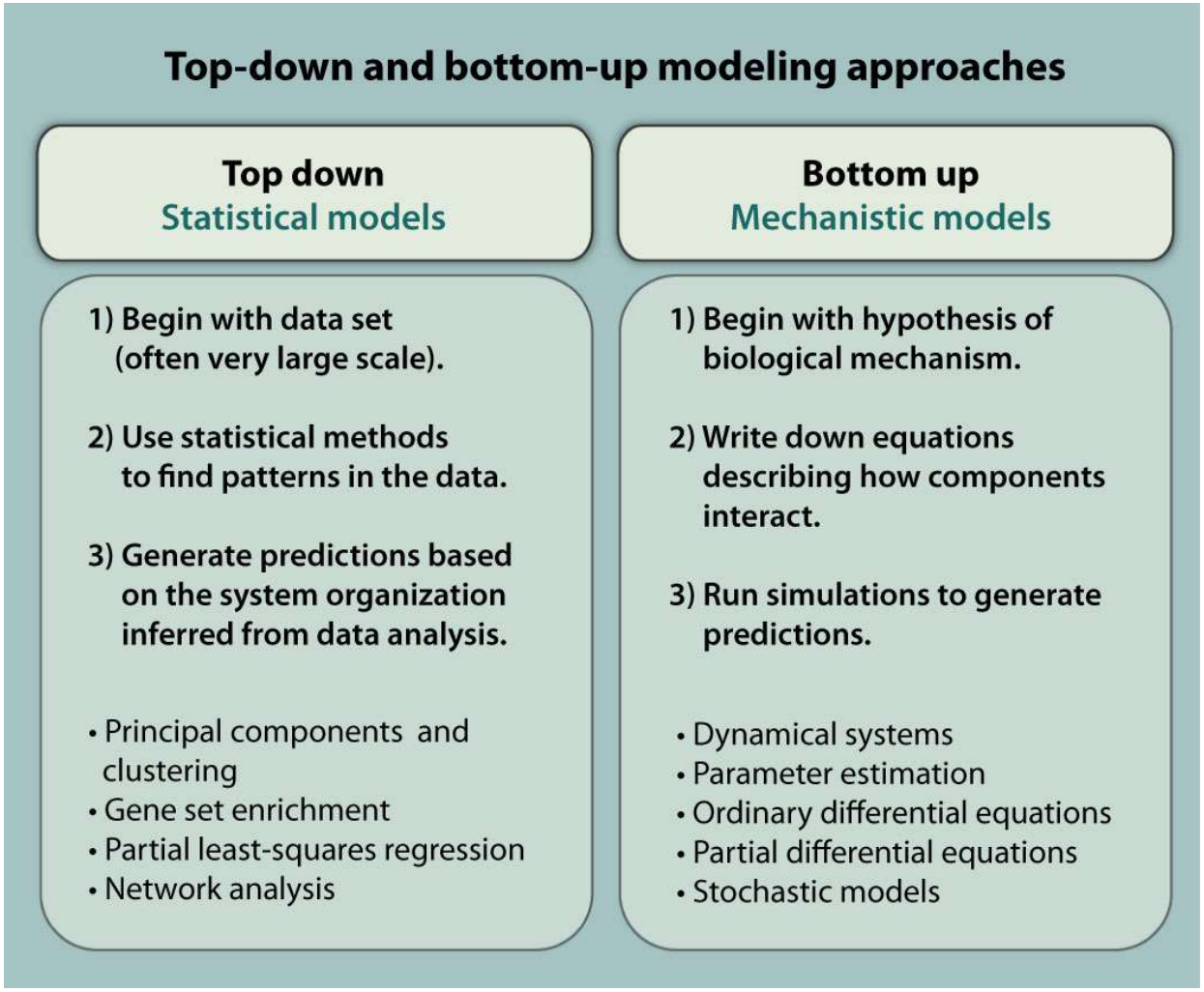
Different Types of Mathematical Models



Eric A. Sobie et al., Sci. Signal. 2011;4:tr2

Focus of next 3 lectures is on Dynamical/Mechanistic Modeling

Statistical Analysis vs. Dynamic Models



Eric A. Sobie et al., Sci. Signal. 2011;4:tr2

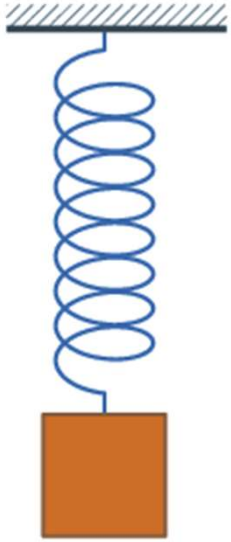
Focus of next 3 lectures is on Dynamical/Mechanistic Modeling

What is a mathematical model?

Uses mathematical language to describe a system

A mathematical model consists of a collection of variables and rules governing their values.

Models are **based on assumptions** inspired by observing some real phenomena in the hope that the model behavior resembles the real behavior.

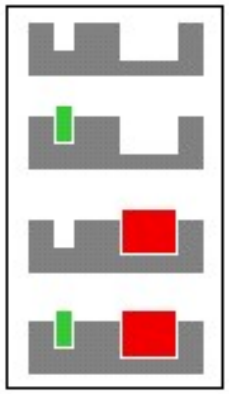


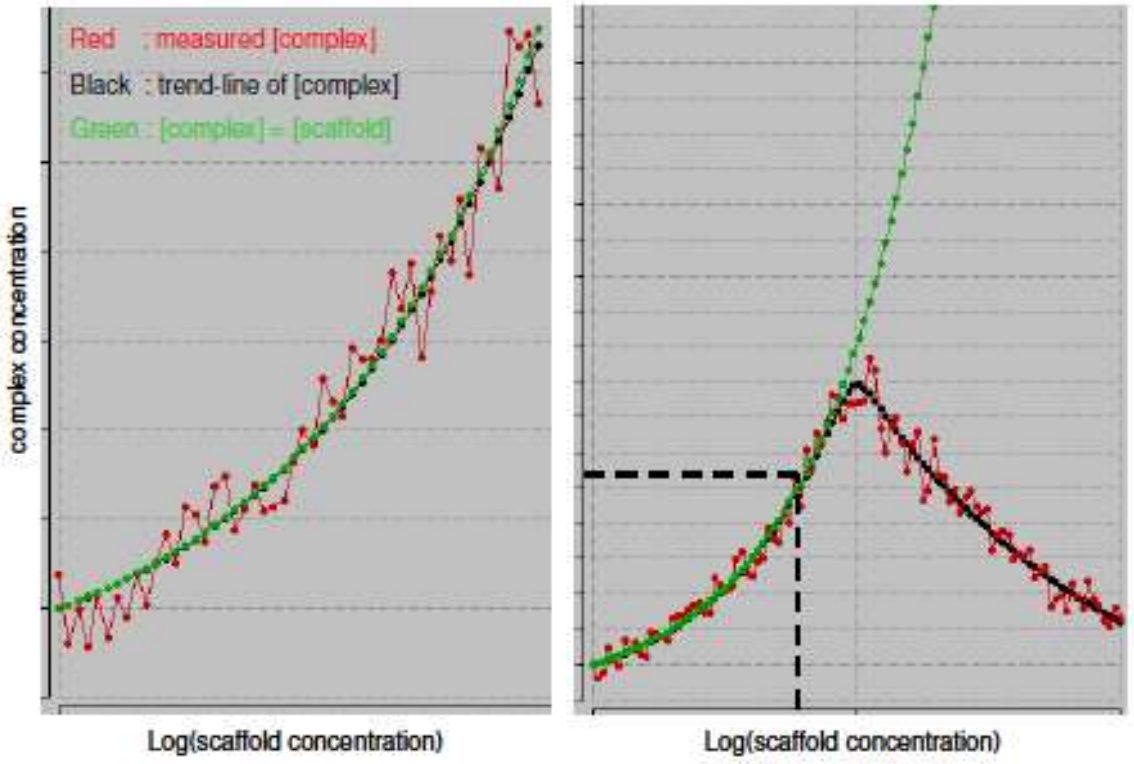
Mathematical modeling is process of constructing, testing, and improving mathematical models

Dynamical (mechanistic) modeling vs. Statistical modeling (curve fitting)

Only mechanistically correct models extrapolate reliably

Gene transcriptionally activated by complex of three proteins, and one acts as scaffold





Figures from: Hamid Bolouri

Interpolation (i.e. within sample predictions) vs.
Extrapolation (i.e. out of sample predictions, as in the right panel)

Advantages of the modeling approach in biology

"Essentially, all models are wrong, but some are useful."
-George Box, University of Wisconsin

- Concise summary of present knowledge of operation of a particular system
- Predict outcomes of modes of operation not easily studied experimentally in a living system
- Provide diagnostic tools to test theories about the site of suspected pathology or effect of drug treatment
- Clarify / simplify complex experimental data
- Suggest new experiments to advance understanding of a system

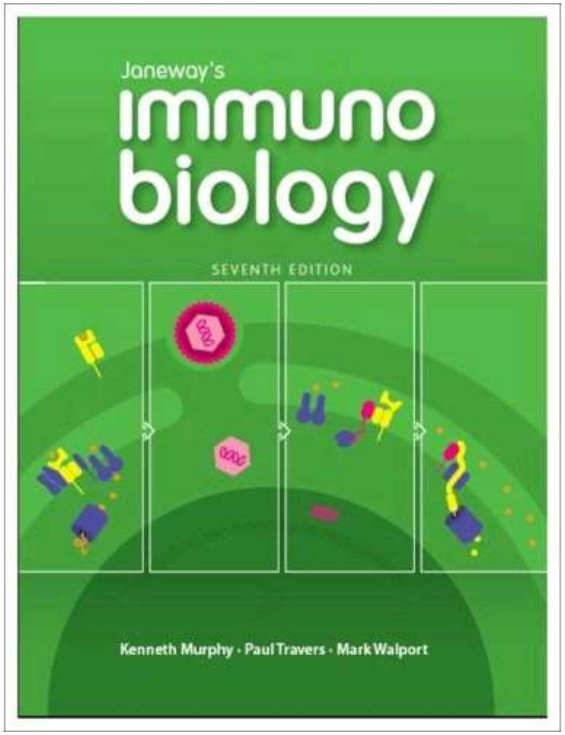
Limitations of the modeling approach

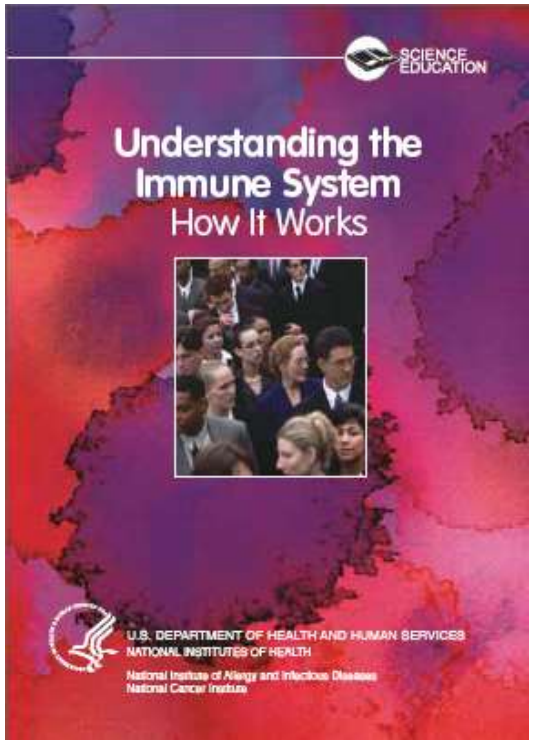
"Essentially, all models are wrong, but some are useful."
-George Box, University of Wisconsin

- Models often require many simplifying assumptions
  - beware of garbage in, garbage out
- Validation of model predictions is essential
  - examination of behavior under known limiting conditions
  - experimental validation
  - limits of model can point out what we don't understand

Modeling the immune response

If you want more information on the biology…





Janeway's Immunobiology

- or -

https://www.niaid.nih.gov/research/immune-system-overview

The Immune System

Science that began with Jenner in 1796

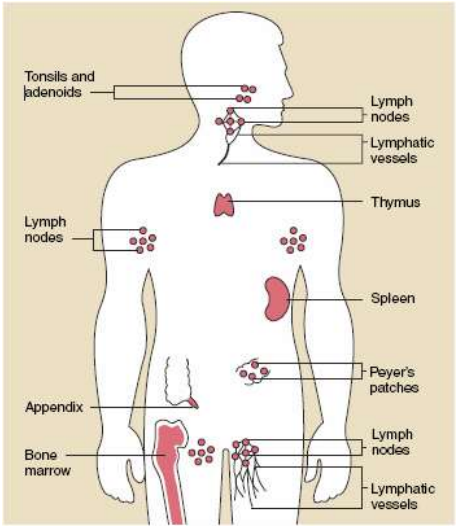
- A network of cells, tissues, and organs that work together to defend the body against attacks by "foreign" invaders (antigens).
  - Primarily microbes (germs)—tiny, infection-causing organisms such as bacteria, viruses, parasites, and fungi.
- Provides basis for vaccines (e.g., flu shot)
- But also implicated in disease:
  - Autoimmune (Lupus, MS, Rheumatoid Arthritis)
  - Respond to harmless foreign substance (ragweed pollen) produces allergy
  - Sepsis, Cancer
- Understanding will lead to better diagnostics & therapies

Organs of immune system = "lymphoid organs", since home to lymphocytes (small white blood cells that are key players in the immune system)

Why Model the Immune System?

Experiments provide only a static window onto the real dynamics of immunity

- Immune response involves the collective and coordinated response of $\approx 10^{12}$ cells and molecules
- Spatially-distributed system
  - blood, lymph nodes, spleen, thymus, bone marrow, etc.



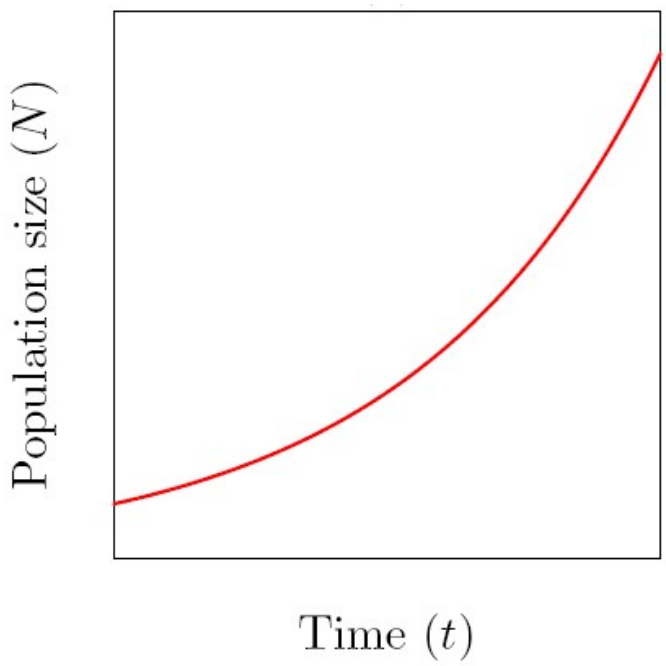
- Feedback loops and non-linear dynamics
- Experiments often require artificial constructs

Models can help understand the source(s) of variability between experiments

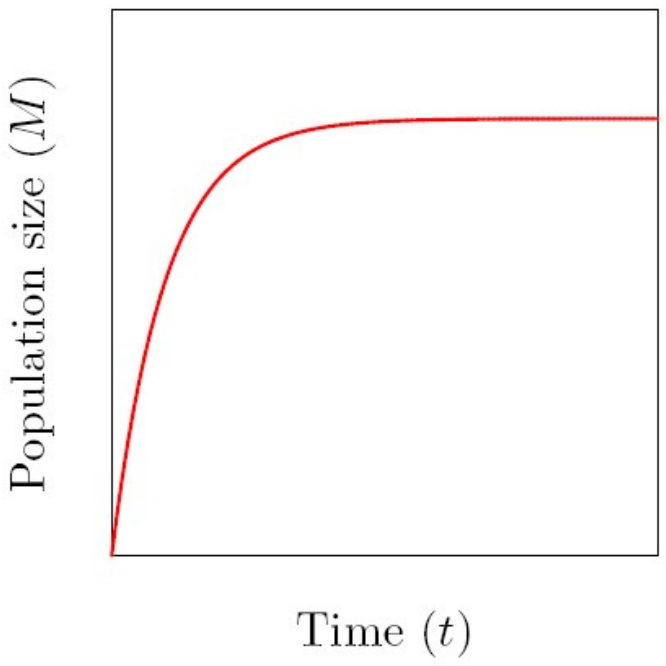
Dynamic vs. Static modeling

A dynamic model accounts for the element of time,
while a static model does not

Exponential growth of virus



White blood cells produced
by bone marrow



http://bioinformatics.bio.uu.nl/rdb/books/ti.pdf

Dynamic equations can be simulated to study system behavior

Types of Dynamic Models

Choosing the type of model is an important first step

- **Continuous**: time or state variables (often called 'density')
  - Ordinary differential equations
- **Discrete**: time or state variables
  - assume a small set of qualitative states e.g. active or inactive
  - changes in state are given by discrete (logical) rules

- **Deterministic**: no randomness is involved in the development of future states of the system
  - Given model structure, parameter values, and initial conditions, there is no variation in output
- **Stochastic**: the next state of is not fully determined by the previous state – probability is involved
  - can take into account the fluctuations in mRNA/protein/cell numbers and external noise

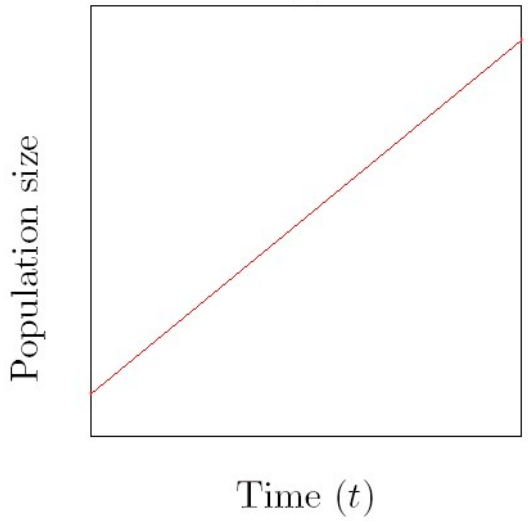
Spatial structure can also important

Ordinary Differential Equations (ODEs)

Continuous and Deterministic

Production rate

$$\frac{dB}{dt} = p \quad \Rightarrow \quad B(t) = B(0) + pt \quad \Rightarrow$$



Change in number of B cells per time

$$\frac{dB}{dt} = \lim_{t \to 0} \frac{B(t + \Delta t) - B(t)}{\Delta t}$$

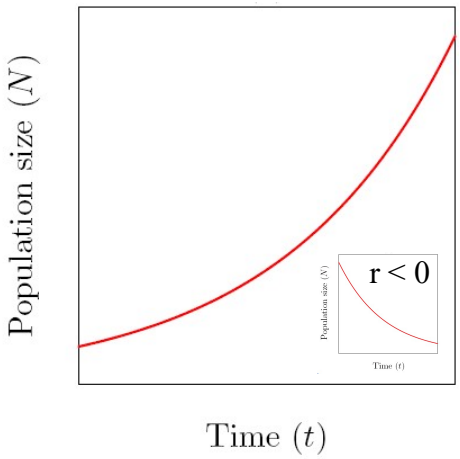
http://bioinformatics.bio.uu.nl/rdb/books/ti.pdf

Most models used in practice not solvable → **simulate**

Exponential growth (and decay)

Continuous and Deterministic

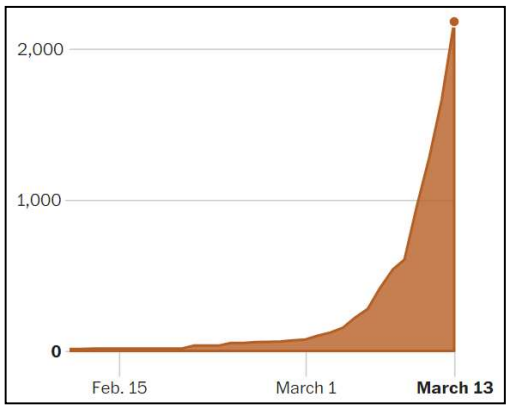
$$\frac{dN}{dt} = rN$$



$$N(t) = N(0)e^{rt}$$

Infectious Disease Spread

(not indefinite, see SIR model later)



How long for population to double?

$$2N(0) = N(0)e^{rt}$$

$$\ln 2 = rt$$

$$t = \ln[2]/r$$

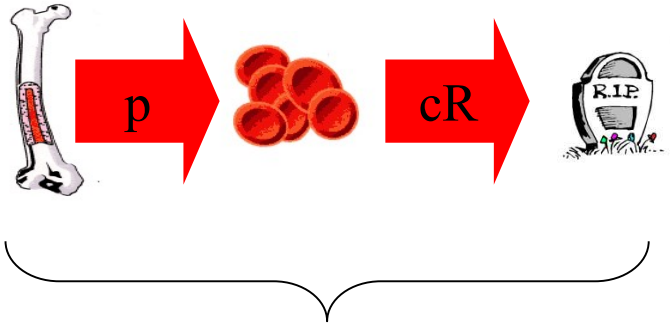
http://bioinformatics.bio.uu.nl/rdb/books/ti.pdf

Doubling time: time for population to reach 2x initial value
Half-life: time for population to reach 50% of initial value

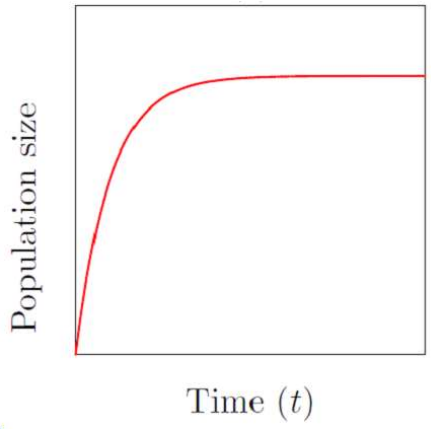
Steady-state

Population sizes remain constant at steady-state

Red Blood Cell production



$$\frac{dR}{dt} = p - cR$$



How many cells at steady-state?

$$0 = p - cR$$

$$R = p/c$$

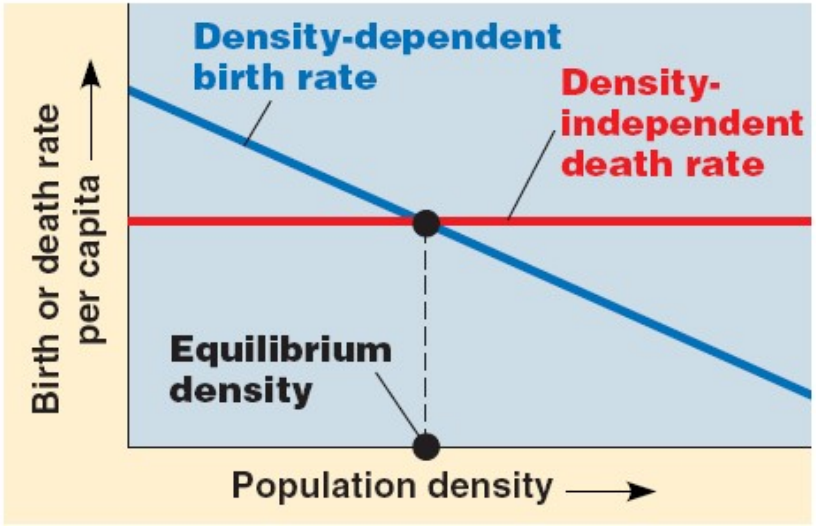
http://bioinformatics.bio.uu.nl/rdb/books/ti.pdf

Solve for steady state by setting derivatives equal to zero

Density dependence

Birth (or death) rate may depend on population size

$$\frac{dN}{dt} = bN \qquad -dN$$



$$N = K\left(1 - \frac{d}{b}\right)$$

http://bioinformatics.bio.uu.nl/rdb/books/ti.pdf

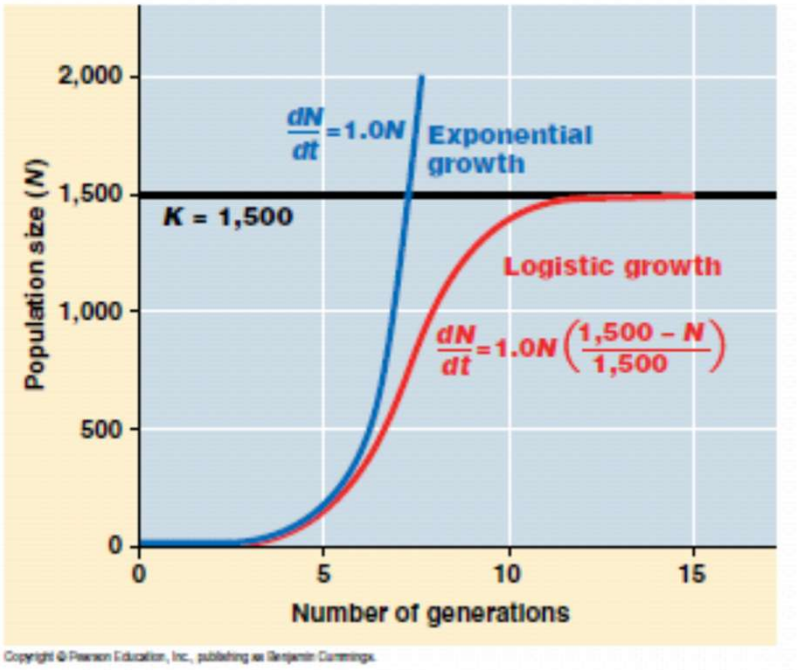
Stable steady-state: small perturbations return to same state

Logistic Model (S-shaped curve)

Includes density-dependent birth and death (r = b - d)

$$\frac{dN}{dt} = rN\left(1-\frac{N}{K}\right)$$

Initial stage of growth is approximately exponential; growth slows as saturation begins, and then stops at maturity.



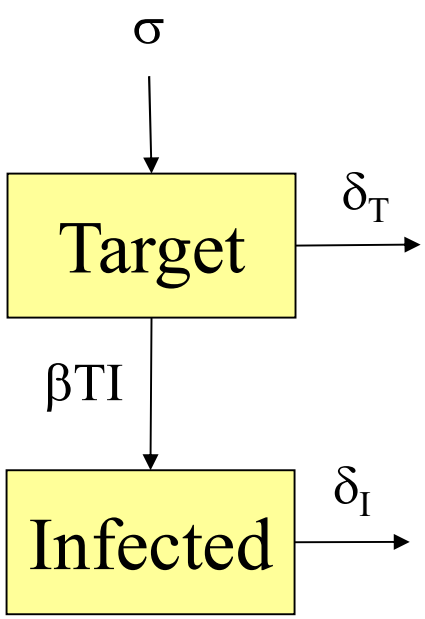
http://bioinformatics.bio.uu.nl/rdb/books/ti.pdf

Is this a "model" if can't explain why birth/death rate *r~N/K?*
phenomenological model

Carrying capacity (K): population size that can be sustained indefinitely

Modeling Interactions

Law of mass action (also called the mean-field assumption):
Entities encounter each other according to their relative abundance across space -- the rate of an elementary reaction is proportional to product of concentrations of participating entities



Target cells (T) become infected cells (I)

Target $\frac{dT}{dt} = \sigma - \delta_T T - \beta T I$

Infected $\frac{dI}{dt} = \beta T I - \delta_I I$

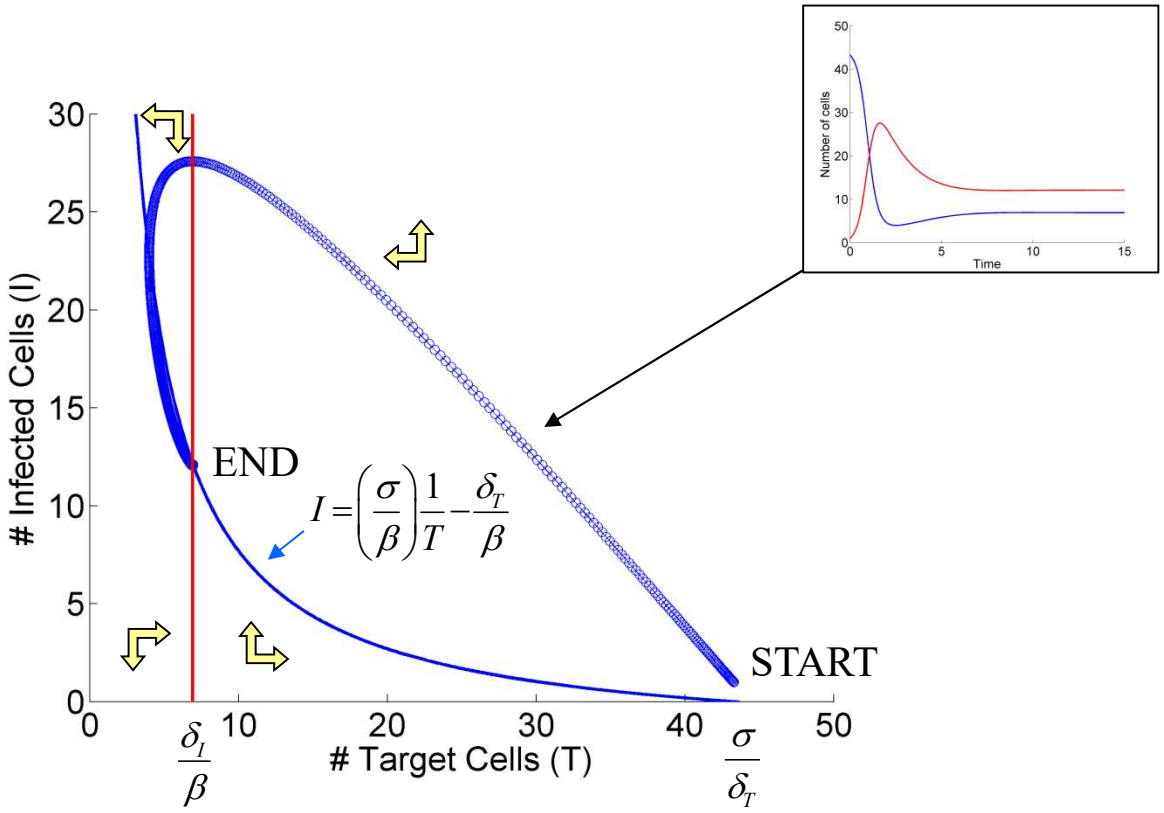
Other approaches are needed to account for spatial structure

Phase Plane Analysis

Nullclines plot where derivatives are zero (cross at steady-state)

Target $\frac{dT}{dt} = \sigma - \delta_T T - \beta T I$

Infected $\frac{dI}{dt} = \beta T I - \delta_I I$

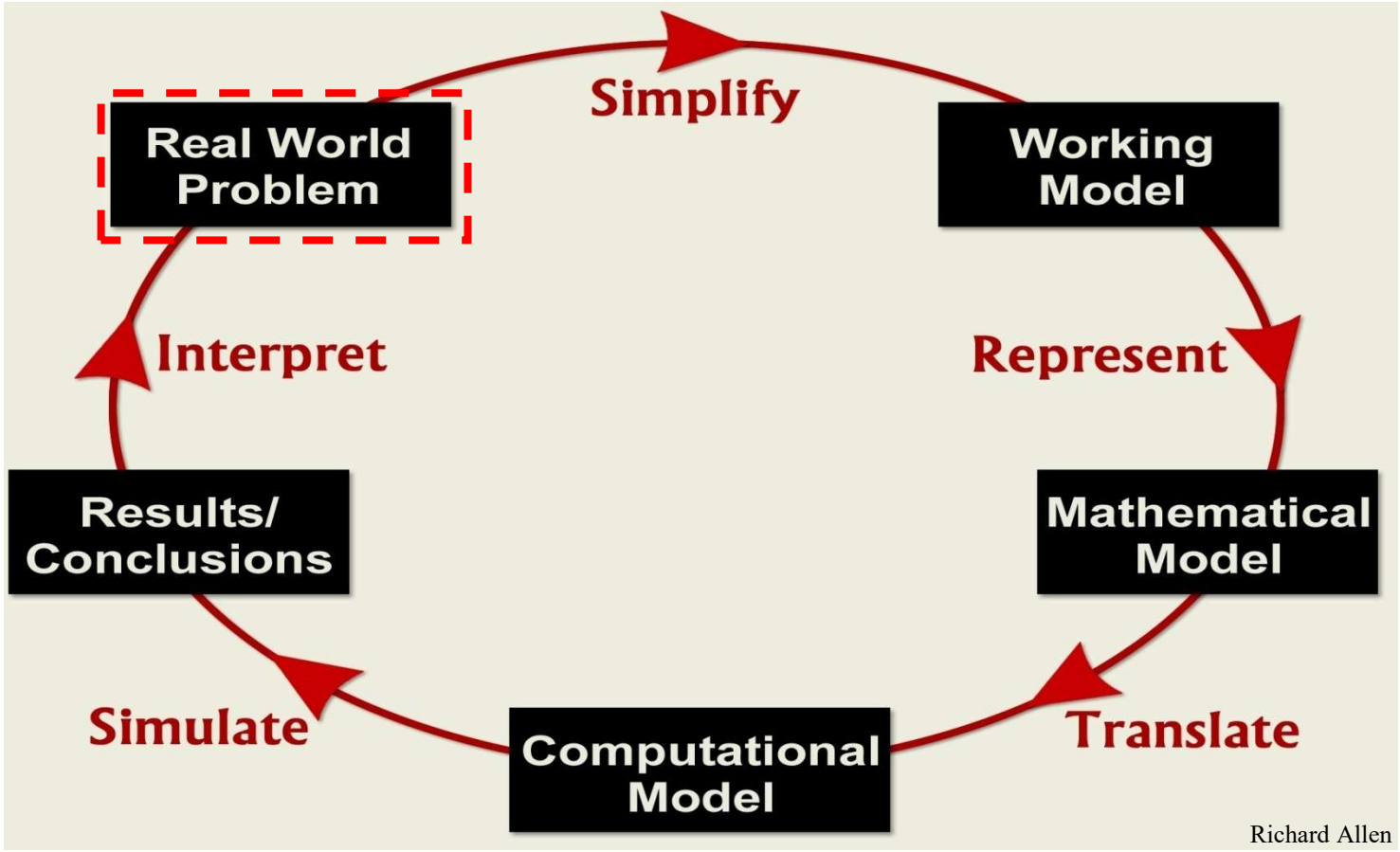


http://bioinformatics.bio.uu.nl/rdb/books/ti.pdf

Phase portraits plot typical trajectories in the state space

The Modeling Process

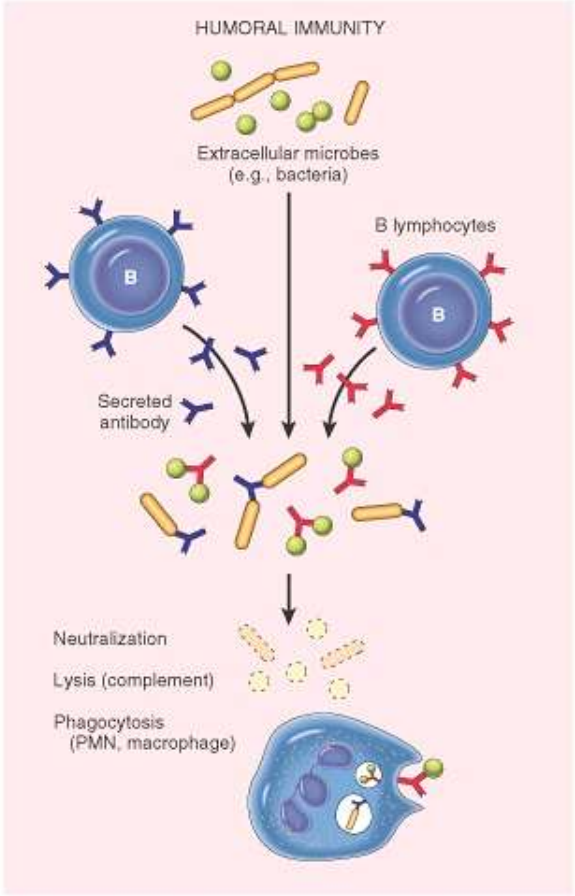
Starts with a specific scientific question



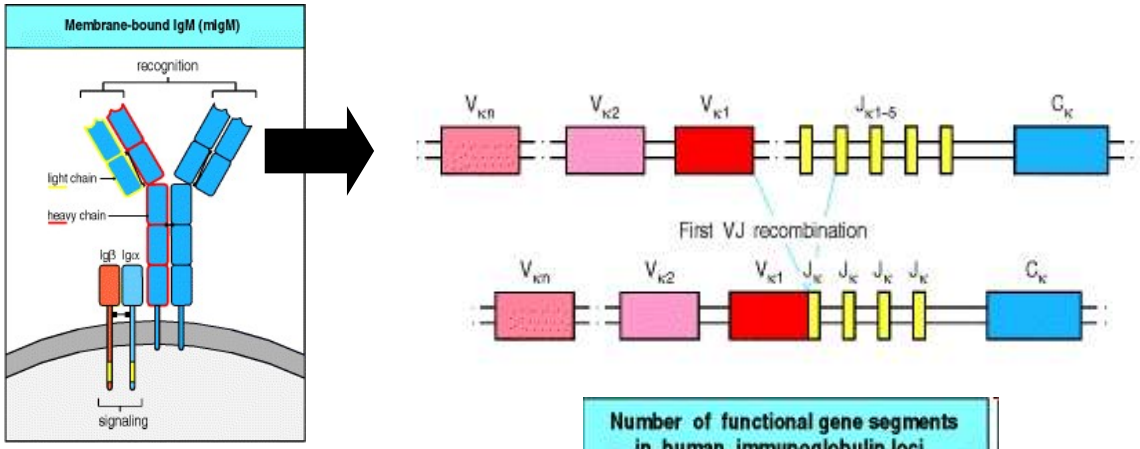
Model should produce predictions that suggest new experiments

B cells "recognize" antigens thorough antibody receptor

First phase of diversification occurs in bone marrow while cell is maturing



Rearrangement generates diverse receptors:



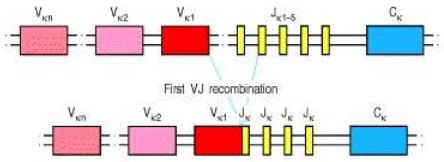
Number of functional gene segments in human immunoglobulin loci

Segment	Light chains		Heavy chain
	κ	λ	H
Variable (V)	40	30	65
Diversity (D)	0	0	27
Joining (J)	5	4	6

Second phase of diversification (by somatic hypermutation) follows activation

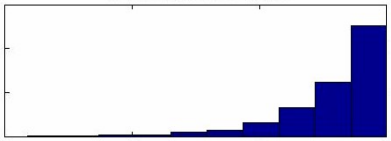
The Modeling Process: V(D)J Recombination

How are VJ segments chosen to generate an Ig light chain?



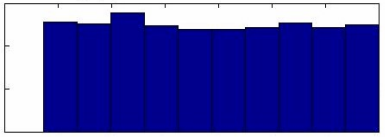
Hypothesis: VJ chosen randomly with equal probability

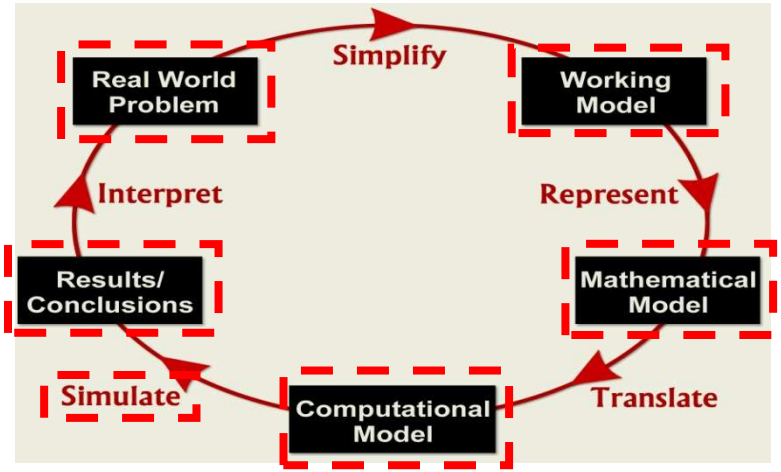
Observed V usage



?

Predicted V usage





$Pr[V_n] = 1/N$; $P[J_m] = 1/M$

randInteger(N) = floor(N * rand()) + 1

Model should produce predictions that suggest new experiments

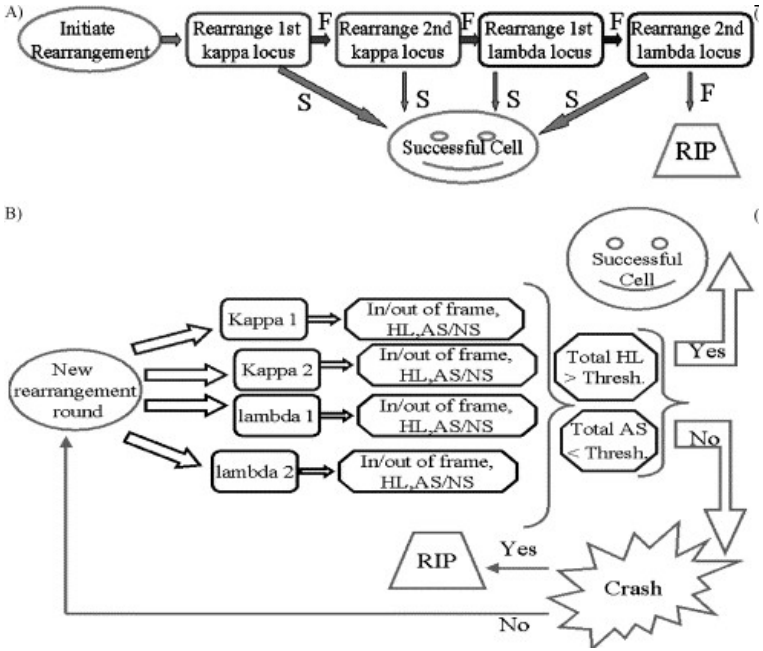
The Modeling Process: V(D)J Recombination

Extend rearrangement model to cover different alleles

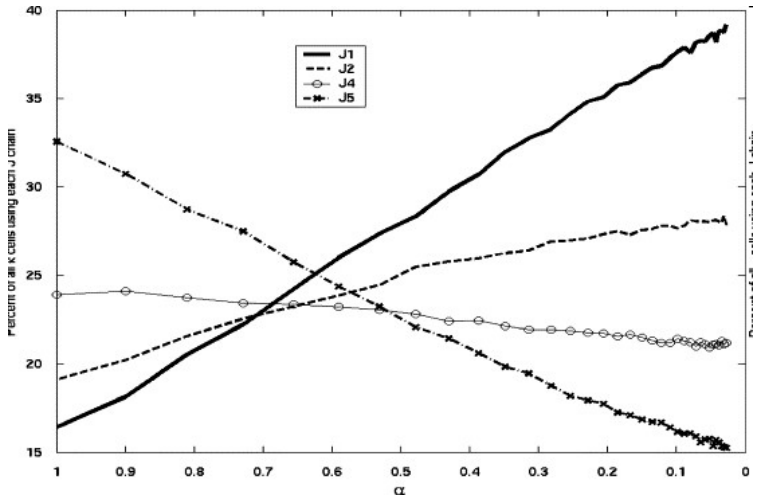
seminars in **IMMUNOLOGY**, Vol. 14, 2002: pp. 169–190
doi:10.1016/S1044–5323(02)00041-6, available online at http://www.idealibrary.com on IDEAL®

Analysis of B cell receptor production and rearrangement Part I. Light chain rearrangement☆

Yoram Louzoun[a,*], *Tzivia Friedman*[a], *Eline Luning Prak*[b], *Sam Litwin*[c] *and Martin Weigert*[a]



A probabilistic model of allelic exclusion fails to explain the status of receptor genes and the receptor phenotype of most B cells… we have revived the purely probabilistic approach in a model that now includes receptor editing and allows for some multi-receptor B cells. **We find that this model can explain the observed properties of B cells when the frequency of self-reactive B cells is high…**



Alpha reflects degree of sequentiality for Jκ rearrrangement.

Revised model of rearrangement suggest new experiments

Things to ask before any modeling study

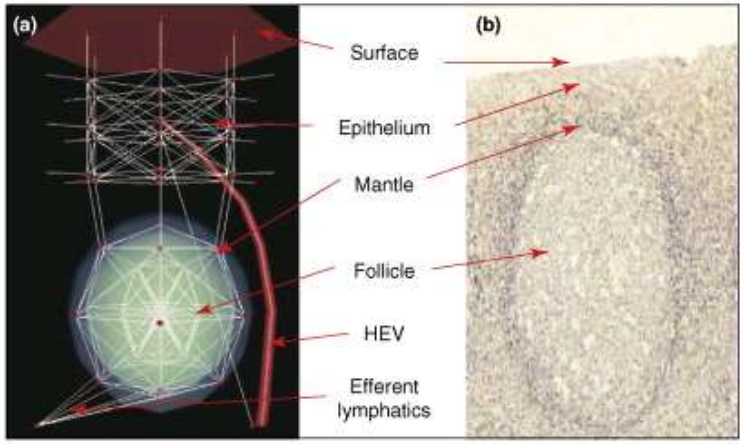
Frank Tobin (2009): Modeling is Powerful BUT Has Far to Go

Bio·IT World.com

1. Why do you want to do modeling?
2. How will you know if you succeed?
3. What will you do with the model once you have it? For what decisions will it be used or what confirmatory experiments will get performed?

Beware motivation: “We want to create a model of process X…”

Forward Modeling



(Thorley-Lawson et al, 2008)

- Detailed mathematical model designed to incorporate a desired level of anatomic or physiologic features
  - Can have arbitrary complexity as desired
  - Parameter values often obtained from published literature
  - Ex: tissue structure formation, cell signaling networks
- Used for simulating realistic experimental data under precisely defined conditions to test hypotheses *in silico*
- Can help design better experiments and reduce animal use
- Generally too complicated for fitting to experimental data

Allows generation of synthetic data sets with prescribed noise characteristics (Monte Carlo simulation) for evaluating parameters obtained by inverse modeling

Inverse Model

- A mathematical model designed to fit experimental data so as to explicitly quantify physical or physiological parameters of interest
- Values of model elements are obtained using parameter estimation techniques aimed at providing a “best fit” to the data
- Generally involves an iterative process to minimize the average difference between the model and the data
- Evaluating the quality of an inverse model involves a combination of established mathematical techniques as well as intuition and creative insight