Agent-Based and Aggregate Modeling: Tradeoffs & Limitations

Nathaniel Osgood

Using Modeling to Prepare for Changing Healthcare Needs
Duke-NUS
April 16, 2014

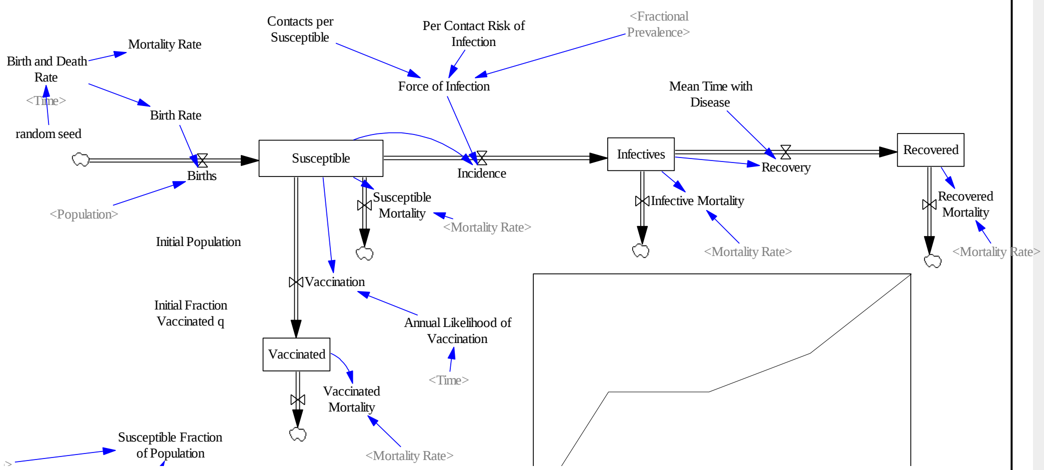
Distinctions

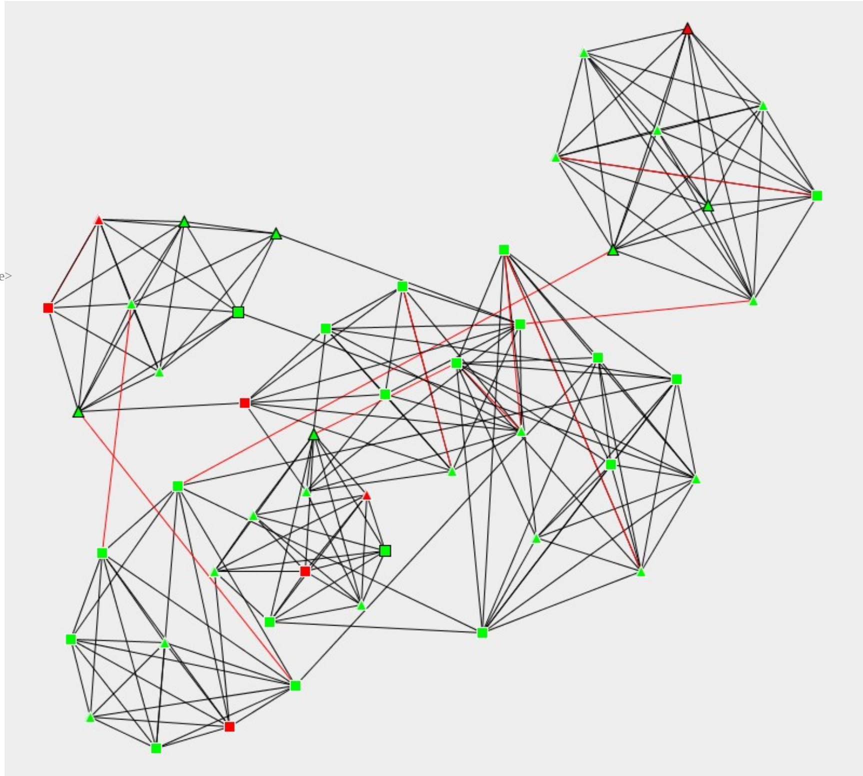
- Inherent, e.g.
  - Qualitative vs. Quantitative
  - Static vs. Dynamic
  - Stochastic vs. Deterministic
  - Capacity to understand single scenario vs. range of scenarios
  - Magnitude of computational resources required
    - Interactive or not
  - Under vs. over-determined calibration
  - Ability to calibrate to/make behaviour depend on individual history
- Important software skills mediation
  - Required level of software development sophistication

Dynamic Models for Health

- Classic: Aggregate Models
  - Differential equations
  - Population classified into 2 or more state variables according to attributes
  - |State Variables|,|Parameters| << |Population|
- Recent: Individual-Based Models
  - Governing equations approach varies
  - Each individual evolves
  - |State Variables|,|Parameters| $\propto$ |Population|

Contrasting Model Granularity





Granularity Selection: Problem Specific

- Selection of granularity is a function of question that are asking – not of the "*true nature of the system*"
  - *Modeling for learning/qualitative insight (requires "caricature model") vs.*
  - *Modeling to quantitatively predict (requires detailed characterization)*
- Quanta of most obvious system components may not align with needs for insight
  - May gain benefits from higher-level representation
    - Many high-level qualitative behaviors of complex systems can be explained with very simple models
    - Often gain greater insight from simpler model: C.f. Gas laws vs. lattice gas model
  - May wish to seek lower level model
    - Small infection spread model : Characterization at level of immune response rather than monolithic person

Myth of Individual-Based Models as "Modeling from the Bottom Up"

- A single person is a natural locus of description
  - Presents for care
  - Lives
  - Dies
  - Coupled internal systems
- But the world has no natural "bottom"
  - It is frequently desirable to include within a person a great deal of "within the skin" detail
- The issues of model depth & breath are just as pressing in individual-based models as in aggregate modeling

Contrasting Benefits

Aggregate Models

- Frequently, easier
  - Construction
  - Calibration
  - Parameterization
  - Formal analysis (Control theoretic & Eigenspace techniques)
  - Understanding
- Performance
  - Lower baseline cost
  - Population size invariance
- Less pronounced stochastics
  - Less frequent need for Monte Carlo ensembles
- Quicker construction, runtime ⇒More time for understanding, refinement

Individual-Based Model

- Better fidelity to many dynamics
- Stronger support for highly targeted policy planning
- Ability to calibrate to & validate off of longitudinal data
- Greater heterogeneity flexibility
- Better for examining finer-grained consequences
  - e.g. transfer effects w/i pop.
  - Network spread
- Simpler description of

Key Needs Motivating Individual-Based Modeling

- Need to calibrate against information on **agent history**
- Need to capture **progression** of agents **along multiple pathways** (e.g. co-morbidities)
- Wish to characterize **learning by and/or memory** of agents based on experience, or **strong history dependence** in agents
- Need to capture distinct **localized perception** among agents
- Seeking to intervene at points in, change behavior on, explain phenomena over or explain dynamics **across networks**
- Seek **distinct interventions for** many **heterogenous categories**
- Need to **capture impact** of intervention across **many categories**
- When it is much simpler to **describe behavior at indiv. level**
- Seek **flexibility** in exploring different **heterogeneity dimensions**
- **Needs of stakeholders** to engage with individual-based models
- Want to describe behaviour at **multiple scales**
- We care about **stochastics/uncertainty** caused by indiv variabilit

Key Needs Motivating Aggregate-Based Modeling

- Need to **execute quickly** (e.g. for user interaction)
- Understand/describe system behaviour across **all possible values for parameters**
  - Seeking to mathematically **analyze the model** (e.g. to determine location or stability of equilibria) for insight
  - To determine shape of **all possible trajectories**
- Want to **use mathematical tools** (e.g. control theory) to identify high-leverage parameters, optimal policies
- Need to **extensively calibrate to much historic data**
- **Desire of stakeholders** to work at higher level
- Behavior for different subgroups **differs only in degree**
- No recourse to **software engineering knowledge**
- **Lack of detailed knowledge** of network structure/ individual-level behaviour/Individual-level data

Individual Descriptions are Sometimes Simpler

- Understanding of *individual* behavior sometimes exceeds that of collective behavior
  - Response to locally visible incentives
  - Company’s response to competition
  - Young person’s response to peer pressure
  - Individual’s response to scarcity of good
- Sometimes it is very difficult to derive *a priori* the aggregate dynamics resulting from individual behavior
- Individual model can be simpler, more transparent

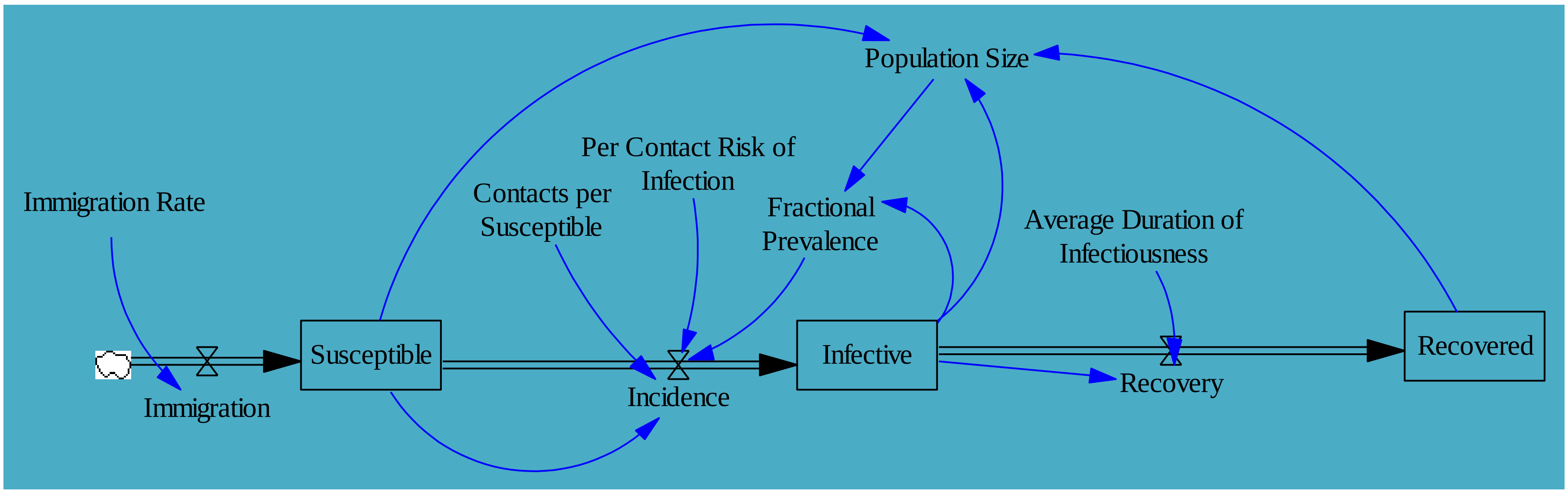
Aggregate Descriptions are Sometimes Simpler

- Aggregate descriptions frequently allow us to abstract away from myriad lower-level hypotheses
  - May afford us an easier mode of description without the need to explicitly posit involved lower-level hypotheses
  - Can be readily formulated from partial data & applied globally
- Consider
  - Using a mixing matrix computed from partial mixing data
  - Formulating population-wide
    - Hypothesized contact networks
    - Mobility patterns driving contact

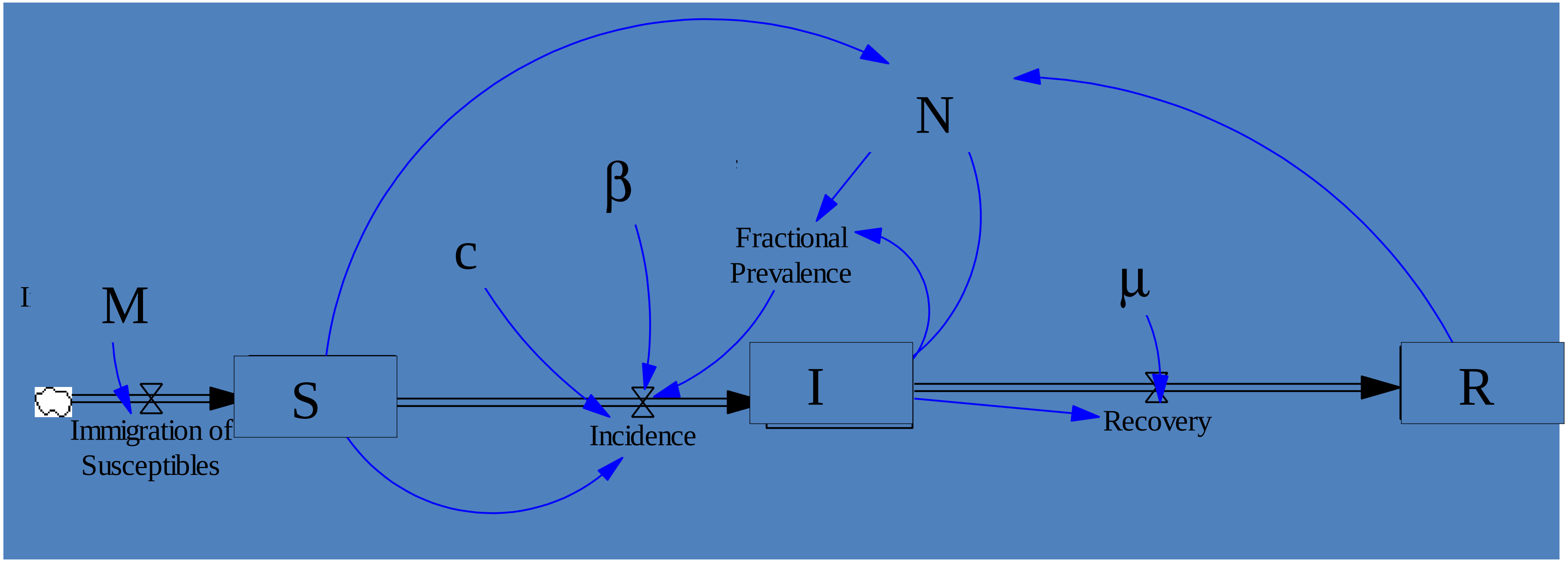
Some Uses of Formal Approaches

- Explaining observed behavior patterns
- Identifying possible behavior modes over a wide variety of possible scenarios (e.g. via eigenspace & phase plane analysis)
- Identifying how behavior depends on parameters (stability, location of equilibria)
- Creating “self-correcting” models (via control theory)
  - Individual-based models are typically not identifiable
- Formal calibration methods

Example Aggregate Model Structure



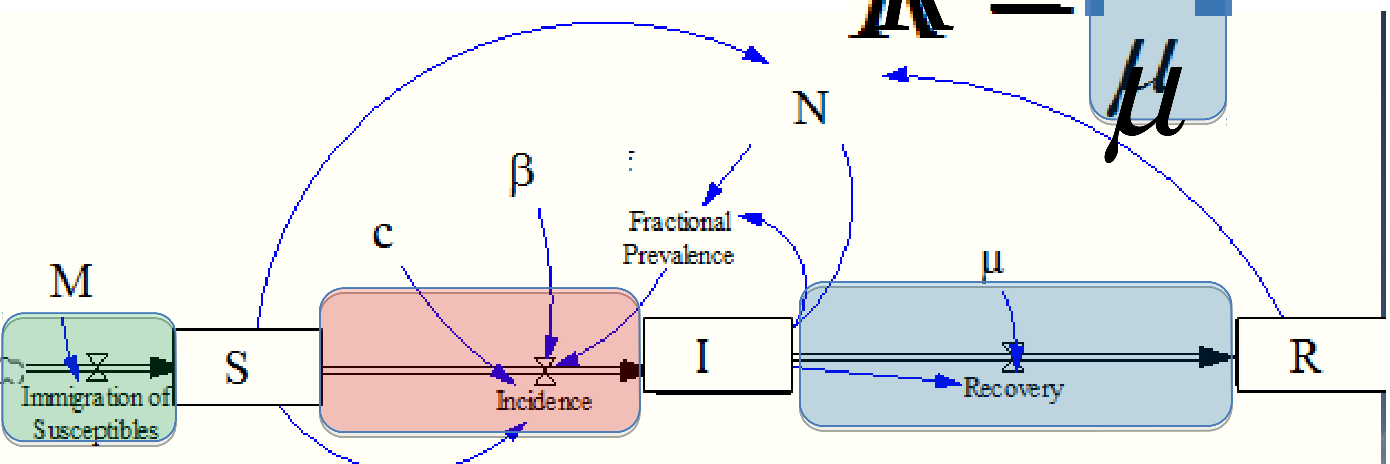
Mathematical Notation



$$\dot{S} = M - c\left(\frac{I}{N}\right)\beta S$$

$$\dot{I} = c\left(\frac{I}{N}\right)\beta S - \frac{I}{\mu}$$

$$\dot{R} = \frac{I}{\mu}$$



Model Mathematical Analysis

System Linearization (Jacobian)

$$\begin{bmatrix} -\beta\, Infectives - \delta & -\beta\, S - \delta \\ \beta\, Infectives & \beta\, S - \dfrac{1}{\mu + \dfrac{\tau\, Infectives}{h}} + \dfrac{Infectives\,\tau}{\left(\mu + \dfrac{\tau\, Infectives}{h}\right)^2 h} \end{bmatrix}$$

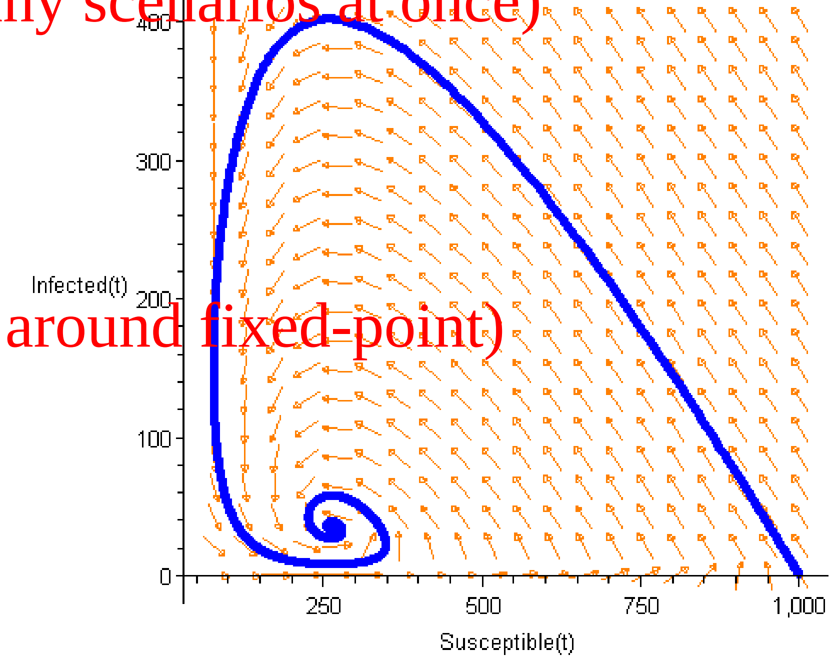
Fixed-Point Criteria

$$\dot{S} = -c\left(\frac{I}{N}\right)\hat{\beta} S + R\delta = 0$$

$$\dot{I} = c\left(\frac{I}{N}\right)\hat{\beta} S - \frac{I}{\mu + \tau\dfrac{I}{h}} = 0$$

$$\dot{R} = \frac{I}{\mu + \tau\dfrac{I}{h}} - R\delta = 0$$

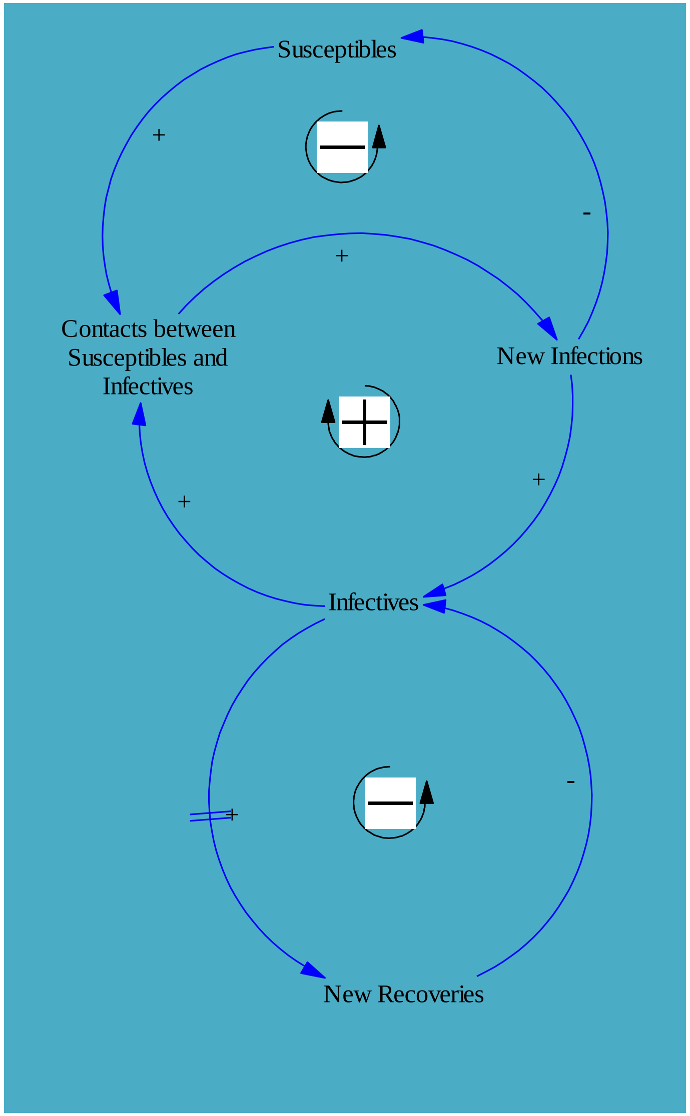
State space diagram (reasoning about many scenarios at once)



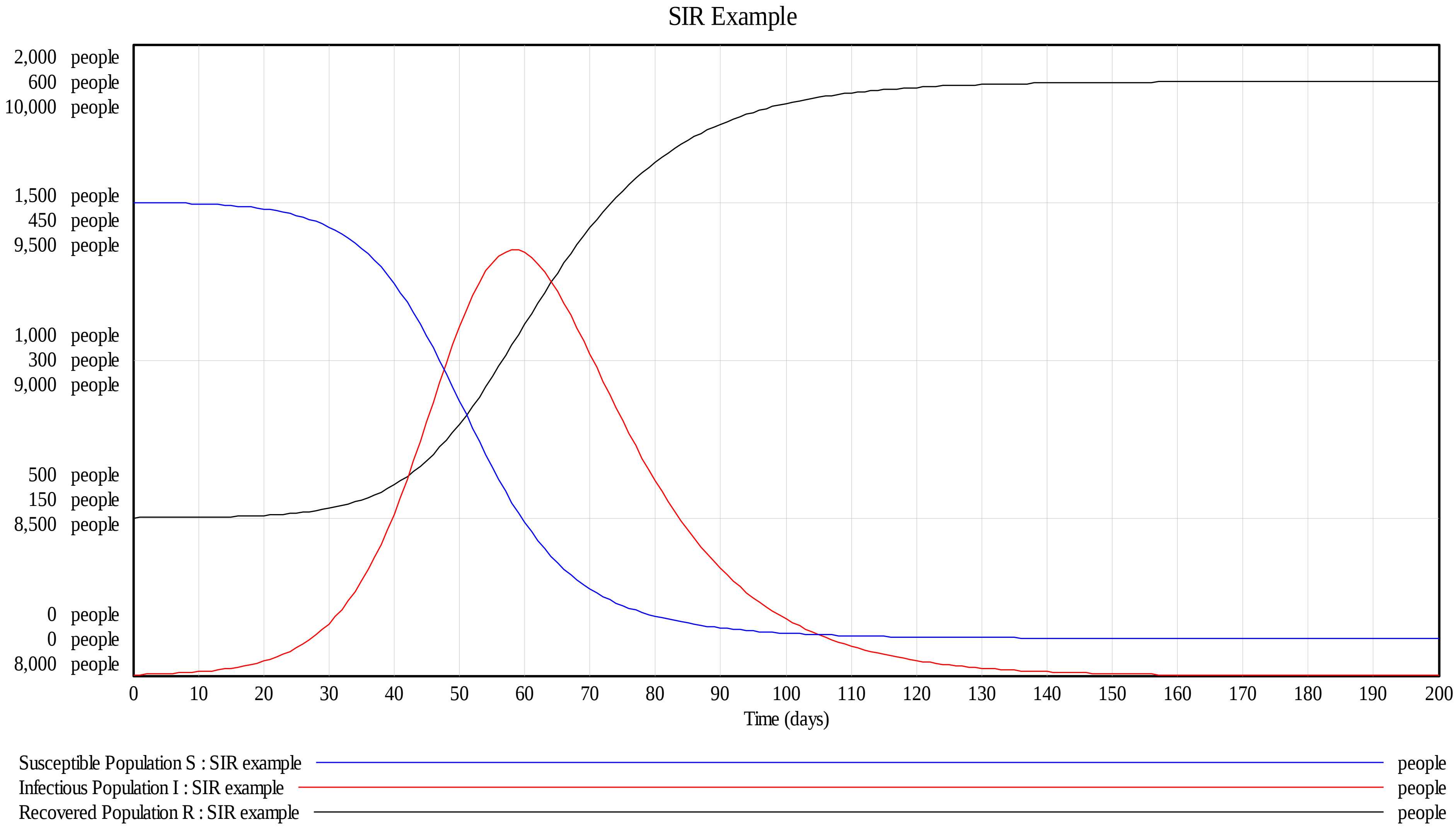
Eigenvalues (e.g. for stability analysis around fixed-point)

$$\frac{1}{2}\beta S - \frac{1}{2}\frac{1}{\mu + \frac{\tau\, Infectives}{h}} + \frac{1}{2}\frac{Infectives\,\tau}{\left(\mu + \frac{\tau\, Infectives}{h}\right)^2 h} - \frac{1}{2}\beta\, Infectives - \frac{1}{2}\delta + \frac{1}{2}\left(\left(\beta S - \frac{1}{\mu + \frac{\tau\, Infectives}{h}} + \frac{Infectives\,\tau}{\left(\mu + \frac{\tau\, Infectives}{h}\right)^2 h}\right)^2 - 2\left(\beta S - \frac{1}{\mu + \frac{\tau\, Infectives}{h}} + \frac{Infectives\,\tau}{\left(\mu + \frac{\tau\, Infectives}{h}\right)^2 h}\right)(-\beta\, Infectives - \delta) + (-\beta\, Infectives - \delta)^2 + 4\beta\, Infectives\,(-\beta S - \delta)\right)^{\frac{1}{2}}$$

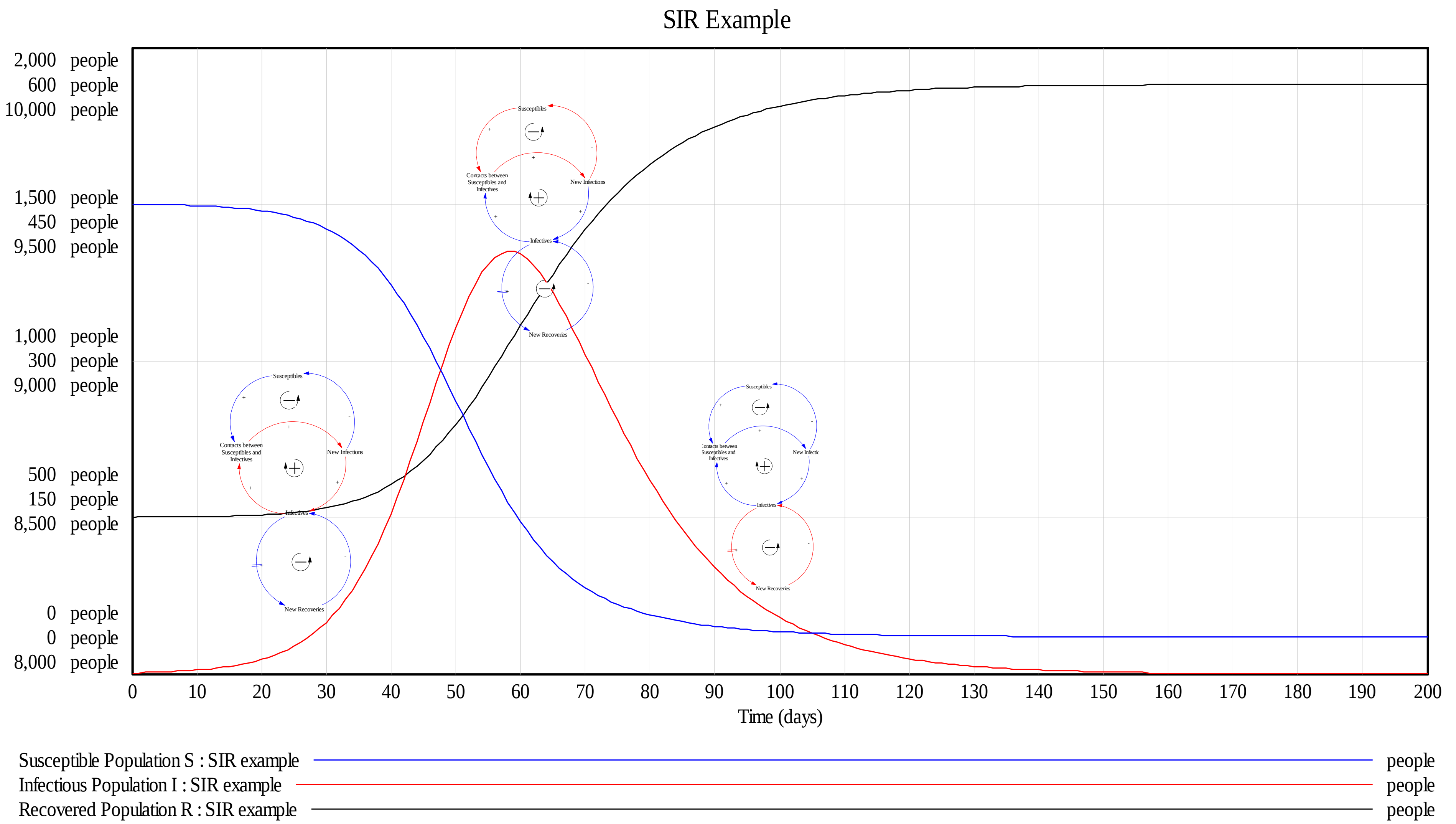
Feedbacks Driving Infectious Disease Dynamics



Example Dynamics of SIR Model (No Births or Deaths)



Shifting Feedback Dominance



Fidelity to Dynamics

- Adequate characterization of system’s causal processes may require fine-grain representation
  - Rich heterogeneity
  - Learning and adaptation
  - Response to local incentives
  - Memoryful processes
  - Behavior over persistent networks
- Aggregate behavior is *not* necessarily the same as |Population|*(Behavior of “average” individual)
- May be able to calibrate an aggregate

Example of Concern: History Information

- Heterogeneity with respect to individual history can be highly important for future health
  - Whether vaccinated
  - *in utero* exposure
  - Degree of glycemic control over the past decade
  - Exposure to adiposity
  - Previous exposure to a pathogen
- In some areas of health, we have access to longitudinal data that provides

Capturing History Information

- Individual based model
  - Both discrete & continuous history information can be readily captured
    - Categorical/discrete: State (in statechart) or variable
    - Continuous: Variable
  - Readily able to capture records of trajectories
- Aggregate model
  - Categorical/discrete: Limited discrete history information can be captured by disaggregating stocks
    - Curse of dimensionality provides tight limits on # of aspects of history can be recorded
  - Continuous: Almost always infeasible
  - Very complex to provide distributions of trajectories (via convolution of potentially changing PSFs of stocks)

Longitudinal Fidelity: Aggregate Models

- An aggregate model provides an ongoing series of *cross-sectional* descriptions of system state
  - In Calibration & validation, we can do rich comparison of these cross-sectional descriptions against available point or time-series data
  - Because the model does not track individuals, we generally cannot explicitly extract model longitudinal trajectories from the model for comparison with empirical giving longitudinal trajectories

Longitudinal Fidelity: Individual-Based Models

- An individual-based model provides easily accessible *cross-sectional* and longitudinal descrip. of system state
  - The system state at a particular moment in time is cross-sectional
  - By following & recording the trajectories of particular individuals, we can obtain longitudinal description
- In Calibration & validation, we can do rich comparison of both longitudinal and cross-sectional descriptions against available point or time-series data
  - It is in principle possible to have a model that accords with cross-sectional data, but which is at odds longitudinally

Comparisons of Model & History that are Difficult in an Aggregate Model

- Proportions of people with certain history characteristics (e.g. fraction of women who develop T2DM who have had 2 or more bouts of gestational diabetes, those with a certain duration of time separating TB infection sand active TB)
  - Can be very valuable for calibration
  - This is critical for assessing model accord with observed effect size (Relative Risk/Odds ratio)
- Model vs. historic trajectories (e.g. for timing of some transitions) for people with certain history characteristics

Example of Additional Information from Longitudinal Data

- Consider trying to distinguish pairs of situations
- e.g.: Smoking
  - Situation 1: One set of people quit & stay quit as former smokers, another set remain as current smokers
  - Situation 2: The entire set of people cycle through situations where they quit, relapse & repeat
- These two situations have very different health consequences
- We’d probably choose vary different sets of interventions for these two situations
- Similar examples are easy to imagine for obesity, STIs, TB, glycemic control & diabetes, etc.

Trajectories Summary

- If any of the following are true….
  - You have significant longitudinal information you'd strongly like the model to match
  - You have good reason to think that trajectory history has important consequences for health
  - You seek to examine the effect of policies that make use of information on individual history (e.g. # previous treatments)
- Then you should strongly consider building a model that captures this history information
  - By disaggregating stocks, you can capture limited *discrete* history information in an aggregate model (e.g. whether a person was exposed *in utero*, Time Since Quit for FS, whether a woman has had a history of gestational diabetes)
  - There is significantly greater flexibility for collecting continuous or discrete history information for guiding individual dynamics & for calibration/validation comparison to historic longitudinal data

Importance of Heterogeneity

- Heterogeneity often significantly impacts policy effectiveness
  - Policies preferentially affect certain subgroups
    - Infection may be maintained within certain subgroups even though would tend to go extinct with random mixing in the entire population
  - Policies alter balance of heterogeneity in population
    - Shifts in the underlying heterogeneity can change aggregate population statistics
  - Given a non-linear relationship, inaccurate to use the mean as a proxy for whole distribution
- **Assessing policy effectiveness often requires representing heterogeneity**
- ***Flexibility*** **in representing heterogeneity is hard to achieve in aggregate (coarse-grained) models**

Impacts of Heterogeneity on Policy Effectiveness

- Value of breast cancer detection (Park & Lees)
- Impact of airbags on deaths (Shepherd&Zeckhauser)
- Value of hernia operations (Neuhauser)
- Impact of cardiovascular disease interventions (Chiang)
- Controlling blood pressure (Shepherd&Zeckhauser)
- Effectiveness of mobile cardiac care unit (Shepherd&Zeckhauser)
- Value of breast cancer treatment (Fox)

Heterogeneity & Equity Considerations

- Failure to disaggregate (to represent heterogeneity) can impose implicit value judgements! e.g.
  - Treating situation as net zero cost if favouring group A while disadvantaging group B

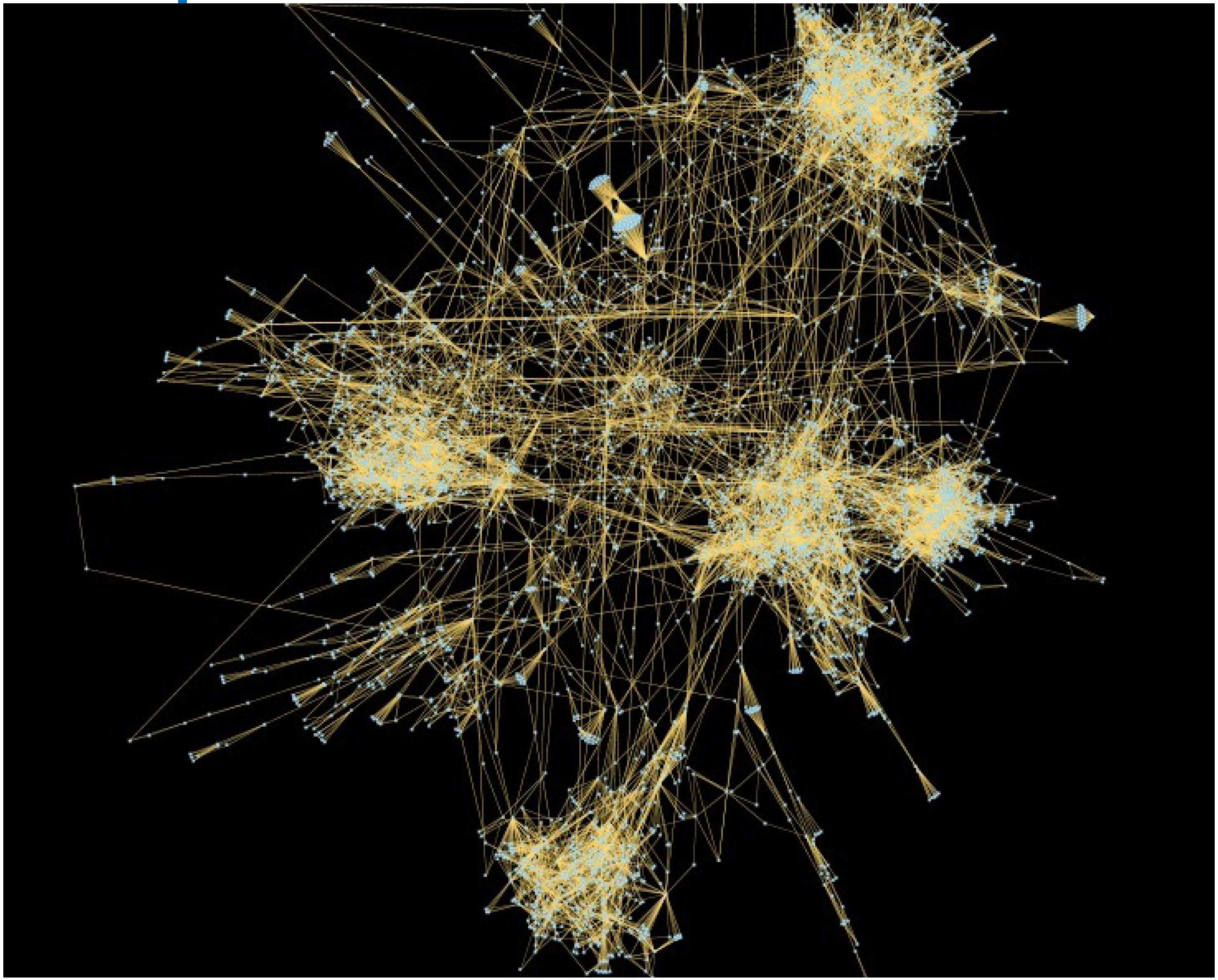
Slides With Elements Adapted from External Source

Redacted from Public PDF for Copyright Reasons

Importance of Core Groups

- Someone with high # of partners is both
  - More likely to be infected by one of the partners
    - Connect to lots of partners
    - More likely than the average individual to be connected with another high-contact person (in turn more likely to be connected)
  - Likely to pass on the infection more susceptible persons
- Often high-contact individuals connect in networks
- We may see very different infection rates in high contact-rate individuals
  - Core groups may be the key factor sustaining the infection
- Via targeted interventions on high contact people, may be able to achieve great "bang for the buck"
- ***Because of all of these considerations, we often seek to explicitly represent & reason about interventions targeting these individuals & their networks***

Example of Network Clustering



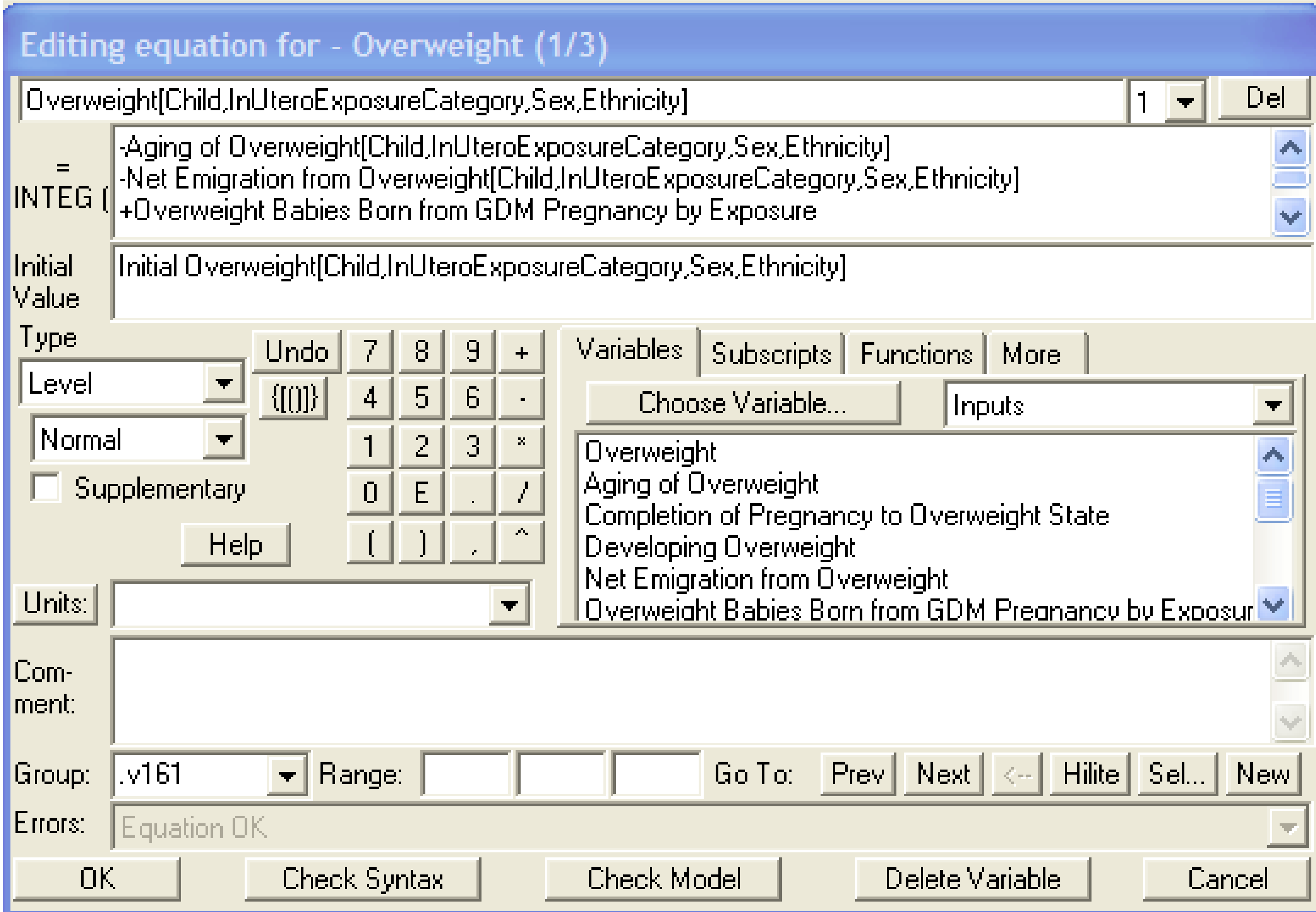
Frequent Heterogeneity Concerns

- No clear boundaries at which to divide people up into discrete categories
- Many dimensions of heterogeneity simultaneously
  - Capturing state with respect n factors requires n dimensions of heterogeneity!
- Need to consider progression along many dimensions simultaneously

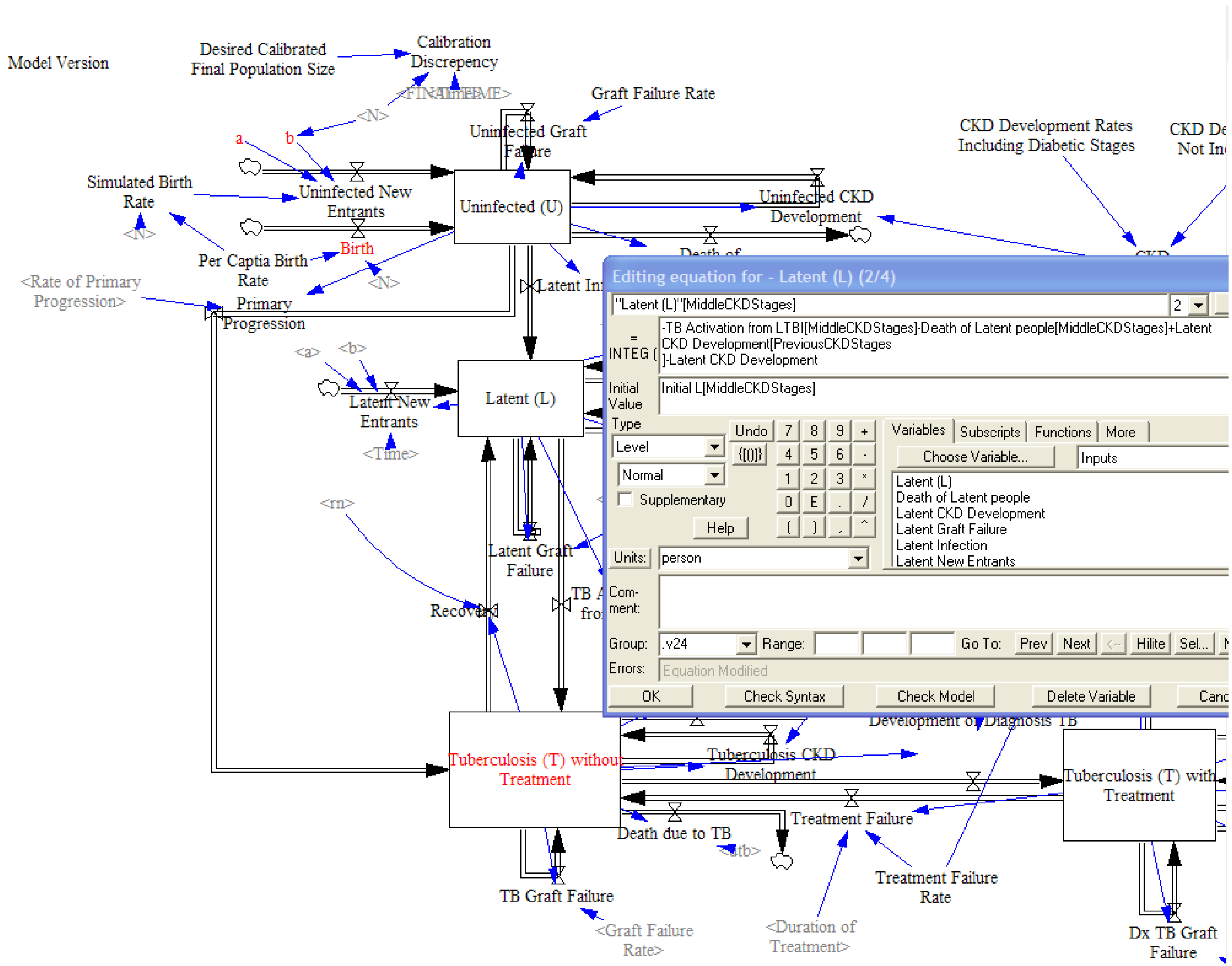
Challenges for Aggregate Model Formulation: Heterogeneity

- Two aggregate means for representing heterogeneity are limited:
  - Attribute-based disaggregation (e.g. via subscripts)
    - Need *n* dimensions to capture individual state with respect to *n* factors of heterogeneity
    - **Poor (geometric) scaling to large # dimensions**
    - Global structural, equation changes required to incorporate new heterogeneity dimensions

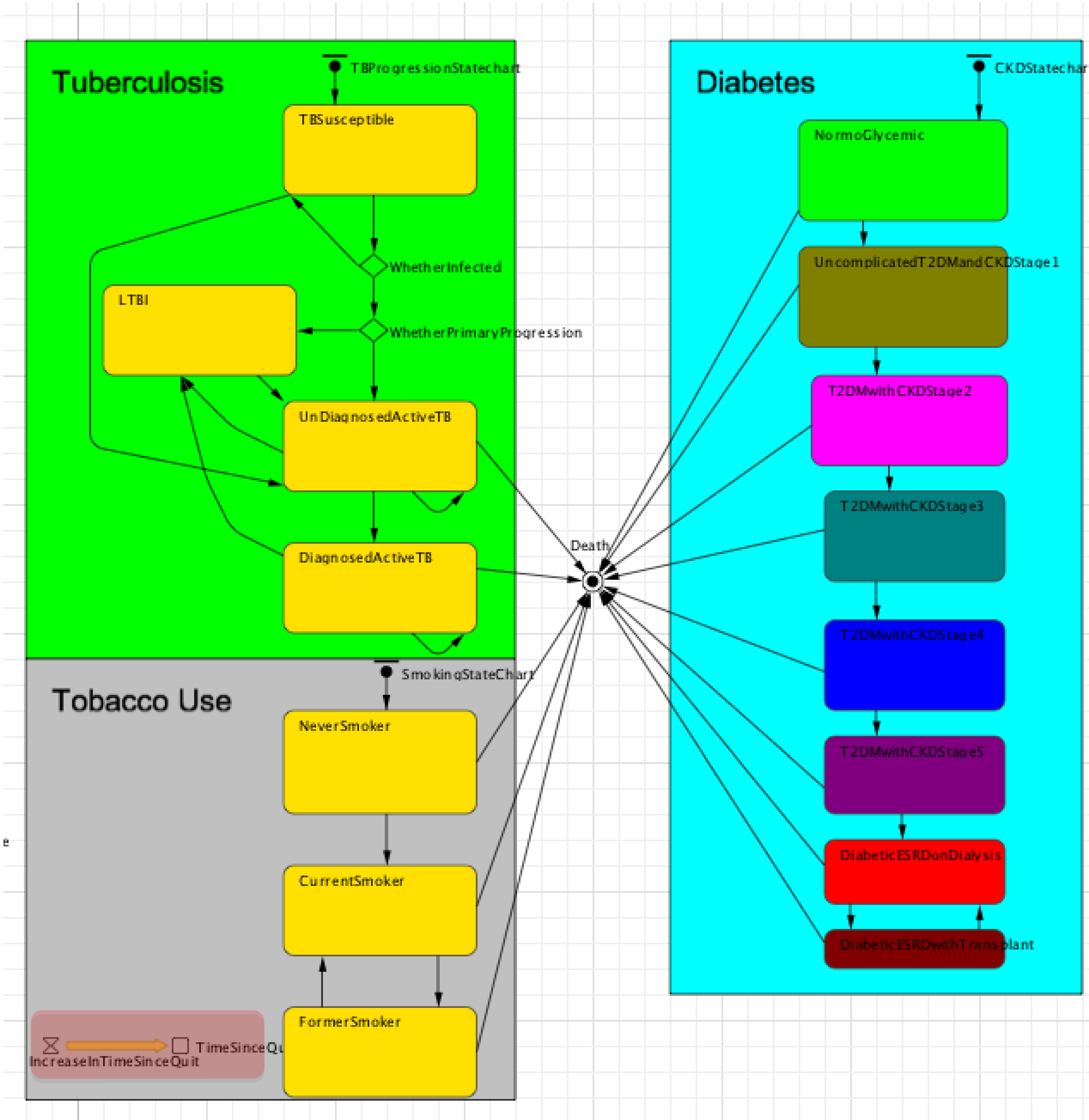
Fragility of Multi-Dimensional Subscripting



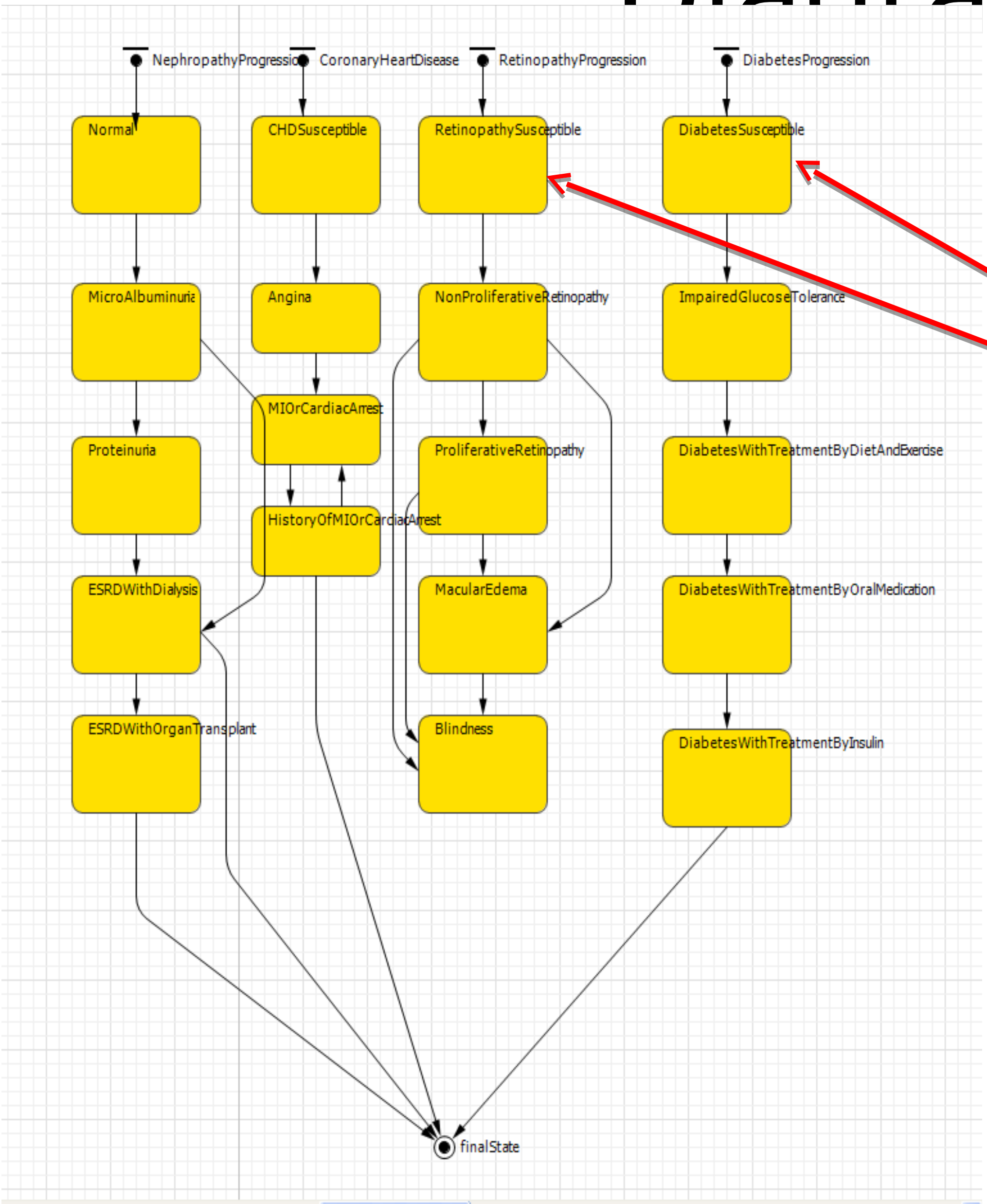
Combinatorial Subscripting: Multi-Dimensional Progression



Parallel Transitions



Parallel State Transition Diagrams



A person is in some particular state with respect to each of these (condition specific) state transition diagrams

This requires representing combinations of possibilities in an aggregate model

Capturing Heterogeneity in Individual-Based vs. Aggregate Models

- Consider the need to keeping track a new piece of information for each person (with d possible values)
  - E.g. age, sex, ethnicity, education level, strain type, city of residence, etc.
- Aggregate Model: Add a subscript
  - This multiplies the model size (number of state variables into which we divide individuals) by *d*!
- Individual based model: Add field (variable/param)

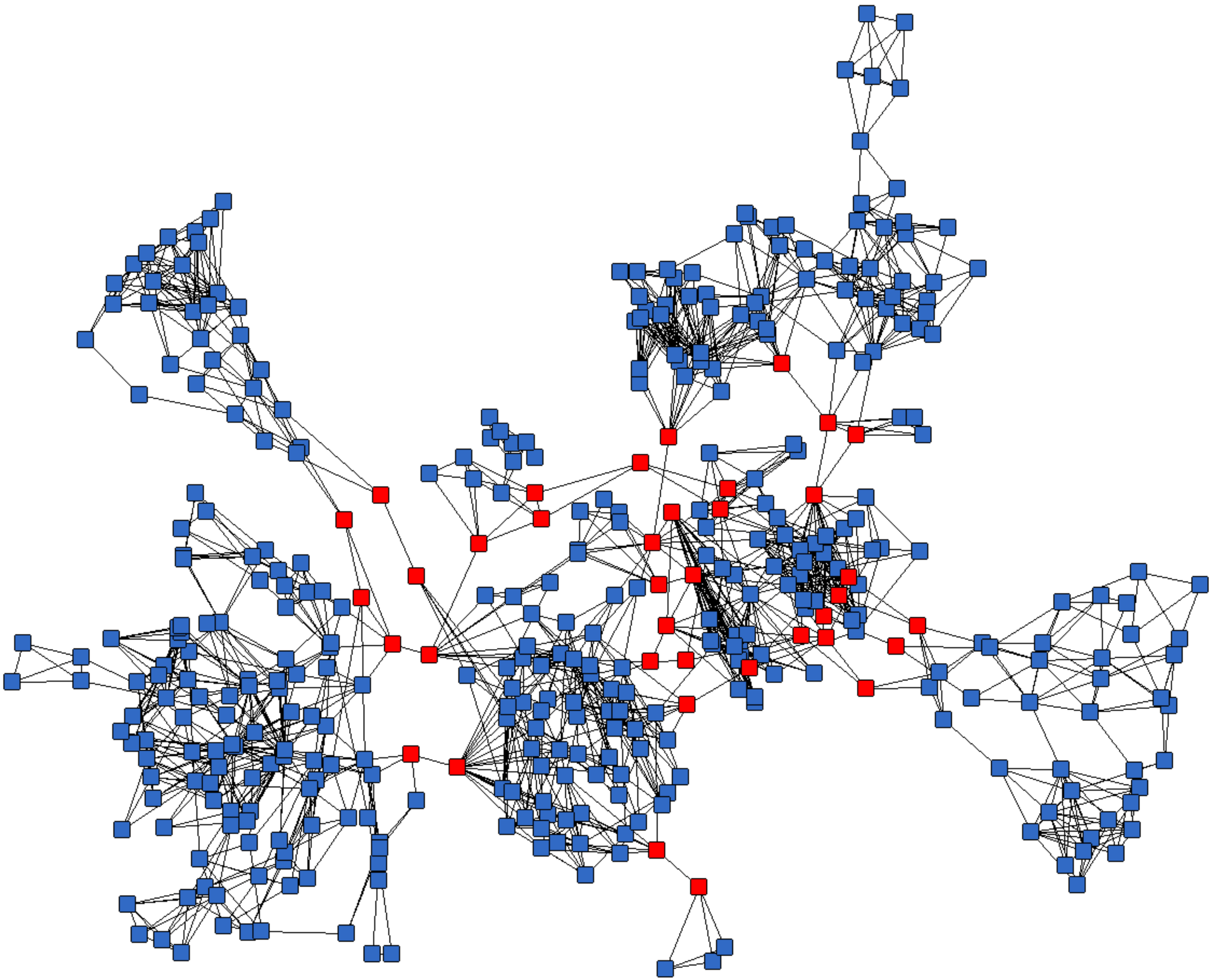
Desired: Flexibility in Representing Heterogeneity

- It is desirable to capture heterogeneity in a flexible fashion.
- More judicious exploration of whether to represent heterogeneity
  - Examine whether some observed covariation might simply be due to colinearity
    - Represent added heterogenity dimensions with no causal interaction, see if model covariations matches what is seen in external world
      - e.g. represent age in a TB model, see if rates of LTBI by age in the model match age-specific infection rate observations
  - Try adding in new dimension of heterogeneity & effects, and see if has impact that is both substantive & plausible

Challenges for Model Formulation: Persistent Interaction

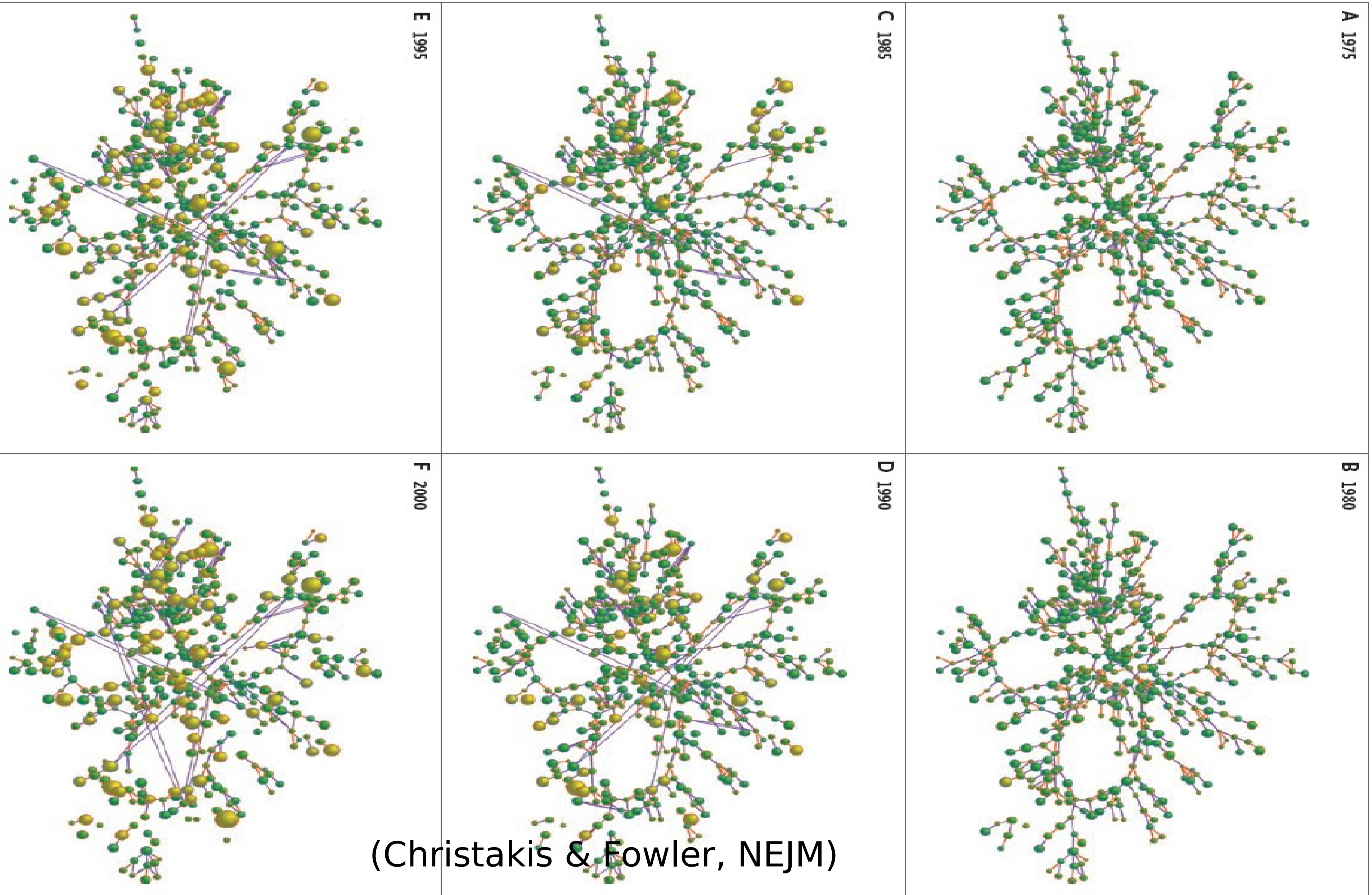
- Network topologies can affect qualitative behavior
- Aggregate representations of network structure are expensive and awkward
- IBM permit expressive, efficient characterization of both dense & sparse networks
- While percolation over many topologies can be simulated in aggregate models, parameter calibration often requires finer-grained simulation

Identifying Bridging Individuals



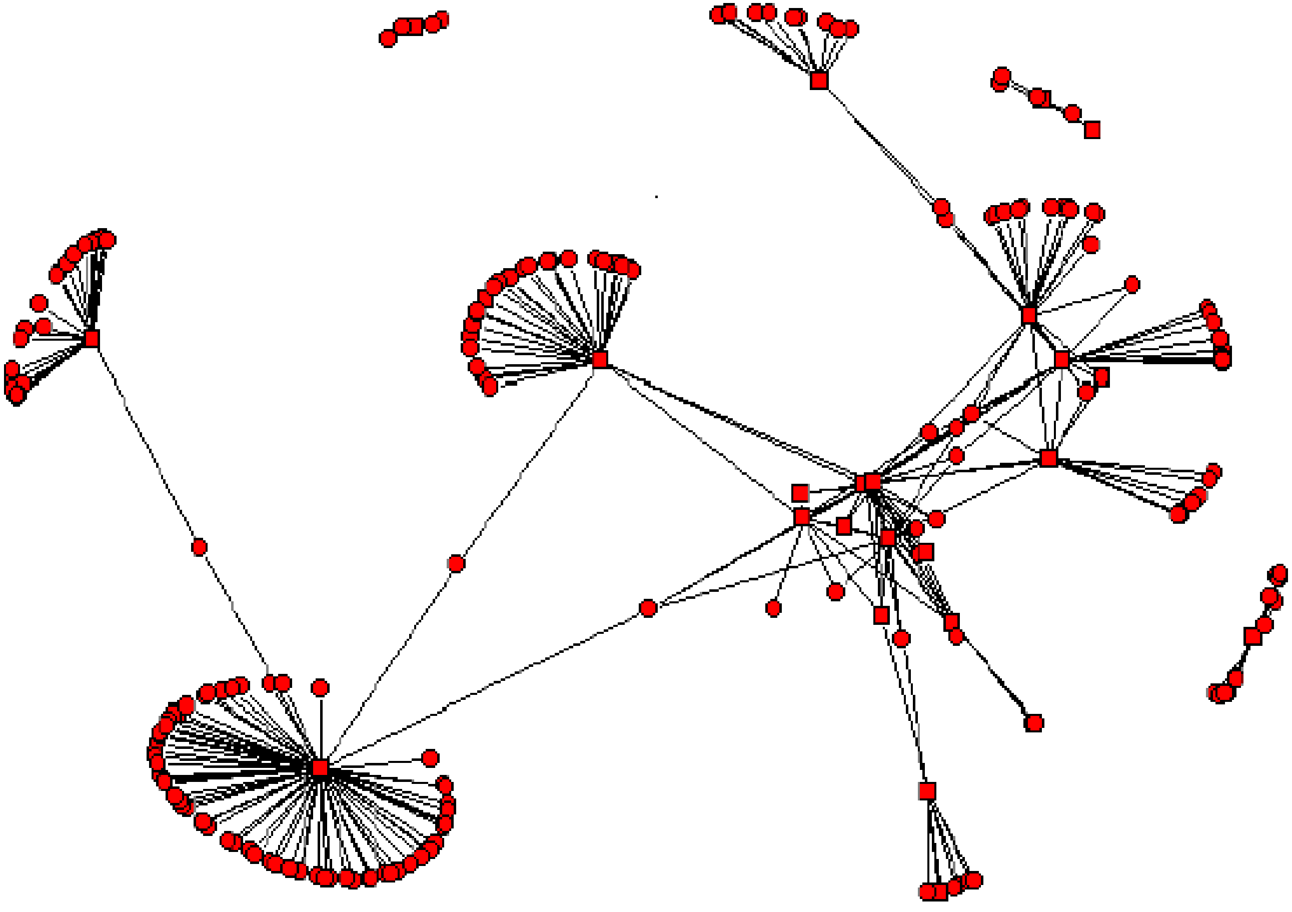
- Preliminary case contact network
- Restricted to nodes of degree 2+
- Data analysis & image: A. Al-Azem

Network Spread of Obesity



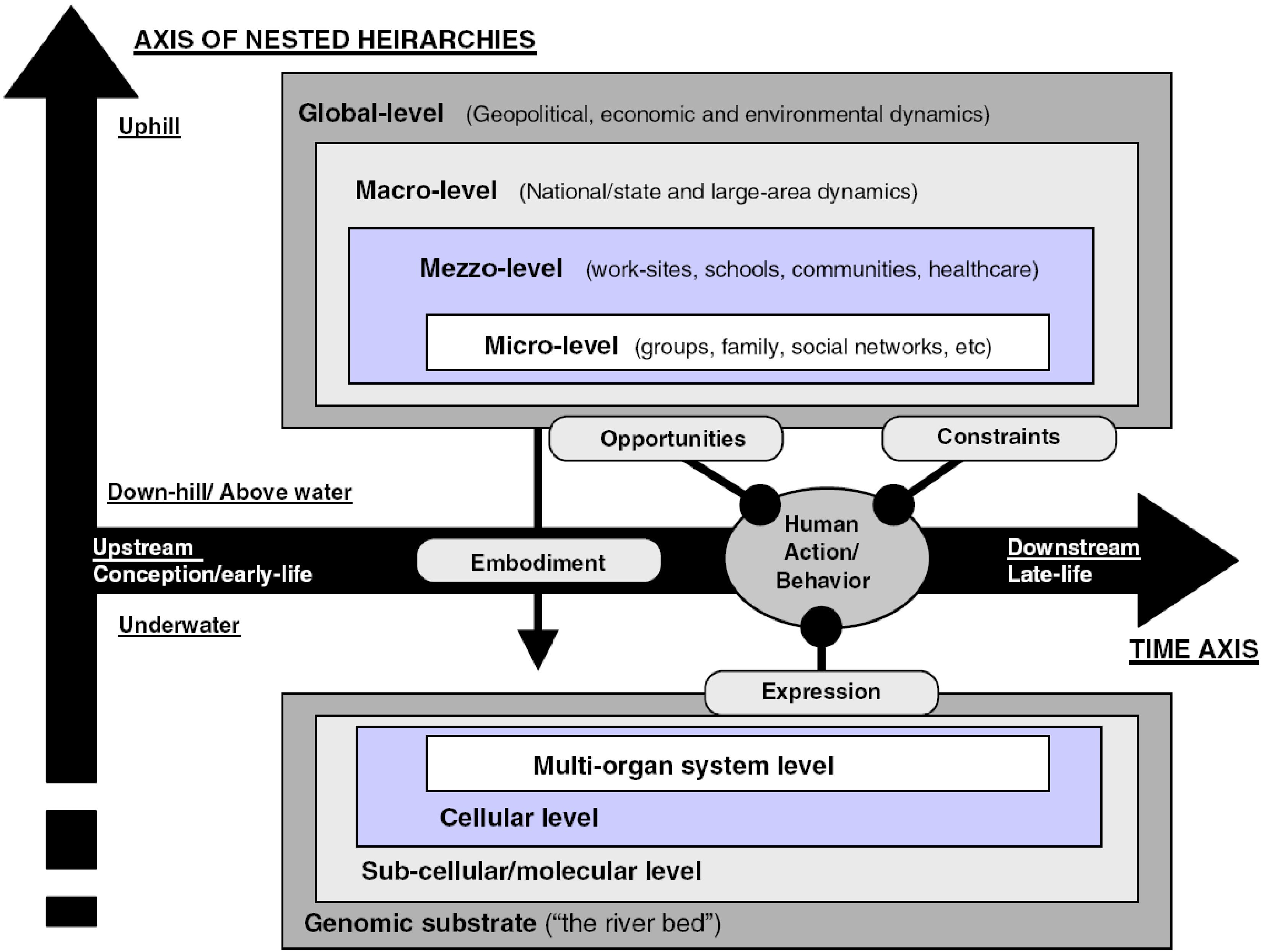
(Christakis & Fowler, NEJM)

TB Infection and Contact Network



Multi-scale Phenomena

- Frequently we are concerned about phenomena on a variety of scales
  - Aggregate societal & policy level
  - Institutional level
  - Individual level
  - Intra-institutional level


AXIS OF NESTED HEIRARCHIES
Uphill
Global-level (Geopolitical, economic and environmental dynamics)
Macro-level (National/state and large-area dynamics)
Mezzo-level (work-sites, schools, communities, healthcare)
Micro-level (groups, family, social networks, etc)
Opportunities
Constraints
Down-hill/ Above water
Upstream
Conception/early-life
Embodiment
Human Action/ Behavior
Downstream
Late-life
Underwater
TIME AXIS
Expression
Multi-organ system level
Cellular level
Sub-cellular/molecular level
Genomic substrate ("the river bed")

Network Medicine — From Obesity to the “Diseasome”

Albert-László Barabási, Ph.D.

N ENGL J MED 357;4 WWW.NEJM.ORG JULY 26, 2007

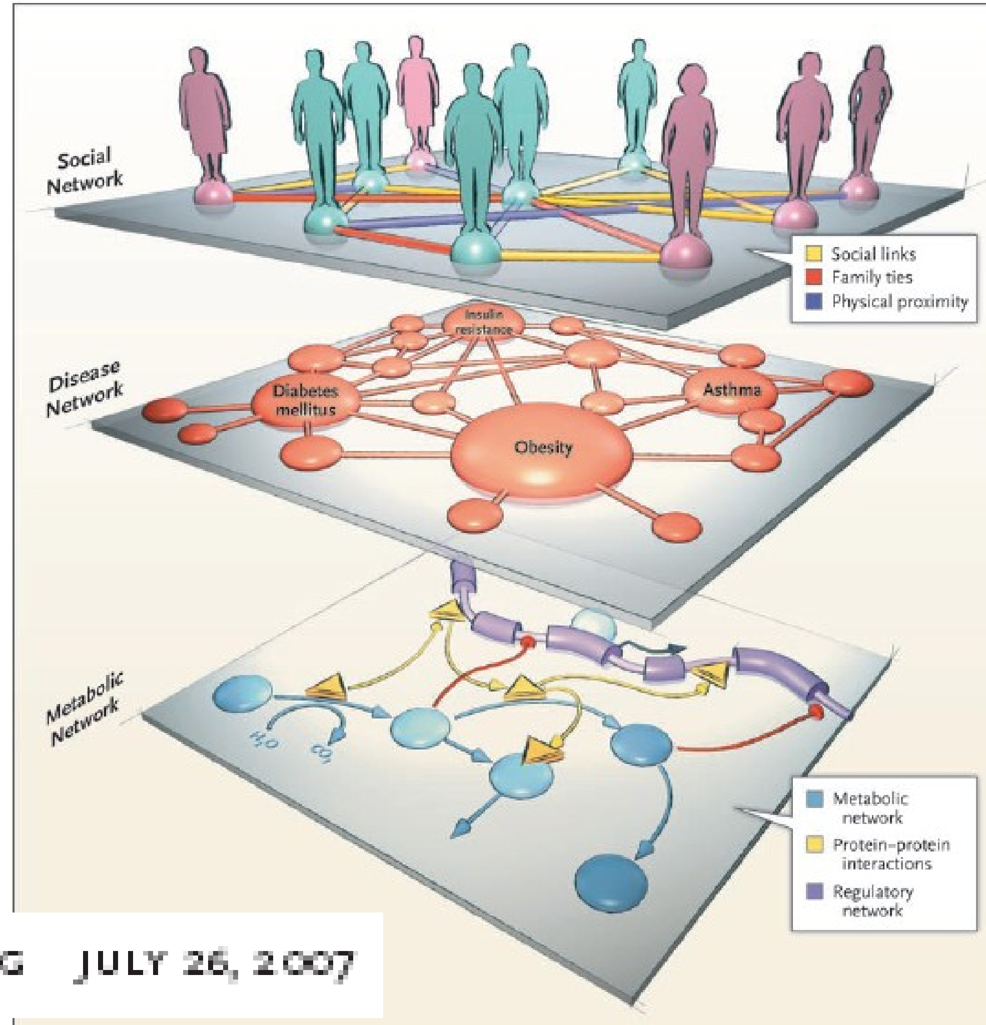
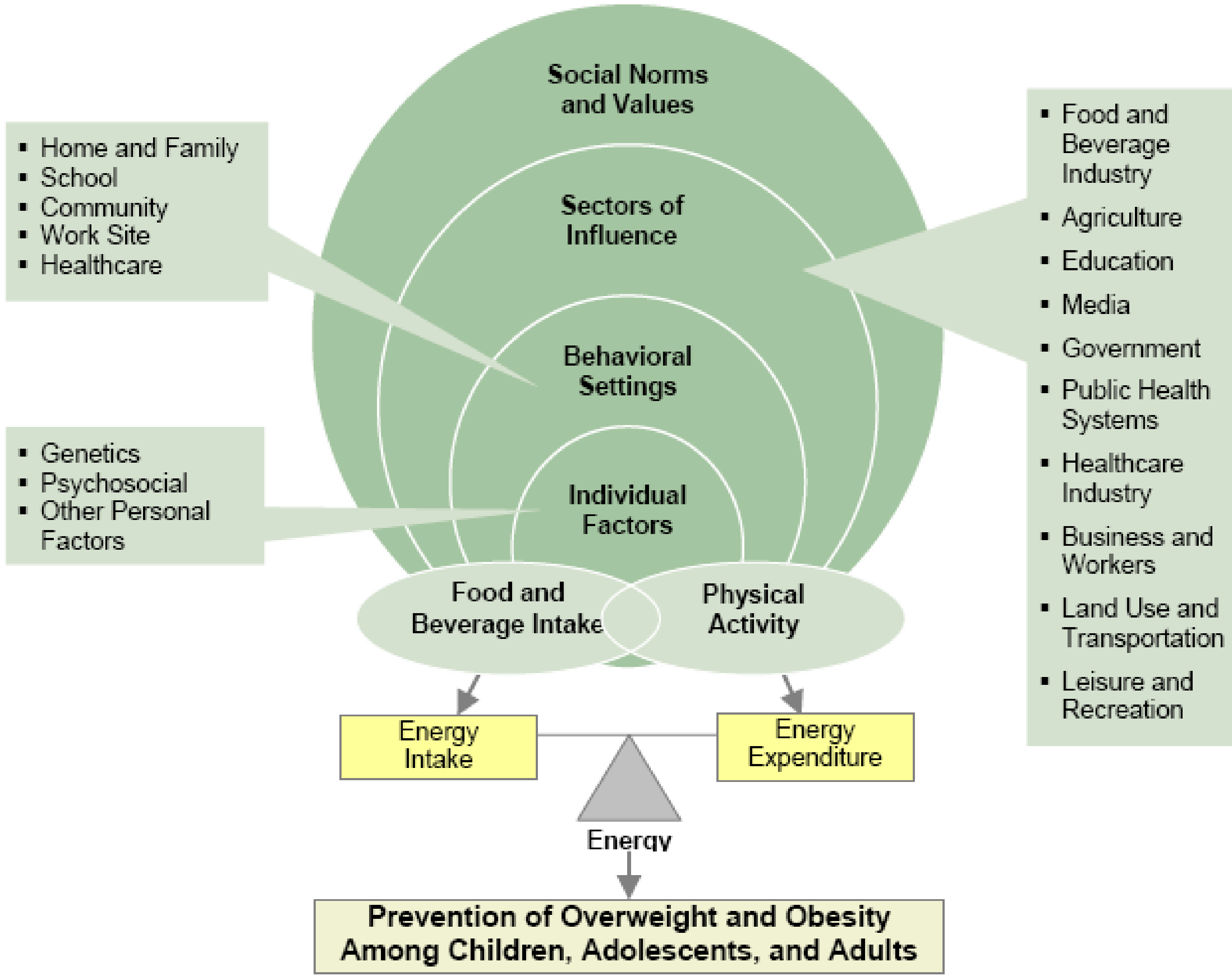


Figure 1. Complex Networks of Direct Relevance to Network Medicine.

Although they are often treated separately, most human diseases are not independent of each other. Many diseases are associated with the breakdown of functional modules that are best described as subnetworks of a complex network connecting many cellular components. Therefore, an understanding of the functionally relevant genetic, regulatory, metabolic, and protein–protein interactions in a cellular network will play an important role in understanding the pathophysiology of human diseases (bottom layer). One way to visualize the ensuing potential interrelationships among human diseases is to construct a disease network (middle layer) in which two diseases are connected if they have a common genetic or functional origin. For example, on the basis of our current knowledge of disease genes, obesity is connected to at least seven other diseases such as diabetes, asthma, and insulin resistance, since genes associated with these diseases are known to affect obesity as well. The third network of key importance to human disease is the social network, which encompasses all human-to-human interactions (e.g., familial, friendship, sexual, and proximity-based contacts) that play a role in the spread of pathogens (top layer). These networks also have an important role in the spread of obesity. Efforts to understand the interactions between the cellular, disease, and social networks are part of network medicine, which aims to quantify the complex interlinked factors that may contribute to individual diseases.



Note: Adapted from IOM 2005 (Figure 3-2, p.85)

Science

Homer et al 2006

Finer Grained Policy Planning

- In the presence of networks or non-well-mixed populations, big difference in effects of targeted interventions
- e.g.
  - Targeted intervention within scale-free network
  - Impact of incentives on competition and cooperation
  - Impact of road structure on traffic jams

Stochastics

- Aggregate models are most commonly deterministic
  - There are important exceptions – models with stochastic flows
- Frequently, capturing the stochastics is advantageous
  - Compare degree of variability to that seen in historic empirical observations
  - See degree of spread in policy results

Parameterization & Calibration

- Individual-based models have many parameters
  - Estimating all of the parameters can require much effort
  - Calibration generally underdetermined (large # of possible sets of parameter values that could calibrate well)
  - May need to make simplifying assumptions
- Pronounced individual-level stochastics frequently require Monte-Carlo calibration

Individual-Based Model Performance Scaling

- Performance varies with population size
  - Large populations impose high computational resource demands
  - Scaling can be superlinear (e.g. O(n2) connections to consider)
  - This can frequently lead to simulations taking minutes at the least, commonly hours or even days
- Desire to characterize stochastic nature of individual-level behavior typically requires Monte Carlo approaches
  - This can lead to days or weeks to complete

Memoryless vs. Memoryful Processes

- ODE models can adequately capture only *memoryless transition processes out of a stock*
  - Stocks treated as “well-mixed”: Transition probability does not depend on residence time
  - Memoryful processes can be approximated, but requires changing model structure to reflect a simple functional relationship (e.g. *n*th order delays)

Individual vs. Aggregate Models: Necessary Tradeoffs

	Transition Generality	Network Representation	Ease of Calibration	Performance Issues				Capturing Learning Adapation	Using Policies & Calibration based on History Information	Capacity for Multi-Scenario Reasoning
				Basal	Scaling with Population	Scaling with Heterogeneity	Need for Stochastics/Monte Carlo			
Individual Based Models	++	++	Varies			++		++	++	
Aggregate Models		+	++	+++	++		+			++

- Both individual-level and aggregate modeling have *inherent* and non-trivial *tradeoffs*
- Both approaches likely to retain strong

The (Current) Package Deal

- **ABM (AnyLogic)**
  - Supports individual-based or aggregate
  - Trajectory files have limited support
  - Both discrete & continuous rules & states
  - Primarily imperative specification
  - Algorithmic (imperative)
  - Little/No explicit mathematical

- **Traditional system dynamics packages**
  - Supports individual-based or aggregate
  - Trajectory files well supported
  - Poor discrete rule support
  - Declarative specification
  - Equational notation & reasoning
  - Explicit mathematical semantics
  - Monolithic
  - Limited metadata (unit checks)

Specifications: How vs. What

- Post-dynamo system dynamics package uniformly support *declarative* specifications
  - User is shielded from many “how” mechanisms
    - Ordering of equation calculation
    - Temporary storage
    - Timestep-to-timestep iteration
- Agent-based packages traditionally focused on “how”
  - Exceptions: AnyLogic (substantial declarative mechanisms), SDML

Implications
Specifications: How vs. What

- Declarative model specifications generally facilitate
  - Initial specification
  - Transparency
  - Communication
  - Easier modification
  - Reasoning about
  - Machine performance optimization
- Cost: Less flexible if outside of normally supported features

Equational vs. Imperative Specification

- This issue is related to – but distinct from – the “what vs. how” issue
  - Somewhat less fundamental in implications
  - Equation-based specifications will generally be declarative
  - Algorithmic specifications can be either declarative or non-declarative
    - Functional languages overlap the “what” and “how”
      - More transparent: Supporting equational reasoning
- Computationally equivalent
- Algorithmic expressiveness higher for some complex rules

Explicit vs. Implicit Mathematical Semantics

- All specifications can be formalized mathematically (e.g. language semantics)
- Only some frameworks are explicitly specified in mathematical terms
  - Traditional system dynamics is explicit
  - Agent-based packages typically have incomplete (or no) explicit math. formulation
    - Imperative code (e.g. Java, Objective C, C++, stateful LISP etc.) rule out equational reasoning
- Equational algorithmic mathematical frameworks possible
- Subissue: Choice of mathematical framework (ODE, PDE, etc.)
  - NB: individual-level semantics will generally include discrete components that are represented continuously in agg. models

Implications
Explicit Mathematical Specification

- Explicit mathematical formulation helps
  - Transparency & understandability
  - Formal analyses (e.g. linearization)
    - Aids in formal analysis of behavior, policy effectiveness
  - Equilibration (and sometimes calibration)
  - Building confidence in model realization
  - Generalize results
  - By Permitting translation to closed-form analysis

Means of Abstraction/Decomposition

- Modeling frameworks use different means of abstraction
  - Most: Object-oriented class hierarchies
  - Matlab: Procedural function hierarchies
  - Most SD packages do not support model modularity

Implications

Support for Abstraction and Modularity

- Simplify reasoning during model formulation
- Foster reuse
- Simplify model modification

Support for Data Semantics

- Data used in models is not merely numbers – it has meaning that reflects its context e.g.
  - Source
  - Age
  - Degree of uncertainty
  - Measurement protocol
- Meta-data specific algebras can aid in model building, validation by capturing aspects of domain not expressed in data
- Traditional system dynamics packages support basic unit metadata

Implications
Support for Data Semantics

- Often useful in formal reasoning
  - Unit & dimension information
    - Unit & dimensional checks
    - Design of scale models
    - Cross-checking analytic results
  - Uncertainty
  - Sampling information
- Valuable for the modeler
  - Understanding varying data pedigrees
  - Bookkeeping

Support for Continuous and Discrete Dynamics & State

- Continuous dynamics express *flow*
- Discrete dynamics can express sudden (non-analytic) changes in behavior
- Continuous&discrete dynamics required for description of many dynamic systems
  - Physical constraints: Exhaustion of a stock, boundary conditions
  - Human rules (do *x* if reservoir level below *y*)

Implications
Support for Continuous and Discrete Dynamics & State

- Support for hybrid modeling helps
  - Avoid simulation artifacts
  - Simplify model expression
- NB: In an *aggregate* model, many discrete states may be aggregated into corresponding continuous state variables

Methodological Implications of Choices (From my experience)

	Transparency	Performance	Ease of Creation	Generality	Analyzability/ Understanding	Ease of Parameterization	Ease of Calibration& Validation	Accuracy	Model Breadth	Scalability (Population)	Scalability (Heterogeneity)	Modifiability
Aggregate	+-	++	+	-	++	+	++	--	++	++	-	+
Discrete & Continuous	+		+	++	+			+				
Declarative	++	+	++		+				++			++
Equational	++		+		+		+		++			+
Explicit Math	++		+		++		++					++
Modular	++		+						+			++
Longitudinal reporting/ visualization					++							
Metadata Support	++		-		+					++		+

Current Package Deal: Modeling Implications (From my Perspective)

	Transparency	Performance	Ease of Creation	Generality	Analyzability/ Understanding	Ease of Parameterization	Ease of Calibration	Accuracy	Model Breadth	Scalability (Population)	Scalability (Heterogeneity)	Modifiability
TSD	++	++	++	+	++	++	++	+	++	++	-	+
ABM	+		+	++		+	+	++			++	+

Current ABM and TSD packages both have important advantages

Central Points: Present

- There are currently many differences between agent-based and traditional SD (TSD) approaches
- The differences have significant impact on model results & the modeling process
- Both traditions have strong advantages when addressing different types of problems (see diagrams)
- Painting broad-brushed dichotomies between traditional system dynamics and agent-based modeling obscures the fundamental commonalities

Suggestions: Making Modeling Choices in the Present

- Individual-level modeling a good option if it matches problem characteristics: e.g.
  - Understanding complex dynamic implications of networks, heterogeneity
  - Flexibility in desired model heterogeneity
  - Clear description of causal mechanisms, memoryful processes
  - Characterizing fine-grained interventions
  - Understanding heterogeneity in intervention effects
  - Need to match/explain longitudinal data
- Be aware that these advantages presently do come with additional tradeoffs

Agenda

- ✓ Motivations & Context
- ✓ Comparing Aggregate & Individual Based Models
- ✓ Granularity Tradeoffs
- ✓ Tools for individual-based modeling
  - Individual-Based Modelers in SD
  - Individual-based models in Agent-Based tools
- ✓ Other tradeoffs
- Looking forward

Central Points: Looking Forward

- Most current differences reflect important but non-essential methodological choices / tool characteristics
- In the long run, these differences will likely lessen and the choice that will remain is that of model granularity
- Both individual-based models and aggregate models will play important roles in system dynamics
- There are good reasons to use all of individual-based models, aggregate models, and hybrid systems

Hybrid Modeling

- Single model with multiple levels
  - Aggregate environment
  - Disaggregate population(s)
- Calibration of aggregate model from disaggregate
- *Framework flexibility is key advantage*

Iterative Mutual Modeling: What

- Overlapping modeling domains permits cross-validation
  - e.g. Ki allow simulation of well-mixed population
  - Agent-based transitions can be simulated with e.g. binomial transitions
- Refine models based on cross-framework observations
- Create high-level models *for qualitative insight* from lower-level detailed models
  - This allows developing confidence that are capturing important features in high-level models

Iterative Mutual Modeling: Why

- Cross-validation
  - Formulation errors
  - Implementation errors
- Model simplification
- Use of tools with relative strengths
- Cross-calibration

Suggestions: Research

- Diversification of model space
  - Newer systems (e.g. AnyLogic) are big steps in the right direction
- Hybrid modeling
- Semantic enrichments of fine-grained models
- Metadata and semantic algebras
- Improved languages for algorithmic specifications
- Model reuse mechanisms
- Open systems for extensibility
- Metalinguistic mechanisms for customization