Ross: A Networked-Epidemiology WebApp

August 25, 2014

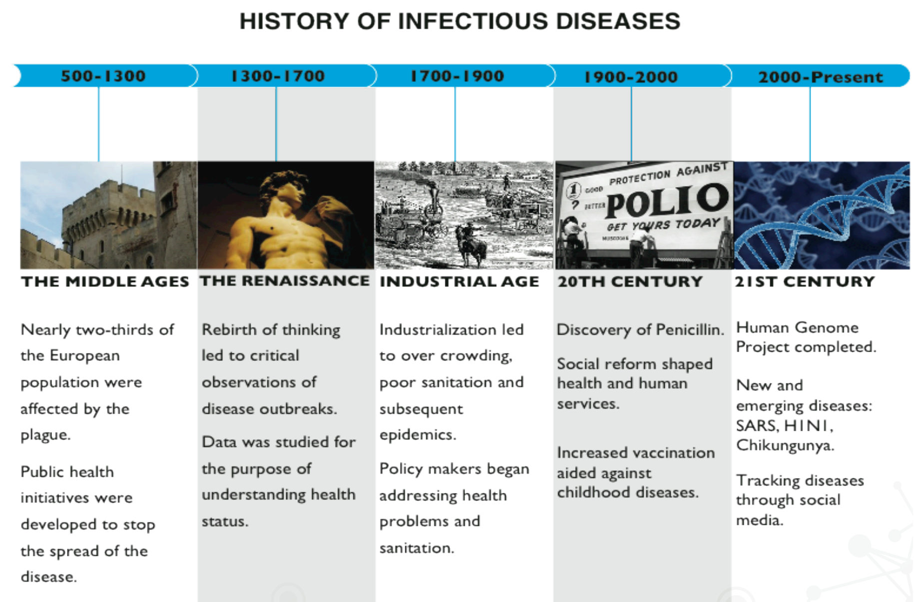
Richard Beckman, Keith Bisset, Jiangzhuo Chen, Bryan Lewis, **Madhav Marathe**, Paula Stretz

Network Dynamics and Simulation Science Laboratory, VBI & Dept. of Computer Science

Virginia Tech

Epidemics and epidemiology in history

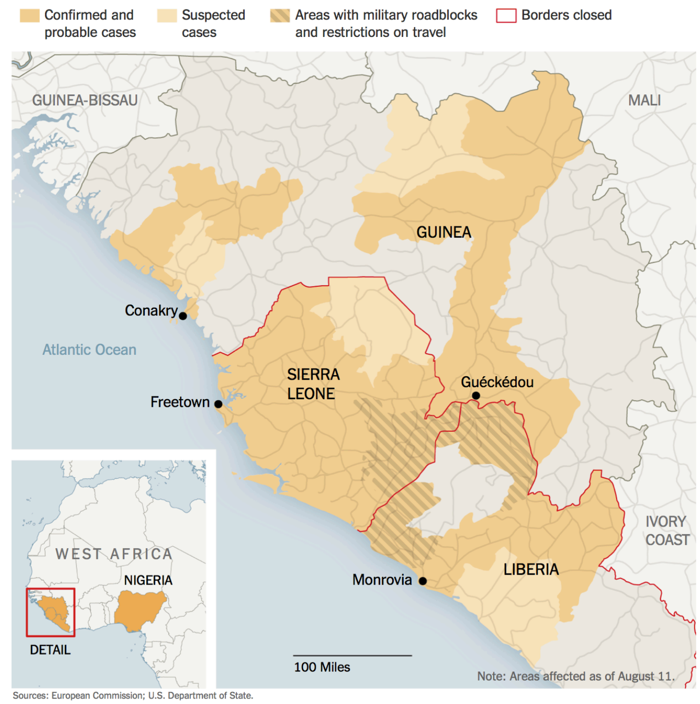
- Good news: Pandemic of 1918 lethality is currently unlikely
  Governments better prepared and coordinated : e.g. SARS epidemic
  But ..
- Planning & response to even a moderate outbreak is challenging: inadequate vaccines/anti-virals, unknown efficacy, hard logistics issues
- Modern trends complicate planning: increased travel, immuno-compromised populations, increased urbanization

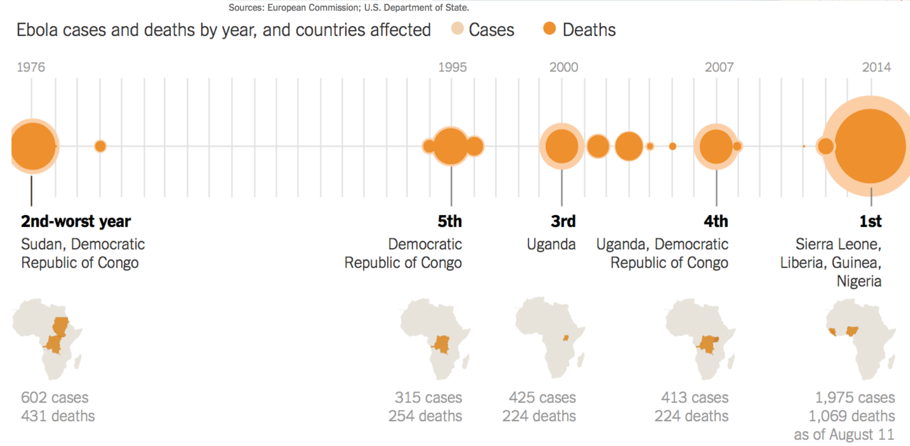


Need for modeling environments to support epidemiologists

Ebola outbreak in Africa

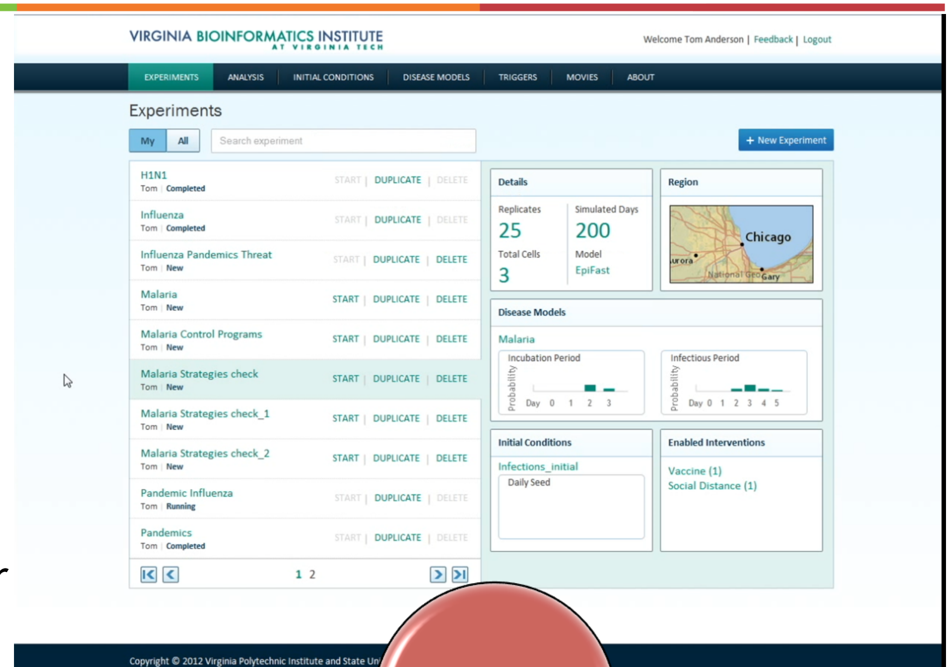
- Largest Ebola outbreak yet: 3 countries; 2000 cases; 1000 deaths.
- Beautifully done NY Times webpage: http://www.nytimes.com/interactive/2014/07/31/world/africa/ebola-virus-outbreak-qa.html
- Questions:
  - *How many people have been infected?*
  - *Where is the outbreak?*
  - *How did it start; tracing the first few cases.*
  - *Chances of getting Ebola in the US?*
  - *How does this compare to past outbreaks?*
  - *How contagious is the virus?*
  - *Why is Ebola so difficult to contain?*
  - *How does the disease progress?*
  - *How is the disease treated?*
  - *Where does the disease come from?*

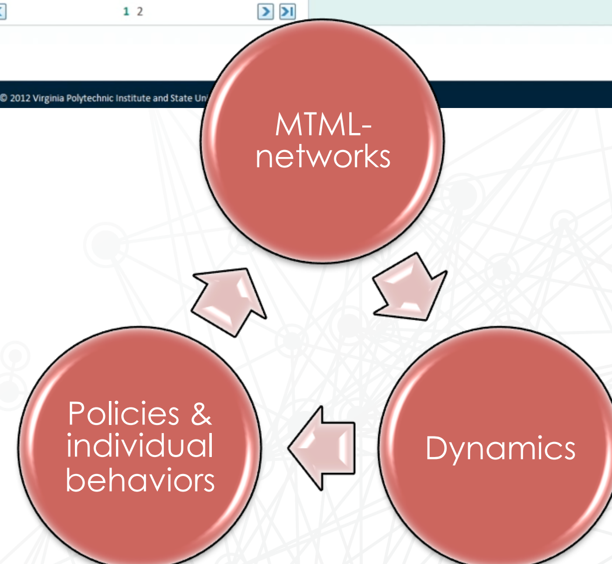




What is Ross

- Ross (Interface to Synthetic Information Systems) is a web based tool to support <u>*real-time epidemic science.*</u>
  - Fineberg and Harvey, Science 2009.
- *Philosophy:* Sophisticated modeling environments should also be easy to by subject matter experts.
  - SMEs have no interest in becoming computer scientists !





Ross supports real-time epidemic science

- Modeling before an epidemic
  - (i) Determine the (non)medical interventions required, (ii) feasibility of containment, (iii) optimal size of stockpile, (iv) best use of pharmaceuticals once a pandemic begins
- Modeling during an epidemic
  - (i) Quantifying transmission parameters, (ii) Interpreting real-time epidemiological trends, and (iii) assessing impact of interventions.

EDITORIAL

Harvey V. Fineberg is president of the Institute of Medicine.

Mary Elizabeth Wilson is associate professor of Global Health and Population at the Harvard School of Public Health and associate clinical professor at Harvard Medical School, Boston, MA.

Epidemic Science in Real Time

FEW SITUATIONS MORE DRAMATICALLY ILLUSTRATE THE SALIENCE OF SCIENCE TO POLICY THAN AN epidemic. The relevant science takes place rapidly and continually, in the laboratory, clinic, and community. In facing the current swine flu (H1N1 influenza) outbreak, the world has benefited from research investment over many years, as well as from preparedness exercises and planning in many countries. The global public health enterprise has been tempered by the outbreak of severe acute respiratory syndrome (SARS) in 2002–2003, the ongoing threat of highly pathogenic avian flu, and concerns over bioterrorism. Researchers and other experts are now able to make vital contributions in real time. By conducting the right science and communicating expert judgment, scientists can enable policies to be adjusted appropriately as an epidemic scenario unfolds.

In the past, scientists and policy-makers have often failed to take advantage of the opportunity to learn and adjust policy in real time. In 1976, for example, in response to a swine flu outbreak at Fort Dix, New Jersey, a decision was made to mount a nationwide immunization program against this virus because it was deemed similar to that responsible for the 1918–1919 flu pandemic. Immunizations were initiated months later despite the fact that not a single related case of infection had appeared by that time elsewhere in the United States or the world (www.iom.edu/swinefluaffair). Decision-makers failed to take seriously a key question: What additional information could lead to a different course of action? The answer is precisely what should drive a research agenda in real time today.

In the face of a threatened pandemic, policy-makers will want real-time answers in at least five areas where science can help: pandemic risk, vulnerable populations, available interventions, implementation possibilities and pitfalls, and public understanding. Pandemic risk, for example, entails both spread and severity. In the current H1N1 influenza outbreak, the causative virus and its genetic sequence were identified in a matter of days. Within a couple of weeks, an international consortium of investigators developed preliminary assessments of cases and mortality based on epidemic modeling.*

Specific genetic markers on flu viruses have been associated with more severe outbreaks. But virulence is an incompletely understood function of host-pathogen interaction, and the absence of a known marker in the current H1N1 virus does not mean it will remain relatively benign. It may mutate or acquire new genetic material. Thus, ongoing, refined estimates of its pandemic potential will benefit from tracking epidemiological patterns in the field and viral mutations in the laboratory. If epidemic models suggest that more precise estimates on specific elements such as attack rate, case fatality rate, or duration of viral shedding will be pivotal for projecting pandemic potential, then these measurements deserve special attention. Even when more is learned, a degree of uncertainty will persist, and scientists have the responsibility to accurately convey the extent of and change in scientific uncertainty as new information emerges.

A range of laboratory, epidemiologic, and social science research will similarly be required to provide answers about vulnerable populations; interventions to prevent, treat, and mitigate disease and other consequences of a pandemic; and ways of achieving public understanding that avoid both over- and underreaction. Also, we know from past experience that planning for the implementation of such projects has often been inadequate. For example, if the United States decides to immunize twice the number of people in half the usual time, are the existing channels of vaccine distribution and administration up to the task? On a global scale, making the rapid availability and administration of vaccine possible is an order of magnitude more daunting.

Scientists and other flu experts in the United States and around the world have much to occupy their attention. Time and resources are limited, however, and leaders in government agencies will need to ensure that the most consequential scientific questions are answered. In the meantime, scientists can discourage irrational policies, such as the banning of pork imports, and in the face of a threatened pandemic, energetically pursue science in real time.

– Harvey V. Fineberg and Mary Elizabeth Wilson

10.1126/science.1176297

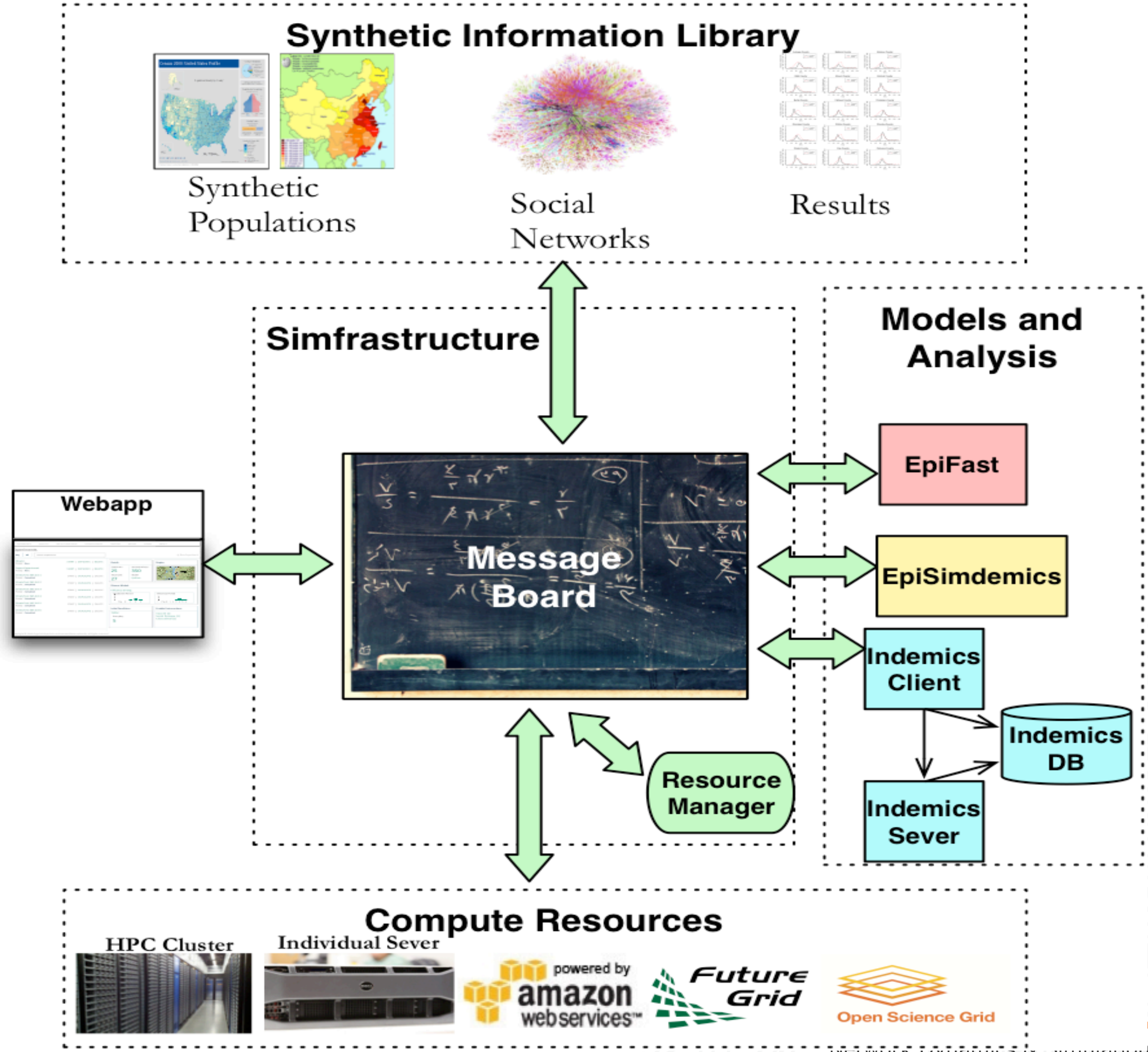
*C. Fraser *et al.*, *Science* 11 May 2009 (10.1126/science.1176062).

CREDITS: (LEFT) ED LAUSCH; (RIGHT) JUPITERIMAGES

Ross Features

- Design, execute, analyze multi replicate, multi cell epidemiology experiments
- Specify social network, disease models, multiple parameterized interventions
- Real-world interventions:
  - Pharmaceuticals (e.g., Vaccination, Antivirals)
  - Work/School closure
  - Sequestration
  - Can be applied to specific subpopulations (school age children, critical workers)
  - Dynamic triggers (e.g., % infected children for school closure)
- Interactive Analysis
  - Automatic ordering of interventions by effect size

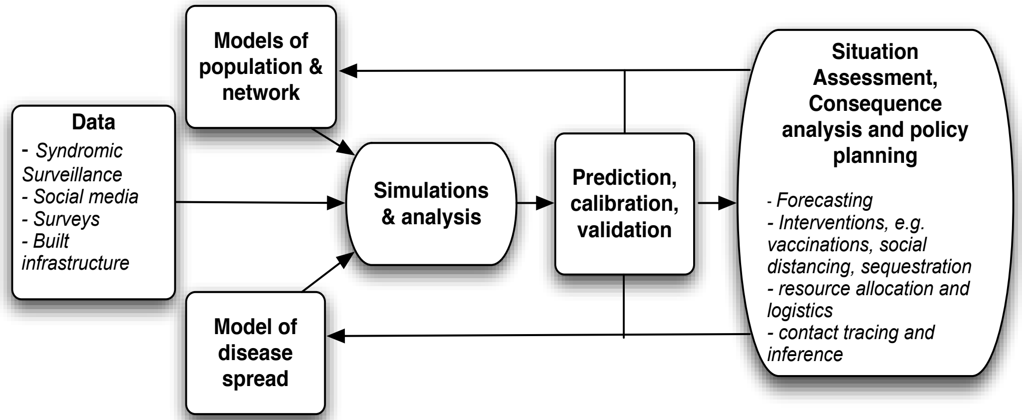
Ross Architecture

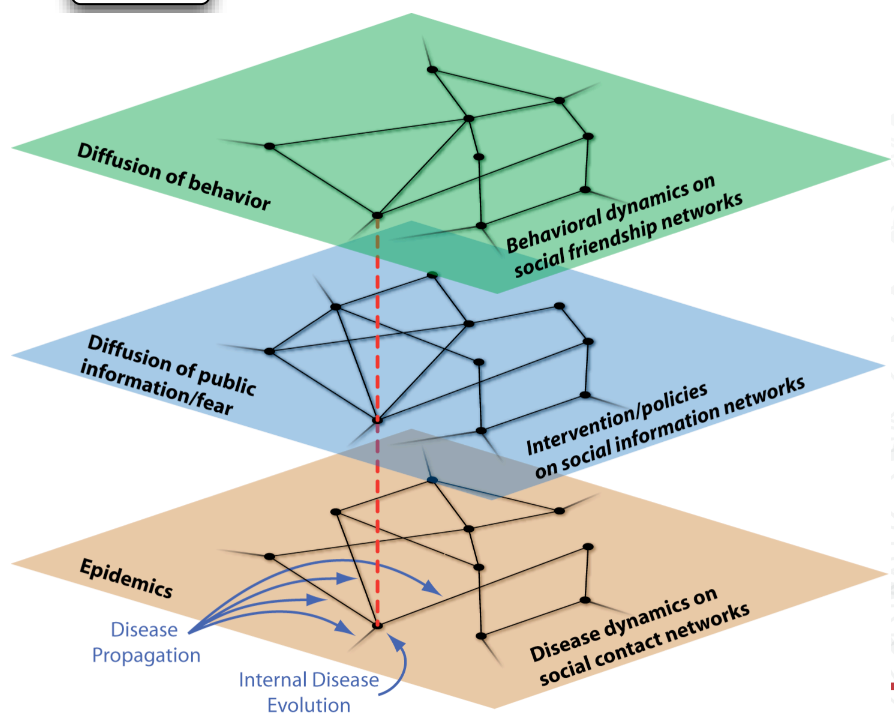


VirginiaTech
Virginia Bioinformatics Institute

Ross supports networked epidemiology

- Supports complex experiments pertaining to epidemic dynamics, control and optimization over realistic but synthetic national scale social contact networks
- Supports multi-theory multi-network representations of the complex system – <u>*co-evolution*</u> of networks, epidemics, behavior and policies can be undertaken



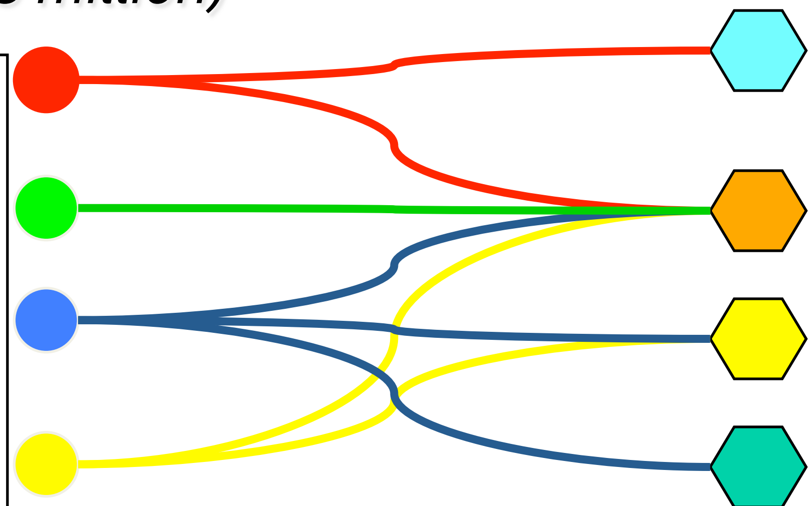


Supports analysis over dynamic social contact networks

People (8 million)

Locations (1 million)

Vertex attributes:
- age
- household size
- gender
- income
- …



Vertex attributes:
- (x,y,z)
- land use
- …

Edge attributes:
- activity type: shop, work, school
- (start time 1, end time 1)
- (start time 2, end time 2)
- …

Supports large and realistic social contact networks

Input files

- Population data
- Scenario
- Disease model
- Runtime configuration

Output files

- Disease
- Dendrogram
- Social net

Population data selected for study

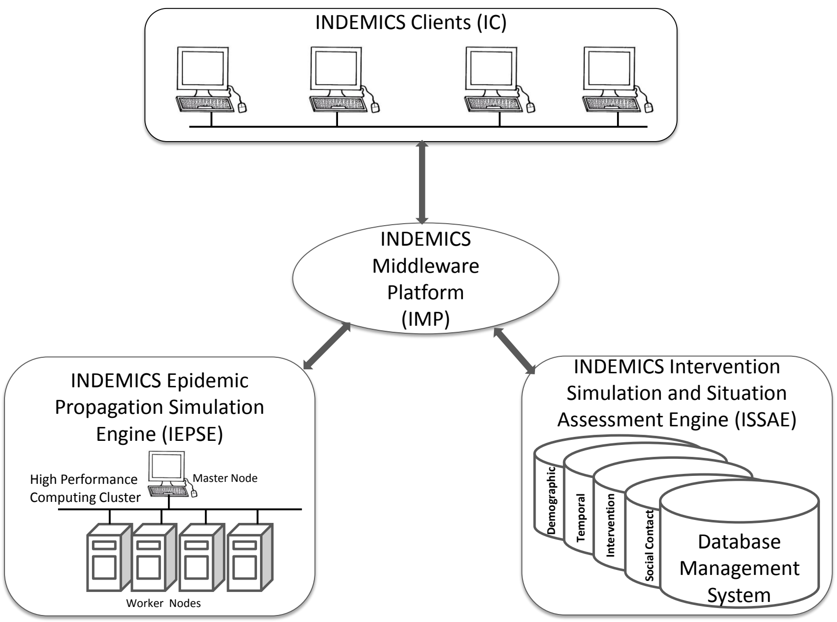
data	visits	people	locations
US	1,541,367,574	280,397,680	71,705,723
CA	183,858,275	33,588,339	7,178,611
NY	98,350,857	17,910,467	4,719,921
MI	52,534,554	9,541,140	2,490,068
NC	47,130,620	8,541,564	2,289,167
IA	15,280,731	2,766,716	748,239
AR	14,803,256	2,685,280	739,507
WY	2,756,411	499,514	144,369

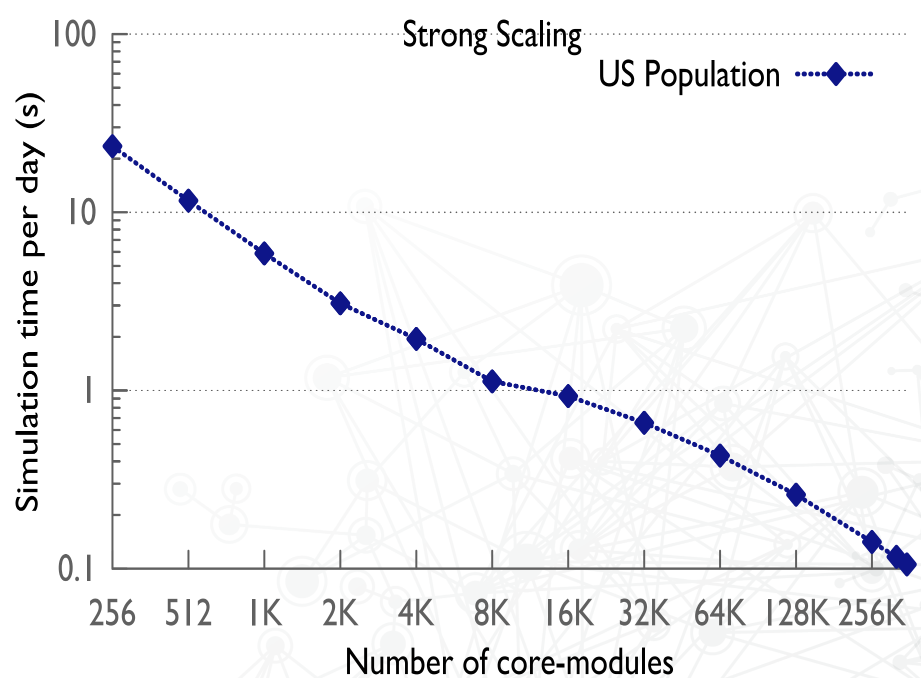
Table : Population data[3] based on various sources including 2009 American Community Survey. US includes 48 contiguous states and DC.

[3] C. Barrett et al., "Generation and Analysis of Large Synthetic Social Contact Networks," WSC09.

Ross supports HPC simulation engines

- *EpiSimdemics*: highly scalable HPC simulation: scales gracefully over 350K cores
- *Epifast*: a consumer-producer style simulation
- *Indemics*: database driven interactive simulation

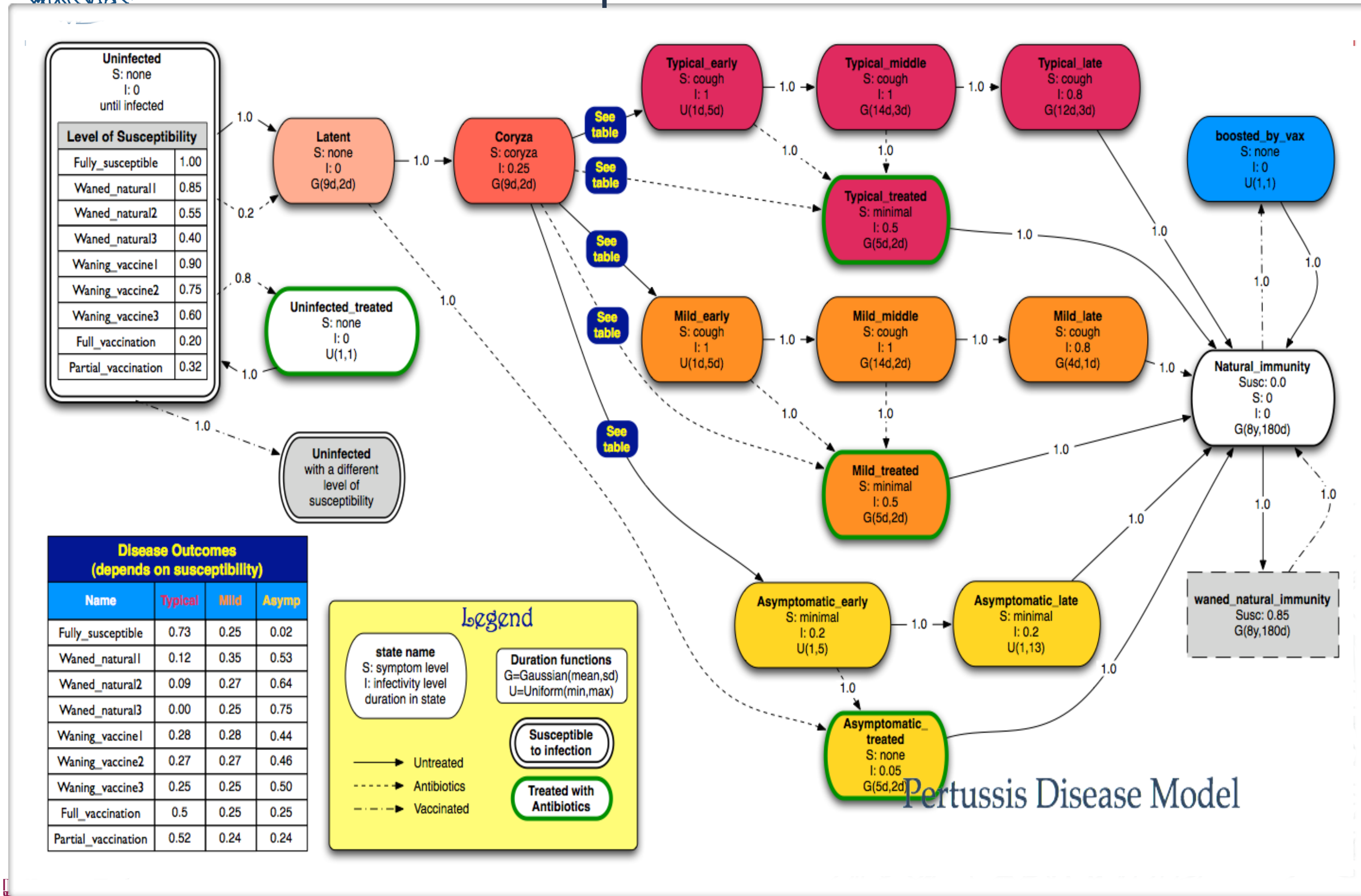




Supports complex interventions

component	specification	example
trigger	scheduled threshold	winter break Dec. 21–Jan. 3 prevalence > 0.05
subpopulation	demographic geographic social health combination of above	seniors people in Fairfax county, Virginia critical workers diagnosed (flu) people diagnosed seniors in Fairfax county
action	 vaccination antiviral school closure work closure isolation combination of above	effects: enhance immunity reduce symptoms or infection probability remove school activities and in-school contacts remove work activities and at-work contacts remove all contacts combined effects

Supports detailed disease representation



Level of Susceptibility	
Fully_susceptible	1.00
Waned_natural1	0.85
Waned_natural2	0.55
Waned_natural3	0.40
Waning_vaccine1	0.90
Waning_vaccine2	0.75
Waning_vaccine3	0.60
Full_vaccination	0.20
Partial_vaccination	0.32

Name	Typical	Mild	Asymp
Fully_susceptible	0.73	0.25	0.02
Waned_natural1	0.12	0.35	0.53
Waned_natural2	0.09	0.27	0.64
Waned_natural3	0.00	0.25	0.75
Waning_vaccine1	0.28	0.28	0.44
Waning_vaccine2	0.27	0.27	0.46
Waning_vaccine3	0.25	0.25	0.50
Full_vaccination	0.5	0.25	0.25
Partial_vaccination	0.52	0.24	0.24

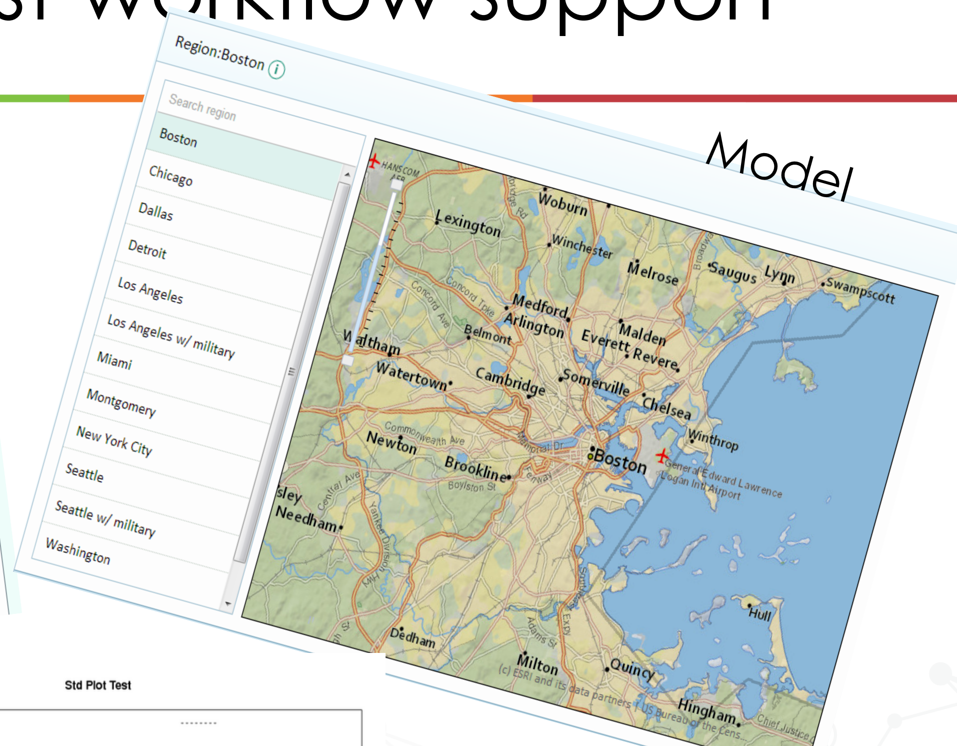
Bigdata challenge: Example of a 20 Cell, 20 replicate design

Area	Population (Nodes in Millions)	Interactions (Edges in Millions)	Input Data (GB)	Output Data (GB)	Compute time days (2000 cores)	Compute Time, Days (300K cores)
West Coast	43	2500	35	3	6	.1
East Coast	75	3750	60	5	10	.2
Entire USA	280	14,000	225	20	43	.8
North America	518	29,000	416	37	80	1.6
Global Population	6500	325,000	5223	464	1000	20

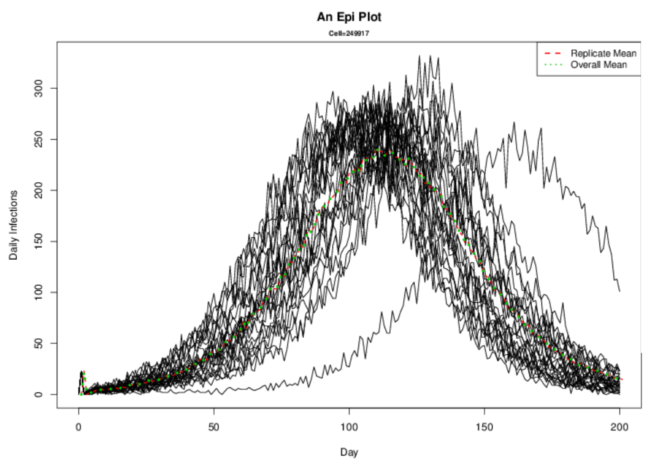
Complete analyst workflow support

Hypothesize

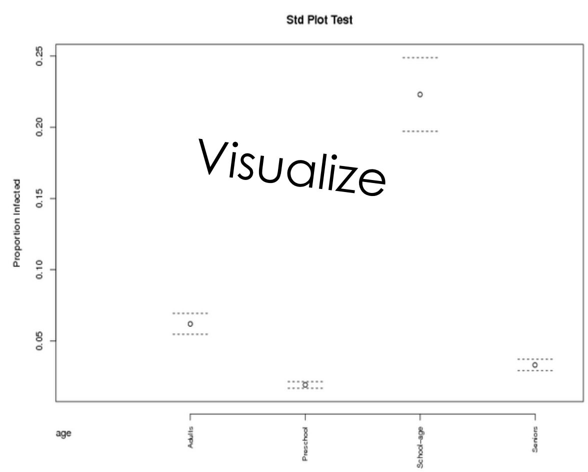
Model

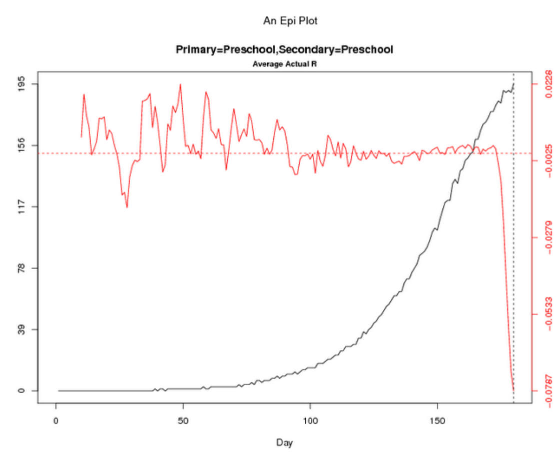


Predict, Decide



Visualize





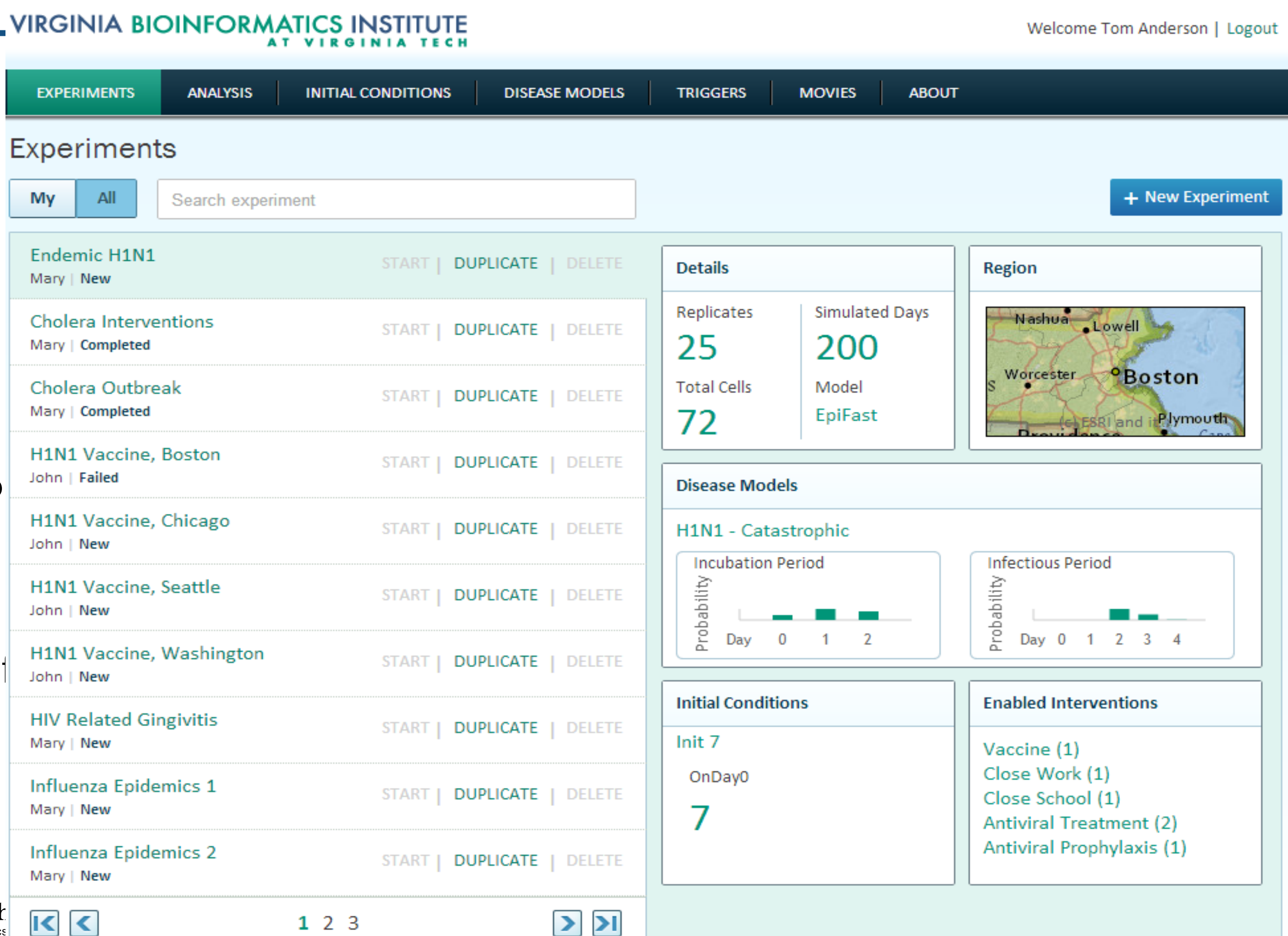
Experiments dashboard

Filter

Search

Executio
Status

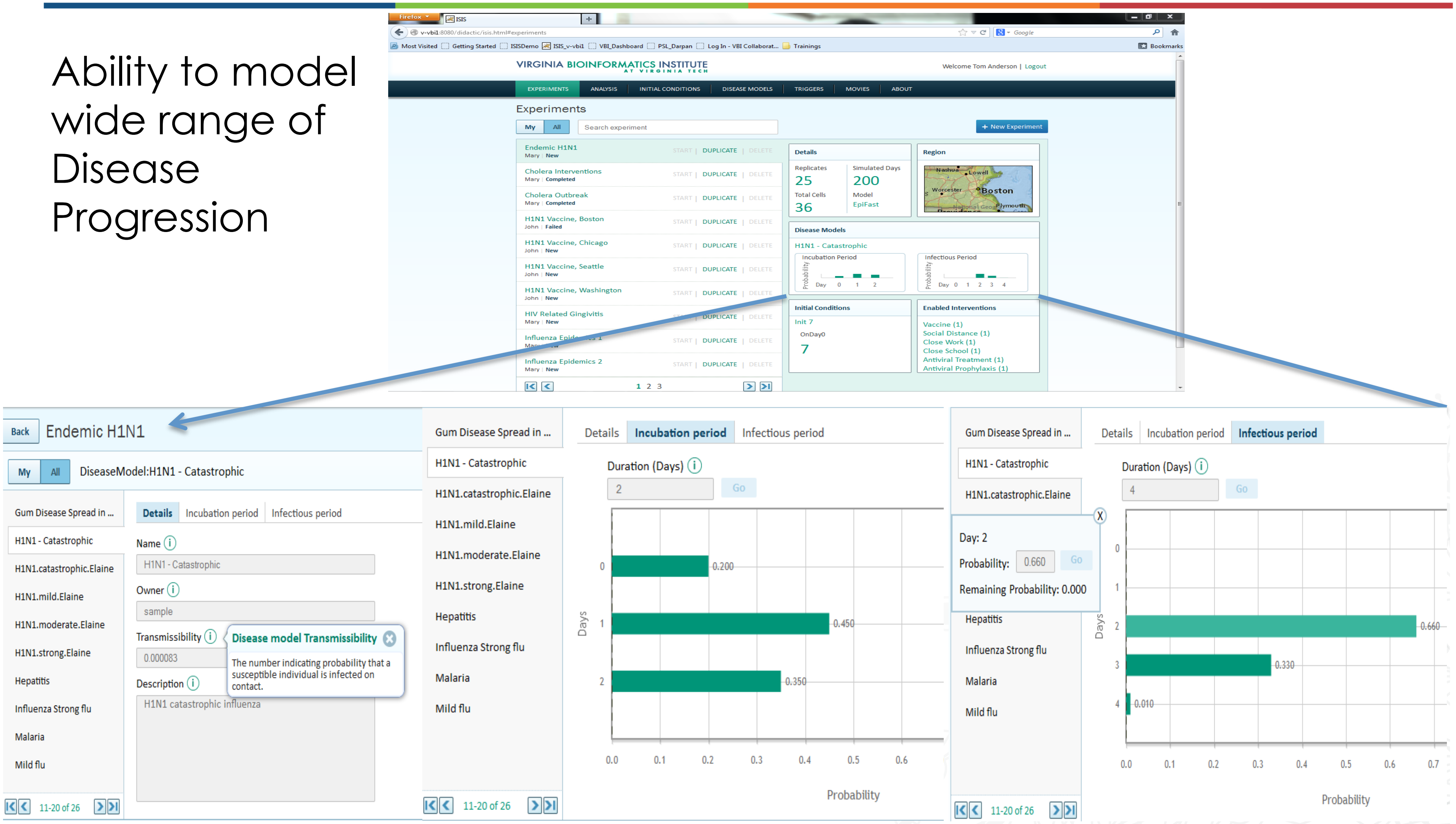
History o
Sessions



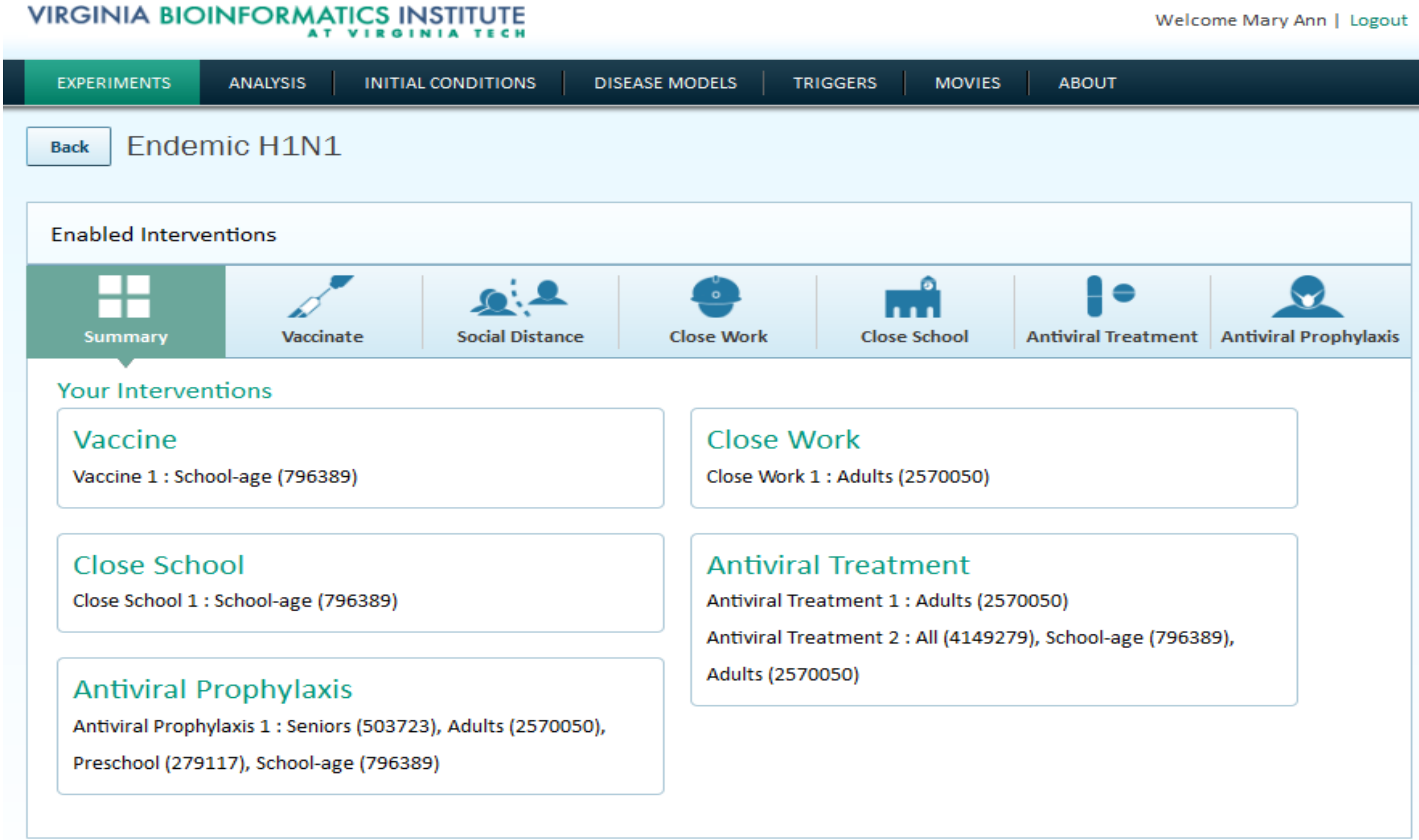
VirginiaTech
Virginia Bioinformatics

Disease models

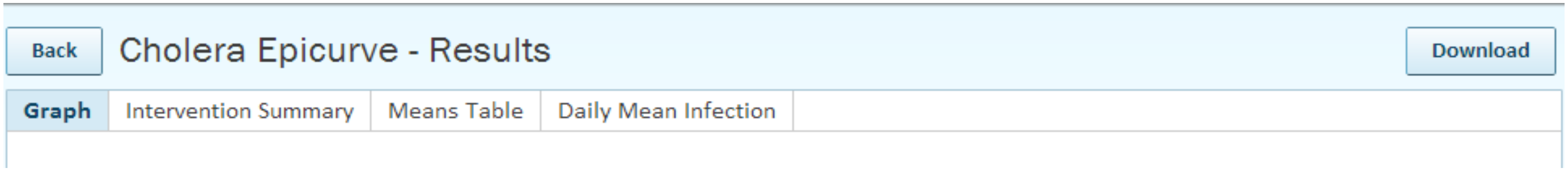
Ability to model wide range of Disease Progression



Interventions

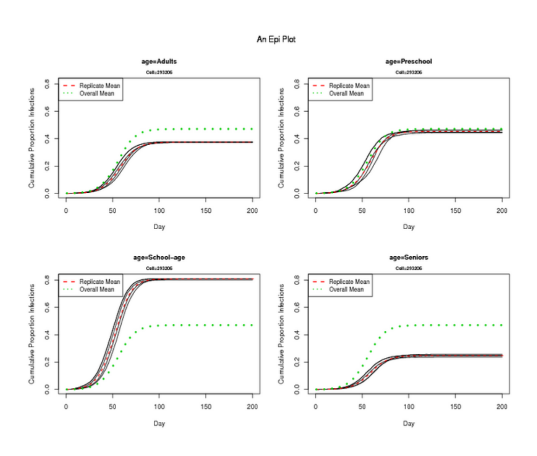


Analysis

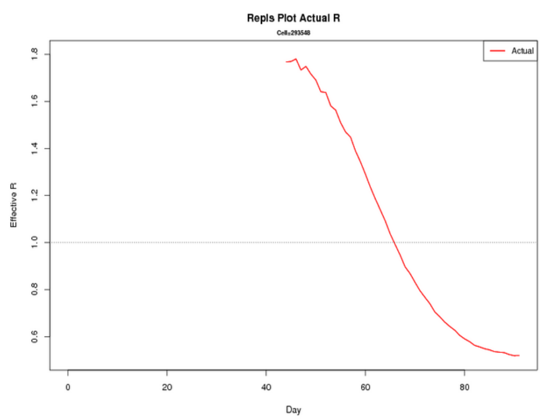


Preview and also Download Tables and Graphs to use in other applications

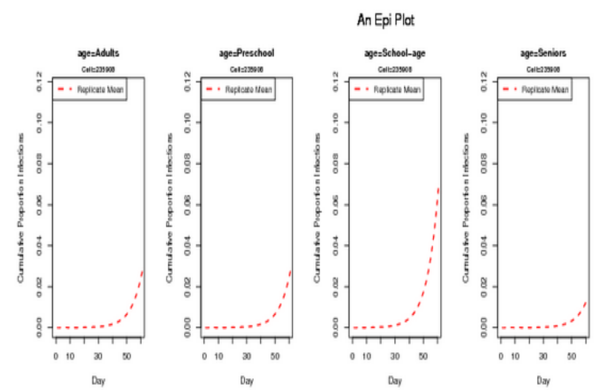
Age-wise Epicurves



Reproductive numbers

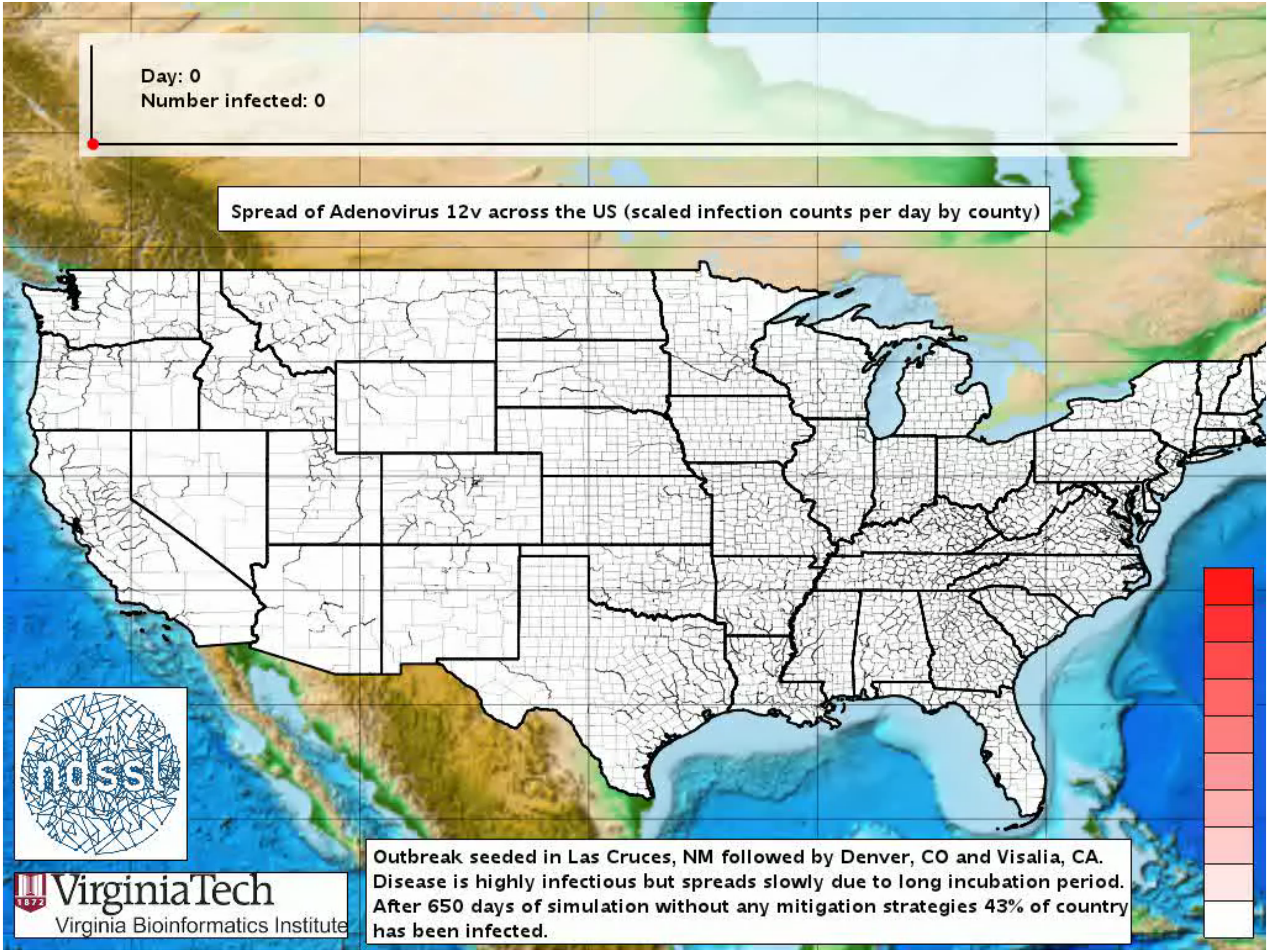


TIGS Analysis Support

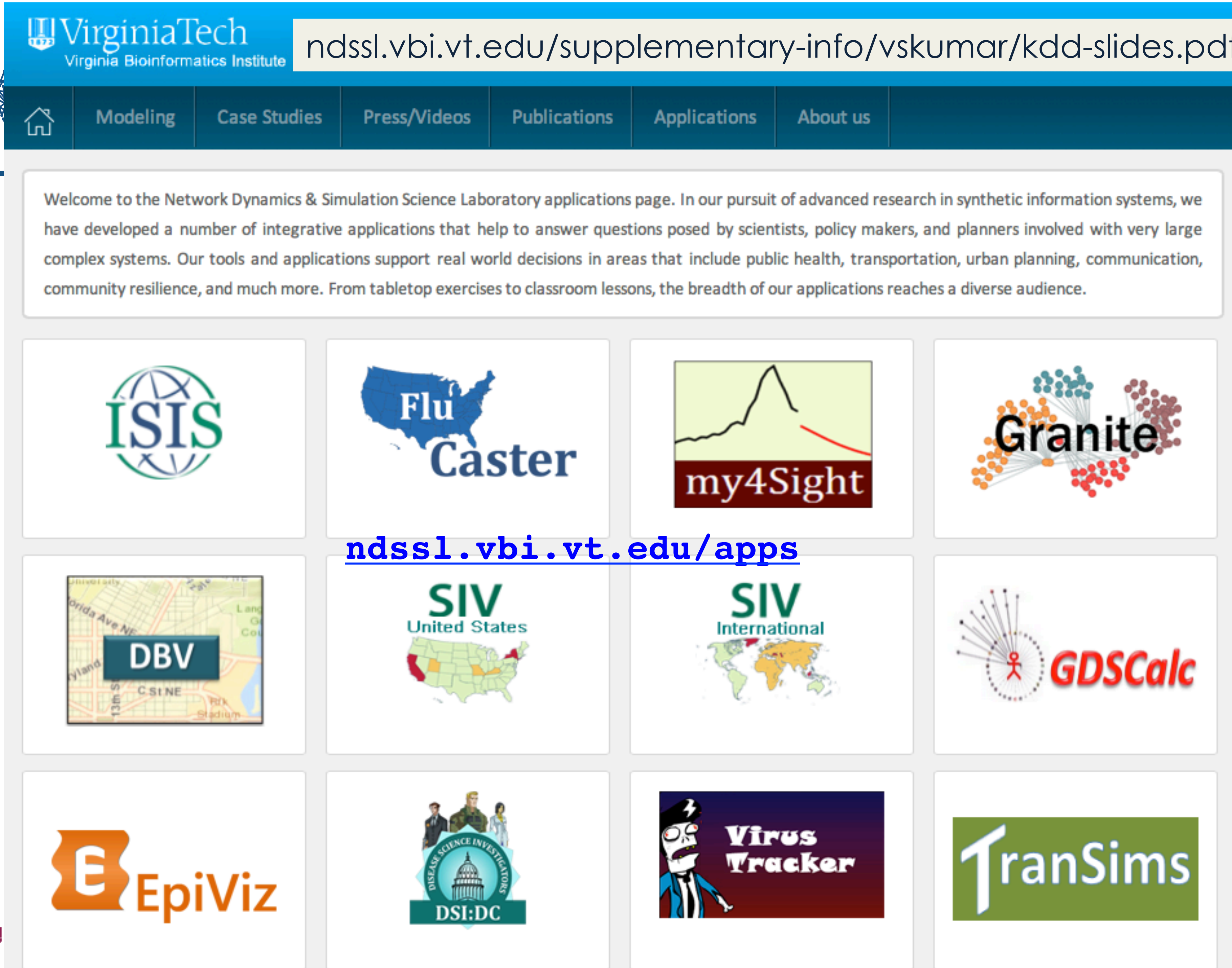


Selected Case Studies

- Pandemic Influenza Planning – Fall 2006
  - How can we prepare for a likely influenza pandemic?
- Emergence of H1N1 Influenza – Spring - Fall 2009
  - What are the characteristics of this influenza strain and their likely impact on US populations?
- Optimizing Household Response to Epidemics – Spring 2011
  - Do behaviors within households, while harder to simulate, have a significant impact on the course of an epidemic?
- SLE: Adenovirus Pandemics Simulation and Analysis – August 2013
  - How can decision makers become familiar with the challenges and decisions they are likely to encounter during a national pandemic, mainly centered on the allocation of scarce resources?
- Emergence of Middle East Respiratory Syndrome – May 2014
  - We are exploring how simulation combined with a synthetic informatics analytic approach might shed light on what may be driving this outbreak.
- EBOLA – Ongoing

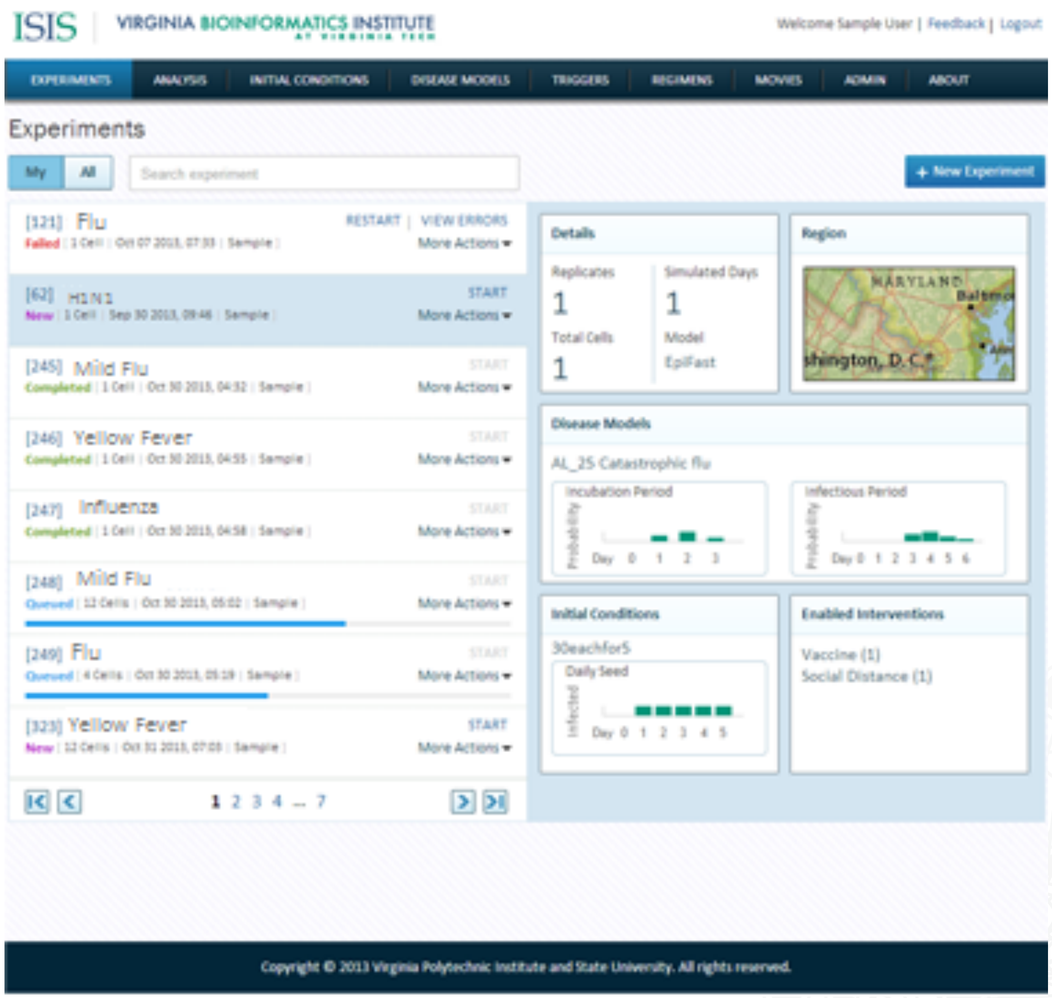

Day: 0
Number infected: 0
Spread of Adenovirus 12v across the US (scaled infection counts per day by county)
ndssl
VirginiaTech
Virginia Bioinformatics Institute
Outbreak seeded in Las Cruces, NM followed by Denver, CO and Visalia, CA. Disease is highly infectious but spreads slowly due to long incubation period. After 650 days of simulation without any mitigation strategies 43% of country has been infected.

ndssl.vbi.vt.edu/supplementary-info/vskumar/kdd-slides.pdf



ndssl.vbi.vt.edu/apps

Demo: Ross V2.3



ndssl.vbi.vt.edu/apps/Ross