Special agents offer modeling upgrade

After playing a key role in the fight against Ebola, agent-based models are poised to help decision-makers tackle other disease outbreaks, economic turbulence, and more.

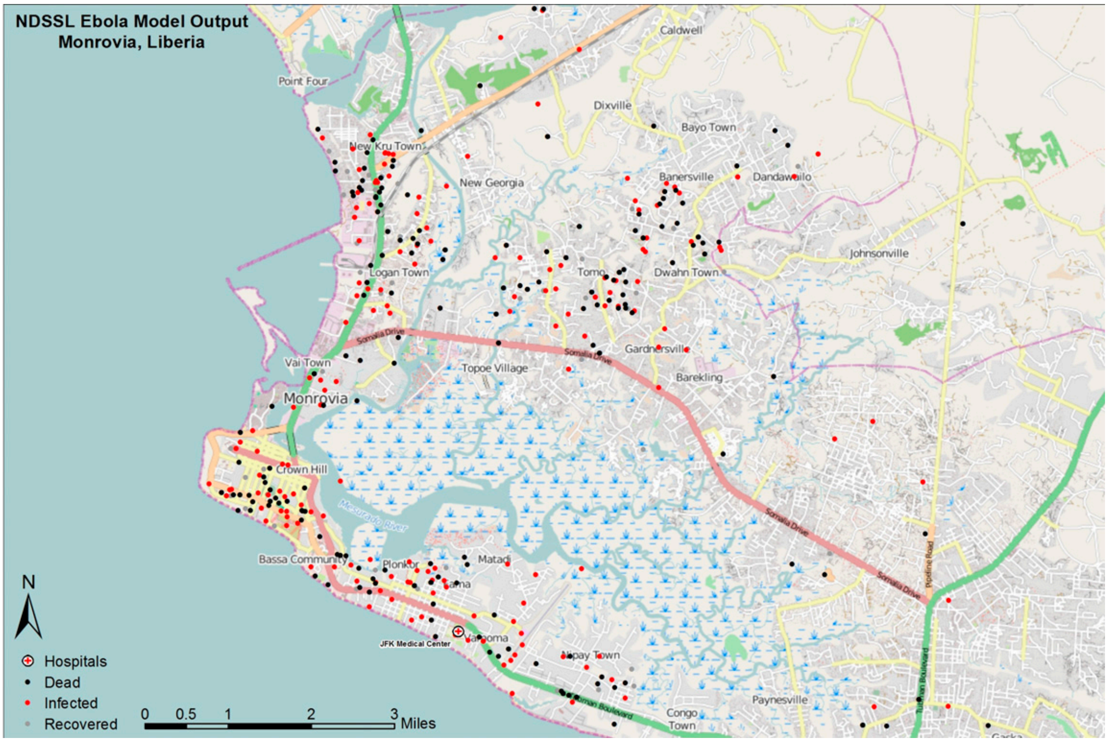
M. Mitchell Waldrop, *Science Writer*

When news of West Africa's Ebola outbreak first appeared in the spring of 2014, "the world wasn't paying a lot of attention," says Bryan Lewis. After all, previous flare-ups of the virus had burned out quickly, mainly because victims tend to die of fever, shock, and spontaneous bleeding before they can spread the virus very far.

Except this time was different. In the hardest hit nations of Guinea, Liberia, and Sierra Leone, the Ebola toll just kept climbing. By early summer, the region was seeing about 100 new cases every week, and alarm bells were going off in public health organizations around the world. So in July, Lewis wasn't too surprised when the United States government's Defense Threat Reduction Agency (DTRA) asked him to develop a model of how the outbreak might progress.

"I didn't even hesitate," says Lewis, a computational biologist at the Virginia Polytechnic Institute and State University (Virginia Tech) in Blacksburg. He and his Virginia Tech colleagues had spent a decade building sophisticated computer simulations of how people and pathogens behave during epidemics, but this was a rare chance to help save people in real time. "It fired up my public health juices," he says.

It was the beginning of a 7-month sprint. Lewis and his colleagues found themselves working in tandem with modeling teams around the world (1), all feverishly racing to answer relief agencies' questions fast enough to make a difference on the ground: How bad will it get? How many mobile hospitals should we send? And how do we field-test a vaccine?



In September 2014, researchers at Virginia Tech in Blacksburg built one of the first large-scale agent-based models of West Africa's Ebola epidemic. As seen here in a close-up of Liberia's capital city, Monrovia, the simulation modeled individual people, or "agents," as they moved back and forth to school, workplaces, or markets and came into contact with other agents who were infected. Image courtesy of The Biocomplexity Institute of Virginia Tech (Blacksburg, VA).

www.pnas.org/cgi/doi/10.1073/pnas.1710350114

The modelers used every tool at their disposal, including mathematical equations that provided the simplest possible model for the spread of infection, and more detailed network models that described how the complex web of human relationships would affect Ebola's trajectory through the population. But for the toughest questions, researchers increasingly turned to cutting-edge "agent-based" models. These artificial societies contain thousands or even millions of subroutines, each representing individual people moving through virtual landscapes while behaving in more or less human ways. Modelers can simulate real-world geography and transportation networks; realistic populations that have the right distribution of ages, genders, and jobs; and details about who comes into contact with whom: "the entire social structure of households, hospitals, and workplaces," says Ira Longini, a statistician and modeler at the University of Florida in Gainesville.

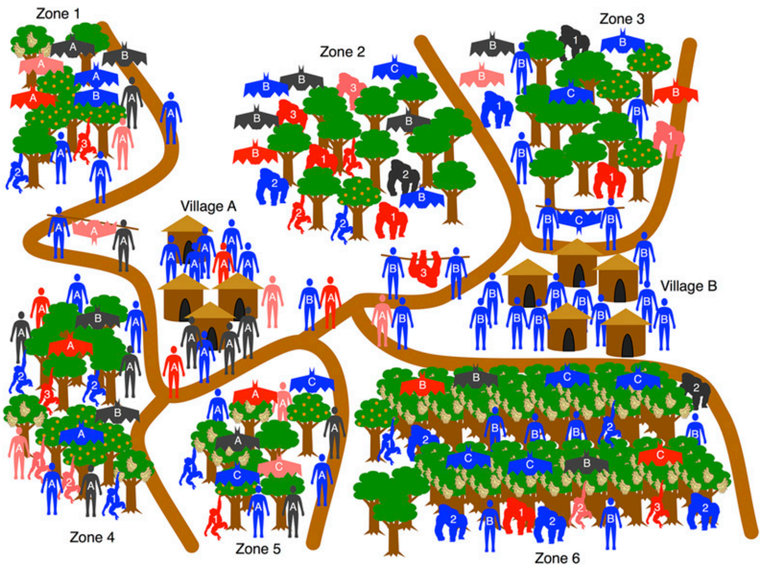
No model can tell what is going to happen to a specific person, any more than a weather forecast can predict exactly where lightning will strike. But agent-based models can predict the chances of stormy weather, so to speak, which is why decision-makers tackling the Ebola crisis embraced such models as never before, desperate to know whether one course of action would do more to temper the storm than another.

That trend seems set to continue. Some modelers are applying the lessons they learned from the Ebola crisis to newly emerging threats, such as the Zika virus; others are designing equally sophisticated artificial societies to help inform real-world decisions in economics, transportation, public health, urban planning, and more. Policymakers are increasingly using these models to try out policy options before making a commitment, says Nigel Gilbert, a sociologist at the University of Surrey, United Kingdom, and founding editor of the *Journal of Artificial Societies and Social Simulation*. And that presents a daunting challenge: "People are going to be relying on our models being right. That's a big responsibility."

Under Pressure

In a sense, agent-based modeling is almost as old as computing itself. The idea of locally interacting bits of software dates back at least to the 1940s, when mathematicians Stanislaw Ulam and John von Neumann started exploring how to simulate crystals and self-reproducing systems with "cellular automata": chessboard-like grids in which each square, or cell, constantly updates itself based on what the neighboring cells are doing (2). Perhaps the first successful effort to do agent-based modeling of human societies was SOCSIM (Social Simulation), a program developed in the 1970s by the late English historian Peter Laslett and his coworkers (3). And a pioneering application of agent-based ideas to disease control was ONCHOSIM, a model devised in the late 1980s to help health officials in Africa control the parasitic worm responsible for onchocerciasis or river blindness (4).

Still, agent-based modeling didn't really start to take off until the mid-1990s, with the advent of increasingly powerful computers. One success from that era was



Although mathematical models capture some of the major features of epidemics' rise and fall, they often only offer up limited labels (healthy or infected, for example) and assume that everyone is in contact with everyone else. Agent-based models can account for real-world factors, such as geography, transportation, family and social relationships, as well as other individual differences. Figure courtesy of ref. 11.

Transims (ndssl.vbi.vt.edu/apps/transims/), which used agents that represented individual vehicles moving through a city-scale traffic simulation. Another was Sugarscape (5), in which agents' quest for a resource (dubbed "sugar") that was abundant in some places and scarce in others generated surprisingly complex group behaviors, including migration, combat, and neighborhood segregation.

Agent-based models have blossomed since then, bolstered by "computers large enough to do a big-ass, giant model; the neuroscience to populate it with cognitively plausible agents; and the big-data sources to calibrate it with lots of real data," says Joshua Epstein, an economist at the Johns Hopkins University in Baltimore and a codeveloper of Sugarscape. By 2014, the simulations had matured into systems that decision-makers could take seriously. "They're the most flexible and detailed models out there," says Longini, "which makes them by far the most effective in understanding and directing policy."

The downside is that these models need plenty of computational muscle. (Virginia Tech runs some of the biggest on supercomputers at places like the Lawrence Livermore National Laboratory in California.) They also depend on lots of data to calibrate the agents' behavior, so that they match the way people respond to events in the real world. And it can take months—or even years—for researchers to build and fine-tune their models.

Unfortunately, data and time were in extremely short supply in the case of the Ebola outbreak, a reality that pushed many of the modelers way out of their academic comfort zone. Response agencies needed answers in a week, if not days; and overwhelmed medical facilities in West Africa could only provide sketchy data about how many people were getting sick, and who recovered.

"The situation was really very frightening," says Richard Hatchett, former deputy director of Biomedical Advanced Research and Development Authority (BARDA), part of the US Department of Health and Human Services. "People were afraid of it spreading to a megalopolis like Lagos—maybe exploding across Africa," he says. Or the world.

Even before attempting an agent-based modeling approach, the modelers were forced to try a new way of working as they entered crisis mode: scrape together whatever data they could find, fill in the rest with educated guesses, give the agencies the best answers they could, and then do it all over again the next day. The Virginia Tech team initially assigned a high-school summer intern, Katie Dunphy, to get up early every morning, gather the latest numbers from the Facebook page of Sierra Leone's public health ministry—the only place where the data were being reported at the time—and then pass the results to an undergraduate intern who plotted them in a geographical information system. After a few weeks, however, as the severity of the epidemic became apparent, these duties were taken over by a graduate student, Caitlin Rivers, who posted everything online for other researchers to access.

That provided enough information to model the epidemic in classic mathematical fashion, using decades-old equations that projected future infection rates based on simple statistics: how many people are susceptible to Ebola right now, how many are already infected and contagious, and how many have survived and are immune. In August 2014, the most influential of these models predicted that if nothing more was done to stop the epidemic, Ebola cases in West Africa would explode from about 3,500 to somewhere between half a million and 1.4 million cases by January 2015. Many critics argued that this was much too pessimistic, but it helped spur US President Barack Obama to send United States military personnel to assist medical response teams in West Africa. "It really galvanized US officials to understand the urgency of moving with great haste to get this under control," says Hatchett.

The big question was: how?

A "What if?" Tool

The struggle to find an answer soon led the Ebola modelers into the second, much tougher phase of their work: planning interventions. "At that point the model becomes something more important than a forecasting tool," says Alessandro Vespignani, who headed an extensive Ebola modeling effort at Northeastern University in Boston. "It becomes a 'what if?' tool."

The good news was that health workers in West Africa were finally beginning to provide more data about the outbreak, enabling modelers to tackle those questions with fully fledged agent-based models. The bad news was that the clock was as tyrannical as ever. On a Friday afternoon in early October 2014, for example, Pyrros Telionis, a Virginia Tech graduate student in computational biology, got a phone call from DTRA: could his models identify the best places to put mobile Ebola treatment units in southern Liberia by Monday morning, before the cargo planes took off?

Telionis and fellow graduate student James Schlitt promptly scrapped their weekend plans and plunged in. They couldn't just recommend putting the treatment units in the most heavily infected areas; they had to project where the epidemic would be when the treatment units actually reached their destinations, and then factor in the muddy or impassible roads that people would use to reach them. The two graduate students worked around the clock, piecing together algorithms and data from colleagues, using agent-based models, mathematical models, whatever worked. But they met their deadline. "DTRA would neither confirm nor deny that they had used our suggestions," says Schlitt. "But the locations they used were very similar to what we gave them."

In the end, though, what really tamed the Ebola epidemic wasn't just international aid—as essential as that was—but the changing behaviors of the people of West Africa themselves: behaviors that could, to some extent, be captured in agent-based models. One of the first signs of those changes had come in September 2014, when infection rates began to fall in Lofa County, Liberia. The region's traditional burial practices—which included rituals such as caressing and washing the deceased—were virtually guaranteed to spread an Ebola infection. But after a local health official urged people to change those rituals, it slowed the transmission of the virus. At a BARDA meeting that month, Lewis used agent-based models to argue that if these behavioral changes were repeated across West Africa, the resulting decrease in new infections would be enough to stop Ebola in its tracks. The attendees were dubious that such a widespread culture change would ever take hold.

However, it did. West Africans embraced sanitation, quarantine, and safe burial, and infection rates started to fall. By the time West Africa was finally declared Ebola-free in January 2016, the toll stood at roughly 28,000 cases, and 11,000 dead, a fraction of what had originally been feared.

From Zika to Macroeconomics

Even when the Ebola epidemic was beginning to abate in early 2015, modeling still had a vital role to play. The outbreak provided health officials with their first opportunity to carry out clinical trials of an experimental Ebola vaccine, called rVSV-ZEBOV, so Longini and others used agent-based models to develop the strategy for the field trials, which took place in Guinea between April and July 2015 (6, 7). For example, officials needed to know if fresh pockets of the disease could spread so rapidly that the trials would have to vaccinate everyone in the region. Apparently not, said the models: the research team could safely follow a classic "ring" strategy, which would target only the people who had come into contact with an Ebola victim, plus the people who had had contact with them. And it worked: the final results of the trial, published in December 2016 (7), concluded that the vaccine could be up to 100% effective at preventing Ebola.

Then, hard on the heels of Ebola came Zika: a mosquito-borne virus that can lead to catastrophic birth defects, such as microcephaly, which has affected Brazil and surrounding countries. Many groups are now modeling the spread of the virus to help officials anticipate where the worst outbreaks might be and where they should concentrate their resources.

But Zika is considerably more complicated than Ebola. Whereas Ebola spreads almost exclusively from human to human, Zika is mostly transmitted via mosquito bites (although it can also spread through unprotected sex). That means the models not only have to incorporate data on human sexual behavior, but everything that affects mosquito populations, from seasonal temperature swings and rainfall to the availability of old tires and other breeding sites. Epstein's group is now collaborating with New York University on an agent-based model of Zika in that city, which will include 8.5 million agents—the number of people in the five boroughs—plus a separate set of agents representing the entire population of individual mosquitoes, as estimated from traps.

To prepare for new crises, many modelers are now calling for a permanent infrastructure for simulating epidemics. That might include libraries of precomputed models and data that could be dusted off and quickly applied to a real-world outbreak, an approach that Epstein calls a "Petabyte Playbook." "For future outbreaks," says Sara Del Valle, a modeler at Los Alamos National Laboratory in New Mexico, "we believe that we need something similar to the National Hurricane Center" (the United States government's tropical-storm forecasting organization in Miami). Such a disease forecasting center would be a worldwide initiative, pulling in data from clinics and medical records around the globe, while running multiple simulations to give people the most likely range of possible outcomes.

Although an international center is not even in the planning stages, a lot of modelers note that government officials are growing more receptive to the guidance it could provide. Agent-based models have become standard tools in transportation planning, thanks to Transims and its descendants, and are fast achieving that status among public health officials. "The respect and esteem given to modelers has changed radically over where it was in the 2000–2001 timeframe," says Hatchett.

A similar process is beginning to unfold in the realm of economics. In November 2016, Andy Haldane, chief economist at the Bank of England, took his discipline to task (8) for failing to anticipate major events, such as the financial collapse of 2007–2008: the economic analog of a global pandemic. A big part of the problem, he said, was that his colleagues typically modeled swings in the economy using equations that focused only on global factors such as interest rates, while ignoring critical details, such as geography or individual human differences and interactions. Instead, they should be building on decades of academic research that has produced agent-based models of issues such as innovation, inequality, and the stability of the financial system.

The Bank's own research program has developed two different agent-based models. In one model (9), a dynamic United Kingdom housing market emerges from the interaction of agents representing renters, landlords, homeowners, mortgage lenders, and regulators. In the other model (10), a bond market emerges from agents representing buyers, fund managers, and traders. In both, said Haldane, the ups and downs seen in the model markets fit those found in real-world markets much better than the predictions of standard equations.

Doyne Farmer, who helped create the Bank's housing model and leads a modeling group at University of Oxford, United Kingdom, says that more (and bigger) agent-based economic models are under development, and that many more institutions are warming to the idea of modeling. In the aftermath of the financial crisis, he says, "there is a recognition among central banks that they have to get a better handle on risk."

And what if the models' predictions are not what senior politicians want to hear? "This is actually a good problem," says Farmer, because if they're listening carefully enough to disagree, "we're close to success."

1 Chretien J-P, Riley S, George DB (2015) Mathematical modeling of the West Africa Ebola epidemic. *eLife* 4:e09186.

2 Von Neumann J (1966) *Theory Of Self Reproducing Automata*, ed Burks AW (Univ of Illinois Press, Urbana, IL), 1st Ed.

3 Zhao Z (2006) Computer microsimulation and historical study of social structure: A comparative review of SOCSIM and CAMSIM. *Rev Demogr Hist* 24:59–88.

4 Plaisier AP, van Oortmarssen GJ, Habbema JD, Remme J, Alley ES (1990) ONCHOSIM: A model and computer simulation program for the transmission and control of onchocerciasis. *Comput Methods Programs Biomed* 31:43–56.

5 Epstein JM, Axtell RL (1996) *Growing Artificial Societies: Social Science From the Bottom Up (Complex Adaptive Systems)* (Brookings Institution Press & MIT Press, Cambridge, MA), 1st Ed.

6 Ebola ça Suffit Ring Vaccination Trial Consortium (2015) The ring vaccination trial: A novel cluster randomised controlled trial design to evaluate vaccine efficacy and effectiveness during outbreaks, with special reference to Ebola. *BMJ* 351:h3740.

7 Henao-Restrepo AM, et al. (2017) Efficacy and effectiveness of an rVSV-vectored vaccine in preventing Ebola virus disease: Final results from the Guinea ring vaccination, open-label, cluster-randomised trial (Ebola Ça Suffit!). *Lancet* 389:505–518.

8 Haldane A (2016) The Dappled World. Available at www.bankofengland.co.uk/publications/Pages/speeches/2016/937.aspx. Accessed June 17, 2017.

9 Baptista R, et al. (2016) Staff Working Paper No. 619: Macroprudential policy in an agent-based model of the UK housing market. Available at www.bankofengland.co.uk/research/Pages/workingpapers/2016/swp619.aspx. Accessed June 17, 2017.

10 Braun-Munzinger K, Liu Z, Turrell A (2016) Staff Working Paper No. 592: An agent-based model of dynamics in corporate bond trading. Available at www.bankofengland.co.uk/research/Pages/workingpapers/2016/swp592.aspx. Accessed June 17, 2017.

11 Alexander KA, Lewis BL, Marathe M, Eubank S, Blackburn JK (2012) Modeling of wildlife-associated zoonoses: Applications and caveats. *Vector Borne Zoonotic Dis* 12(12):1005–1018.