JOHNS HOPKINS BLOOMBERG SCHOOL *of* PUBLIC HEALTH | Center for Health Security

OUTBREAK SCIENCE INITIATIVE

Caitlin Rivers, PhD, MPH
Senior Scholar, Johns Hopkins Center for Health Security
CRIVERS6@JHU.EDU

Every few years, a new pathogen emerges from the shadows to threaten global public health. And every few years, the global public health community struggles to mount a timely and effective response. The incremental advances that have defined public health in the last half century have been repeatedly outpaced by fast-moving epidemics. From HIV to Ebola, most innovations in preventing and containing outbreaks have come from medical countermeasures like vaccines, diagnostics, and therapeutics rather than advancements in epidemiological practice.

This gap is exemplified in the outdated and inefficient use of computational tools and resources. Even though the very core of the discipline is the iterative collection and analysis of data to inform decisions, most epidemiology students are not trained in data management, programming, or advanced statistics. They may attain some fluency in Excel or perhaps SAS, taught in conjunction with a small number of statistics courses. This brief introduction to computational skills is inadequate at best and misleading at worst; these tutorials do not even hint at the scope and power of modern computer science in epidemiology.

When students trained in this paradigm enter state and local health departments, they codify their skill gaps into institutional policy and practice. In any given public health department, collected data is likely to be poorly structured and stored, limiting its usefulness. These data are often analyzed superficially, if at all. Repeat analyses, like those used to produce regular reports, are performed manually, which wastes time and creates errors. The existing systems are ineffective during 'peacetime', while data and analysis problems are magnified several times over during an outbreak.

None of these practices support the aforementioned iterative collection and analysis of data to inform decisions. It is perhaps no wonder then that the decision-making process in an outbreak response is often disconnected from the analytical insights needed to determine how best to plan for and contain outbreaks. The recent Ebola outbreak offers an illustrative example. Among numerous problems with the epidemiological data, problems in the data collection pipeline ultimately led Ebola czar Ron Klain to declare that it was useless in directing the response.

Perversely, the dearth of quality primary data during the Ebola outbreak led to an increased reliance by decision makers on model results and other 'data extenders'. One influential model published by the Centers for Disease Control and Prevention forecast a worst-case scenario of more than one million cases if the epidemic continued unabated. It's thought that the CDC model galvanized the international response that ultimately contributed to control of the epidemic. Other models were used to direct interventions, plan clinical trials, and assess progress on outbreak containment. This partnership between modelers and decision makers hints at a possible future where better integration of infectious disease modeling, data science and visualization, and modern data practices into public health practice supports more effective and efficient response operations.

On the other hand, many of the models produced during the Ebola outbreak have rightly been criticized for being some combination of inaccurate, irrelevant, or untimely. These pitfalls are at least partly consequent to the aforementioned challenges that epidemiology has faced in effectively incorporating data and analytics best practices into prevention, preparedness, and response. Although highly trained and technically excellent, the scientists who produce infectious disease models are primarily researchers at academic institutions without experience in either applied epidemiology or the US government, where decisions about responding to major outbreaks are made. Because no formal mechanism exists to surge external modeling support during a public health event, these scientists respond to outbreaks without compensation or formal integration into response operations. This disconnect between outbreaks and the timeliness and applicability of most existing models limits the effectiveness of the response.

Further, the scientists who do respond have no choice but to rely on the patchy, poor quality, and irregular data that is commonly produced in outbreaks. Model outputs are only as good as their inputs; data that are missing, out of date, or irregular will invariably lead to suboptimal results. Data inadequacies are compounded by pervasive issues around sharing, which limit data availability to only immediate members of the outbreak response team. Secondary responders, including most US government departments and agencies, NGOs, and scientists, are without the data they need to make decisions. The result is data that are simply not available enough or robust enough to support detailed and accurate analytics.

None of these problems are intractable, but they do point to the need of a strategic, coherent modernization of outbreak epidemiology. The Outbreak Science Initiative (OSI) is an effort to remediate the gaps that have prevented epidemiology from making the most effective use of computational tools by establishing outbreak science, a new subfield focused on these issues. With sufficient investment, outbreak science can transform the United States response to outbreaks through disease modeling, data science and visualization, and modern data practices. OSI at the Johns Hopkins Center for Health Security will take an innovative, comprehensive approach to modernize outbreak preparedness and response with better data, better analyses, and better decisions.