Infectious diseases remain among the top contributors to human illness and death worldwide. While some infectious disease activity appears in consistent, regular patterns within a population, many diseases produce less predictable epidemic waves of illness. Uncertainty and surprises in the timing, intensity, and other characteristics of these epidemics stymies planning and response of public health officials, health care providers, and the general public. Accurate forecasts of this information with well-calibrated descriptions of their uncertainty can assist stakeholders in tailoring countermeasures, such as vaccination campaigns, staff scheduling, and resource allocation, to the situation at hand, which in turn could translate to reductions in the impact of a disease.

Domain-driven epidemiological models of disease prevalence can be difficult to fit to observed data while incorporating enough details and flexibility to explain the data well. Meanwhile, more general statistical approaches can also be applied, but traditional modeling frameworks seem ill-suited for irregular bursts of disease activity, and focus on producing accurate single-number estimates of future observations rather than well-calibrated measures of uncertainty on more complicated functions of the data. The first part of this work develops variants of simple statistical approaches to address these issues, and a way to incorporate features from certain domain-driven models.

Epidemiological surveillance systems commonly incorporate a data revision process, whereby each measurement may be updated multiple times to improve accuracy as additional reports and test results are received and data is cleaned. The second part of this work discusses how this process impacts proper forecast evaluation and visualization. Additionally, it extends the models above to "backcast" how existing measurements will be revised, which in turn can be used to improve forecast accuracy. These models are then expanded further to include auxiliary data from other surveillance systems.

The preceding sections describe several prediction algorithms, and many more are available in existing literature and deployed in operational systems. The final part of this work demonstrates one method to combine output from multiple such prediction systems with consideration of the domain, which on average tends to match or outperform its best individual component.

**Thesis Committee:**
Roni Rosenfeld, Chair
Ryan Tibshirani
Zico Kolter
Jeffrey Shaman, Columbia University