Introduction to the Special Section on Inference for Infectious Disease Dynamics

Theodore Kypraios and Vladimir N. Minin

1. INTRODUCTION

The past three decades have seen significant growth in the field of mathematical modeling of infectious diseases, leading to substantial increase in our understanding of epidemiology and control of these diseases. Ability to quickly unravel the dynamics of the spread of infectious diseases is important for effective prevention of future outbreaks and for control of ongoing ones. Recent interest in infectious disease modeling was initially stimulated by the discovery of HIV in the early 1980s and has been maintained by the need to respond to other infectious disease related crises such as, for example, foot-and-mouth disease and SARS outbreaks, healthcare-associated infections, and elevated risks of human influenza pandemics (e.g., risks of global spread of avian and swine flu). Recent Ebola and Zika outbreaks further underscored importance of mathematical and statistical analyses of epidemic dynamics. As a result, mathematical infectious disease modeling remains high on the global scientific agenda.

To respond to challenges posed by infectious diseases the epidemic modeling community has become very engaged in public health policy development, further stimulating interest in models of disease transmission. This involvement in practical applications of infectious disease modeling was enabled by algorithmic advances and the continuing increase in computing power. For example, it is now possible to perform simulations based on parameter rich and realistic *agent-based models* that generate individual-level behavior, infections, and recoveries of millions of individuals. Moreover, analyses on infectious disease outbreak data can now be performed using computationally intensive methods such as maximum likelihood estimation, Markov chain Monte Carlo (MCMC), sequential

Theodore Kypraios is Associate Professor, School of Mathematical Sciences, University of Nottingham, NG7 2RD, United Kingdom (e-mail:
theodore.kypraios@nottingham.ac.uk). Vladimir N. Minin is Professor, Department of Statistics, University of California, Irvine, California 92617, USA (e-mail: vminin@uci.edu).

Monte Carlo (SMC), and approximate Bayesian computation (ABC).

The transmissible nature of infectious diseases makes them fundamentally different from noninfectious diseases, and therefore it is difficult to analyze disease outbreak data using off-the-shelf statistical methods. This is mainly due to (i) *presence of strong and complex dependencies in the data* and (ii) *missing data significantly surpassing observed data in size*, because the actual transmission process cannot be directly observed. Therefore, specialized and problem-specific methods are required. Despite recent significant advances in fitting stochastic epidemic models to data, there are still a number of challenges to overcome.

In view of the ever growing research activity in the area and the practical importance of effectively analyzing disease outbreak data, we have organized this special section of *Statistical Science* in which the aim is to provide an overview of the most recent developments as well as the current research challenges facing the epidemic modeling community. We hope that readers of the journal will get a broad idea of the field as well as of its current research directions.

2. STOCHASTIC EPIDEMIC MODELS

There exist many forms of models to represent the dynamics of infectious diseases. Most models are concerned with a population consisting of individuals who are potentially able to transmit the disease to one another. In this section we focus on stochastic models. Although in some settings deterministic models can approximate stochastic dynamics well, behavior of stochastic and deterministic models can be qualitatively different in certain parameter regimes (e.g., when the disease is spreading in a small population). This suggests that stochastic models should be preferred to deterministic ones if one wishes to stay true to the stochastic nature of infectious disease dynamics.

Stochastic epidemic models are generally defined at the level of individuals, for instance, specifying probability distributions that describe how long an individual remains infectious. A key aspect of any disease