



Estimation in emerging epidemics: biases and remedies

Tom Britton* Gianpaolo Scalia Tomba†

May 2018

Abstract

When analysing new emerging infectious disease outbreaks, one typically has observational data over a limited period of time and several parameters to estimate, such as growth rate, the basic reproduction number R_0 , serial interval or generation time distribution, latency and incubation times or case fatality rates. Also parameters describing times between onset of symptoms, notification, death and recovery/discharge will be of interest. These parameters form the basis for predicting the future outbreak, planning preventive measures and monitoring the progress of the disease outbreak.

We study inference problems during the emerging phase of an outbreak, and point out potential sources of bias related to exponential growth during the emerging phase and to the unobservability of the moment of transmission, with emphasis on: contact tracing backwards in time, replacing generation times by serial intervals, multiple potential infectors and to censoring effects amplified by exponential growth. These biases directly affect the estimation of e.g. the generation time distribution and the case fatality rate, but can then propagate to other estimates such as R_0 and growth rate. We propose methods to remove or at least reduce bias using statistical modelling. We illustrate the theory by numerical examples and simulations based on the recent 2014-15 West Africa Ebola epidemic to quantify possible biases, which may be up to 20% underestimation of R_0 , if the epidemic growth rate is fitted to observed data or, conversely, up to 62% overestimation of the growth rate if the correct R_0 is used in conjunction with the Euler-Lotka equation.

*Stockholm University, Sweden; tom.britton@math.su.se

†University of Rome Tor Vergata, Italy, scaliato@mat.uniroma2.it