Case study Ebola outbreak

- LFI is well suited for infectious disease epidemiology
- Simulating the dynamics of the disease at individual-level enables flexibility not achievable in traditional models
- In this case study we demonstrate modern LFI using the Ebola haemorrhagic fever outbreak in West Africa in 2014 as the example
 - We estimate the basic reproduction number R_0
 - the mean value of secondary infections caused by infectee when no countermeasures are in place

Case study Ebola outbreak

The NEW ENGLAND JOURNAL of MEDICINE

ESTABLISHED IN 1812

OCTOBER 16, 2014

VOL. 371 NO. 16

Ebola Virus Disease in West Africa — The First 9 Months of the Epidemic and Forward Projections

WHO Ebola Response Team*

ABSTRACT

On March 23, 2014, the World Health Organization (WHO) was notified of an out- Address reprint requests to Dr. Christle break of Ebola virus disease (EVD) in Guinea. On August 8, the WHO declared the Donnelly at c.donnelly@imperial.ac.uk epidemic to be a "public health emergency of international concern."

By September 14, 2014, a total of 4507 probable and confirmed cases, including 2296 deaths from EVD (Zaire species) had been reported from five countries in West Africa — Guinea, Liberia, Nigeria, Senegal, and Sierra Leone. We analyzed a detailed subset of data on 3343 confirmed and 667 probable Ebola cases collected in Guinea, Liberia, Nigeria, and Sierra Leone as of September 14.

The majority of patients are 15 to 44 years of age (49.9% male), and we estimate that the case fatality rate is 70.8% (95% confidence interval [CI], 69 to 73) among persons with known clinical outcome of infection. The course of infection, including signs and symptoms, incubation period (11.4 days), and serial interval (15.3 days), is similar to that reported in previous outbreaks of EVD. On the basis of the initial periods of exponential growth, the estimated basic reproduction numbers (R,) are 1.71 (95% CI, 1.44 to 2.01) for Guinea, 1.83 (95% CI, 1.72 to 1.94) for Liberia, and 2.02 (95% CI, 1.79 to 2.26) for Sierra Leone. The estimated current reproduction numbers (R) are 1.81 (95% CI, 1.60 to 2.03) for Guinea, 1.51 (95% CI, 1.41 to 1.60) for Liberia, and 1.38 (95% CI, 1.27 to 1.51) for Sierra Leone; the corresponding doubling times are 15.7 days (95% CI, 12.9 to 20.3) for Guinea, 23.6 days (95% CI, 20.2 to 28.2) for Liberia, and 30.2 days (95% CI, 23.6 to 42.3) for Sierra Leone. Assuming no change in the control measures for this epidemic, by November 2, 2014, the cumulative reported numbers of confirmed and probable cases are predicted to be 5740 in Guinea, 9890 in Liberia, and 5000 in Sierra Leone, exceeding 20,000 in

These data indicate that without drastic improvements in control measures, the numbers of cases of and deaths from EVD are expected to continue increasing from hundreds to thousands per week in the coming months.

or Dr. Christopher Dye at dyec@who.int.

*The authors (members of the World Health Organization [WHO] Ebola Response team who contributed to this article) are listed in the Appendix.

This article was published on September 23,

N Engl J Med 2014;371:1481-95 Copyright @ 2014 World Health Organization.

INTERFACE

royalsocietypublishing.org/journal/rsif

Research



Cite this article: Britton T, Scalia Tomba G. 2019 Estimation in emerging epidemics: biases and remedies. J. R. Soc. Interface 16: 20180670.

http://dx.doi.org/10.1098/rsif.2018.0670

Received: 5 September 2018 Accepted: 17 December 2018

Subject Category: Life Sciences - Mathematics interface

Subject Areas biomathematics

statistics, emerging epidemic, estimation bias, basic reproduction number

Author for correspondence Gianpaolo Scalia Tomba e-mail: scaliato@mat.uniroma2.it

Electronic supplementary material is available online at https://dx.doi.org/10.6084/m9. figshare.c.4347491.

THE ROYAL SOCIETY

Estimation in emerging epidemics: biases and remedies

Tom Britton1 and Gianpaolo Scalia Tomba2

Department of Mathematics, Stockholm University, 10691 Stockholm, Sweden. ²Department of Mathematics, University of Rome Tor Vergata, 00133 Rome, Italy

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 R₀, the case fatality rate and distributions of serial intervals, generation times, latency and incubation times and times between onset of symptoms, notification, death and recovery/discharge. These parameters form the basis for predicting a 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, with emphasis on: contact tracing backwards in time, replacing generation times by serial intervals, multiple potential infectors and censoring effects amplified by exponential growth. These biases directly affect the estimation of, for example, the generation time distribution and the case fatality rate, but can then propagate to other estimates such as Ro 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.

Introduction

During the last decades, several new disease outbreaks have caused worldwide alarm, e.g. SARS, foot and mouth disease, H1N1 influenza, and, more recently, Ebola. What these outbreaks have in common is the need for estimation of key parameters early on, in order to plan interventions and monitor the progress of the disease. Thus estimation must be performed in the emerging phase of an outbreak, when the number of infected individuals is in the hundreds or at most thousands, while the community fraction of infected is still small. Typically, the early numbers grow exponentially, as also predicted by mathematical epidemic models [1].

There may be many complicating or limiting factors related to incompleteness of data, lack of detailed knowledge about the disease and other issues when analysing data from the early phase of an outbreak. Despite these complicating factors, the conclusions drawn from early analyses, often based on simple models, are usually highly valuable. The aim of the present paper is to identify and highlight some of the potential biases in the statistical analysis of emerging outbreaks and to illustrate how they can propagate to parameter estimates and predictions. A further aim is to give some fairly simple suggestions for how to reduce, or even remove, such biasing effects.

The typical available data consist of reported numbers of cases per day or week, some case histories illustrating the course of the disease and some contact tracing data containing information about possible durations between onset of symptoms of infected individuals and their infectors. The epidemic models used in the statistical analyses are often of simple form, neglecting various hetities. The use of simple models in these situations is motivated by the lack of detailed information but has also recently been studied [2], showing that neglecting population structures when making inference in emerging

© 2019 The Authors, Published by the Royal Society under the terms of the Creative Commons Attribution License http://creativecommons.org/licenses/by/4.0/, which permits unrestricted use, provided the original