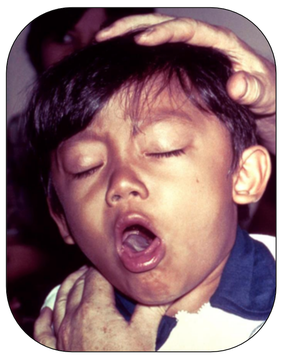
Using paired tests to recover trends of infection

Gustaf Rydevik[123], Mike R Hutchings[1], Piran C L White[3], Ross S Davidson[12], Glenn Marion[2], Giles Innocent[2]

[1]SRUC Scotland [2]BioSS Scotland [3]University of York England

Multiple testing to improve surveillance

Infectious diseases are a major cause for concern in both humans and animals. Infectious disease epidemics can cause substantial damage to health, economy and welfare. Despite this, disease surveillance is often limited in its ability to detect epidemics and track trends in incidence and infection



Boy with B. Pertussis infection

In this poster we demonstrate how knowledge of the host-pathogen dynamic can be combined with surveillance data to give a more detailed picture of the state of an epidemic. By exploiting dynamic characteristics of diagnostic tests, it is possible to infer the population level distribution of time since infection in epidemics even though only cross-sectional data may be available.

We combine results from two or more diagnostic tests in order to back-estimate the time of infection for a population of tested, previously infected individuals. Combining the estimated infection times produces an estimate of the overall trend of infection up until the time of testing.

A bayesian framework to recover trends of infection

We want to evaluate the posterior distribution of infection times $\boldsymbol{T}$ given the collected test data $\boldsymbol{Y}$, incorporating information on the latent state $\boldsymbol{L}$ of infection response (modelled deterministically) given a time since infection $P(\boldsymbol{L}|\boldsymbol{T},\boldsymbol{q})$ and the distribution $P(\boldsymbol{Y}|\boldsymbol{L},\boldsymbol{p})$ of the observed measurements given the latent state. $\boldsymbol{p}$ and $\boldsymbol{q}$ are vectors of parameters for the distribution.

Finally, we consider the distribution of infection times $P(\boldsymbol{T}|\boldsymbol{\Theta})$, where $\boldsymbol{\Theta}$ is a parameter vector. This is to be parametrized according to the expected shape of the infection curve. For the examples below, $P(\boldsymbol{T}|\boldsymbol{\Theta})$ was given a beta distribution, combined with a scaling factor to estimate the overall timespan of the epidemics.

The full data likelihood can be formalised as

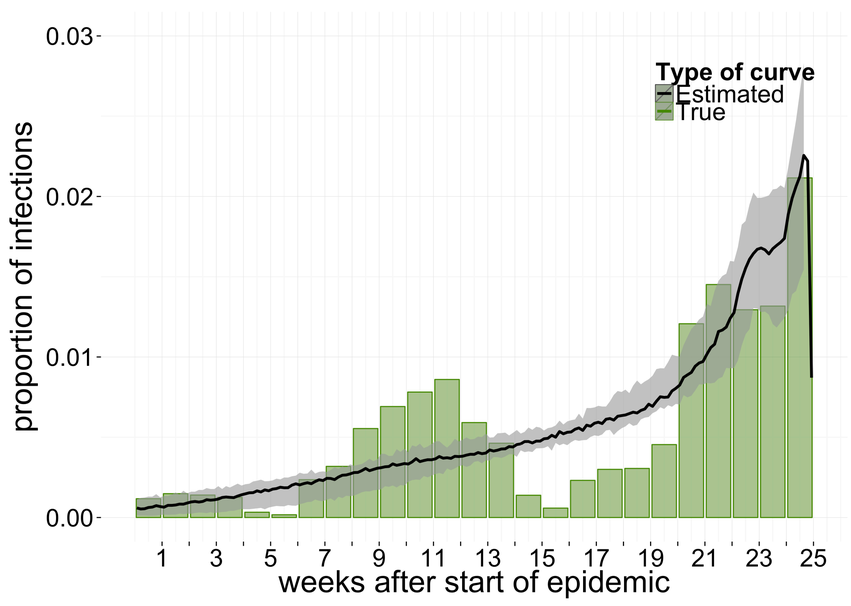
$$P(\boldsymbol{T},\boldsymbol{\Theta},\boldsymbol{L},\boldsymbol{p},\boldsymbol{q}|\boldsymbol{Y}) = P(\boldsymbol{Y}|\boldsymbol{L},\boldsymbol{p})P(\boldsymbol{L}|\boldsymbol{q})P(\boldsymbol{T}|\boldsymbol{\Theta})$$

From this, we can obtain the posterior distribution for our parameters as

$$P(\boldsymbol{L},\boldsymbol{T},\boldsymbol{\Theta},\boldsymbol{q},\boldsymbol{p}|\boldsymbol{Y}) \propto P(\boldsymbol{Y},\boldsymbol{L},\boldsymbol{T}|\boldsymbol{\Theta},\mathbf{q},\mathbf{p})$$

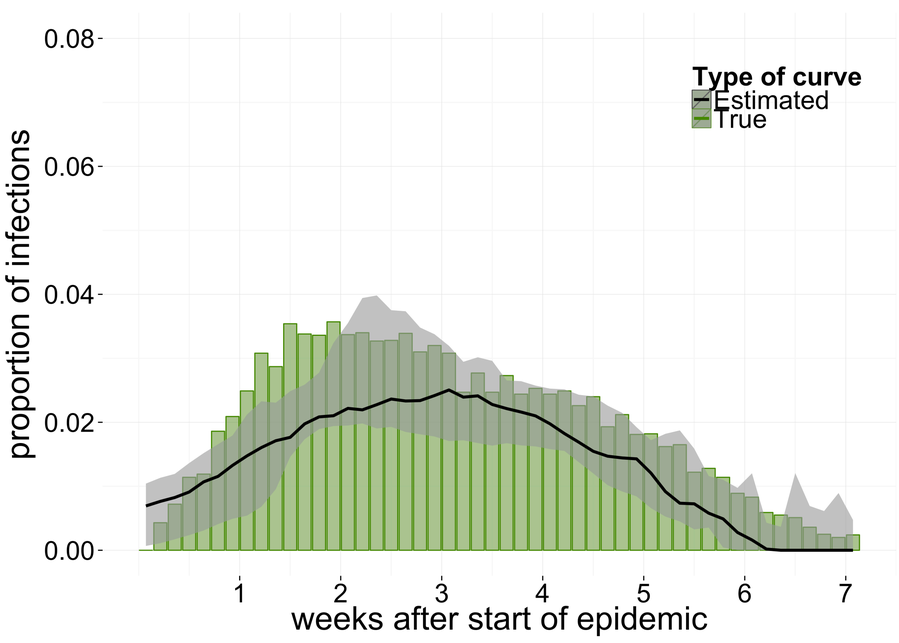
In order to generate estimates of the historical pattern of the epidemic, we use a Metropolis-Hastings MCMC algorithm combined with uninformative priors to draw samples of the distribution.

Example 1: Pertussis outbreak in Wisconsin 2003



Pertussis (Bordetella Pertussis) is a bacterial pathogen that causes whooping cough in humans. In Fond du Lac County, Wisconsin, USA in 2003-2004, there was a countywide outbreak of Pertussis primarily among adolescents and adults. We simulated test data from this outbreak, assuming that all cases were tested with both antibody and nucleic acid tests 25 weeks after the start of the outbreak. The log-normal variation in test results was about 46%. From these simulated datasets, we then estimated the trend of infection up until testing time (black line shows median estimated trend). The mean R-squared value for the estimated trend was 0.54.

Example 2: Bluetongue virus outbreak in UK 2007

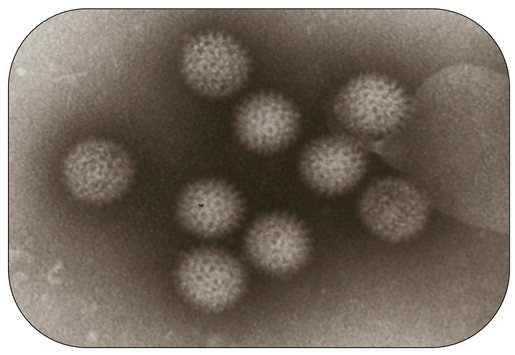


Bluetongue virus (BTV) is a midge-borne virus that can cause Bluetongue disease in ruminants such as sheep, cattle, deer and camelids, causing swelling and lesions in the mouth, and in some species death. The UK had its first outbreak of BTV in 2007. We simulated testing of all reported cases seven weeks after the start of the outbreak, using antibody and virological tests with a log-normal variance of 14%, and estimated the trend until testing time. The mean R-squared for the estimated trend was 0.70

Discussion

Our results indicate that recovering the trend of infection from multiple test data in this way could be possible with a fairly high rate of accuracy. This has the potential to complement otherwise lacking surveillance data with information on the spread of pathogens – generating information of great usefulness to the response and management of infectious diesase epidemics.

The approach of combining knowledge of disease with surveillance data also highlights the value of longitudinal studies of host-pathogen disease dynamics and diagnostic test behaviour to inform surveillance and response to epidemics.



Bluetongue virus particles

The University of York, Biomathematics and Statistics Scotland, and SRUC supported this work. The project was funded in part by the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 222633, and is being conducted as part of the WildTech project.

Presenting author
Gustaf Rydevik