Time-Series Based Estimation of the Effective Reproduction Number (R_t)

IEMAG Epidemiology Modelling subgroup

May 21, 2020

The method to estimate R_t as proposed by Wallinga and Teunis (2004) has been applied to a number of time series relating to the current covid-19 epidemic in Ireland. This method - devised in 2004 to estimate R_t for the SARS outbreak of the early 2000's - was proposed to provide estimates when there is little (or no) information relating to contacts of individuals, and can provide estimates when the only information available is a time series of daily counts of a statistic used to monitor the daily numbers of people who become infected.

If full contact information were available - and full knowledge of those infected existed and the dates that infection occurred, computing R_t on a given day t would be straightforward. For each infected person, we would be aware of the number of people they infected, and averaging this number of all infected people would give R_t. However, in reality this information is not available, and several aspects of this process must be estimated.

Given we do not know whether person i infected person j on a given day, an estimate of the probability of this occurring is given instead. Without contact tracing, we rely on knowledge of the probability distribution of the *serial interval* - the time difference between an individual becoming infected, and when they pass on the infection to another person to estimate this probability. Currently the serial interval is modeled as having a gamma distribution with a mean of 6.5 - although alternatives are considered to investigate the sensitivity of the approach to changes in this assumption.

Summing these probabilities for each individual i gives an estimate of R for that individual - averaging these over all people on a given day gives the running estimate of R_t for that day. In practice, the above calculation can be formulated in terms of daily case counts (rather than individuals), and so time series of such counts may be used as inputs.

Data

The method has been applied to four distinct time series from Ireland:

1. *HPSC 24 hour external tests backdated* - the reported number of daily cases from returned test results. A back-dating estimation is applied for delayed results - such as those from samples sent to Germany
2. *Confirmed cases - incidence by specimen collection date* - number of positive test results by date of specimen collection.
3. *Confirmed cases - incidence by specimen collection date - excl LTRC and HCW* - as above but excluding cases for health care workers and those in long term residential care
4. *Hospital Admissions* - Daily reported counts of covid-19 cases admitted to hospital.

Results

The results of estimating the values of R_t during the current epidemic - using data up to May 19th - and truncating after May 11th for the specimen collection-based counts (to allow for results not yet returned) - gives the set of curves below. To investigate sensitivity to assumptions regarding the serial interval, calculations using mean values of 4.5, 5.5 and 6.5 were used.
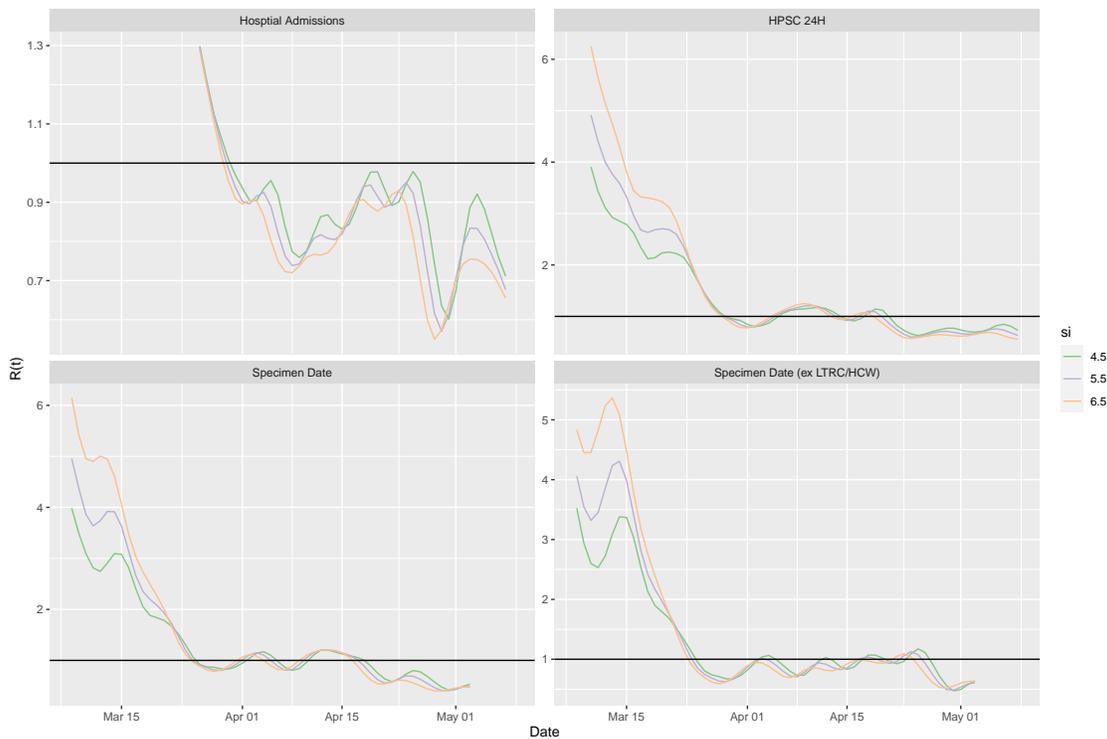
Effects of lockdown policies are apparent - as are differences between the different time series used to estimate the curve. However, all estimates show very similar shaped curves. Note that the y-scale on the hospital admissions plot has a smaller range than the others.

References

Wallinga, Jacco, and Peter Teunis. 2004. "Different Epidemic Curves for Severe Acute Respiratory Syndrome Reveal Similar Impacts of Control Measures." *American Journal of Epidemiology* 160 (6): 509–16.