

Real-time estimation of the effective reproduction number

CNVAC 2022 Seminar series

Sam Abbott
@seabbs
samabbott.co.uk



centre for the
mathematical
modelling of
infectious diseases

LONDON
SCHOOL of
HYGIENE
& TROPICAL
MEDICINE



Talk overview

Who am I

Overview

Challenges

Tools

Case studies

What is left to do?

Summary

Real-time estimation of the effective reproduction number?

Real-time analysis case studies

Who am I?

Sam Abbott

@seabbs

samabbott.co.uk



centre for the
mathematical
modelling of
infectious diseases

LONDON
SCHOOL of
HYGIENE
& TROPICAL
MEDICINE

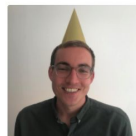


Effective reproduction number estimation

I'm an infectious disease researcher interested in real-time analysis, forecasting, semi-mechanistic modelling, and open source tool development.

- Research Fellow at the London School of Hygiene and Tropical Medicine
- Member of [epiforecasts.io](#)
- PhD in modelling Tuberculosis and the optimal use of the BCG vaccine
- Background in Mathematical modelling and mathematics

Sam Abbott



Sam Abbott

Twitter Posts Notes Research Ideas Code Papers

Posts Notes Research Ideas Code Papers Talks About

Hi there 🙋

I'm an infectious disease researcher interested in real-time analysis, forecasting, semi-mechanistic modelling, and open source tool development. More on my research interests [here](#).

NOW

- Working at the London School of Hygiene and Tropical Medicine in the [EpiForecasts](#) group:
- 🛠️ Crafting extensions to [forecast.vocs](#)
- 📄 Crafted last [epinowcast](#)
- 🏗️ Currently working on:
 - Estimation of the test to test distribution as a proxy for generation interval distribution for the Omicron variant in England
 - Real-time estimation of the time-varying transmission advantages of Omicron in England using S-Genie Target Status as a Proxy
 - Evaluating the use of real-time sequences for short-term forecasting
 - Evaluating a new method for nowcasting eight truncated count data.

BIO

- 🏢 I'm currently working at London School of Hygiene and Tropical Medicine
- 🎓 I did my PhD at the University of Bristol
- 📅 I use daily: `rs`, `stan`
- 🧠 I like to perform analysis using novel models on interesting data and generalise those approaches into software 🤖
- 🗣️ I'm mostly active within the R Community
- 📖 Learning all about Julia and Turing.jl
- 📖 Reading all of [China Méville's](#) work.
- 📧 Ping me about statistical modelling of infectious diseases, real-time analysis of infectious diseases, estimating transmission dynamics in real-time, and team science opportunities
- 📧 Reach me: sam.abbott@lshm.ac.uk

Real-time estimation of the effective reproduction number?

Real-time analysis case studies

Overview

Sam Abbott

@seabbs

samabbott.co.uk



centre for the
mathematical
modelling of
infectious diseases

LONDON
SCHOOL of
HYGIENE
& TROPICAL
MEDICINE



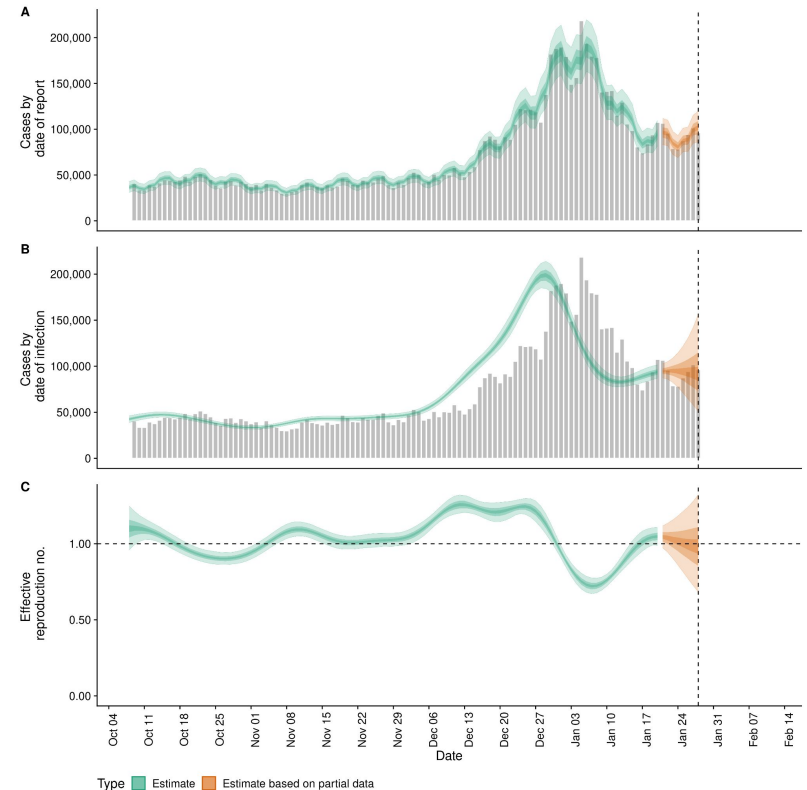


An 80% right paper before a policy decision is made is worth ten 95% right papers afterwards, provided the methodological limitations imposed by doing it fast are made clear.

Effective reproduction number estimation

The average number of secondary infections produced by a single infected person.

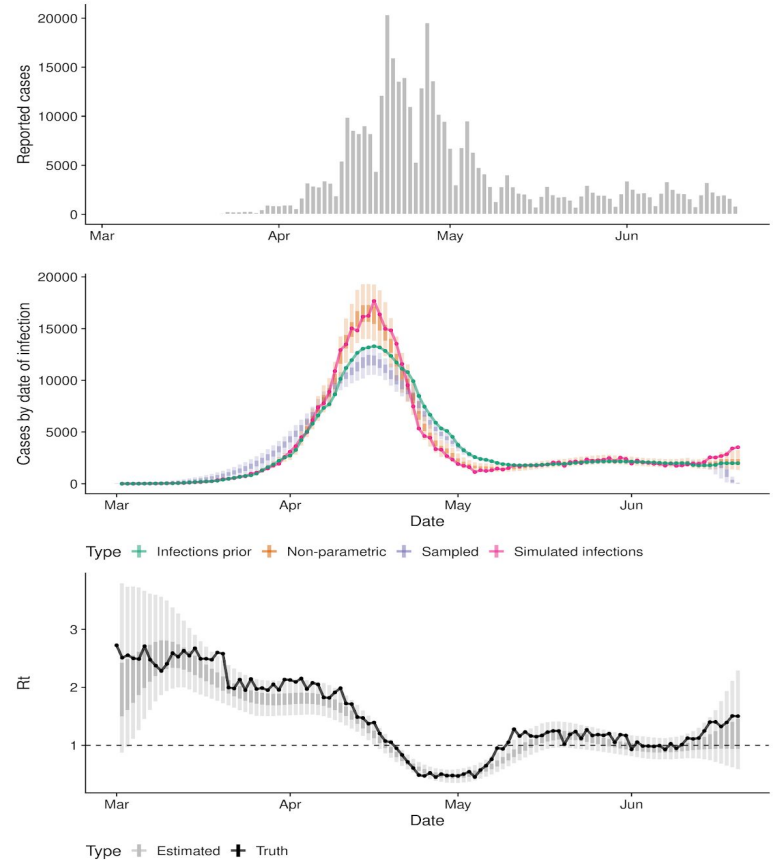
- A helpful metric to track transmission.
- Difficult to estimate as depends on unobserved infections and on the interval between primary and secondary infections.
- Estimating using various methodologies since February 2020.
- Estimates submitted as part of the SPI-M consensus estimate each week.
- Estimates were also published each day for over 1000 locations since April 2020 on epiforecasts.io/covid with over 500k users including the WHO, ECDC, PH agencies.



Effective reproduction number estimation

Key Challenges

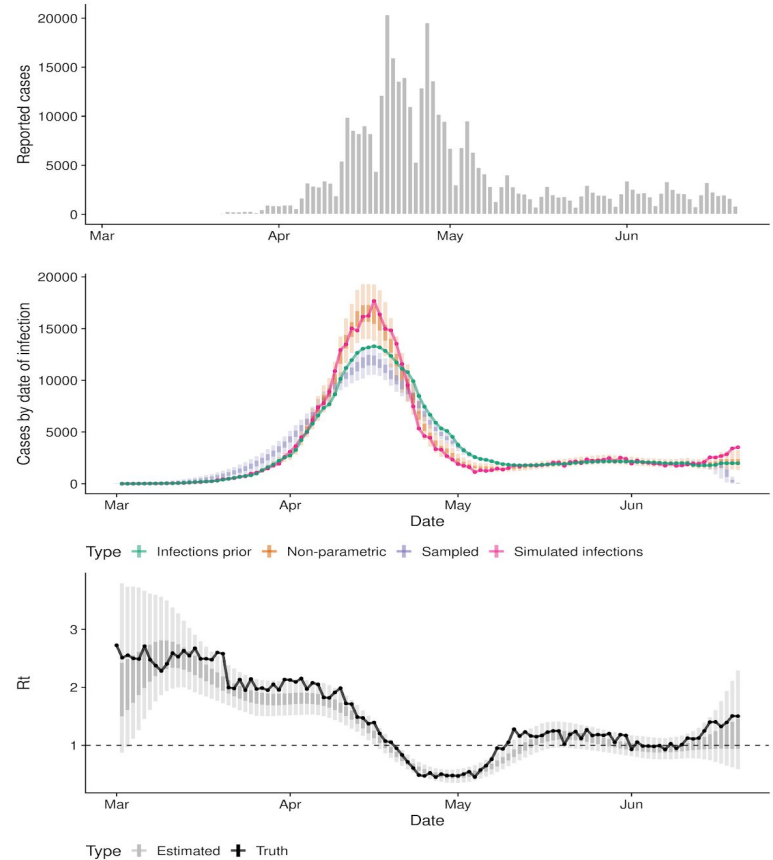
- We care about linking policy changes with changes in transmission but only observe delayed proxies like reported cases and deaths.
- Surveillance data subject to a range of difficult to account for biases.
- Estimation at scale in real-time is computationally challenging.



Effective reproduction number estimation

Objectives

- Develop a model that can be used for real-time surveillance, nowcasting, and short-term forecasting.
- Include known epidemiological structure of the infection and reporting process.
- The model should include a parameter that is referenced to the infection process and that can be used to compare disparate surveillance data source.
- The model should ideally capture changes in trend as quickly as possible to make it useful for situational awareness.



Effective reproduction number estimation

EpiNow2 1.3.3.10 Home News Functions

EpiNow2: Estimate real-time case counts and time-varying epidemiological parameters

License MIT contributors 12 PRs welcome commits since v1.3.2 85 DOI 10.5281/zenodo.5036949

This package estimates the time-varying reproduction number, growth rate, and doubling time using a range of open-source tools ([Abbott et al.](#)), and current best practices ([Gostic et al.](#)). It aims to help users avoid some of the limitations of naive implementations in a framework that is informed by community feedback and is under active development.

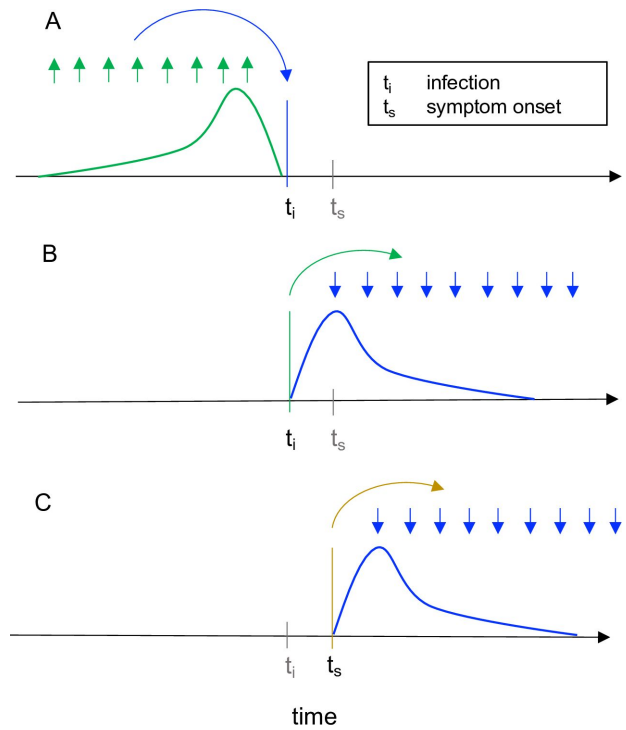
How do we define the reproduction number?

Instantaneous reproduction number

- Transmission at a specific point in time
- Backwards looking
- Can be derived from compartmental models (i.e SEIR)
- Transmission rate * fraction population susceptible * mean duration of infectiousness
- Real-time method and useful for understanding how transmission changes at a given time-point.

$$R_t^{inst} = \beta(t)S(t)D$$

$$R_t = \frac{I_t}{\sum_{s=1}^t I_{t-s}W_s}$$



Case reproduction number

- Transmission by a cohort of individuals
- Forwards looking
- Answer for the question how do individuals now contribute to future transmission.

$$R_t^{case} = \int_{u=t}^{\infty} R_u^{inst} w(u-t) du$$

$$p_{ij} = \frac{w(t_i - t_j)}{\sum_{i \neq k} w(t_i - t_k)}$$

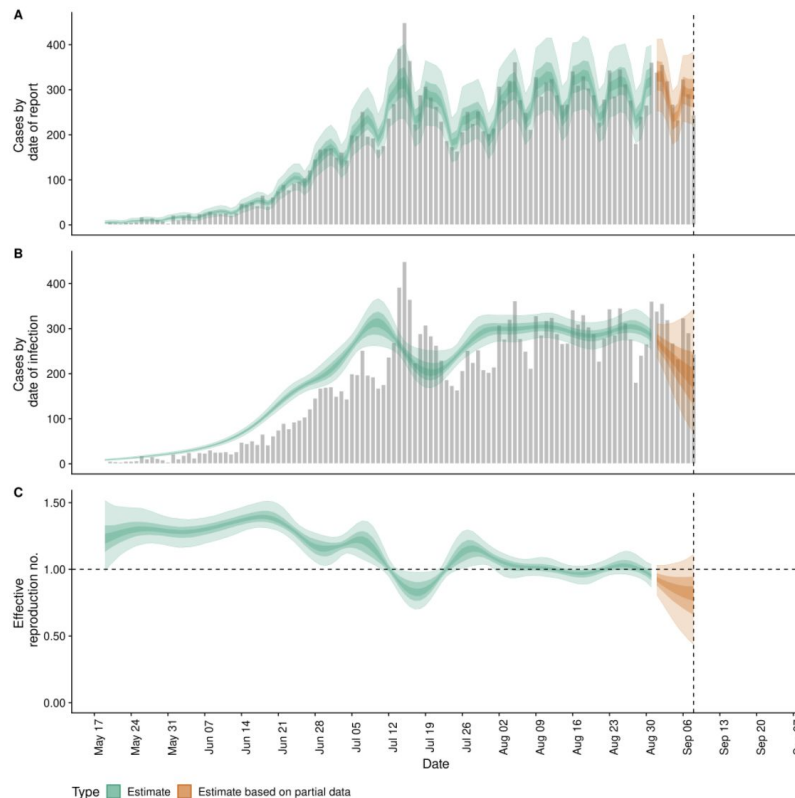
$$R_j = \sum_i p_{ij}$$

Practical considerations for measuring the effective reproductive number, R_t

Katelyn M. Gostic , Lauren McGough, Edward B. Baskerville, Sam Abbott, Keya Joshi, Christine Tedijanto, Rebecca Kahn, Rene Niehus, James A. Hay, Pablo M. De Salazar, Joel Hellewell, Sophie Meakin, James D. Munday, [...], Sarah Cobey [view all]

Our general approach

- Bayesian approach combining nowcasting and R estimation
- Uncertain generation interval estimates allowing for negative serial interval
- Latent process for estimating I_t
- Negative binomial reporting with multiplicative day-of-the-week effect
- R_t estimates with correlation between R_{t+1} and R_t based on a Gaussian Process prior (with other options).



The model

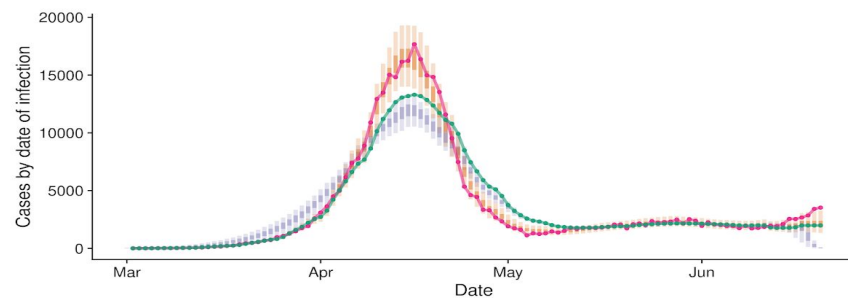
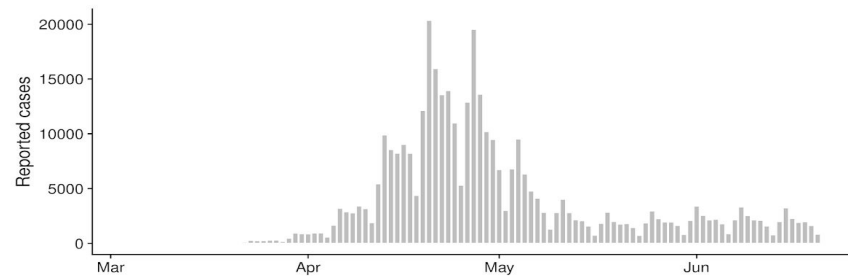
$$\log R_t = \log R_{t-1} + GP_t$$

$$I_t = R_t \sum_{\tau=1}^{15} w(\tau | \mu_w, \sigma_w) I_{t-\tau}$$

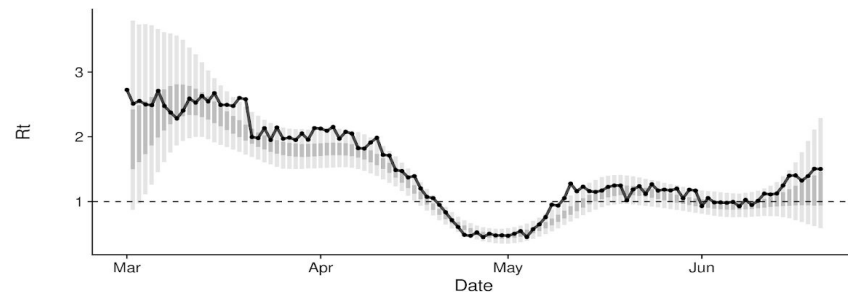
$$O_t = \sum_{\tau=0}^{15} \xi_O(\tau | \mu_{\xi_O}, \sigma_{\xi_O}) I_{t-\tau}$$

$$D_t = \alpha \sum_{\tau=0}^{15} \xi_D(\tau | \mu_{\xi_D}, \sigma_{\xi_D}) O_{t-\tau}$$

$$C_t \sim \text{NB}(\omega_{(t \bmod 7)} D_t, \phi)$$



Type + Infections prior + Non-parametric + Sampled + Simulated infections



Type + Estimated + Truth

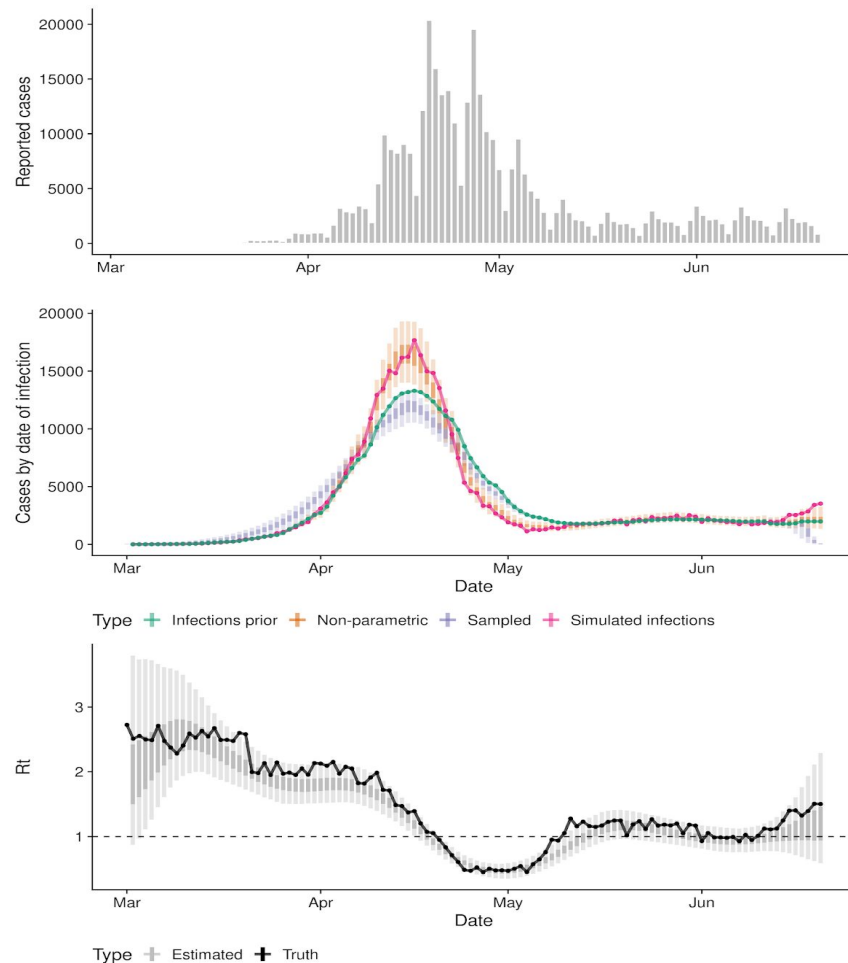
Method: doi.org/10.12688/wellcomeopenres.16006.2

Stan code: git.io/JUxRt

The model

$$I_t = I_0 \exp(rt)$$
$$I_0 \sim \mathcal{LN}(\log I_{obs}, 0.2)$$
$$r \sim \mathcal{LN}(r_{obs}, 0.2)$$

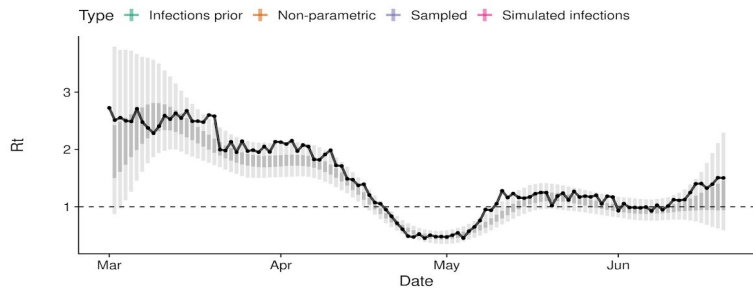
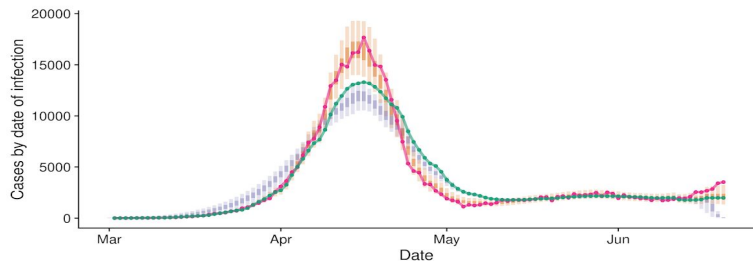
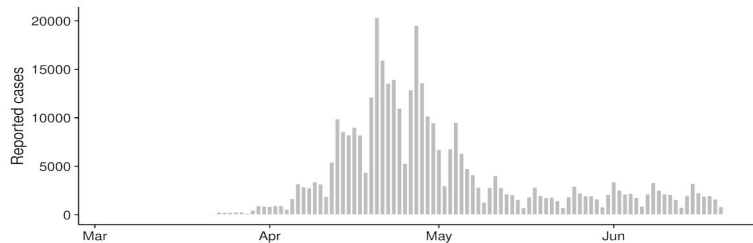
$$w \sim \mathcal{G}(\mu_w, \sigma_w)$$
$$\xi_O \sim \mathcal{LN}(\mu_{\xi_O}, \sigma_{\xi_O})$$
$$\xi_D \sim \mathcal{LN}(\mu_{\xi_D}, \sigma_{\xi_D})$$



Method: doi.org/10.12688/wellcomeopenres.16006.2

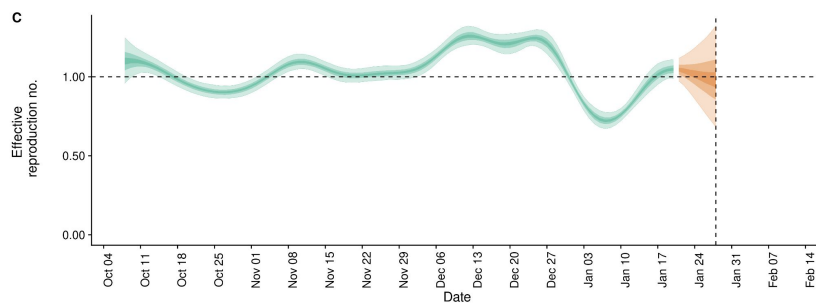
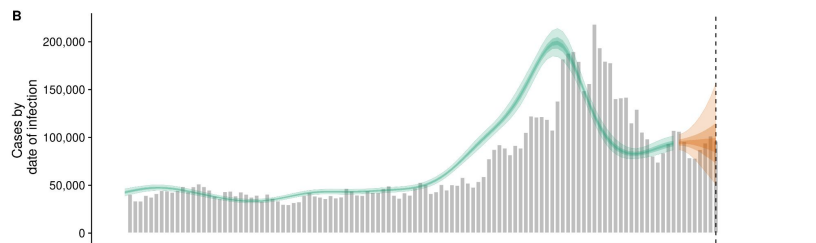
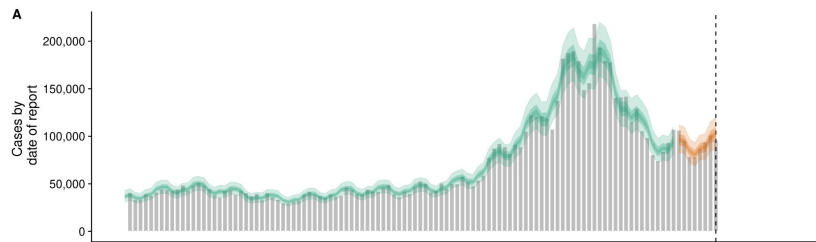
Stan code: git.io/JUxRt

Effective reproduction number estimation



Type + Infections prior + Non-parametric + Sampled + Simulated infections

Type + Estimated + Truth



Type Estimate Estimate based on partial data

Real-time estimation of the effective reproduction number?

Real-time analysis case studies

Challenges

Sam Abbott

@seabbs

samabbott.co.uk



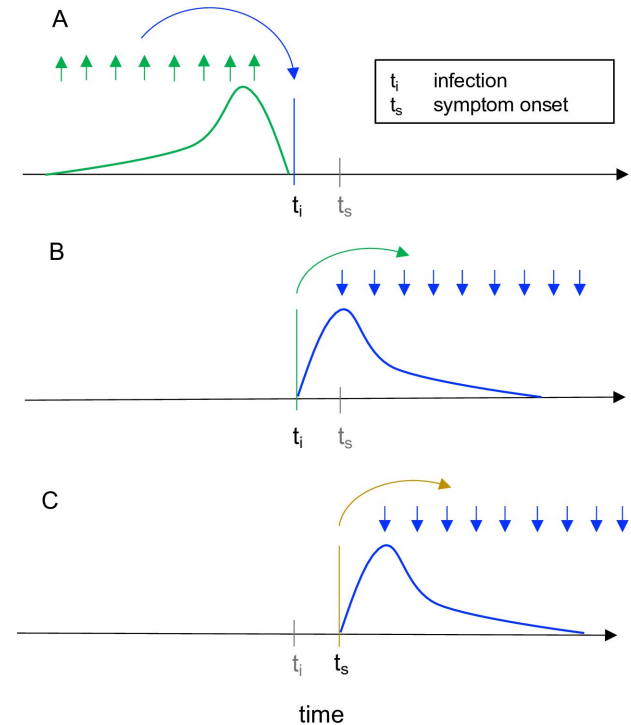
centre for the
mathematical
modelling of
infectious diseases

LONDON
SCHOOL of
HYGIENE
& TROPICAL
MEDICINE



What are we actually trying to do?

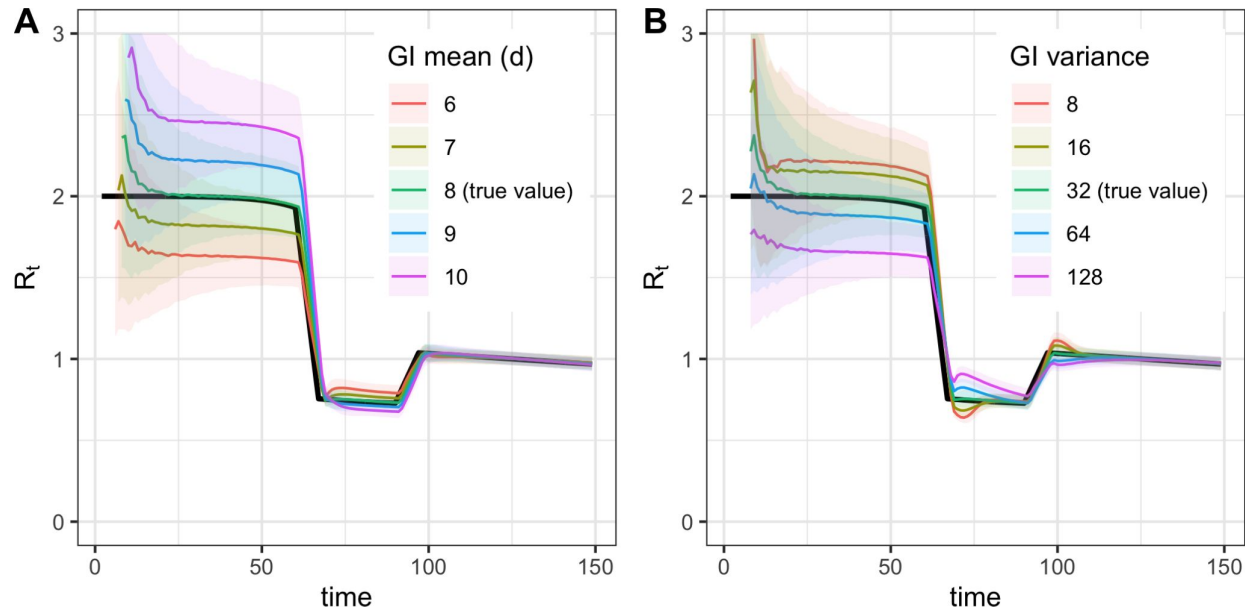
- What kind of reproduction number are we estimating (case, instantaneous etc).
- Are we trying to estimate the reproduction number for the true latent infection process or using observed data as a proxy?
- Do we want a non-parametric estimate, a forecast of future transmission, or an understanding of the impact of NPIs etc.
- What do we a priori believe about how the reproduction number and generation time evolve over time.



Practical considerations for measuring the effective reproduction number, R_t

Katelyn M. Gostic , Lauren McGough, Edward B. Baskerville, Sam Abbott, Keya Joshi, Christine Tedijanto, Rebecca Kahn, Rene Niehus, James A. Hay, Pablo M. De Salazar, Joel Hellewell, Sophie Meakin, James D. Munday, [...], Sarah Cobey [view all]

Generation interval misspecification

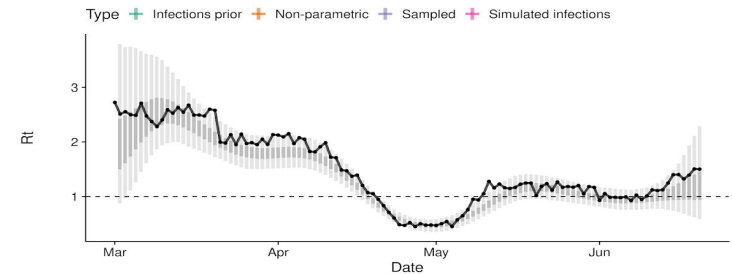
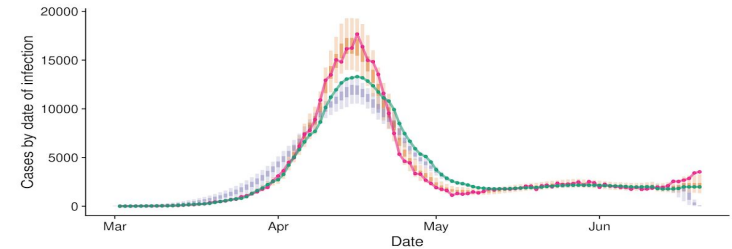
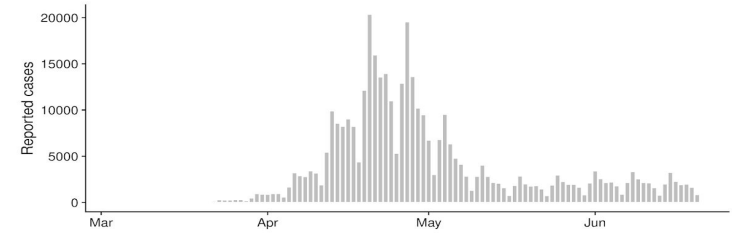
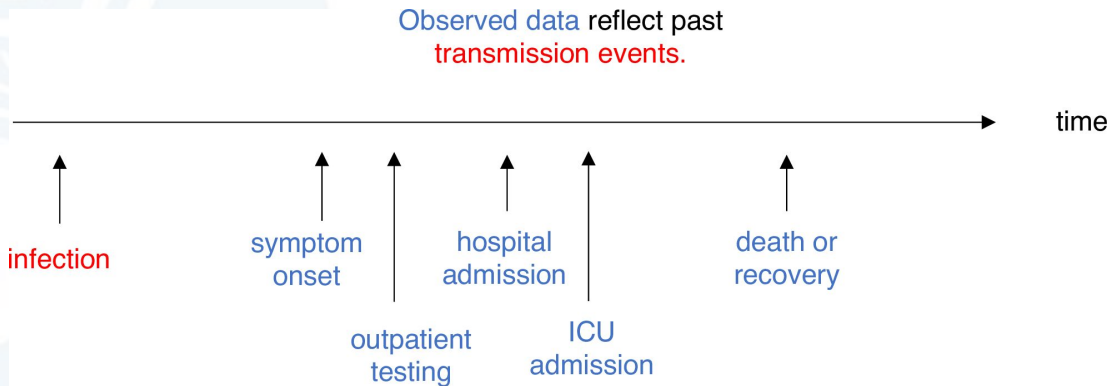


Practical considerations for measuring the effective reproductive number, R_t

Katelyn M. Gostic  Lauren McGough, Edward B. Baskerville, Sam Abbott, Keya Joshi, Christine Tedijanto, Rebecca Kahn, Rene Niehus, James A. Hay, Pablo M. De Salazar, Joel Hellewell, Sophie Meakin, James D. Munday, [...], Sarah Cobey [view all]

Adjusting for delays

We observe some later proxy for infections blurred by a delay distribution.



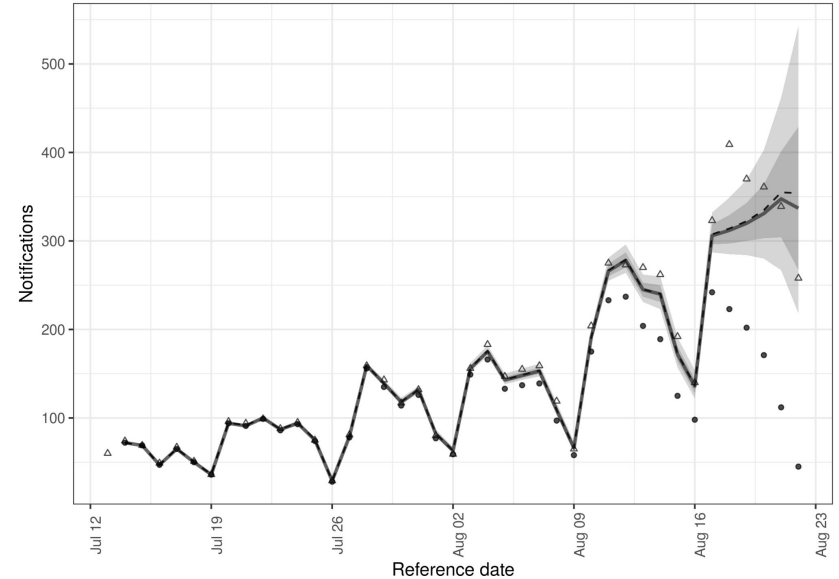
Practical considerations for measuring the effective reproductive number, R_t

Katelyn M. Gostic  Lauren McGough, Edward B. Baskerville, Sam Abbott, Keya Joshi, Christine Tedijanto, Rebecca Kahn, Rene Niehus, James A. Hay, Pablo M. De Salazar, Joel Hellewell, Sophie Meakin, James D. Munday, [...], Sarah Cobey [view all]

Adjusting for right truncation

Incoming data can be delayed meaning what we currently observed isn't what we will observe.

Equally as infections are observed via a delay after reconstructing them they are also not fully observed.



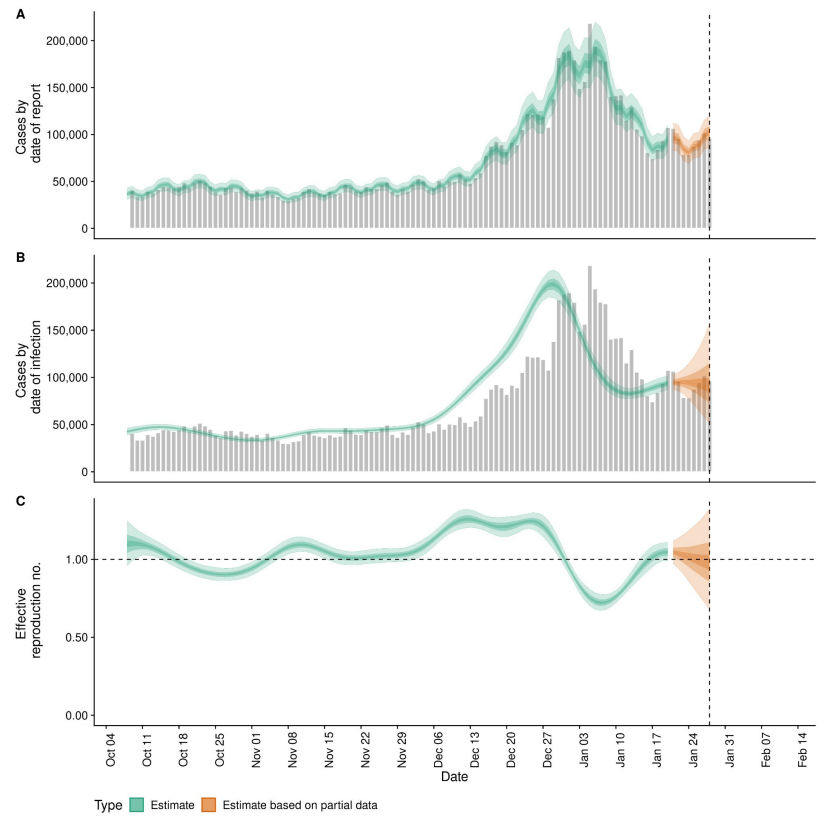
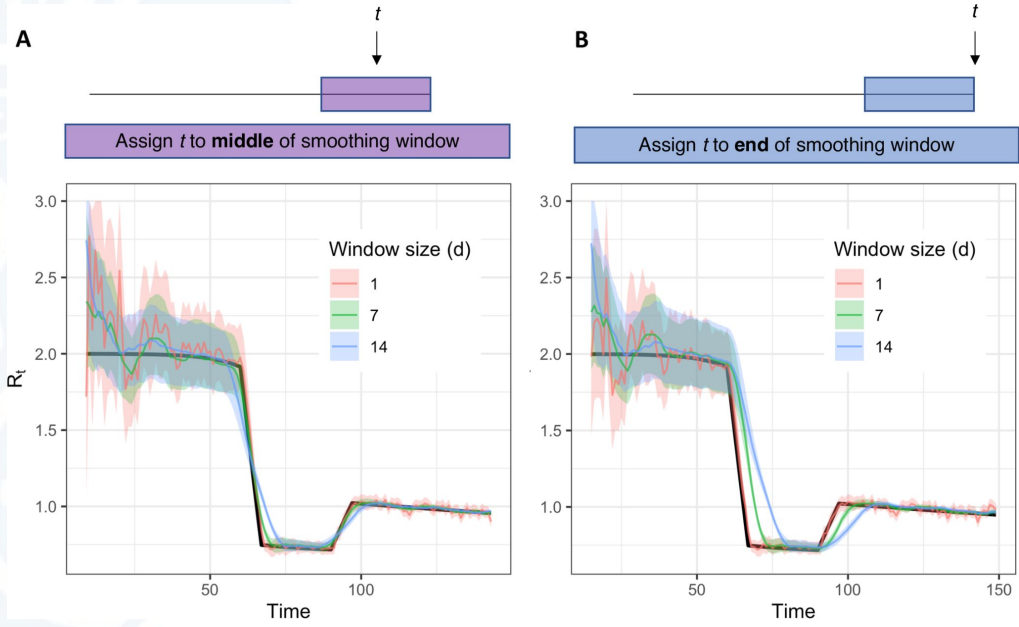
Practical considerations for measuring the effective reproductive number, R_t

Katelyn M. Gostic  Lauren McGough, Edward B. Baskerville, Sam Abbott, Keya Joshi, Christine Tedijanto, Rebecca Kahn, Rene Niehus, James A. Hay, Pablo M. De Salazar, Joel Hellewell, Sophie Meakin, James D. Munday, [...], Sarah Cobey [view all]

Published: December 10, 2020 • <https://doi.org/10.1371/journal.pcbi.1008409>

How to smooth

How does the reproduction number evolve over time and how much of that do we want to include in our estimates?



Practical considerations for measuring the effective reproductive number, R_t

Katelyn M. Gostic  Lauren McGough, Edward B. Baskerville, Sam Abbott, Keya Joshi, Christine Tedijanto, Rebecca Kahn, Rene Niehus, James A. Hay, Pablo M. De Salazar, Joel Hellewell, Sophie Meakin, James D. Munday, [...], Sarah Cobey [view all]

Retrospective or real-time?

Real-time = Orange. This is expensive to care about.

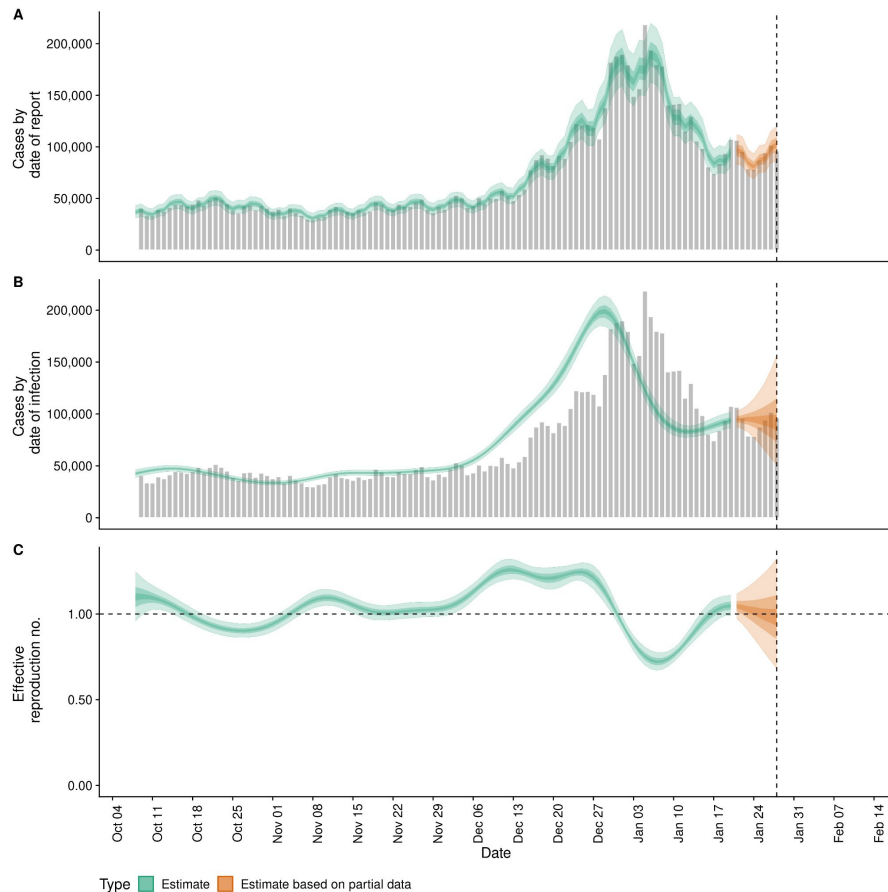
$$\log R_t = \log R_{t-1} + GP_t$$

$$I_t = R_t \sum_{\tau=1}^{15} w(\tau | \mu_w, \sigma_w) I_{t-\tau}$$

$$O_t = \sum_{\tau=0}^{15} \xi_O(\tau | \mu_{\xi_O}, \sigma_{\xi_O}) I_{t-\tau}$$

$$D_t = \alpha \sum_{\tau=0}^{15} \xi_D(\tau | \mu_{\xi_D}, \sigma_{\xi_D}) O_{t-\tau}$$

$$C_t \sim \text{NB}(\omega_{(t \bmod 7)} D_t, \phi)$$



Method: doi.org/10.12688/wellcomeopenres.16006.2

Stan code: [git.io/JUxRt](https://github.com/juuxrt)

Real-time estimation of the effective reproduction number?

Real-time analysis case studies

Tools

Sam Abbott

@seabbs

samabbott.co.uk



centre for the
mathematical
modelling of
infectious diseases

LONDON
SCHOOL of
HYGIENE
& TROPICAL
MEDICINE



EpiNow2 - Real-time non-parametric estimation

EpiNow2 1.3.3.10 Home News Functions

EpiNow2: Estimate real-time case counts and time-varying epidemiological parameters

License MIT contributors 12 PRs welcome commits since v1.3.2 85 DOI 10.5281/zenodo.5036949

This package estimates the time-varying reproduction number, growth rate, and doubling time using a range of open-source tools ([Abbott et al.](#)), and current best practices ([Gostic et al.](#)). It aims to help users avoid some of the limitations of naive implementations in a framework that is informed by community feedback and is under active development.

Epidemia - Hierarchical regression + flexible model options

[Install](#)[Model ▾](#)[Tutorials ▾](#)[Reference](#)[News](#)

epidemia

Flexibly specify and fit Bayesian, regression-oriented models for infectious diseases. The implemented models define a likelihood for all observed data while also explicitly modeling transmission dynamics: an approach often termed as *semi-mechanistic*. Multiple regions can be modeled simultaneously with multilevel models. Key epidemiological quantities, including reproduction numbers and latent infections, may be estimated within the framework. The models may be used to evaluate the determinants of changes in transmission rates, including the effects of control measures. Epidemic dynamics may be simulated either from a fitted model or a "prior" model; allowing for prior/posterior predictive checks, experimentation, and forecasting.

[Source Code](#)[License : GPL-3](#)[Citation](#)[Authors](#)










Getting Started

After [installing](#) the software, the best way to get started is to become familiar with the [model](#) and its [implementation](#). Once familiar with the model, there are a few articles showcasing the main features of the package using real example data. The simplest such article is [here](#).

EpiEstim - The core model, well implemented

EpiEstim: Estimate Time Varying Reproduction Numbers from Epidemic Curves

Tools to quantify transmissibility throughout an epidemic from the analysis of time series of incidence as described in Cori et al. (2013) <[doi:10.1093/aje/kwt133](https://doi.org/10.1093/aje/kwt133)> and Wallinga and Teunis (2004) <[doi:10.1093/aje/kwh255](https://doi.org/10.1093/aje/kwh255)>.

Version: 2.2-4
Depends: R (≥ 2.10)
Imports: [coarseDataTools](#) (≥ 0.6-4), stats, graphics, [reshape2](#), [ggplot2](#), [gridExtra](#), [fitdistrplus](#), [coda](#), [incidence](#) (≥ 1.7.0), [scales](#), grDevices
Suggests: [testthat](#), utils, [vdiff](#), [covr](#), [knitr](#), [rmarkdown](#)
Published: 2021-01-07
Author: Anne Cori  [aut, cre], Simon Cauchemez [ctb], Neil M. Ferguson  [ctb], Christophe Fraser  [ctb], Elisabeth Dahlqwist  [ctb], P. Alex Demarsh [ctb], Thibaut Jombart  [ctb], Zhian N. Kamvar  [ctb], Justin Lessler  [ctb], Shikun Li [ctb], Jonathan A. Polonsky  [ctb], Jake Stockwin [ctb], Robin Thompson  [ctb], Rolina van Gaalen [ctb]
Maintainer: Anne Cori <a.cor@imperial.ac.uk>
BugReports: <https://github.com/mrc-ide/EpiEstim/issues>
License: [GPL-2](#) | [GPL-3](#) [expanded from: GPL (≥ 2)]
URL: <https://github.com/mrc-ide/EpiEstim>
NeedsCompilation: no
Materials: [README NEWS](#)
In views: [Epidemiology](#)
CRAN checks: [EpiEstim results](#)

Documentation:

Reference manual: [EpiEstim.pdf](#)
Vignettes: [EpiEstim demonstration](#)

Downloads:

Package source: [EpiEstim_2.2-4.tar.gz](#)
Windows binaries: r-devel: [EpiEstim_2.2-4.zip](#), r-release: [EpiEstim_2.2-4.zip](#), r-oldrel: [EpiEstim_2.2-4.zip](#)
macOS binaries: r-release (arm64): [EpiEstim_2.2-4.tgz](#), r-oldrel (arm64): [EpiEstim_2.2-4.tgz](#), r-release (x86_64): [EpiEstim_2.2-4.tgz](#), r-oldrel (x86_64): [EpiEstim_2.2-4.tgz](#)
Old sources: [EpiEstim archive](#)

Reverse dependencies:

Reverse imports: [covid19india](#), [earlyR](#), [EpiLPS](#)
Reverse suggests: [epidemia](#), [projections](#)

Linking:

Please use the canonical form <https://CRAN.R-project.org/package=EpiEstim> to link to this page.

Real-time estimation of the effective reproduction number?

Real-time analysis case studies

Case studies

Sam Abbott

@seabbs

samabbott.co.uk



centre for the
mathematical
modelling of
infectious diseases

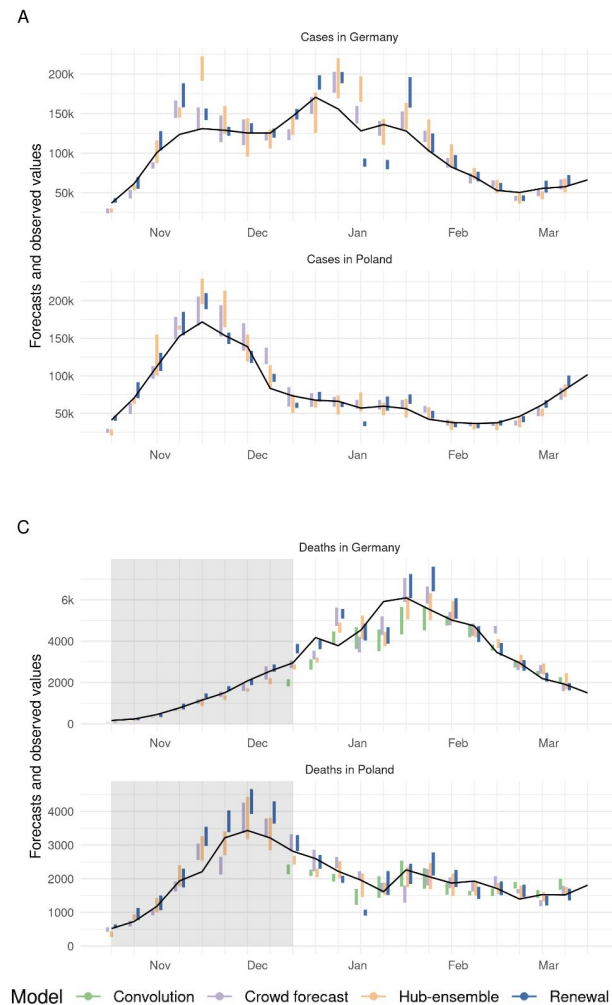
LONDON
SCHOOL of
HYGIENE
& TROPICAL
MEDICINE



Short-term forecasts

What do we think will happen to reported metrics over the next 1-4 weeks

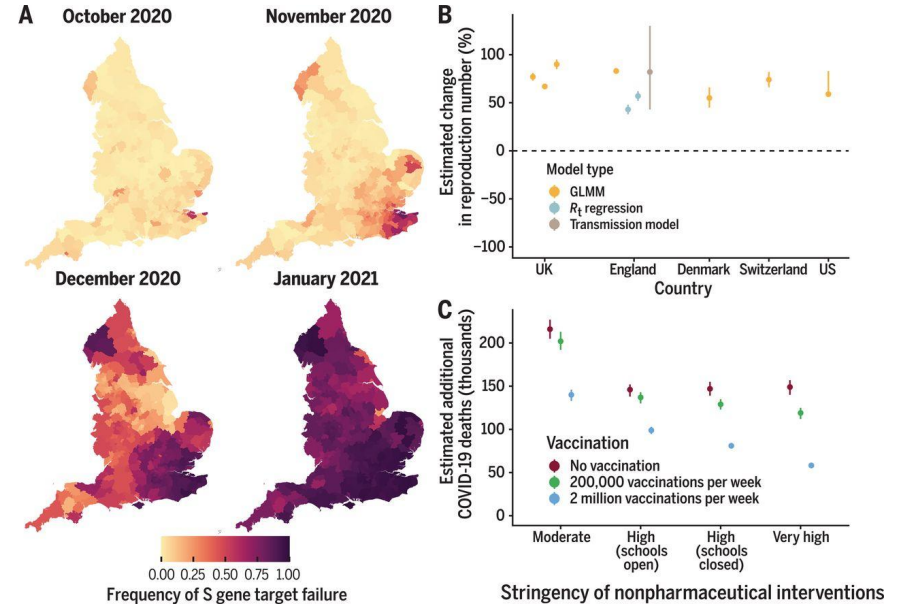
- Using the reproduction number model and similar discrete convolutions models.
- Submitting to the ECDC and CDC forecasting hubs weekly as well as to SPI-M 3 times a week until mid 2020.
- Performs well compared to other models but outperformed by an all-model ensemble.
- Also outperformed by a human judgement model.
- All forecasts struggle to account for policy changes and behavioural changes.



Estimating the transmission advantage of Alpha

How much more transmissible is Alpha than wild-type?

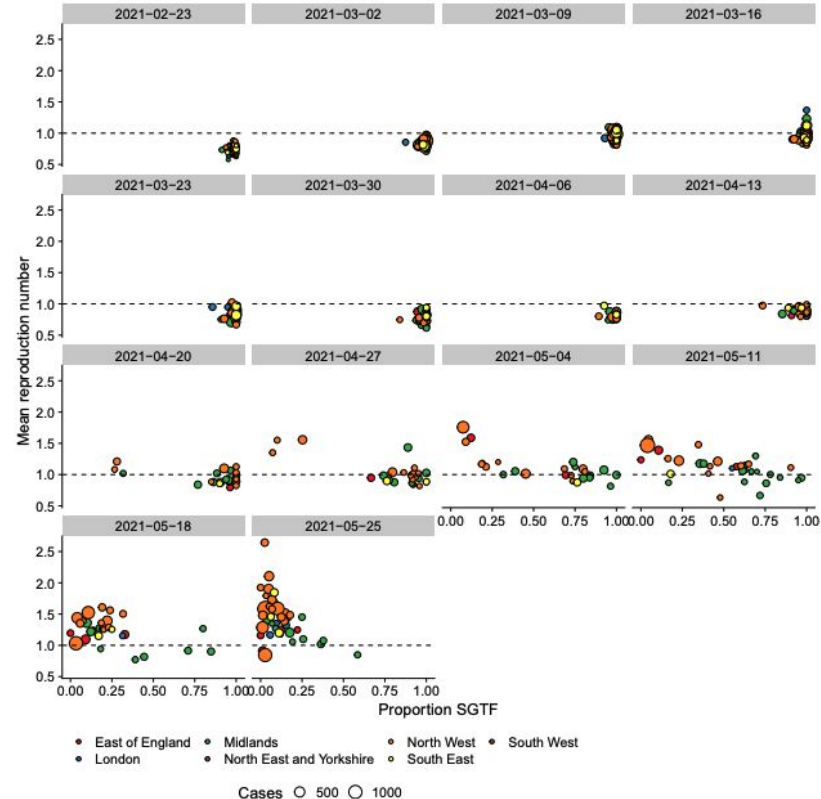
- Part of a multi-method approach by Davies et al.
- Used **reproduction number estimates** by Lower-Tier local authority as “data” + S-gene target failure status (SGTF) as a proxy for variant status.
- Estimated the transmission advantage using an extended regression model adjusted for confounders
- Work done between December 20th 2020 and January 1st 2021.



Estimating the transmission advantage of Delta

How much more transmissible is Delta than Alpha?

- Repurposed the approach used for Alpha using **reproduction number estimates** as data.
- Extended the methodology to include uncertainty for the reproduction number estimates and SGTF status.
- Estimates combined with others as part of the SPI-M consensus statement.

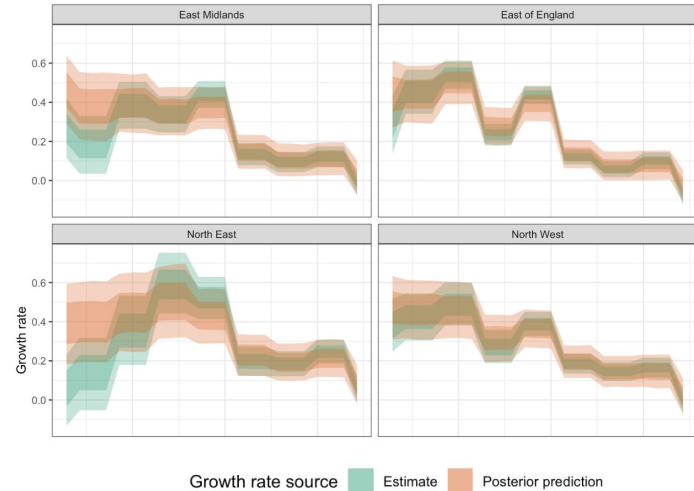


Estimating the generation time of Omicron

Is Omicron's generation time shorter than Delta's?

- Observed reduction in transmission advantage could indicate a shorter generation time.
- This is due to the relationship between the daily growth rate and the reproduction number.
- Used growth rates for Omicron and Delta to explore this and found that a shorter generation time was plausible.
- Findings supported by a study from UKHSA using household contact data.
- Results available in early January and formed part of the SPI-M consensus statement.

$$R = \left(1 + \bar{G}kr\right)^{\frac{1}{k}}$$



Real-time estimation of the effective reproduction number?

Real-time analysis case studies

What is left to do?

Sam Abbott

@seabbs

samabbott.co.uk



centre for the
mathematical
modelling of
infectious diseases

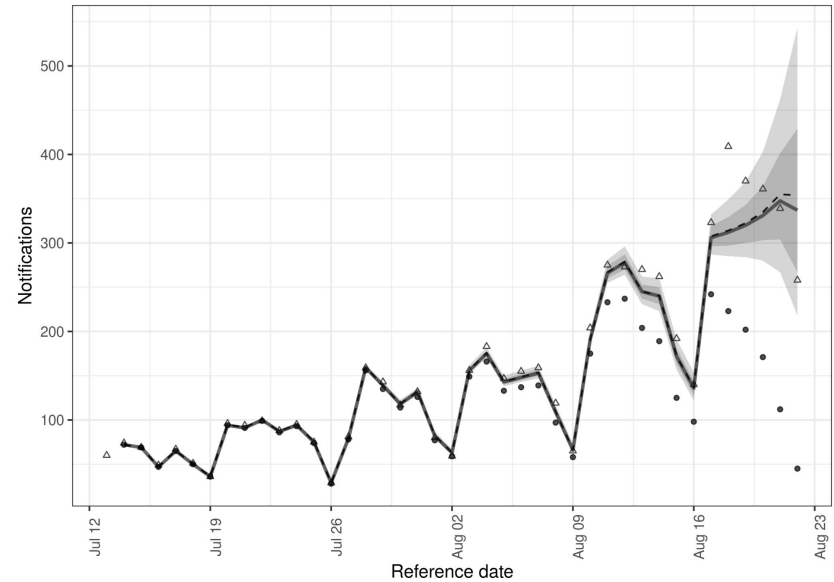
LONDON
SCHOOL of
HYGIENE
& TROPICAL
MEDICINE



Nowcasting

What is happening now to metrics we partially observe

- **International collaboration** estimating 7-day hospital admissions by date of positive test in Germany (the key metric used by decision makers).
- Statistical approach is to decompose the model into forecast and reporting delay components.
- Multi-method ensemble outperforms any single model.
- Nowcasting is conceptually difficult to understand and so potentially under used.
- Open access tools, such as **epinowcast**, make access easier.



Validation - which method works best when

Latest Estimate of R-effective is:

0.95

Spread of COVID-19 is likely stable

What does a R-eff of this size mean

Low/High Estimates of R-effective:

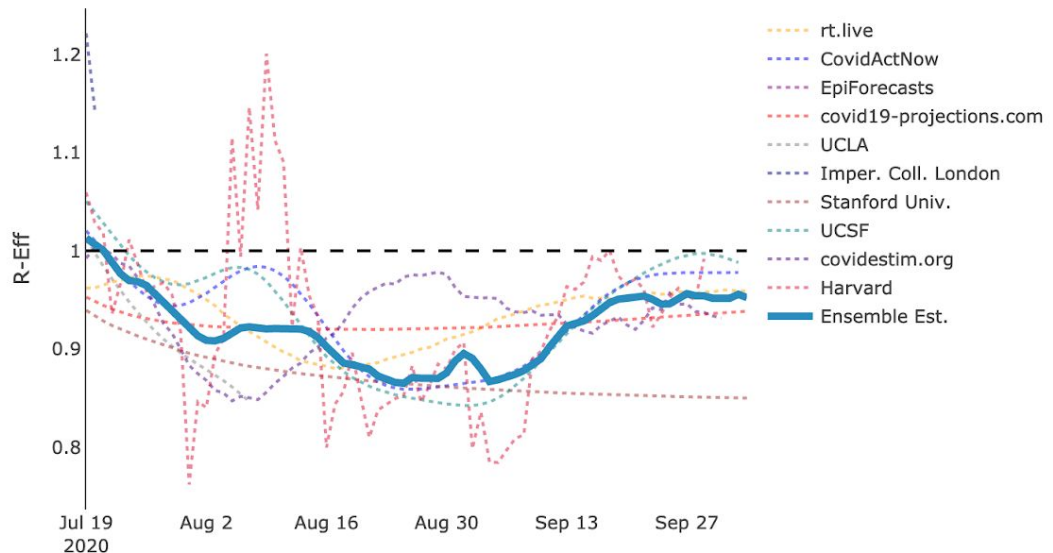
0.85 - 0.96

Stanford - rt.live

Download R-eff Values

Statewide Estimates of R-effective

The effective reproductive number (R) is the average number of secondary infected persons resulting from a infected person. If $R > 1$, the number of infected persons will increase. If $R < 1$, the number of infected persons will decrease. At $R = 1$, the number of infected persons remains constant.



Novel data sources

Estimating epidemiological quantities from repeated cross-sectional prevalence measurements

 Sam Abbott,  Sebastian Funk

doi: <https://doi.org/10.1101/2022.03.29.22273101>

Estimating epidemiologic dynamics from cross-sectional viral load distributions

[JAMES A. HAY](#) , [LEE KENNEDY-SHAFFER](#) , [SANJAT KANJILAL](#) , [NIALL J. LENNON](#) , [STACEY B. GABRIEL](#), [MARC LIPSITCH](#) , AND [MICHAEL J. MINA](#) 

[Authors Info & Affiliations](#)

SCIENCE · 3 Jun 2021 · Vol 373, Issue 6552 · [DOI: 10.1126/science.abh0635](https://doi.org/10.1126/science.abh0635)

Other

- Documentation, testing, case studies, modularity.
- Joint estimation of the time-varying generation time.
- Estimation and support for time-varying delays.
- Better understanding of how the reproduction number should be modelled to evolve over time.
- Hybrid models producing optimal retrospective and real-time estimates
- Linking stochastic and deterministic models
- Linking to phylodynamic methods
- Faster methods without sacrificing robustness

Real-time estimation of the effective reproduction number?

Real-time analysis case studies

Summary

Sam Abbott

@seabbs

samabbott.co.uk



centre for the
mathematical
modelling of
infectious diseases

LONDON
SCHOOL of
HYGIENE
& TROPICAL
MEDICINE



Summary


Overview

Challenges

Tools

Case studies

What is left to do?



Thanks to the epiforecast.io group and my [collaborators](#). Please see individual slides for links containing more details of each case study.