Local area reproduction numbers and S-gene positivity

Work in progress - not peer reviewed

Sam Abbott, Sebastian Funk on behalf of the CMMID Covid-19 Working Group

1 June, 2021

For correspondence: sebastian.funk@lshtm.ac.uk

Abstract

Background: Local estimates of the time-varying effective reproduction number (Rt) of Covid-19 in England have become increasingly heterogeneous in England in the past few weeks. This variation may be attributed to the spread of novel Covid-19 variant "delta." In this report we aimed to investigate the changes in the proportion of positive cases that were S-gene positive, an indicator of the novel variant in front of the background of the previously predominant variant "alpha," and to correlate them with changes in the estimated reproduction number at the level of upper tier local authorities (UTLA).

Method: We explored the relationship between the proportion of samples that were S-gene negative and the effective reproduction number of test positive cases over time. Effective reproduction numbers were estimated using the `EpiNow2` R package independently for each local authority using two different estimates of the generation time. The `brms` R package was used to fit a range of regression models to estimate a multiplicative relationship between S-gene positivity and weekly mean R. Models were compared using the expected log-predictive density. Results are shown as 95% credible intervals.

Results: We found evidence of an association between increased mean Rt and the proportion of S-gene positives across all models evaluated with the magnitude of the effect increasing as model flexibility was decreased. Models that adjusted for either national level or NHS region level residual variation over time were found to fit the data better and yielded consistent results of the association of Rt with S-gene posistivity.

Conclusions: Our results indicate that even after adjusting for between NHS region residual variation over time S-gene positivity was associated with an increase in the effective reproduction number of Covid-19. These findings were robust across a range of models and generation time assumptions though the specific effect size was variable depending on the assumptions used. The lower bound of the estimated effect indicated that the reproduction number will remain above 1 at current levels of restrictions.

Limitations: Associations are purely observational and therefore subject to bias due to confounders not included in the model. Surge testing and infections brought in by travellers may have affected the case numbers and, consequently, reproduction number estimates. Our analysis does not distinguish between increased reproduction numbers due to transmission advantage and transmission advantage due to vaccine escape and therefore cannot distinguish between the two. Whilst we did investigate two different generation interval distributions we did not consider a scenario where the generation times were different between S-gene positives and negatives.

Method

Data

We used 4 main sources of data: test positive Covid-19 notifications by UTLA,[1] S-gene status from PCR tests by local authority provided by Public Health England (PHE), Google mobility data stratified by context,[2]

and data on the timings of national easing. We aggregated the data at the weekly level and restricted the analysis to the period beginning Tuesday, 23 February.

Statistical analysis

We calculated the weekly proportion of positive tests that were S-gene positive over time by local authority. We estimated reproduction numbers using the method described in[3] and[4] and implemented in the `EpiNow2` R package.[5] Daily updated estimates can be downloaded at https://github.com/epiforecasts/covid-rt-estimates/blob/master/subnational/united-kingdom-local/cases/summary/rt.csv. We used two sets of estimates, obtained using uncertain, gamma distributed, generation interval distributions with a mean of 3.6 days (standard deviation (SD): 0.7), and SD of 3.1 days (SD: 0.8)[5,6] or with a mean of 5.5 days (SD: 0.5 days), and SD of 2.1 days (SD: 0.25 days),[7] respectively.

We then built a separate model of the expected reproduction number in UTLA i during week t starting in the week beginning 23 February, 2021, as a function of local restrictions, mobility indicators, residual temporal variation, and proportion of positive tests S-gene positive:

$$R_{i,t} = (1 + \alpha f_{it}) \exp\left(s(t) + \sum_j \beta_j T_{ijt} + \sum_k \gamma_k G_{ikt} + \log R_i \right)$$

where R_t is an UTLA-level intercept corresponding to Rt during national lockdown in February/March, T_{ijt} is 1 if intervention j (out of: January lockdown, reopening phase 1/2) is in place and 0 otherwise, G_{ikt} is the relative mobility in context k (home, workplace, public transport) at time t in UTLA i as measured by Google, and $s(t)$ is a time-varying component, modelled either as a region-specific thin-plate regression spline ("Regional time-varying"), the sum of a static regional parameter and a national spline ("National time-varying"), or only a static regional parameter ("Regional static"). We lastly fitted the model for each t separately, i.e. at fixed time slices ("Time-sliced").

The key parameter is α, the relative change in reproduction number in the presence of s-gene positivity that is not explained by any of the other variables, where f_{it} is the proportion out of all positive tests for SARS-CoV-2 where the S-gene was tested negative, and the reproduction number in any given UTLA is

$$R_{t,i} = (1 - f_{it}) R^+_{t,i} + f_{it} R^-_{t,i}$$

where $R^-_{t,i}$ is the S-gene negative reproduction number, $R^+_{t,i}$ is the S-gene positive reproduction number, and it is assumed that $R^+_{t,i} = (1 + \alpha) R^-_{t,i}$.

We used a Student's t-distribution observation model with a single variance parameter and a single degrees of freedom parameter. We restricted the analysis to UTLA/week combinations where more than 20% of reported cases and more than 20 in total had an S-gene status reported. All models were implemented using the `brms`[8] package in `R`. All code required to reproduce this analysis is available from https://github.com/epiforecasts/covid19.sgene.utla.rt/ (`may-update` branch).

Results

We found consistent evidence of an association between S-gene positivity and increased UTLA level reproduction number estimates. The association became more apparent over time. The outer 90% credible intervals of the estimates in the time-sliced analysis (Fig. 2) ranged from a 1% increase to a 80% increase, in April/May, depending on the assumption about generation times and the model used, as the proportion of of tests that were S-gene positive increased heterogeneously across NHS regions. Out of the models fitted to all time, ones that adjusted for residual variation over time on both a national and NHS region level fit the data better than those that did not but all models had evidence of increased Rt with S-gene positivity with the best fitting model yielding a lower bound of 20% higher Rt of S-gene positive cases with a short generation time (Table 1), higher than the model that only adjusted for national level residual variation over time (lower bound: 21%) and lower than the model that did not adjust for residual variation over time (lower bound:

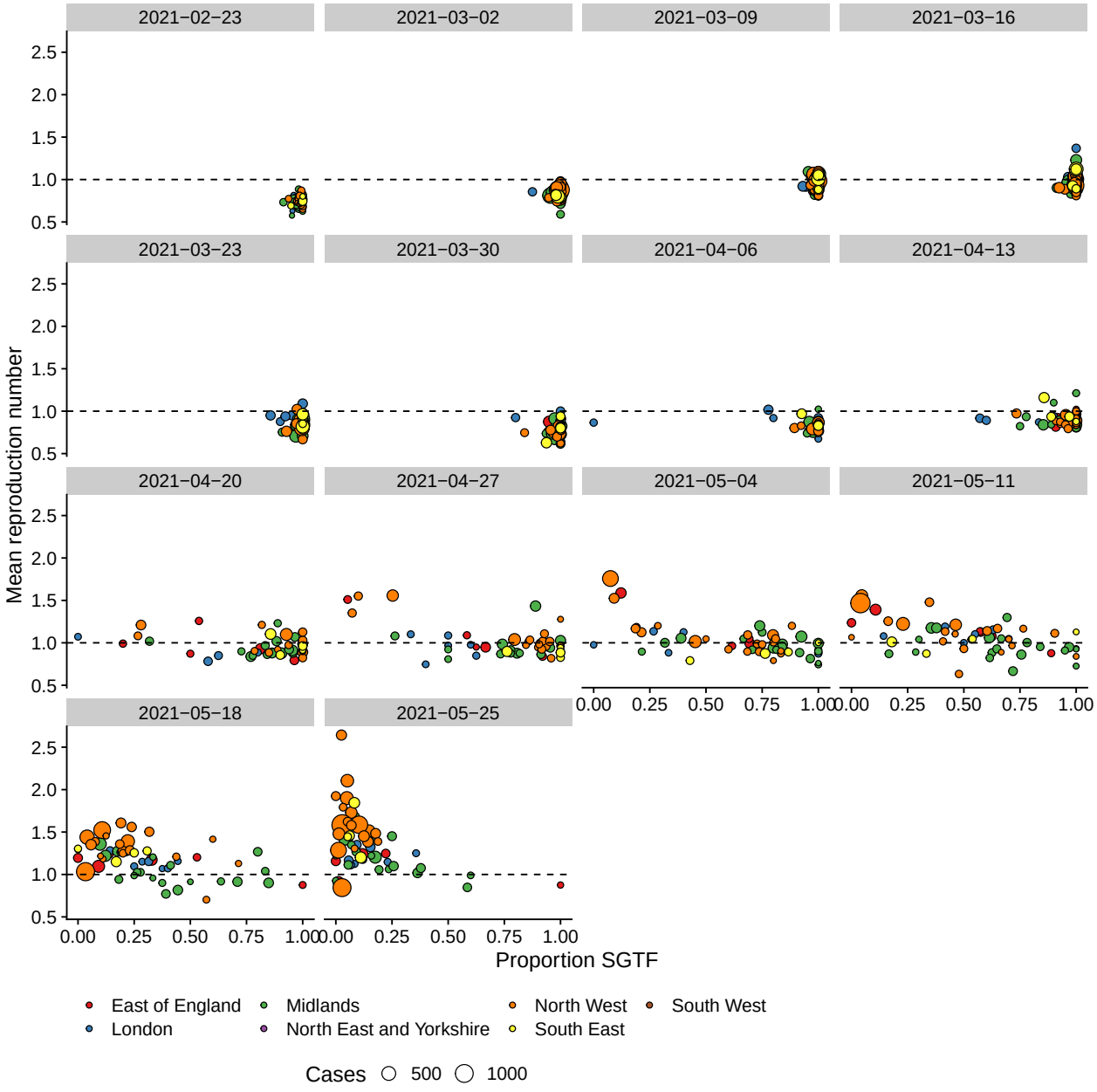

Figure 1: *Mean reproduction numbers using a generation time with a mean of 5.5 days since the week beginning 23 February, 2021, compared to the proportion of all test-positives tested for S-gene that tested S-gene positive/negative that week. Each point represents one UTLA, with the size given by the number of cases in the week following the week of the given reproduction number to account for the delay from infection to testing. Only UTLAs with sufficient coverage of S-gene results at least 20% of cases tested and at least 20 results in total)*

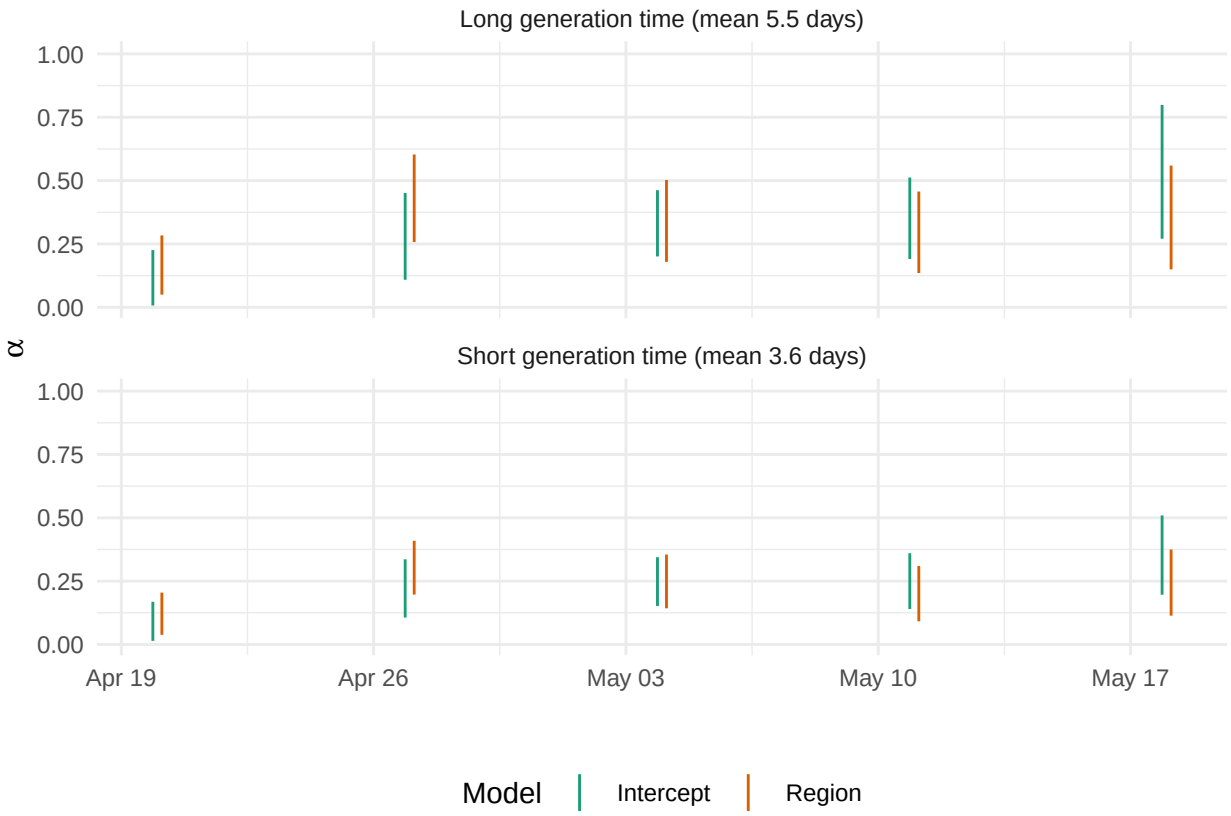

Figure 2: *Estimates of α from the time-sliced model.*

Table 1: *Parameter α with 95% credible intervals for the three different models of $s(t)$ for short (3.6 days mean) and long (5.5 days mean) generation intervals. The estimate corresponds to the multiplicative increase in reproduction number estimated for S-gene positive cases.*

Model	Estimate (short GT)	Estimate (long GT)
Regional static	0.28–0.4	0.38–0.55
National time-varying	0.21–0.35	0.27–0.47
Regional time-varying	0.2–0.34	0.27–0.46

28%). With a longer generation time, these lower bounds changed to 27%, 27%, and 38%, respectively. The upper bound of the increase in Rt varied from 34% to 55% in models that fit the data comparably (but had different assumed generation times).

The best fitting model appeared to reproduce estimated reproduction numbers well over time (Fig. 3) although there are notable outliers in some of the weeks. Alternative model parametrisations fit the data less well than those presented here whilst producing comparable effect size estimates.

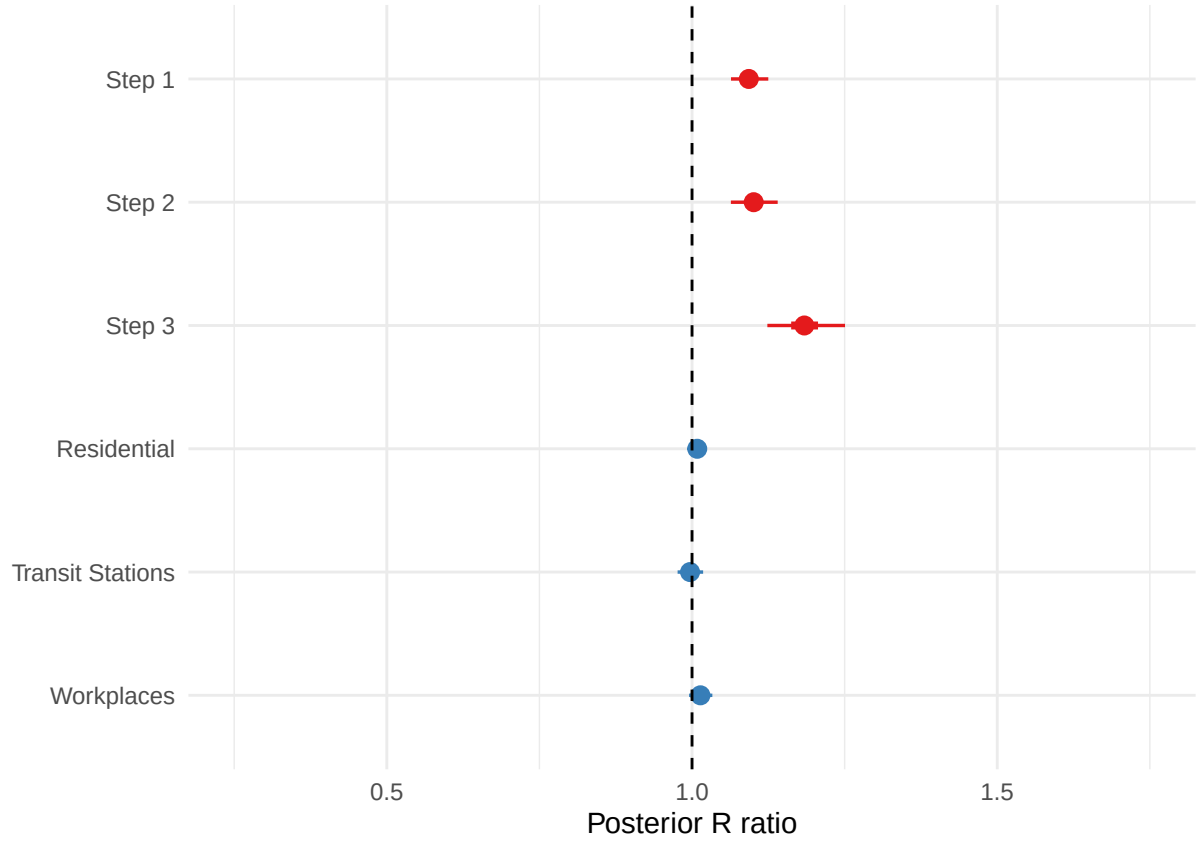

Figure 3: *Parameter estimates (R ratios) of the Regional static model. These correspond to a multiplicative effect on a baseline R of 0.76–0.83). Intervention indicators are coloured in red, and mobility indicators in blue.*

References

1. *Coronavirus (covid-19) in the UK.* (2021). https://coronavirus.data.gov.uk/details/healthcare.

2. *Covid-19 community mobility reports.* (2021). https://www.google.com/covid19/mobility/.

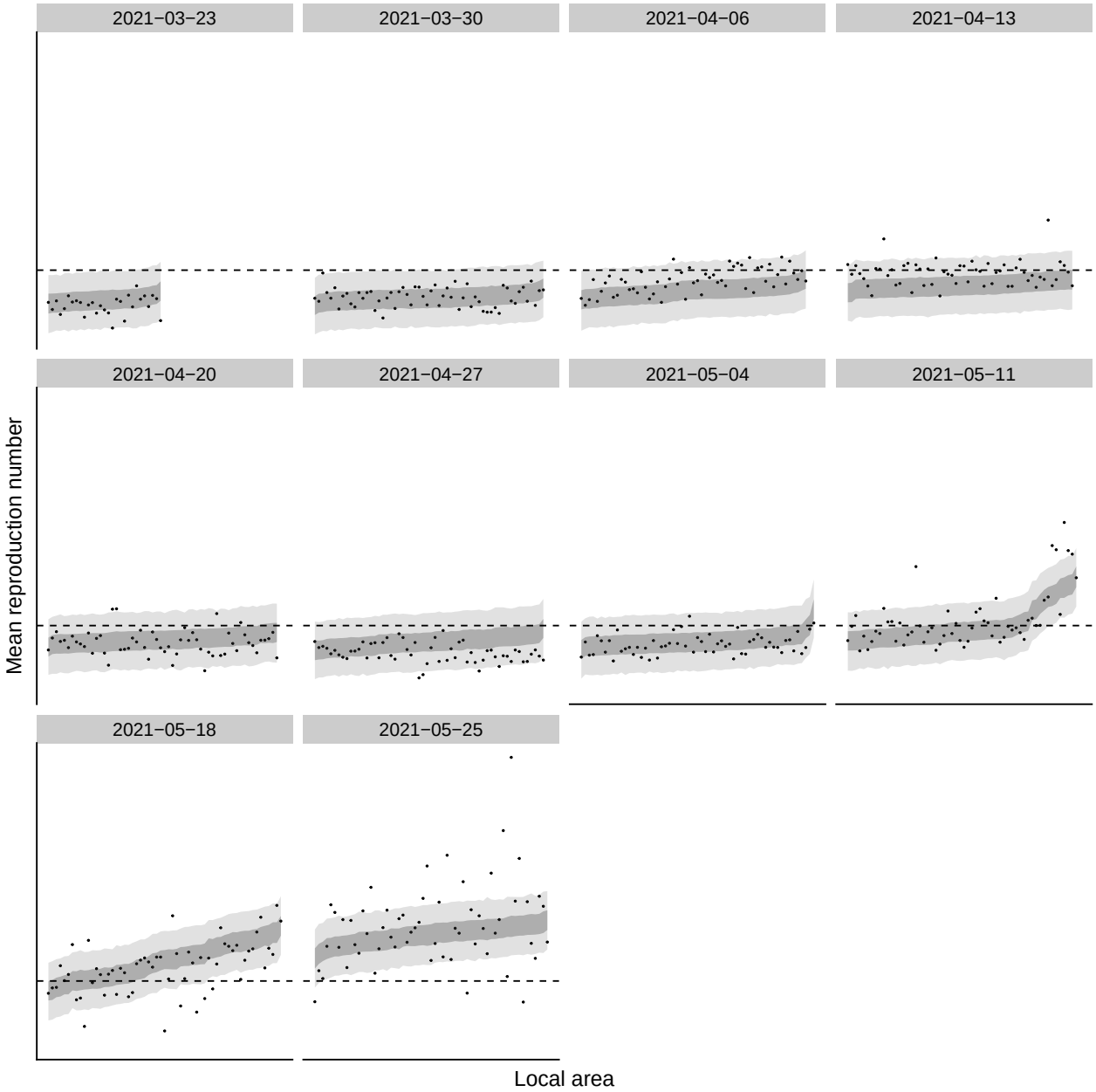

Figure 4: *Predictions of the regional static model with long generation interval (per-region spline) compared to the data (solid dots). Dark grey: central 50% prediction interval; light grey: central 90% prediction interval. Areas are ordered each week according to predicted median.*

3. Abbott, S., Hellewell, J., Thompson, R., Sherratt, K., Gibbs, H., Bosse, N., Munday, J., Meakin, S., Doughty, E., Chun, J., Chan, Y., Finger, F., Campbell, P., Endo, A., Pearson, C., Gimma, A., Russell, T., null, null, Flasche, S., ... Funk, S. (2020). Estimating the time-varying reproduction number of SARS-CoV-2 using national and subnational case counts. *Wellcome Open Research*, *5*(112). https://doi.org/10.12688/wellcomeopenres.16006.2

4. Sherratt, K., Abbott, S., Meakin, S. R., Hellewell, J., Munday, J. D., Bosse, N., Jit, M., & Funk, S. (2020). Evaluating the use of the reproduction number as an epidemiological tool, using spatio-temporal trends of the covid-19 outbreak in england. *medRxiv*. https://doi.org/10.1101/2020.10.18.20214585

5. Abbott, S., Hellewell, J., Sherratt, K., Gostic, K., Hickson, J., Badr, H. S., DeWitt, M., Thompson, R., Epi-Forecasts, & Funk, S. (2020). *EpiNow2: Estimate real-time case counts and time-varying epidemiological parameters*. https://doi.org/10.5281/zenodo.3957489

6. Ganyani, T., Kremer, C., Chen, D., Torneri, A., Faes, C., Wallinga, J., & Hens, N. (2020). Estimating the generation interval for coronavirus disease (COVID-19) based on symptom onset data, march 2020. *Eurosurveillance*, *25*(17). https://doi.org/10.2807/1560-7917.ES.2020.25.17.2000257

7. Ferretti, L., Ledda, A., Wymant, C., Zhao, L., Ledda, V., Abeler-Dorner, L., Kendall, M., Nurtay, A., Cheng, H.-Y., Ng, T.-C., Lin, H.-H., Hinch, R., Masel, J., Kilpatrick, A. M., & Fraser, C. (2020). The timing of COVID-19 transmission. *medRxiv*. https://doi.org/10.1101/2020.09.04.20188516

8. Bürkner, P.-C. (2018). Advanced Bayesian multilevel modeling with the R package brms. *The R Journal*, *10*(1), 395–411. https://doi.org/10.32614/RJ-2018-017