Association between S-gene target failure and the case fatality rate of Covid-19

Preliminary analysis - not yet peer reviewed

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

12 January, 2021

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

Aim

Explore the association between S-gene target failure and the case fatality rate of Covid-19 using test positive Covid-19 notifications by UTLA and Covid-19 deaths by UTLA.

Method

Data

We used 3 main sources of data: test positive Covid-19 notifications by UTLA,[1] deaths linked to Covid-19 notification within 28 days of notification,[1] and S-gene status from PCR tests by local authority provided by Public Health England (PHE)[2]. We aggregated the data at the weekly level and restricted the analysis to the period beginning Monday, 5 October.

Statistical analysis

We calculated the weekly proportion of positive tests that were S-gene negative over time by local authority and adjusted to date of infection from date of specimen by shifting all estimates back by a week. We adjusted cases and deaths at the local authority level from date of notification to date of infection using the deconvolution method described in in[3] and[4] and implemented in the `EpiNow2` R package[5] (without any adjust for the fraction of latents reported in each dataset). Daily updated estimates can be downloaded at https://github.com/epiforecasts/covid-rt-estimates/blob/master/subnational/united-kingdom-local/cases/summary/cases_by_infection.csv. We used central estimates without uncertainty and aggregated by week of infection.

We assumed that the observed number of Covid-19 deaths (D) within 28 days by date of infection were a function of Covid-19 notifications (D) by date of infection scaled by the case fatality rate of S-gene positive cases (c^+) and S-gene negative cases (c^-),

$$D_{i,t} \sim \text{NB}\left(c^+ \left(1 - f_{it}\right) C_{i,t} + c^- f_{it} C_{i,t}, \phi\right)$$

where i indicates UTLA, t week of infection, and f_{it} is the fraction of cases that were found to be S-gene negative by UTLA each week. The case fatality rate of S-gene negative cases then assumed was then assumed to be a function of static local variation, normalised Covid-19 notifications by date of infection, and residual temporal variation.

$$c^+ = \text{logit}^{-1}\left(\gamma_i + \beta C_{i,t} + s(t)\right)$$

where γ_i is a UTLA-level intercept corresponding to the baseline case fatality rate, β is the contribution of case load (represented by normalised notifications by date of infection), and $s(t)$ is a time-varying component, modelled either as a region-specific thin-plate regression spline, the sum of a static regional parameter and a national spline.

The S-gene negative case fatality rate was then assumed to be related to the S-gene positive case fatality rate via a multiplicative relationship,

$$c^- = \alpha c^+$$

or an additive relationship

$$c^- = \alpha + c^+$$

where α represents either the multiplicative change in case fatality rate or the additive change. These alternative parameterisations represent either a population wide effect for the former parameterisation or a subpopulation effect in the latter parameterisation.

All models were implemented using the `brms`[6] package in `R`. All code required to reproduce this analysis is available from https://github.com/epiforecasts/covid19.sgene.utla.rt/.

Results

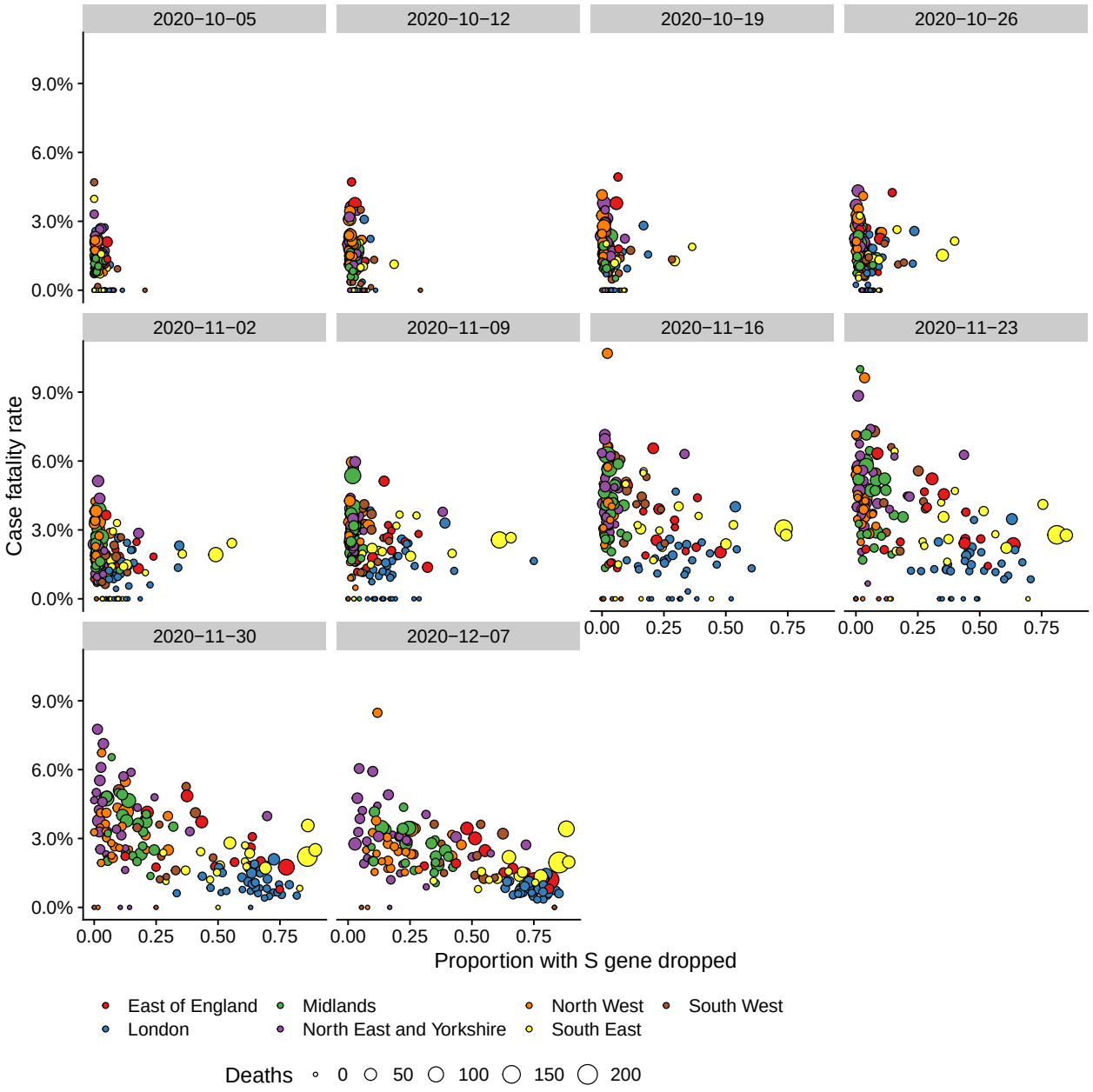

Figure 1: Proportion with S gene dropped compared to the adjusted case fatality rate each week beginning Monday the 5th of October. Each point represents one UTLA, with the size given by the number of deaths by date of infection.

Type	Model	ELPD difference
Additive	All covariates with regional time-varying	0.00
Multiplicative	All covariates with regional time-varying	-37.55
Additive	All covariates with national time-varying	-63.96
Multiplicative	All covariates with national time-varying	-82.53
Additive	All covariates	-426.24
Multiplicative	All covariates	-440.19
Additive	UTLA	-605.58
Multiplicative	UTLA	-616.85
Multiplicative	Time	-804.45
Additive	Time	-809.80
Multiplicative	Region	-839.40
Additive	Region	-842.25
Multiplicative	Cases	-1060.87
Additive	Intercept	-1060.91
Multiplicative	Intercept	-1060.93
Additive	Cases	-1060.96

Table 1: Model comparison by difference in expected log-predictive density. Models are identified by the covariates they include.

Model	Additive	Multiplicative
All covariates with regional time-varying	0.5% (0.37%, 0.63%)	-63% (-75%, -49%)
All covariates with national time-varying	0.34% (0.21%, 0.46%)	-37% (-51%, -21%)
All covariates	1.1% (0.94%, 1.3%)	210% (180%, 260%)
UTLA	0.75% (0.56%, 0.96%)	110% (78%, 150%)
Time	-2.3% (-2.6%, -2%)	-77% (-83%, -70%)
Region	1.2% (0.84%, 1.5%)	120% (80%, 160%)
Cases	-0.76% (-1.1%, -0.39%)	-36% (-50%, -20%)
Intercept	-0.83% (-1.2%, -0.46%)	-37% (-51%, -22%)

Table 2: Parameter α with 95% credible intervals for the both additive and multiplicative assumptions across all models evaluated. In the additive model the effect can be interpreted as a direct change in the case fatality rate related to S-gene negativity whilst in multiplicative model the effect can be interpreted as a scaling of the S-gene positive case fatality rate.

Discussion

We studied the relationship between SGTF (as a proxy for the new variant of concern) and the case fatality rate of Covid-19 adjusted using both an additive and a multiplicative relationship. We considered a range of confounders and explored residual variation on both a national and regional level. Models that adjusted for residual variation fitted the data best but may make interpreting effect sizes difficult. In models that did not adjust for residual variation the best fitting model, that included all covariates, indicated that S-gene negativity was associated with an increase in the baseline case fatality rate though the presence of unmeasured confounders is likely. Models that included the SGTF as an additive effect fit the data better than those that treated it as multiplicative. Additional work is needed before these results can be interpreted further.

References

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

2. England, P. H. (2020). *Investigation of novel sars-cov-2 variant: Variant of concern 202012/01.* https://www.gov.uk/government/publications/investigation-of-novel-sars-cov-2-variant-variant-of-concern-20201201.

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., EpiForecasts, & Funk, S. (2020). *EpiNow2: Estimate real-time case counts and time-varying epidemiological parameters*. https://doi.org/10.5281/zenodo.3957489

6. 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