Exploring the transmission advantage of Omicron in England

Epiforecasts

14 December 2021

Contents

Background

We aimed to assess competing explanations of the transmission advantage of a variant of concern (VoC), Omicron, compared to the existing dominant strain, Delta, in England. We explored the likelihood of increased transmissibility compared to immune escape, using S-gene target failure as a proxy for infection with Omicron.

- It is highly likely that there is extensive transmission of Omicron.
- It is not yet clear if Omicron is more transmissible than other variants, including the Delta variant.
- Omicron has the potential to cause a new wave of infections in the UK; the magnitude of any wave is uncertain.

We use a model framework where we vary only the relationship between variants while holding all other parameters constant. We compare leave-one-out validation among the models and explore the estimated transmission advantage.

Methods

We used case data from England that included information on S-gene target negativity. These are the results of PCR tests following a positive lateral flow test result. We excluded data before 2021-11-01, to reduce overfitting to stochastic imports. Data are included up to 2021-12-05 and modelled at a 7 day resolution.

To model a single strain, we model the mean of reported cases as an order 1 autoregressive (AR(1)) process on the log scale by time unit. The model is initialised by assuming that the initial reported cases are representative with a small amount of error (2.5%). This uses an initial growth rate, where the exponential is the effective reproduction number (Rt). We scaled the reproduction number to the mean generation time, set at 5.1 days. The growth rate is then itself modelled as a differenced AR(1) process. We then assume a negative binomial observation model with overdispersion for reported cases.

We expanded the single strain model to explore dynamics between two strains. This model used the single strain model as a starting point with the addition of strain specific auto-regressive AR(1) variation, and a beta binomial observation process for SGTF data. Mean reported cases are again defined using a AR(1) process on the log scale for each strain and then combined for overall mean reported cases.The mean reported cases due to the variant of concern (VoC) is derived by calculating the mean proportion of cases that had the VoC for the first time point using the overall number of reported cases and the number of cases that were positive for the VoC.

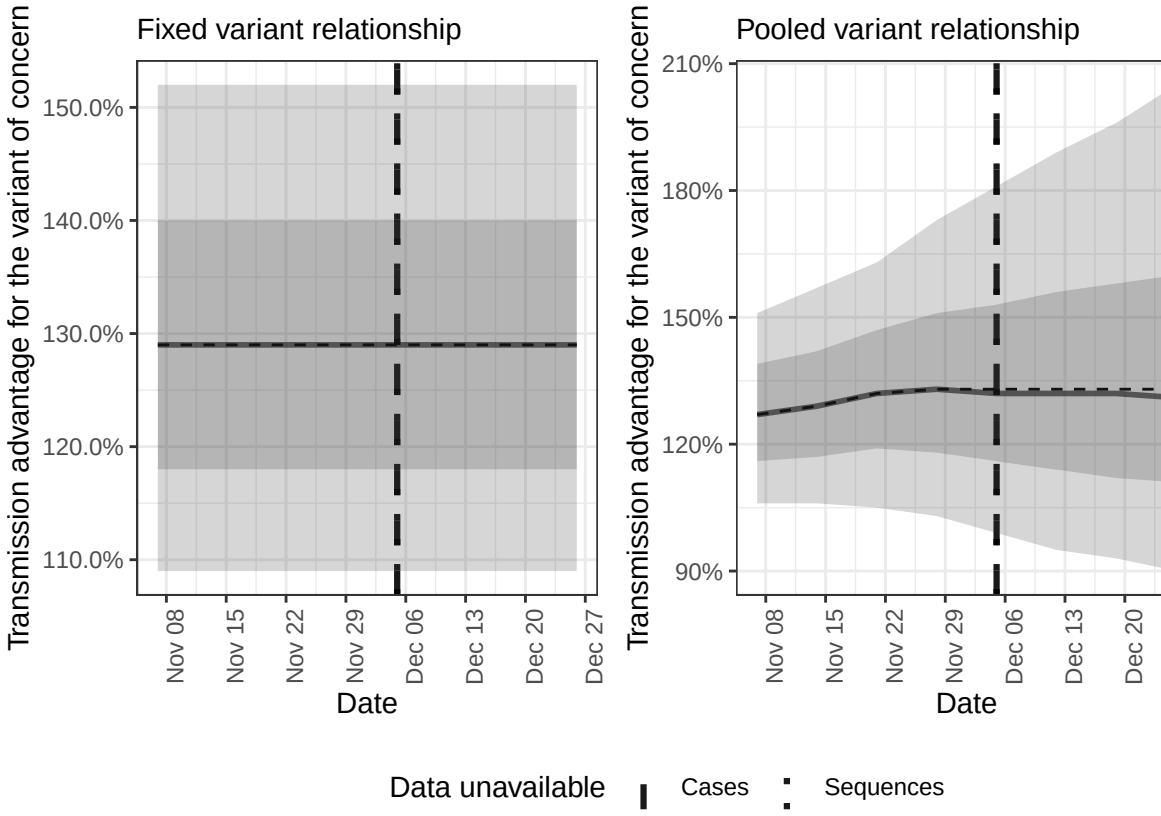
The growth rate for VoC and non-VoC cases is then modelled as a combination of an overall growth rate (as defined for the single strain model), a strain specific modifier, and an AR(1) error term post introduction.

We used a weakly informative prior for a transmission advantage for the VoC vs non-VoC cases of mean 0.21 (standard deviation 0.2), based on early work from South Africa[1]. We defined the relationship between variants as either fixed or pooled. The fixed relationship meant the VoC differs only by the transmissibility difference to the non-VOC strain and any variation over time is shared between strains. The pooled relationship allowed the VoC to differ with dependence determined from the data.

Finally, the mean proportion of samples that have the VoC is then estimated using the mean reported cases with the VoC and the overall mean reported cases. We assume a beta binomial observation model for the number of sequences that are positive for the VoC with overdispersion.

We tested the difference between the two models by comparing out of sample predictive fit using leave-one-out cross-validation with Pareto smoothed importance sampling, and model scoring.

Results

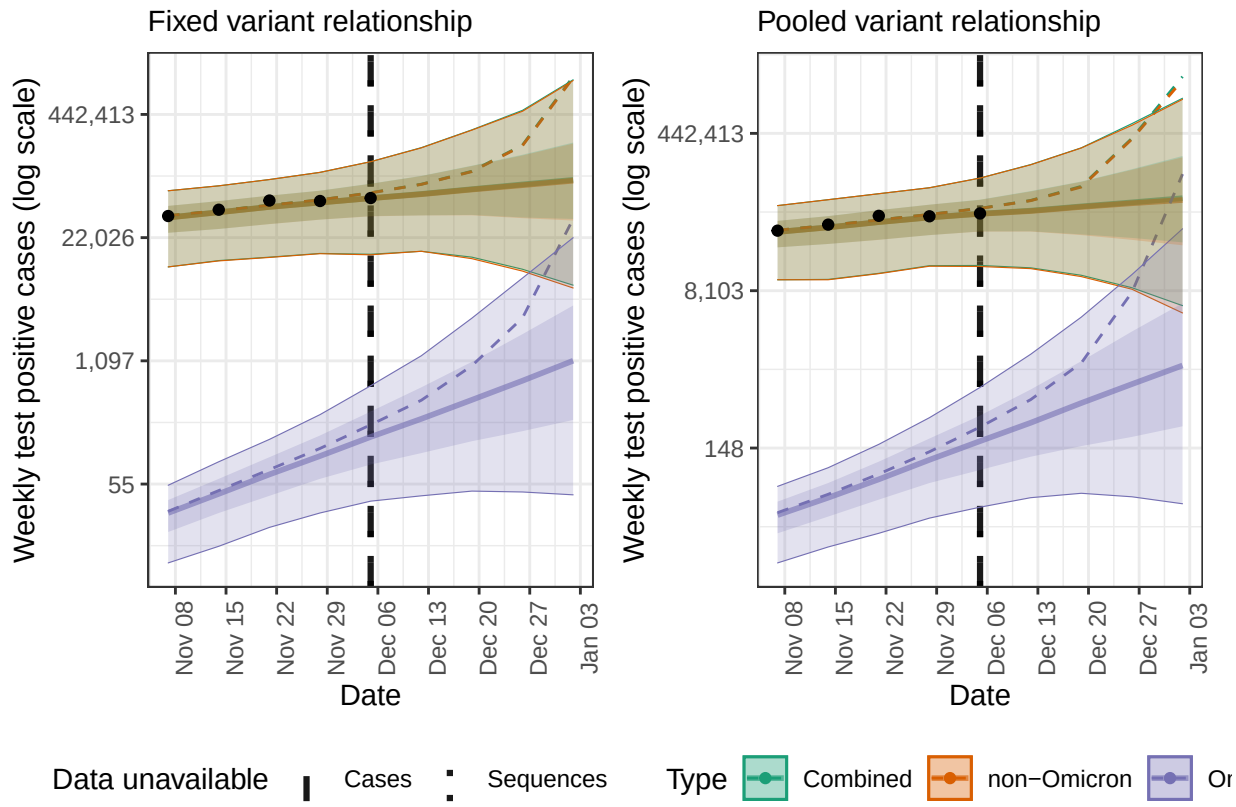


Model estimation

Transmission advantage is shown where 100% is equivalent to the current dominant strain, Delta (figure **??**). Both models indicated a stronger transmission advantage for Omicron.

As a baseline, we assumed a fixed relationship between variants to model transmission advantage without variation over time. Here, estimated Omicron advantage is 1.41 (95% credible interval 1.12 - 1.78). In comparison, we used a pooled relationship drawn from data to allow transmission advantage to vary over time. In this setting, the apparent transmission advantage of Omicron rose over time before stabilising at a higher transmission advantage compared to Delta. On average over time, we found an Omicron advantage of 1.45 (95% CrI 1.06 - 2.01).

[1] 2021-12-03, Carl Pearson and others, "Omicron spread in South Africa", Epidemics8



We modelled cases attributable to each of the two variants under each assumption, with the posterior prediction shown in figure **??** from the two strain model ("Combined") and the unobserved estimates for each strain.

Model comparison We explored the difference in expected predictive performance between the two models. Comparing the models on PSIS-LOO indicated an estimated difference in expected log pointwise predictive density of -0.49 (with a standard error of 0.04) for the pooled model compared to the scaled model.

Variant relationship	Sharpness
scaled	5290
pooled	5550

Discussion

We found that transmission advantage varied depending on the choice of relationship between the two variants. We compared between a model that uses a fixed relationship between variants, against a model that allowed the relationship to vary over time.

This can be interpreted in terms of variant transmissibility compared to immune escape. If the two models show little difference in growth rate and have similar predictive performance, this suggests that increased transmissibility largely accounts for the growth of Omicron. However, if results from each model differ this might suggest transmissibility is not sufficient to explain the spread, and the advantage of Omicron is more likely to be due to immune escape.

- our findings
- limitations

– SGTF poor proxy for Omicron

Appendix: Model parameters and code

- Model parameters

```
## $voc_scale
## [1] 0.21 0.20
##
## $strains
## [1] 2
##
## $r_init
## [1] 0.00 0.25
##
## $scale_r
## [1] 0.7285714
##
## $overdispersion
## [1] TRUE
```

- All code

```
knitr::opts_chunk$set(echo = FALSE, include = FALSE,
                      warning = FALSE, message = FALSE,
                      root.dir = here::here())
# Load packages
library(here)
library(dplyr)
library(tidyr)
library(forecast.vocs)
library(ggplot2)
library(patchwork)
library(purrr)
library(data.table)
library(loo)
library(scoringutils)
library(knitr)
options(mc.cores = 4)

# Load data
source(here("code", "load-data.R"))
obs <- data.table(weekly)
timespan <- 7

# Cut off data before November and <3 days old
start_date <- as.Date("2021-11-01")
end_date <- Sys.Date() - 3
obs <- filter(obs, between(date, start_date, end_date))

# Load parameters
source(here("code", "load-parameters.R"))
voc_label <- "Omicron"
horizon <- 4
output_loglik <- TRUE
```

```
adapt_delta <- 0.99
max_treedepth <- 15
refresh <- 0
show_messages <- FALSE

variant_relationships <- c("scaled", "pooled")

# Model with 1) scaled and 2) time-dependent relationship between variants
forecast_fits <- map(variant_relationships,
                           ~ forecast(obs,
                                # scaled or pooled relationship
                                variant_relationship = .x,
                                # variant options
                                voc_scale = parameters$voc_scale,
                                scale_r = parameters$scale_r,
                                strains = parameters$strains,
                                r_init = parameters$r_init,
                                overdispersion = parameters$overdispersion,
                                timespan = timespan,
                                horizon = horizon,
                                voc_label = voc_label,
                                # processing options
                                output_loglik = output_loglik,
                                adapt_delta = adapt_delta,
                                max_treedepth = max_treedepth,
                                refresh = refresh,
                                show_messages = show_messages))
names(forecast_fits) <- variant_relationships
forecasts <- map(forecast_fits,
                       ~ unnest_posterior(.x) %>%
                         mutate(across(mean:q95, round, 2)))

# Plot VOC advantage
advantage <- map(forecasts,
                 ~ filter(.x, value_type == "model" &
                            variable == "avg_voc_advantage") %>%
                   select(variant_relationship, variable, clean_name,
                          exponentiated:ess_bulk))

voc_scaled <- plot_voc_advantage(forecasts$scaled) +
  labs(subtitle = "Fixed variant relationship")
voc_pooled <- plot_voc_advantage(forecasts$pooled) +
  labs(subtitle = "Pooled variant relationship")

voc_scaled +
  voc_pooled +
  plot_layout(nrow = 1, guides = "collect") +
  plot_annotation(theme = theme(legend.position = "bottom"))

# Plot cases
cases_scaled <- plot_cases(forecasts$scaled, obs, log = TRUE) +
  labs(subtitle = "Fixed variant relationship")
cases_pooled <- plot_cases(forecasts$pooled, obs, log = TRUE) +
```

```
  labs(subtitle = "Pooled variant relationship")
cases_scaled +
  cases_pooled +
  plot_layout(nrow = 1, guides = "collect") +
  plot_annotation(theme = theme(legend.position = "bottom"))

# Compare LOO
models <- c("scaled", "pooled")
loo_scaled <- forecast_fits$scaled$fit[[1]]$loo()
loo_pooled <- forecast_fits$pooled$fit[[1]]$loo()
loo_comp <- as_tibble(loo_compare(loo_scaled, loo_pooled), rownames = "model") %>%
  mutate(model = recode(model, "model1" = "scaled", "model2" = "pooled"),
         across(c(elpd_diff,se_diff), round, 2)) %>%
  filter(elpd_diff != 0)

# Compare model scores
scores <- map2_dfr(forecasts,
                   variant_relationships,
                   ~ summary(.x, target = "forecast", type = "cases") %>%
                     fv_score_forecast(., obs, summarise_by = "strains",
                                       metrics = NULL) %>%
                     mutate("Variant relationship" = .y))

scores <- scores %>%
  select("Variant relationship",
         "Sharpness" = sharpness)
kable(scores)

# Display model parameters
print(parameters)
```