**Severity of the SARS-CoV-2 Omicron variant (B.1.1.529) in England**

**Version:** v0.2

**Date:** 4 Jan 2022

|  |  |  |
| --- | --- | --- |
| Version history | Date | Comment |
| 0.1 | 8 Dec 2021 | Initial draft created |
| 0.2 | 4 Jan 2022 | Refining of statistical methods prior to running first comparative analyses |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

**Contents**

[Background 3](#_Toc92306075)

[Objectives 3](#_Toc92306076)

[Study Design and Population 3](#_Toc92306077)

[Inclusion criteria 4](#_Toc92306078)

[Exclusion criteria 4](#_Toc92306079)

[Causal framework 4](#_Toc92306080)

[Study Measures 5](#_Toc92306081)

[Exposure 5](#_Toc92306082)

[Outcomes 5](#_Toc92306083)

[Covariates 5](#_Toc92306084)

[Statistical Methods 8](#_Toc92306085)

[Baseline characteristics 8](#_Toc92306086)

[28-day all-cause mortality 8](#_Toc92306087)

[AE attendance 8](#_Toc92306088)

[Covariate adjustment 8](#_Toc92306089)

[Defining regions 9](#_Toc92306090)

[*A priori* subgroup analyses 9](#_Toc92306091)

[Sensitivity and additional analyses 9](#_Toc92306092)

[Software and reproducibility 10](#_Toc92306093)

[Strengths and Limitations 10](#_Toc92306094)

[Non-random availability of SGTF data 10](#_Toc92306095)

[Ethics 10](#_Toc92306096)

[Conflicts of Interests 10](#_Toc92306097)

[References 10](#_Toc92306098)

## 

Background

The SARS-CoV-2 (COVID-19) Omicron variant B.1.1.529 (Omicron) was first identified in South Africa in late 2021. Analysis suggests that Omicron is more transmissible than the predominant Delta variant and it has since become the dominant strain throughout the UK. Only a small number of Omicron cases are identified by whole-genome sequencing. Spike gene target failure (SGTF) has been adopted as a proxy for identifying Omicron and has been shown to have excellent sensitivity.

Early studies from the UK Health Security Agency (UKHSA) using line listing data and hospital admissions have assessed the relative severity of Omicron compared to the Delta variant and have demonstrated a decrease in the risk of hospital attendance associated with the Omicron. Although these studies were able to account for age, sex, ethnicity, deprivation index, time period and geographical area, they are unable to account for comorbidities, which have been shown to be strongly associated with case severity among those diagnosed with COVID-19.1

### Objectives

This study will follow the principles set out in previous OpenSAFELY publications describing the case severity of the SARS-CoV-2 Alpha variant.2,3

To estimate the relative severity of SARS-CoV-2 infection with the Omicron variant compared to the Delta variant. Comparing the risk of accident and emergency (AE) attendance and death by 28-days among those infected with Omicron to those infected with Delta, after accounting for both demographic factors and comorbidities. The primary method of analysis will be a Cox proportional hazards regression model stratified by region, as defined by upper local authority area (UTLA).

Study Design and Population

We will use a cohort study nested within the OpenSAFELY platform. Using test result data from the UKHSA Second Generation Surveillance System (SGSS), we will select all those people who are:

1. positive for SARS-CoV-2 based on PCR swab test results in the time window 5th December 2021 to 1st January 2022 and
2. have data on SGTF

The study will focus on the comparison between those with SGTF and those without. Inconclusive SGTF results are considered in the sensitivity and additional analyses section.

The analysis has two main outcomes, AE attendance and 28-day all-cause mortality.

All-cause mortality will be determined from ONS death data and is expected to have full ascertainment of deaths with a 2-week delay. Therefore, the complete 28-day all-cause mortality analysis will include all individuals with at least 42-days follow-up from the date of COVID-19 diagnosis to the date of last ONS death data upload (28-days plus 14-days to account for the delay in the ONS death data).

However, in order to answer this pressing public health concern in a timely manner we will also perform an early analysis of AE attendance following SARS-CoV-2 infection as a proxy for case severity.

In early analysis all individuals testing positive for SARS-CoV-2 up to the 25th December with data on SGTF status will be included. Follow-up will be censored one week prior to the date of emergency care data upload.

### Inclusion criteria

* A positive SARS-CoV-2 PCR swab test result in SGSS within the window 5th December 2021 to 1st January 2022
* Data available on SGTF in SGSS.
* Registered with a primary care practice using The Phoenix Partnership (TPP) software on the date of COVID-19 diagnosis, with at least one year of continuous GP registration.

### Exclusion criteria

* Missing age, sex, or index of multiple deprivation, as these are indicators of poor data quality.

### 

### Causal framework

The motivation for adjusting for demographics and comorbidities is not that they impact on the variant of COVID-19 infection *per se*, but that they are likely to be associated with the upstream process of getting a test (e.g. test-seeking behaviour, ability to access testing facilities). Therefore, adjustment attempts to correct for imbalances between the Omicron and Delta exposure groups with respect to factors associated with getting a test. With the study population defined by SARS-CoV-2 positive test and SGTF data available and analysis by regional stratification defined *a priori*, the minimum sufficient adjustment set implied by **Figure 1** is adjustment for SARS-CoV-2 vaccination/prior infection alone. However, adjustment for all covariates also provides a causal interpretation and this will be considered the primary analysis.

**Figure 1 Causal framework DAG**

Chart

Description automatically generated

Study Measures

### Exposure

SGTF on SARS-CoV-2 PCR swab test from SGSS data, referred to as the Omicron exposure group. The comparator group being SARS-CoV-2 diagnoses without SGTF in SGSS data, referred to as the Delta group.

### Outcomes

AE attendance, death from any cause.

### Covariates

Age, sex, SARS-CoV-2 vaccination and prior infection status, deprivation index, ethnicity, smoking status, household size.

Region, defined by middle layer super output area (MSOA) from patient post code, or NHS England region.

Rural and urban location classification, and care home status.

Epidemiological week of the positive test.

Comorbidities: obesity, hypertension, chronic respiratory diseases other than asthma, chronic heart disease, diabetes, non-haematological and haematological cancer, reduced kidney function, chronic liver disease, stroke, dementia, other neurological disease, organ transplant, asplenia, rheumatoid arthritis/lupus, psoriasis, and other immunosuppressive conditions.

1. **Study measures**

|  |  |
| --- | --- |
| Exposures | OpenCodelists Definition |
| SGTF | UKHSA Second Generation Surveillance System (SGSS) |
| Outcomes |  |
| Death | All-cause registered deaths, from ONS |
| AE attendance | Attendance at AE due to COVID-19, from emergency care data |
| Covariates |  |
| Ethnicity | <https://codelists.opensafely.org/codelist/opensafely/ethnicity/> |
| SARS-CoV-2 vaccination status | Vaccination dates are documented from TPP general practise records |
| Previous SARS-CoV-2 infections | Previous SARS-CoV-2 infections are defined by positive test results in SGSS or a TPP general practise record of infection at least 7 days prior to the study index infection date |
| Region | MSOA and STP are extracted from patient post code  UTLA regions are found using these look-up tables:  <https://geoportal.statistics.gov.uk/datasets/lower-layer-super-output-area-2011-to-upper-tier-local-authorities-2019-lookup-in-england-and-wales->  <https://geoportal.statistics.gov.uk/datasets/9f4c270148014f20bf24abff9a7aef62_0> |
| Comorbidities | <https://codelists.opensafely.org/codelist/opensafely/aplastic-anaemia/>  <https://codelists.opensafely.org/codelist/opensafely/asplenia/>  <https://codelists.opensafely.org/codelist/opensafely/asthma-diagnosis/>  <https://codelists.opensafely.org/codelist/opensafely/asthma-inhaler-salbutamol-medication/2020-04-15/>  <https://codelists.opensafely.org/codelist/opensafely/asthma-inhaler-steroid-medication/2020-04-15/>  <https://codelists.opensafely.org/codelist/opensafely/asthma-oral-prednisolone-medication/2020-04-27/>  <https://codelists.opensafely.org/codelist/opensafely/bone-marrow-transplant/2020-04-15/>  <https://codelists.opensafely.org/codelist/opensafely/cancer-excluding-lung-and-haematological/2020-04-15/>  <https://codelists.opensafely.org/codelist/opensafely/chemotherapy-or-radiotherapy-updated/2020-04-15/>  <https://codelists.opensafely.org/codelist/opensafely/chronic-cardiac-disease/2020-04-08/>  <https://codelists.opensafely.org/codelist/opensafely/chronic-respiratory-disease/2020-04-10/>  <https://codelists.opensafely.org/codelist/opensafely/chronic-liver-disease/2020-06-02/>  <https://codelists.opensafely.org/codelist/opensafely/dementia/2020-04-22/>  <https://codelists.opensafely.org/codelist/opensafely/diabetes/2020-04-15/>  <https://codelists.opensafely.org/codelist/opensafely/chronic-kidney-disease/2020-04-14/>  <https://codelists.opensafely.org/codelist/opensafely/gi-bleed-or-ulcer/2020-04-08/>  <https://codelists.opensafely.org/codelist/opensafely/haematological-cancer/2020-04-15/>  <https://codelists.opensafely.org/codelist/opensafely/hiv/2020-07-13/>  <https://codelists.opensafely.org/codelist/opensafely/hypertension/2020-04-28/>  <https://codelists.opensafely.org/codelist/opensafely/inflammatory-bowel-disease/2020-04-07/>  <https://codelists.opensafely.org/codelist/opensafely/lung-cancer/2020-04-15/>  <https://codelists.opensafely.org/codelist/opensafely/other-neurological-conditions/2020-06-02/>  <https://codelists.opensafely.org/codelist/opensafely/permanent-immunosuppression/2020-06-02/>  <https://codelists.opensafely.org/codelist/opensafely/ra-sle-psoriasis/2020-04-14/>  <https://codelists.opensafely.org/codelist/opensafely/sickle-cell-disease/2020-04-14/>  <https://codelists.opensafely.org/codelist/opensafely/smoking-clear/2020-04-29/>  <https://codelists.opensafely.org/codelist/opensafely/smoking-unclear/2020-04-29/>  <https://codelists.opensafely.org/codelist/opensafely/solid-organ-transplantation/2020-04-10/>  <https://codelists.opensafely.org/codelist/opensafely/stroke-updated/2020-06-02/>  <https://codelists.opensafely.org/codelist/opensafely/temporary-immunosuppression/2020-04-24/> |

Statistical Methods

### Baseline characteristics

Participant characteristics, including all covariates listed above, will be described at baseline (the date of positive SARS-CoV-2 test), comparing the two exposure groups (Omicron and Delta). Continuous variables will be summarised by the mean and standard deviation and compared with a t-test, or median and interquartile range and Wilcoxon signed-rank test, as appropriate. Categorical variables will be summarized by the number and proportion in each group (n (%)) and compared with a chi-square test.

The median time-to-event and interquartile range of those who attend AE and die will be presented by exposure group.

The proportion of SARS-CoV-2 positive tests with SGTF will be plotted over the study period by NHS England region.

### 28-day all-cause mortality

The relative hazard of death for Omicron compared to Delta will be estimated from a Cox proportional hazards regression model. Follow-up will be censored two weeks prior to the date of ONS death data upload.

The hazard of death following a SARS-CoV-2 positive test result is expected to vary considerably between regions in England over time. Consequently, adjustment for region is unlikely to satisfy the proportional hazards assumption of a Cox model. To account for this variability, we will stratify the analysis on region, allowing a separate baseline hazard to be estimated for each region, but with covariate effects estimated over the full population – a stratified Cox PH model. The definition of regions is discussed below.

### AE attendance

The relative hazard of AE attendance for Omicron compared to Delta will be estimated from a Cox proportional hazards regression model stratified by region. Follow-up will be censored one week prior to the date of emergency care data upload.

Primary analysis will exclude people who attend AE on the date of testing positive for SARS-CoV-2 as these people may have been diagnosed in hospital. Sensitivity analysis will explore the impact of including people testing positive on the same day as AE attendance by adding 1 day to all survival times.

### Covariate adjustment

Unadjusted, causal minimal adjustment, demographically adjusted, and fully adjusted estimates will be presented for each analysis. The fully adjusted estimate will be primary.

Demographically adjusted models will include adjustment for the following covariates: Age will be included as a cubic spline term. Ethnicity will be grouped into five categories. The primary analysis will exclude patients with missing ethnicity. Sex, deprivation index, household size, and care home status will be included as categorical terms.

Epidemiological week of the baseline SARS-CoV-2 positive test will be included as a categorical variable.

Region will be defined by UTLA, unless data sparsity prevents this level of granularity. In which case region will be defined by STP, or aggregated geographical areas defined by NHS England region. Rural or urban location classification will be included as a categorical variable with 5 levels in line with other work.

Fully adjusted models will additionally adjust for patient comorbidities, smoking status, and obesity status. Comorbidities will be aggregated into a categorical term taking values none, 1, and 2 or more. In line with previous work on the risk of death from COVID-19 on the OpenSAFELY platform. For smoking and obesity, missing values will be categorised as never smoked and no evidence of obesity, in line with previous OpenSAFELY studies.4,5

The causal framework indicates both these adjustment sets result in a causal estimate of the effect of the relative effect of Omicron on mortality. For comparison, we will also fit a model using the minimum sufficient adjustment set implied by the causal DAG (vaccination status).

### Defining regions

Regional stratification will be a key consideration due to variability in the incidence of COVID-19 outcomes over time. Regions will be defined using patient middle super output area (MSOA) codes derived from patient post codes. MSOA data will be aggregated into upper tier local authority areas (UTLA) which will be the primary definition of regions for analysis.

### *A priori* subgroup analyses

Relative case severity will be estimated in subgroups of *a priori* interest, after adjustment for confounding. Differences in risk in these subgroups will be formally tested with a likelihood ratio test for an interaction with SGTF exposure status.

The subgroups of interest to be assessed are:

* Age group (<40; 40-54; 55-64; 65-74; 75-84; 85+)
* Ethnicity (in 5 categories)
* Comorbidity status (none, 1, 2+)
* Vaccination and prior SARS-CoV-2 infection status
* Deprivation index (deciles of deprivation)
* Epidemiological week of positive SARS-CoV-2 test (each week of the study period, fortnightly if data are sparse)
* NHS England region (East, London, South East, South West, Midlands, North East and Yorkshire, North West)

### Sensitivity and additional analyses

Inconclusive SGTF results

SGTF flags will be inconclusive in some cases. SGTF data are expected to take the values yes, no, unknown. The primary analysis will focus on the comparison of the yes group (Omicron) with the no group (Delta). In additional analysis the severity of disease in the unknown group will also be quantified and compared to that of the Omicron and Delta exposure groups.

Multiple imputation of missing ethnicity

Previous work in OpenSAFELY has identified that ethnicity data are missing for up to one quarter of all patients. The primary analysis will use the complete case set with regards to ethnicity. Sensitivity analysis will assess the impact of excluding records with missing ethnicity by imputing missing ethnicity using multiple imputation based on all variables included in the full adjustment set.

### Software and reproducibility

Data management will be performed using Python and Google BigQuery, with analysis carried out using Stata 16.1 / Python. Code for data management and analysis as well as codelists archived online <https://github.com/opensafely/sgtf-cfr-research>.

### 

Strengths and Limitations

### Non-random availability of SGTF data

Although the fact that we adjust for factors associated with getting tested should help account for possible non-random availability of SGTF data, we will also compare the characteristics of people included in the study (who all have SGTF data) with those not included in the study due to lack of SGTF data. This will help us assess whether those with SGTF are representative of all those tested during the time period of the study, and allow us to discuss the implications of this in our write-up as necessary.

Ethics

This study was approved by the Health Research Authority (REC reference 20/LO/0651) and by the LSHTM Ethics Board (reference 21863).

Conflicts of Interests

None to declare

References

1. Williamson, Elizabeth J., Alex J. Walker, Krishnan Bhaskaran, Seb Bacon, Chris Bates, Caroline E. Morton, Helen J. Curtis et al. "Factors associated with COVID-19-related death using OpenSAFELY." *Nature* 584, no. 7821 (2020): 430-436.

2. Grint, Daniel J., Kevin Wing, Elizabeth Williamson, Helen I. McDonald, Krishnan Bhaskaran, David Evans, Stephen JW Evans et al. "Case fatality risk of the SARS-CoV-2 variant of concern B. 1.1. 7 in England, 16 November to 5 February." *Eurosurveillance* 26, no. 11 (2021): 2100256.

3. Grint, Daniel J., Kevin Wing, Catherine Houlihan, Hamish P. Gibbs, Stephen JW Evans, Elizabeth Williamson, Helen I. McDonald et al. "Severity of Severe Acute Respiratory System Coronavirus 2 (SARS-CoV-2) Alpha Variant (B. 1.1. 7) in England." *Clinical Infectious Diseases* (2021).