OpenSAFELY Protocol: Effectiveness of sotrovimab/molnupiravir use vs non-use

Authors: John Tazare, Linda Nab

Version: v1.0

Date: April 2022

Table of Contents

Table of Contents **2**

Background **3**

Objectives **3**
- Primary Objectives 3
- Secondary Objectives 3

Methods **4**
- Data Source 4
- Data Sharing 4
- Information Governance 4
- Study Design and Population 5
- Study Measures 6
  - Exposure 6
  - Outcome 6
  - Covariates 6
- Statistical Analysis 7
  - Primary objective 7
  - Secondary objective 7
  - Sensitivity and supplementary analyses 7
- Software and Reproducibility 8
- Limitations 8

Administrative **10**
- Ethics 10
- Funding 10
- Conflict of Interests 10

References **11**

Background

- In December 2021, COVID-19 medicine delivery units (CMDUs) were launched across England to offer antiviral medicines and neutralising monoclonal antibodies (nMABs) to nonhospitalised COVID-19 patients thought to be at high risk of severe outcomes.[1,2]
- Before February 2022, sotrovimab and molnupirarvir were the most commonly prescribed medications.[3] Approval and adoption of these medications was largely driven by evidence from two phase III randomised placebo-controlled trials.[4,5]
- Amid understandable concerns surrounding early regulatory authorisations,[6] noninterventional evidence surrounding the effectiveness of these medications is needed to guide policy surrounding the use of these medications in routine clinical practice.
- Using the target trial framework,[7] we aim to emulate a randomised trial using observational data to estimate the effectiveness of either sotrovimab or molnupiravir versus no-treatment, amongst nonhospitalised COVID-19 patients in one of the high-risk groups.

Objectives

Primary Objectives

1. To estimate the hazard of COVID-19 related hospital or death within 28 days associated with use of sotrovovimab or molnupiravir versus no-treatment in non-hospitalised high-risk COVID-19 patients.

Secondary Objectives

1. Repeat stratified by individual drugs comparing a) sotrovimab versus no-treatment and b) molnupiravir versus no-treatment.
2. Estimate the hazard of all-cause mortality or all-cause hospitalisation within 28 days of follow-up
3. Estimate corresponding absolute risks for the outcomes studied.

Methods

Data Source

Primary care records managed by the GP software provider, TPP were linked to the Office for National Statistics (ONS) death data through OpenSAFELY, a data analytics platform created by our team on behalf of NHS England to address urgent COVID-19 research questions (https://opensafely.org). OpenSAFELY provides a secure software interface allowing the analysis of pseudonymised primary care patient records from England in near real-time within the EHR vendor's highly secure data centre, avoiding the need for large volumes of potentially disclosive pseudonymised patient data to be transferred off-site. This, in addition to other technical and organisational controls, minimises any risk of re-identification. Similarly pseudonymised datasets from other data providers are securely provided to the EHR vendor and linked to the primary care data. Further details on our information governance can be found on information governance and ethics.

The dataset analysed within OpenSAFELY is based on 24 million people currently registered with GP surgeries using TPP SystmOne software. It includes pseudonymised data such as coded diagnoses, medications and physiological parameters. Patient-level vaccination status is available in the GP records directly via the National Immunisation Management System (NIMS). No free text data are included. The following linked data were also used for this study: accident and emergency (A&E) attendance and in-patient hospital spell records via NHS Digital's Hospital Episode Statistics (HES); national coronavirus testing records via the Second Generation Surveillance System (SGSS); and the "COVID-19 therapeutics dataset", a patient-level dataset on antiviral and nMAbs treatments, newly sourced from NHS England, derived from Blueteq software that CMDUs use to notify NHS England of COVID-19 treatments.[8]

Data Sharing

All data were linked, stored and analysed securely within the OpenSAFELY platform https://opensafely.org/. All code is shared openly for review and re-use under MIT open license. Detailed pseudonymised patient data is potentially re-identifiable and therefore not shared. We rapidly delivered the OpenSAFELY data analysis platform without prior funding to deliver timely analyses on urgent research questions in the context of the global Covid-19 health emergency: now that the platform is established we are developing a formal process for external users to request access in collaboration with NHS England; details of this process will be published shortly on OpenSAFELY.org.

Information Governance

NHS England is the data controller; TPP is the data processor; and the key researchers on OpenSAFELY are acting on behalf of NHS England. This implementation of OpenSAFELY is hosted within the TPP environment, which is accredited to the ISO 27001 information security

standard and is NHS IG Toolkit compliant; patient data have been pseudonymised for analysis and linkage using industry standard cryptographic hashing techniques; all pseudonymised datasets transmitted for linkage onto OpenSAFELY are encrypted; access to the platform is through a virtual private network (VPN) connection; the researchers hold contracts with NHS England and only access the platform to initiate database queries and statistical models; all database activity is logged; and only aggregate statistical outputs leave the platform environment following best practice for anonymisation of results such as statistical disclosure control for low cell counts. The OpenSAFELY research platform adheres to the data protection principles of the UK Data Protection Act 2018 and the EU General Data Protection Regulation (GDPR) 2016. In March 2020, the Secretary of State for Health and Social Care used powers under the UK Health Service (Control of Patient Information) Regulations 2002 (COPI) to require organisations to process confidential patient information for the purposes of protecting public health, providing healthcare services to the public and monitoring and managing the COVID-19 outbreak and incidents of exposure. Together, these provide the legal bases to link patient datasets on the OpenSAFELY platform. GP practices, from which the primary care data are obtained, are required to share relevant health information to support the public health response to the pandemic, and have been informed of the OpenSAFELY analytics platform.

Study Design and Population

The target trial specification observational emulation and challenges surrounding data for observational emulation are outlined below (summarised in Table 1).

We will conduct a population-based cohort study with all adults (≥ 18 years old) registered at a TPP practice (with at least 3-months of continuous GP registration) within OpenSAFELY between 16th December 2021 and 10th February 2022 who:

a) Tested positive with SARS-CoV-2 infection (PCR or lateral flow test)

b) ≥ 1 risk factor for disease progression at time of positive test. Defined as member of high-risk cohort (see Figure 1 as described by Green et al)[3]. Operationally, this will be defined using EHR-derived data only (instead of CMDU).

c) No treatment history (i.e. ever) of any other antivirals or nMABs for COVID-19 prior to positive test

d) Not hospitalised at the date of entry

e) No evidence of COVID-19 infection up to 90 days before current test-positive spell (i.e. evidence of previous COVID-19 spell evident by positive test or hospitalisation)

Patients with missing age, sex, STP or IMD will be excluded.

Cohort entry will be defined as SARS-CoV-2 test positive. Patients will be followed up until the earliest of the outcome (COVID-19 related hospitalisation or death), death date (death other than COVID-19 related death), deregistration or 28 days post-index date .[9] In the primary analysis patients will be excluded if they experience outcomes during the treatment assessment window (i.e. within 5 days of positive test).

Study Measures

Exposure

We will use an intent-to-treat exposure definition; this is, in part, given limitations surrounding accurately ascertaining on-treatment time. We therefore look to compare the following treatment strategies:

1) Initiation of either sotrovimab or molnupiravir within 5 days of SARS-CoV-2 positive test
2) No initiation of either sotrovimab or molnupiravir therapy within 5 days of SARS-CoV-2 positive test

Treatment status and date will be ascertained from the COVID-19 therapeutics dataset.

Note: Median time between positive test and treatment is 2 days for Molnupiravir and 3 days for Sotrovimab.

We will exclude patients who are recorded as receiving both sotrovimab and molnupiravir on the same day. If prescriptions are one day apart (or more) we will use the first prescription recorded (will describe the number of patients this affects).

Outcome

The primary outcome is 28-day COVID-related hospitalisation or death defined as hospitalisation after treatment, excluding hospitalisation (i.e. hospitalised day cases) for treatment (i.e., sotrovimab was often administered in a hospital).

Secondary analyses will focus on the outcome of 28-day all-cause hospitalisation or death.

Covariates

The following potential confounders available in OpenSAFELY will be extracted at index date:

- Age (will be modelled using splines/fractional polynomials)
- Sex
- Ethnicity (in 6 categories: Black, Mixed, South Asian, White, Other, Unknown)
- Deprivation: defines using quintiles of the English Index of Multiple Deprivation, and based on postcode of residence.
- Rurality (potentially influencing vaccine access)
- Smoking status
- Other comorbidities / clinical characteristics: chronic cardiac disease; chronic obstructive pulmonary disease (COPD); obesity (most recent adult body mass index (BMI) ≥30); dialysis; immunosuppressive diagnoses or medications; severe mental illness (psychosis, schizophrenia and bipolar disorder); learning disabilities including Down syndrome; dementia; lung cancer; haematological cancer; other cancers; diabetes.
- Brand of most recent vaccine/number of vaccines
- High risk group

- Time since most recent vaccination date
- STP
- Calendar time (will inspect model fit but will start with ‘week’ in the propensity score)

Statistical Analysis

Primary objective

Distributions of baseline covariates will be compared between the users and non-users. We will use propensity score (PS) weighting to account for confounding bias.[10] Our approach, based on inverse probability of treatment weights (IPTW) will estimate the average treatment effect (ATE).[10–12] The ATE is the average effect, at the population level, of moving the entire population from untreated to treated: This is a two-stage approach:

1. Estimation of PS:
    - We will use logistic regression to model treatment strategy assignment on the set of baseline covariates.[13]
    - Estimated probabilities will be obtained from this model and inverse probability of treatment weights (IPTW) derived.
    - PS diagnostics (e.g. overlap plots and absolute standardised differences) will be used to assess covariate balance in the unweighted and weighted populations.

2. Estimation of treatment effect
    - We will use an IPTW Cox model (follow-up period time scale) to obtain hazard ratios for 28-day COVID-19 related hospitalisation and mortality. Robust standard errors will be applied.

Secondary objectives

The above procedure will be repeated for the all-cause hospitalisation and mortality outcomes as well as in the subgroups of a) matched sotrovimab users and non-users and b) matched molnupiravir users and non-users.

We will derive the cumulative incidence assuming everyone is treated and everyone is untreated (from the cox models) and additionally consider reporting the absolute risk difference.

Sensitivity and supplementary analyses

Sensitivity analyses

1) Patients experiencing outcome events during the initial 5-day treatment window will be excluded from the primary analysis. We will then include these excluded patients. Those

who had received AV/MABs will be classified under the treatment strategy “initate within 5 days”: the others will have treatment assignment imputed (20 times, or reduced if too computationally intensive) and we will combine results across imputed datasets using Rubin’s rules to obtain an overall hazard ratio.[14]

2) Unmeasured confounding is a key concern:
    a) The e-value formula will be used to assess the strength of association that an unmeasured confounder would need to have with exposure or outcome to fully explain the observed association.[15] Options include a) disease severity at treatment initiation (i.e. association between disease severity at 2/3 days post infection and hosp/death). b) Symptomatic COVID-19 potentially based on ZOE App estimates.e.g. Using ZOE estimates for the strength of association between symptoms on day 2/3 and outcome or estimate differential symptomatic prevalence between users (assume 100%) and non-user (from ZOE)
    b) PS trimming will be applied trimming both treated and untreated observations below 2.5 percentile of observed PS in the treated and above the 97.5 percentile of observed PS in the untreated, i.e. Sturmer trimming. The purpose of this is to investigate unmeasured confounding and we acknowledge trimming can complicate interpretation of the causal parameter.[16]

2) We will consider additionally presenting hazard ratios on day 14 and on day 28 after positive test (i.e. 0-14 days, 0-28 days).[17]

4) For the primary analyses, consider additionally require patients not to be pregnant at time of positive test, acknowledging limitations in identifying these patients.

Descriptive analyses

1) Describe the number of patients with treatment records of both sotrovimab and molnupiravir.

2) Describe number of patients with outcome within 5 days

3) Describe number of patients who are treated after between days 5 and 28

4) Symptomatic flag - Explain limitations with data and describe characteristics in those who have it recorded by treatment group

Other analyses

After initial review of our results and in light of possible issues surrounding confounding and selection bias, we will consider:

1) Updating the study period to include periods including new variants of interest
2) Performing the analysis in important clinical subgroups, for example, to investigate effectiveness in the renal population and differences in effectiveness by vaccination history.

Software and Reproducibility

Data management will be performed using Python, with analysis carried out using Stata 16.1 and R. Code for data management and analysis as well as codelists archived online.

Limitations

Eligibility:

- Restriction of study population to people in 'high risk cohort' may mean users and non-users are more comparable. However, this might affect generalizability. Specifically, there are people receiving treatment who are in the data but not recorded to be in the high-risk cohort (we assume this is misclassification due to data recording). If we decide NOT to require this criteria, we are essentially selecting a 'high risk' group in the exposed group but applying the same other criteria in the controls will lead us to have controls across the risk spectrum (with tendency towards low risk -> severe confounding by indication). Therefore we have decided to restrict to the 'high risk' inclusion even at the cost of losing some of the mab group.
- Trials looked at symptomatic COVID-19 patients - we have 'symptomatic flag' but we understand data to be unreliable.

Treatment strategy

- In our data, we can't distinguish easily between non-user because declined versus non-user because treatment wasn't offered.[3] This is important if the reasons to decline in non-users are associated with outcome (e.g. symptoms have already resolved).

Unmeasured confounding

- Possible mechanisms include: healthy initiator bias/trust in healthcare system e.g. who is more likely to try a 'new drug' or even just come forward for treatment.

Table 1

	Target trial specification	Target trial emulation	Challenges in observational emulation
Eligibility criteria	• Symptomatic SARS-CoV-2 infection (non-hospitalised) and ≤5 days since onset of symptoms) • Member of a 'highest' risk group • Not pregnant • Aged ≥ 18 between 16th Dec 2021 and 1st Feb 2022 • No prior treatment of antivirals or nMABS • No evidence of SARS-CoV-2 infection up to 90 days before current symptomatic SARS-CoV-2 infection spell.	• We defined symptomatic SARS-CoV-2 infection as evidence of a positive test (PCR or lateral flow test) • Pregnancy eligibility criteria not applied in emulation • Previous SARS-CoV-2 infection defined using positive test and hospitalisation information • We additionally require all patients to have >3 months registration in TPP in order to ascertain patient factors • Only patients will non-missing data on age, sex, IMD and STP included Otherwise, same as the target trial	• Difficult to reliably identify patients who are pregnant and current methods are likely to substantially overestimate pregnancy. • Second Generation Surveillance System (SGSS) contains information surrounding symptomatic SARS-CoV-2 infection, however, the reliability of this information has not been validated (important in light of substantial missingness). • EHR data will not have complete information on symptoms of COVID-19 recorded. For example, the sotrovimab trial defined the following symptoms: fever, chills, cough, sore throat, malaise, headache, joint or muscle pain, change in smell or taste, vomiting, diarrhea, shortness of breath on exertion.[4] • To ensure consistent identification of 'highest' risk group information we only used information available in primary care records (as opposed to CMDU records; only available for treated patients). Previous work showed recording of this information was reliably identifiable in primary care data alone.[3]
Treatment strategies	Eligible patients are then randomised on the day of diagnosis to either: I. Initiate antiviral or nMAB therapy within 5-days II. Not initiate either therapy within 5-days	We defined the date of medication initiation to be the first date of prescription received. Otherwise, same as the target trial	• Sotrovimab is given as a single infusion so it is likely that treatment is completed. • Molnupiravir is given as 4 capsules twice a day for 5 days. Once prescribed, we do not have data on whether the full course of treatment is completed. • Prescription date might not be the date a patient receives treatment (for

			example: sotrovimab required an appointment at a hospital and molnupiravir was posted to patient's home). Both could lead to a slight lag between prescription date and date the treatment was initiated.
Treatment groups	Patient treatment groups are defined by randomisation to one of two treatment strategies defined above.	Same as the target trial	• Patients can experience the outcomes during the initial 5-day treatment window. In the primary analysis these patients are excluded. • In secondary analyses, we will impute treatment assignment for the patients excluded whose follow-up was consistent with both treatment strategies (20 times). Results will be combined across imputed datasets using Rubin's rules to obtain an overall hazard ratio.[14]
Treatment assignment	Eligble patients are randomly assigned to a strategy	• Patients are classified as receiving a given strategy based on information in their medical records • We will attempt to emulate randomisation by adjusting for baseline confounders (defined at index date) via propensity score methods	• Patients initiating therapy versus not are likely to be different and possible residual confounding is a key issue. Possible mechanisms include initiators being healthier than non-intiators due to coming forward for treatment; initiators experiencing more (severe) COVID-19 symptoms without signs of improvement than non-initiators. Treatment initiation might also signify increased trust in the healthcare system.
Outcomes	COVID-19-related hospital admission or death within 28 days	Same as the target trial • COVID-19 outcomes recorded using ICD-19 codes in hospital records or on death certificates	Misclassification of cause of death on death certificates.
Follow-up	Index date is date of positive SARS-CoV-2	• Primary analysis will	• In the target trial, treatment strategies

	test until end of 28 days, outcome (COVID-19 hospitalisation or death), administrative censoring (i.e. deregistration)	exclude all patients experiencing outcome events within 5 days of onset symptomatic SARS-CoV-2 infection. ● Follow-up approach is the same as target trial, aside from start date. ● Secondary analysis includes patients initially excluded (see Treatment groups for details).	would be assigned at index date. However, in the observational emulation we wanted to ensure that early outcomes did not influence the definition of the treatment groups.
Causal contrasts	Intention-to-treat	Observational analogue to intention-to-treat	
Statistical analysis	● 28-day hazard ratios ● 28-day absolute risk difference	Same as the target trial	● Same as the target trial with adjustment for baseline confounders.

Figure 1: Copied from[3]

Box 1: Patient cohorts considered at higher risk from COVID-19 and to be prioritised for treatment with antivirals and nMABs, as determined by an independent advisory group commissioned by the UK Department of Health and Social Care (DHSC)[7]. For further details on these criteria and how they were applied in the present study see Supplementary Table S1.

- Patients with **down's syndrome**;
- Patients with a **solid cancer,** such as active metastatic cancer, or active solid cancers at any stage;
- Patients with a **haematological disease and stem cell transplant recipients**, such as those with sickle cell disease;
- Patients with **renal disease**, such as those with chronic kidney stage 4 or 5;
- Patients with **liver disease**, such as those on immune suppressive therapy;
- Patients with **immune-mediated inflammatory disorders,** such as those treated with rituximab or other B cell depleting therapy in the past 12 months;
- Patients with **primary immune deficiencies**, such as severe combined immunodeficiency;
- Patients with **HIV/AIDS** with high levels of immune suppression;
- **Solid organ transplant recipients**;
- Patients with **rare neurological conditions** (multiple sclerosis motor neurone disease, myasthenia gravis or huntington's disease)

References

1. Department of Health and Social Care. UK's most vulnerable people to receive life-saving COVID-19 treatments in the community. GOV.UK. Accessed April 26, 2022. https://www.gov.uk/government/news/uks-most-vulnerable-people-to-receive-life-saving-covid-19-treatments-in-the-community
2. Agarwal A, Rochwerg B, Lamontagne F, et al. A living WHO guideline on drugs for covid-19. *BMJ*. 2020;370:m3379. doi:10.1136/bmj.m3379
3. Green A, Curtis HJ, Higgins R, et al. Trends, regional variation and clinical characteristics of recipients of antivirals and neutralising monoclonal antibodies for non-hospitalised COVID-19: a descriptive cohort study of 23.4 million people in OpenSAFELY. Published online March 10, 2022:2022.03.07.22272026. doi:10.1101/2022.03.07.22272026
4. Gupta A, Gonzalez-Rojas Y, Juarez E, et al. Early Treatment for Covid-19 with SARS-CoV-2 Neutralizing Antibody Sotrovimab. *N Engl J Med*. 2021;385(21):1941-1950. doi:10.1056/NEJMoa2107934
5. Jayk Bernal A, Gomes da Silva MM, Musungaie DB, et al. Molnupiravir for Oral Treatment of Covid-19 in Nonhospitalized Patients. *N Engl J Med*. 2022;386(6):509-520. doi:10.1056/NEJMoa2116044
6. Brophy JM. Molnupiravir's authorisation was premature. *BMJ*. 2022;376:o443. doi:10.1136/bmj.o443
7. Hernán MA, Robins JM. Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. *Am J Epidemiol*. 2016;183(8):758-764. doi:10.1093/aje/kwv254
8. Brown M, Saund J, Qureshi A, et al. *Demographics and Outcomes of Covid-19 Medicines Delivery Units Across 4 UK Centres During the Initial B1.1.529 Omicron Epidemic: A Service Evaluation.* Social Science Research Network; 2022. Accessed June 10, 2022. https://papers.ssrn.com/abstract=4123333
9. Andersen PK, Geskus RB, de Witte T, Putter H. Competing risks in epidemiology: possibilities and pitfalls. *Int J Epidemiol*. 2012;41(3):861-870. doi:10.1093/ije/dyr213
10. Williamson E, Morley R, Lucas A, Carpenter J. Propensity scores: from naive enthusiasm to intuitive understanding. *Stat Methods Med Res*. 2012;21(3):273-293. doi:10.1177/0962280210394483
11. Webster-Clark M, Stürmer T, Wang T, et al. Using propensity scores to estimate effects of treatment initiation decisions: State of the science. *Stat Med*. 2021;40(7):1718-1735. doi:10.1002/sim.8866
12. Austin PC, Stuart EA. Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. *Stat Med*. 2015;34(28):3661-3679. doi:10.1002/sim.6607
13. Smeeth L, Douglas I, Hall AJ, Hubbard R, Evans S. Effect of statins on a wide range of health outcomes: a cohort study validated by comparison with randomized trials. *Br J Clin Pharmacol*. 2009;67(1):99-109. doi:10.1111/j.1365-2125.2008.03308.x
14. Kaura A, Sterne JAC, Trickey A, et al. Invasive versus non-invasive management of older patients with non-ST elevation myocardial infarction (SENIOR-NSTEMI): a cohort study based on routine clinical data. *The Lancet*. 2020;396(10251):623-634. doi:10.1016/S0140-6736(20)30930-2
15. Ding P, VanderWeele TJ. Sensitivity Analysis Without Assumptions. *Epidemiol Camb Mass*. 2016;27(3):368-377. doi:10.1097/EDE.0000000000000457
16. Stürmer T, Rothman KJ, Avorn J, Glynn RJ. Treatment effects in the presence of unmeasured confounding: dealing with observations in the tails of the propensity score distribution--a simulation study. *Am J Epidemiol*. 2010;172(7):843-854.

doi:10.1093/aje/kwq198
17. Hernán MA. The hazards of hazard ratios. *Epidemiol Camb Mass*. 2010;21(1):13-15. doi:10.1097/EDE.0b013e3181c1ea43