COVID-19 collateral: Lateral impacts of the coronavirus pandemic on non-COVID outcomes and health inequalities: An international comparison of the UK and Denmark using the OpenSAFELY platform and National Registry Data

Authors:

Ruth Costello, John Tazare, Emily Herrett, Dominik Piehlmaier, Anthony Matthews, COVID Collateral team, Aarhus Team, OpenSafely Collaborative, Laurie A Tomlinson, Sinéad M Langan, Kathryn E Mansfield, Rohini Mathur
*order not finalized

Version: 1.0

Date: 08/02/2023

Version history	Date	Comment
0.1	08/02/2022	Initial draft created by Rohini Mathur and Ruth Costello
0.2	06/05/2022	Draft shared with international collaborators
0.3	24/10/2022	Updated to remove Swedish collaborators from project
1.0	08/02/2023	Removed Swedish collaborators from author list

Table of Contents

Background

Diversion of healthcare resources to pandemic management has negatively affected non-COVID-related healthcare provision, including prevention activities, potentially creating or worsening physical and mental health.[1] While the UK has experienced one of the worst COVID-19 outbreaks in Europe, several Scandinavian countries have experienced better COVID-19 outcomes and faster healthcare system recovery.[2] We do not know which aspects of the pandemic are driving non-COVID health consequences. Comparing countries with different pandemic curves where different measures have been taken at different times will allow us to gain insight into which components, if any, drive between-country differences. This will help inform policy for any future infectious disease outbreaks in the UK. While the direct effects of the COVID-19 pandemic have been found to disproportionately affect older people, minority ethnic groups, and deprived populations, inequalities in the indirect effects of the pandemic have yet to be fully explored.[3,4]

Hypothesis

The impacts of population-wide restrictions on clinical monitoring, hospital admissions and mortality will be greater for people of lower socioeconomic status and from minority ethnic or migrant backgrounds compared with those of higher socioeconomic status or those of the majority ethnic group.

Aims and Objectives

Aim

To determine the impact of population-wide restrictions and the diversion of care-related resources on ethnic and social inequalities in risk factor monitoring, hospital admissions and mortality for non-COVID related conditions in England, Sweden and Denmark.

Objectives

1. To estimate how ethnic and social disparities in **clinical monitoring activities** have changed between the period prior to the pandemic (March 2018-March 2020) and between periods of tightened and relaxed social restrictions (“lockdowns”) during the pandemic (March 2020 onwards)

2. To estimate how ethnic and social disparities in **hospital attendance and mortality** for diabetes, cardiovascular disease, respiratory disease, and mental health conditions have changed between the period prior to the pandemic (March 2018-March 2020) and between periods of tightened and relaxed social restrictions (“lockdowns”) during the pandemic (March 2020 onwards).

3. To determine whether inequalities in clinical monitoring activities, hospital attendance and deaths differ between **type of social restrictions** (ie/ full, partial, regional, mask wearing only etc).

4. To compare descriptively how the pandemic has impacted ethnic and social inequalities in clinical monitoring and outcomes across the UK, Denmark and Sweden to help determine whether certain public health approaches are associated with more equitable health outcomes.

Methods

Database Description: England

For the England based analysis, we will use data from general practice (GP) records, obtained from the GP software provider TPP, linked to the Emergency Care Data Set, Admitted Patient Care and ONS mortality records. The data will be accessed, linked and analysed through openSAFELY.org - a data analytics platform created on behalf of NHS England to address urgent questions relating to the epidemiology and treatment of COVID-19. 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 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. Descriptions of OpenSAFELY have been previously published (REF), and more information can be found on https://opensafely.org/.

Primary care records retrieved from the TPP SystmOne electronic health record system include diagnoses (SNOMED or Read 3 CTV3), prescriptions (dictionary of medicines and devices), basic sociodemographics and vital signs for 22 million individuals – approximately 40% of the English population. Data extracted by SystmOne have previously been used in medical research, as part of the ResearchOne dataset (REFS).

All data are held in a secure research environment hosted by TPP, which is a Tier 3 data centre, accredited to NHS Digital standards for centrally hosted clinical systems (ISO 27001 standard and IG Toolkit version 2). We received ethics approval to conduct the data linkage and analyses by the London - City & East Research Ethics Committee on the 2nd of April 2020 (REC reference: 20/LO/0651) and LSHTM Ethics Board (ref 21863). No further ethical or research governance approval was required by the University of Oxford but copies of the approval documents were reviewed and held on record.

Latest Database Description available here:

https://docs.google.com/document/d/1d6fw9sc80_N_UQO7qib_R8yBZGObLEzPS_xcri222rA/edit]

Database Description: Denmark

Can get information on number of visits, but not diagnoses

Data on hospitalisations, prescriptions, mortality (including psychiatry)

Statistics Denmark - poorer lab data but has socioeconomic data

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. Patient data have been pseudonymized for analysis and linkage using industry standard cryptographic hashing techniques; all pseudonymized 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 anonymization 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 organizations 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 basis 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

This study will comprise two main analytic approaches. Firstly, we will conduct a descriptive time series analysis to estimate differences in the monthly frequency of recorded clinical measurements, hospital attendances and deaths, stratified by ethnic group and deprivation. Secondly, we will conduct a series of survival analyses to compare ethnic and social differences in risks of hospitalisation and mortality between the period prior to the pandemic (March 2018-March 2020) and between periods of tightened and relaxed social restrictions (“lockdowns”) during the pandemic (March 2020 onwards).

Analysis population

The study population will be adults (males and females 18 years and above) registered in a TPP general practice in England between 1st March 2018 and 31st January 2022. Descriptive analyses will include

patients registered with a GP practice at any time point during follow-up as counts of outcomes will be measured monthly. Survival analyses will include people registered with a GP practice up to 31st December 2021. Follow-up will end on April 30th 2022 (end of current data).

We will follow individuals from the study start date (01/03/2018).

And until the earliest of:

1. Death;
2. De-registration from GP practice from TPP;
3. Latest TPP data are available;
4. End of study (30/04/2022).

We will censor datasets at the earliest end date of available linkages to make the linked datasets comparable.

Inclusion Criteria

Adults over the age of 18 alive and under follow-up at the study start date (01/03/2018), registered with a primary care practice using TPP software, with at least three months of continuous GP registration.

Exclusion Criteria

People in households greater than 15 individuals (based on TPP-derived household size), in order to reduce the likelihood that we will be including institutions such as care homes in our analyses. People with missing age, sex, Sustainability and Transformation Partnership (STP) region or individual level index of multiple deprivation, since these are likely to indicate poor data quality. We will exclude individuals identified as living in a care home (based on TPP's care home indicator). The pseudonymised household identifier developed by TPP links people living at the same address on 1 February 2020 (Forbes et al. 2021; Wing, Kevin, n.d.) and care home status is derived from addresses matched to publicly available Care Quality Commission data). Patients who joined a TPP practice after 1 February 2020 will not have a household status and cannot be classified in terms of care home status, and therefore will be excluded.

Subgroup inclusion criteria

The inclusion and exclusion criteria above will generate a general population study population. For disease specific outcomes there will be four disease study populations:

- **Diabetes**: People with a primary care code for type 1 or type 2 diabetes followed from the study start or the first diabetes code, whichever is latest.
- **Cardiovascular disease**: People with a primary care code for coronary heart disease, stroke or transient ischemic attack followed from study start or the first code, whichever is latest.
- **Respiratory disease**: there will be two groups: 1) people with asthma defined as having a primary care code for asthma in the previous 3 years, from the study start or the first asthma

code, whichever is latest. 2) people with chronic obstructive pulmonary disease (COPD) defined as those age >40 with a primary care diagnosis of COPD and with a smoking history followed from study start or first COPD code, whichever is latest.

- **Mental health**: People with a primary care code for severe mental illness: schizophrenia, bipolar disorder and other psychoses, followed from study start or first code, whichever is latest.

Study Measures

Exposures

1. Our first exposure will be the introduction of lockdown in the UK on March 23rd, 2020. Time will be split as follows:

- 1st March 2018 to 22nd March 2020 (pre-pandemic),
- 23rd March 2020 to 30th April 2022 (pandemic).

We will stratify our analysis based on the following time periods:

- 1st March 2018 to 22nd March 2020 (pre-pandemic),
- 23rd March 2020 to 30th June 2020 (wave 1, lockdown),
- 1st July 2020 to 31st October 2020 (easing restrictions),
- 1st November 2020 to 31st March 2021 (wave 2),
- 1st April 2021 to 31st November 2021 (easing restrictions)
- 1st December 2021 to 30th April 2022 (Omicron wave).

Outcomes

Outcomes include

(i) clinical monitoring activities for diabetes, cardiovascular disease, respiratory disease, and mental health conditions (based on those described by the Service Restoration Observatory report). The specific monitoring activities will be as follows:

- Diabetes: glycated haemoglobin A1c (HbA1c) levels, estimated glomerular filtration rate (eGFR), and blood pressure monitoring.
- Cardiovascular disease: blood pressure monitoring in the cardiovascular disease study population.
- Respiratory disease: chronic obstructive pulmonary disease annual review in those with COPD. Asthma annual review in those with asthma.
- Mental health: Blood pressure monitoring in the mental health study population.

(ii) Hospital admissions (Emergency and inpatient) for diabetes, cardiovascular disease, respiratory disease, and mental health conditions

Hospital admissions outcomes will be defined in two ways

- Outcome (see below for definitions) as primary diagnosis
- Outcome as secondary or other diagnosis

(iii) Deaths due to diabetes, cardiovascular disease, respiratory disease, mental health conditions , and any cause (all-cause mortality)

Specific outcomes for hospital admission and death are:

- **Diabetes**: admissions related to type 1 diabetes, type 2 diabetes or diabetetic ketoacidosis in people with type 1 and type 2 diabetes.
- **Respiratory disease**: Asthma exacerbation in patients with asthma, COPD exacerbation in people with COPD.
- **Cardiovascular disease**: admissions for Myocardial infarction, stroke, transient ischaemic attack, unstable angina, heart failure or venous thromboembolism in the general study population.
- **Mental health**: Depression, anxiety, severe mental illness, self harm (fatal and non-fatal), eating disorder, obsessive compulsive disorder, and suicide in the general study population.

For clinical monitoring and hospital admissions, a maximum of one event per individual per month will be included. However, individuals will be allowed to appear in multiple months if they have repeated records of the outcome. Death will be counted once on the date of the event recorded in the ONS Mortality data.

Stratifying Variables

In the UK data, our main stratifying variables will be ethnicity in six categories (White, South Asian, Black, Mixed, Other, Unknown) and deprivation quintiles based on the index of multiple deprivation as of 1st February 2020. The exposures are not time updated.

In the Danish data, our main stratfying variables will be socioeconomic status and migration status.

Potential confounders

Multivariable regression models will adjust for

- Age, categorised in 20 year bands,
- Sex,
- Region (Sustainability Transformation Partnership level (STP, English NHS administrative region)),
- Urban/rural location,
- SES (IMD, in quintiles) (in models where ethnicity is the main exposure of interest)
- Ethnicity (in models where IMD is the main exposure of interest)
- Eligibility for shielding defined using the ‘At high risk of COVID-19’ flag on primary care records https://github.com/opensafely/documentation/discussions/486

Statistical Analysis

Part 1: Descriptive and time-series analysis of monthly trends

Each month (March 2018 to January 2022), the denominator adult population (18+) who meet the inclusion criteria will be extracted. The period prevalence of each outcome will be calculated across the study population each month. This will assume that a person is eligible in the denominator for the whole month if they are eligible on the 1st of the month. Each outcome will be analysed separately in the relevant study population (Diabetes, COPD, or general population)

Each person will be counted only once each month, but people can appear in multiple months if they have repeated records of the outcome. We will calculate 95% confidence intervals for the total proportions each month.

The monthly prevalences will be stratified by ethnic group and deprivation quintile.

For the interrupted time-series analysis, we will use binomial generalised linear models with the monthly prevalence using dynamic population sizes (updated monthly). The interruption will be the onset of the 2020 covid-19 pandemic in England and we will use a binary pre-pandemic (before 23rd March 2020) vs mid pandemic variable (1st July to January 2022) to measure the step change in outcomes. To estimate the change in events as a result of the pandemic (the step change), we will calculate odds ratios (ORs) for the relative difference in outcomes with the restriction period compared with the pre-lockdown period. This OR will represent the average with restrictions compared to the average before the pandemic, adjusted for month, etc. The early months of 2020 will be excluded as a transition period. An interaction term will be included to understand whether the impact of the lockdown differed between ethnic groups and deprivation level.

Comparable pre- and mid- pandemic time periods and periods of social restrictions for Sweden and Denmark will be agreed in consultation with collaborators, with the understanding that public health approaches differ across countries.

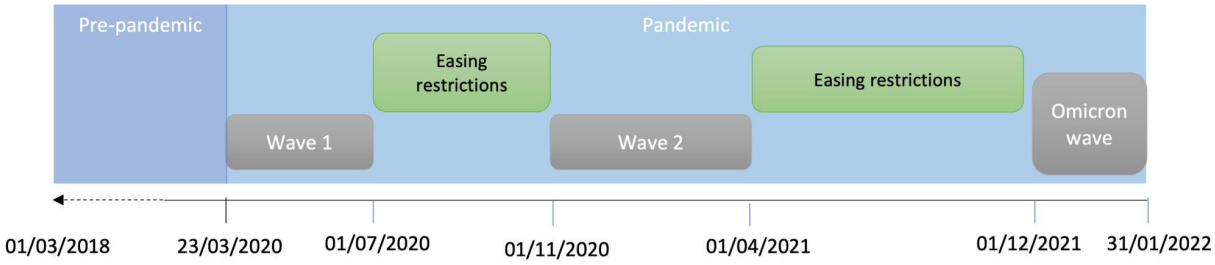
Part 2: Survival Analysis comparing risk of hospitalisation and mortality between pandemic time periods.

A series of cohort studies will be used to estimate differences in hospitalisations and mortality by ethnicity and deprivation for people with diabetes, COPD, CVD, and mental health conditions in the seven time periods of interest:

- 1st March 2018 to 22nd March 2020 (pre-pandemic),
- 23rd March 2020-31st January 2022 (pandemic)
- 23rd March 2020 to 30th June 2020 (wave 1, lockdown),

- 1st July 2020 to 31st October 2020 (easing restrictions),
- 1st November 2020 to 31st March 2021 (wave 2),
- 1st April 2021 to 31st November 2021 (easing restrictions)
- 1st December 2021 to 31st January 2022 (Omicron wave).

Figure 1: Restrictions timeline



The study cohort will be defined at the start of each study period. Individuals will enter the cohort at the start of each of the seven study periods and leave the cohort at the earliest of experiencing the outcome of interest (hospitalisation or death), leaving the OpenSAFELY database, or the end of the study period.

Multivariable Cox-proportional hazards regression will be used to estimate differences in the risk of hospitalisation or death by ethnic group (with white as the reference group) and deprivation quintile (with the most affluent quintile as the reference group). Models will adjust for age, sex, urban/rural location, deprivation (for ethnicity models), ethnicity (for deprivation models), and eligibility for shielding. All models will be clustered by STP region using a cluster term in the regression model.

Individuals who have experienced a hospitalisation in any one period will be eligible to experience the outcome again in any subsequent time period. For individuals experiencing multiple hospitalisations for each outcome of interest in any one time period, we will only count time to the first hospitalisation. Individuals dying in any time period will (by definition) not be included in subsequent cohorts.

Potential 3rd approach:
As a further approach to estimating the effect of social restrictions on hospital admissions, we will conduct a self-controlled case-series study (SCCS). In SSCS, all individuals will contribute follow-up time to exposed periods (tight restrictions) and unexposed periods (relaxed restrictions). The risk of hospital admissions will then be compared between periods of tight and relaxed social restrictions.

Software and Reproducibility

Data management will be performed using Python 3.8 and SQL with analysis carried out using Stata 17. Code for data management and analysis as well as codelists will be archived online.

Power

A recent study examining the representativeness of the OpenSAFELY database indicated that the ethnic breakdown of the study population when using primary care data alone is 10 877 978 (62·9%) White, 1 025 319 (5·9%) South Asian, 340 912 (2·0%) Black, 320 788 (1·9%) other, 170 484 (1·0%) mixed ethnicity, and 4 553 051 (26·3%) unknown.[4]

IMD information is available for all OpenSAFELY participants. Supplementing primary care ethnicity data with secondary care data from the Secondary User Services (SUS) will increase completeness of ethnicity from 75% to 90% (Opensafely internal report). Given these numbers, we do not anticipate an issue with power for analyses stratified by ethnicity or IMD.

Limitations

- We only have ethnicity and deprivation measured as of 1st February 2020. While ethnicity is likely to remain stable over the study period, we are unable to account for any changes in deprivation occurring before or during the pandemic. The index of multiple deprivation may also be a poor proxy for an individual's lived experience, as it is unable to capture individual level information on education, income, employment, household circumstances.
- 25% of patients have missing data for ethnicity in primary care. We will supplement the missing ethnicity in primary care records with ethnicity captured in SUS data, which brings completeness up to 90%. We will include missing ethnicity as a category so that we can stratify for this. We will also consider the use of multiple imputation to deal with missing ethnicity in sensitivity analyses.

Example outputs (generated using dummy data)

References

1. Mansfield, K. E. *et al.* COVID-19 Collateral: Indirect Acute Effects of the Pandemic on Physical and Mental Health in the UK. (2020) doi:10.2139/ssrn.3719882.
2. Griffin, S. Covid-19: Failure to control pandemic and inequalities made England worst affected in Europe, says report. *BMJ* **371**, m4842 (2020).
3. Race Disparity Unit, Cabinet Office. *Quarterly report on progress to address COVID-19 health inequalities*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3618215 (2020) doi:10.1101/2020.05.02.20078642v3.full.pdf.
4. Mathur, R. *et al.* Ethnic differences in SARS-CoV-2 infection and COVID-19-related hospitalisation, intensive care unit admission, and death in 17 million adults in England: an observational cohort study using the OpenSAFELY platform. *Lancet* **397**, 1711–1724 (2021).