

# Ethnic Differences in pathways from COVID-19 infection to hospitalisation, treatment and outcomes

## Study Protocol

**Version: v1.0**

**Date: 10/07/2020**

This is a collaboration between the following institutions as part of OpenSAFELY.org:

- The DataLab, Nuffield Department of Primary Care Health Sciences, University of Oxford, OX26GG
- Electronic Health Records Research Group, Department of Non-communicable Disease Epidemiology, London School of Hygiene and Tropical Medicine

# Table of Contents

<b>Table of Contents</b>	<b>1</b>
<b>Objectives</b>	<b>3</b>
Primary Objectives	3
Secondary Objectives	3
Exploratory Objectives	3
<b>Methods</b>	<b>4</b>
Database Description	10
Study Design and Population	5
Inclusion Criteria	6
Exclusion Criteria	6
Study Measures	6
Exposure	6
Outcomes	6
Table 1. Operational Definition of Outcome Variables	7
Covariates	8
Conceptual diagram	9
Statistical Analysis	10
Descriptive	10
Univariable and multivariable models	10
Software and Reproducibility	11
Issues to consider when interpreting results	11
<b>Administrative</b>	<b>16</b>
Ethics	17
Funding	18
Conflict of Interests	18
<b>References</b>	<b>19</b>

## Background

In the UK, risks of COVID-19 infection and outcomes have been reported to be disproportionately increased amongst Black, Asian, and Minority Ethnic (BAME) groups.[1–7] Because existing studies have focused on single outcomes, or been conducted exclusively within primary care or secondary care settings, It remains unclear whether excess mortality stems from increased risk of infection, or a poorer prognosis following infection.[8]

It is hypothesized that ethnic differences in COVID-19 infection and outcomes are not driven solely by biological differences, but rather by differences in factors such as living in deprived areas, working in high-exposure/frontline occupations, living in overcrowded homes, burden of underlying conditions, experiences of racism, or access to health and community services.[5,9–12]

High level ethnic groupings such as white, south Asian, black can conceal significant heterogeneity with respect to these risk factors.[13,14] For example, while Bangladeshi and African populations are more likely to live in deprived areas, Indian and Chinese groups live in more affluent areas and experience less material deprivation.[15] Therefore it is of vital importance to disaggregate ethnic groups to better model the overlapping contributions of health and social factors on COVID-19 infection, severity, and mortality.

While previous studies have variously accounted for underlying conditions and area level measures of social deprivation, to date, no study of ethnic differences has incorporated measures of household composition. Furthermore, while most studies have examined single outcomes in single settings, none have yet linked data from across the full healthcare system to determine where ethnic disparities are most pronounced or whether ethnic differences in downstream outcomes are due to ethnic differences in risk of infection.

The aims of this study are firstly, to detail ethnic differences across the full trajectory of the COVID-19 pathway - from suspected to confirmed infection, to hospitalisation, treatment and mortality; secondly, to explore whether ethnic differences in downstream outcomes are due to higher risk of infection or poorer prognosis once infected; and thirdly, to explore the role of household composition, deprivation and underlying conditions on ethnic patterning of COVID-19 infection and outcomes.

# Objectives

## Primary Objectives

- 1) To describe ethnic differences in the risk of developing severe COVID-19 by comparing differences in:
  - a) The risk of attending A&E for COVID-19
  - b) The risk of being admitted to intensive care for COVID-19
  - c) The risk of being given invasive mechanical ventilation for COVID-19 once admitted to ICU
  - d) The risk of COVID-19 related mortality in and out of hospital
- 2) To describe ethnic differences in the risk of being infected with COVID-19 by describing differences in:
  - a) The risk of being coded with suspected COVID-19 in primary care
  - b) The risk of being coded with confirmed COVID-19 in primary care records
  - c) The risk of being tested for COVID-19
  - d) The risk of testing positive for COVID-19 (overall and amongst those with evidence of a test being conducted)
- 3) To determine whether ethnic differences in risk of severe COVID-19 are due to ethnic differences in infection by comparing risk of severe COVID-19 outcomes between the general population and those with evidence of COVID-19 infection prior to A&E attendance, ICU admission, invasive ventilation, or death.
- 4) To quantify the contribution of household size and composition to ethnic differences in COVID outcomes by determining:
  - a) How ethnic differences in outcomes are changed after adjustment for the number of people within a household and/or the age-range of individuals within a household.
  - b) How the relationship between household composition and risk of outcomes differs within each ethnic group

## Secondary Objectives

- 1) To estimate sex-specific and age-sex specific rates of each outcome by ethnic group
- 2) To investigate interactions between ethnicity and other individuals-level factors such as deprivation, geographic region, and household composition to better understand how these factors modulate ethnic differences in infection, hospitalisation, treatment, and mortality.

## Exploratory Objectives

- 1) Based on results from the main analysis, to decide whether it is sensible to explore ethnic differences in 'Severe COVID-19' by using a composite score made up of A&E admission, ICU admission, or death.

- 2) To explore the role of co-morbidities on risk of outcomes including type 1 and type 2 diabetes, CVD, hypertension, and COPD.
- 3) To compare ethnic differences over time, and across the periods of 'pre-first wave', 'peak of first-wave' and 'post-peak of first wave' with potential to look at other waves if they occur
- 4) To describe ethnic differences in discharge destination from A&E attendance
- 5) To describe ethnic differences in use of non-invasive ventilation (CPAP), which was used more frequently as the pandemic progressed.
- 6) To describe completeness of ethnicity recording in the TPP database overall and by region and how this is improved when incorporating data from linked sources

## Sensitivity Analyses

1. To estimate all outcomes using complete case analysis for ethnicity
2. To estimate all outcomes using multiple imputation for ethnicity calibrated to the regional distribution of ethnic groups

# Methods

## Database Description

We will use data from general practice (GP) records, obtained from the GP software provider The Phoenix Project (TPP), linked to outcomes data, as described below. The data will be accessed, linked and analysed through OpenSAFELY.org - a new data analytics platform created to address urgent questions relating to the epidemiology and treatment of Covid-19 in England, hosted by TPP. OpenSAFELY provides a secure software interface that allows NHS records to be pseudonymised, linked and analysed in near real-time, without data leaving their designated storage location. More information can be found on <https://opensafely.org/>.

The research dataset analysed through openSAFELY is based on primary care records retrieved from the TPP SystmOne electronic health record system. These data include diagnoses (Read 3 CTV3), prescriptions (dm+d), 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. These records will be linked to data from a number of other organisations who have agreed to make their data available for Covid-19 research as part of the openSAFELY initiative. Currently, linkage is possible to: (1) the NHSE/NHSX Emergency Care Data Set (ECDS), which contains data on emergency presentation at A&E clinics across England for patients with suspected or confirmed SARS-COV-2 infection, (2) the Intensive Care National Audit & Research Centre (ICNARC) Case Mix Programme, which holds data on Covid-19 related Intensive Treatment Units (ICU) admissions, (3) the NHSE/NHSX Second Generation Surveillance System (SGSS) data on SARS-COV-2 test results, (4) the NHSE/NHSX Covid-19 Patient Notification System (CPNS) data on deaths among Covid-19 inpatients occurring in hospitals; and (5) Office for National Statistics (ONS) death data, which includes information on all deaths, including those due to non-Covid-19 causes as well as those occurring outside the hospital setting.

For the purposes of this analysis, we use primary care records linked to data from SGSS, ECDS, ICNARC, CPNS, and ONS. All data is 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 2<sup>nd</sup> of April 2020 (REC reference: 20/LO/0651)

# Study Design and Population

## Inclusion Criteria

All adults registered (males and females aged 18 years and above) and under follow-up in a TPP primary care practice (in England) on February 1st 2020 with at least one year of prior follow-up (to be able to ascertain comorbidity data).

## Exclusion Criteria

Missing age, children aged <18 and adults >110 at the start of follow-up will be excluded.

## Study Measures

### Exposure

Self-reported ethnicity as captured in GP record or linked data (HES, ONS) and categorized into the census groupings of:

- |                 |               |
|-----------------|---------------|
| ● White British | ● African     |
| ● White Irish   | ● Caribbean   |
| ● Other White   | ● Other Black |
| ● Indian        | ● Mixed       |
| ● Bangladeshi   | ● Other       |
| ● Pakistani     | ● Unknown     |
| ● Other Asian   |               |

### Outcomes

- 1) Being coded with suspected COVID in primary care
- 2) Being coded with confirmed COVID in primary care
- 3) Being tested for COVID
- 4) Testing positive for COVID
- 5) COVID-19 related A&E attendance
- 6) COVID-19 related ICU admission
- 7) Invasive mechanical ventilation for COVID-19 in ICU
- 8) COVID-19 related death in or out of hospital

**Table 1. Operational Definition of Outcome Variables**

Outcome variable	Data Source	Definition	Timeframe
Being coded with suspected or confirmed COVID-19 in primary care	TPP	Date of first Read or SNOMED code for suspected or confirmed COVID	On or after 1 Feb 2020 till end of available data
Being tested for COVID-19	TPP SGSS	Date of first test	On or after 1 Feb 2020 till end of available data
Testing positive for COVID-19	TPP SGSS	Date of first positive test	On or after 1 Feb 2020 till end of available data (first positive test ever recorded was 3rd Feb)  <a href="https://github.com/ebmdatalab/opensafely-outcomes-notebook/blob/master/notebooks/explore_sgss.ipynb">https://github.com/ebmdatalab/opensafely-outcomes-notebook/blob/master/notebooks/explore_sgss.ipynb</a>
A&E admission	ECDS	Date of first hospital admission with confirmed or suspected COVID-19 (U07.1 and U07.2) Or Date of first ECDS record during follow-up	On or after 1 Feb 2020 till end of available data
ICU admission	ICNARC	Date of first ICNARC record during follow-up	On or after 1 Feb 2020 till end of available data
Invasive mechanical ventilation	ICNARC	Mechanical ventilation in ICU	On or after 1 Feb 2020 till end of available data
Mortality	ONS and CPNS	Death occurring in or out of hospitals with laboratory confirmed or suspected COVID-19.	On or after 1 Feb 2020 until seven days prior to the end of data availability (to account for week lag in data completeness)



## Covariates

### Demographics

- Age (continuous as spline, and in categories of 18-<40, 40-<50, 50-<60, 60-<70, 70-<80, 80+, based on month and year of birth).
- Sex (male or female)
- Deprivation, (quintile of the index of multiple deprivation (IMD), derived from the patient's postcode at lower super output area level for a high degree of precision)
- LSOA in two categories
- Household composition (Household size, age-range)
- Number of consultations in the 12 months prior to baseline

### Clinical Measures:

- BMI – obese class I/II/III using ethnic specific cut-points vs not obese and as continuous
- Smoking – current vs former vs never
- Systolic and Diastolic BP (continuous and categorised as raised BP)
- HbA1c (continuous)

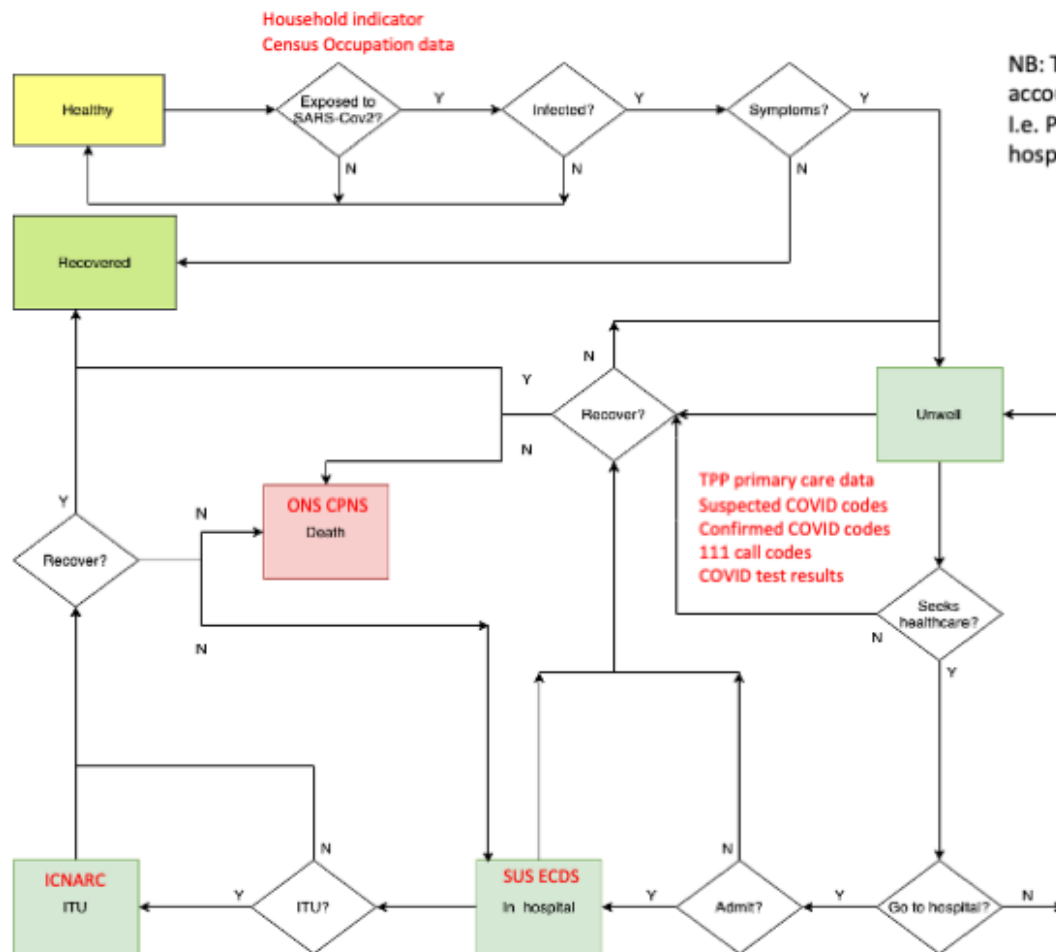
### Comorbidities:

- Diagnosed hypertension - will combine with antihypertensive medication and BP values to generate a variable of treated vs. untreated hypertension
- respiratory disease other than asthma
- asthma
- chronic cardiac disease (CVD)
- Diabetes separated by type (T1 and T2)
- cancer (haematological and non-haematological considered together)
- liver disease
- Stroke
- dementia
- other neurological disease
- kidney disease ( eGFR<60)
- end-stage renal disease (dialysis/transplant)
- organ transplant
- asplenia
- Grouped rheumatoid arthritis, lupus, psoriasis;
- other immunosuppressive condition (permanent (ever on record) and temporary (in last year))
- Time period (calendar month and categories of 'pre-first wave' first wave 'post-first wave')
- Geographic region and sustainability and transformation partnerships (STP) and ICU catchment area (if technically feasible)

### Medications: current medication at least 2 scripts in the 6 months prior to baseline:

- Oral antidiabetic drugs
- Insulin
- Lipid lowering drugs
- Antihypertensive drugs

## Conceptual diagram



NB: This diagram does not account for all possible pathways. I.e. People may go from healthy → hospital → COVID

## Statistical Analysis

Participants will be observed from the 1st of February 2020 and observation for outcome-specific analyses will end at the earliest of: occurrence of the outcome, death, or the relevant end date for the dataset capturing outcome.

Descriptive statistics for all ethnic groups will be reported in the descriptive tables. Subsequent results for White British, White Irish, Indian, Bangladeshi, Pakistani, African and Caribbean groups will be reported in the main text, while results for all other ethnic groups will be reported in the supplementary materials.

For the primary analysis, missing ethnicity will be imputed using standard multiple imputation techniques. Previous methods to calibrate multiple imputation of missing ethnicity by weighting the imputation to match regional ethnicity distributions have yielded comparable findings to standard multiple imputation techniques. Sensitivity analyses using complete case approaches and calibrated multiple imputation will also be conducted. Multiple imputation of other covariate data will be considered in order to make best use of the available data.

### **Start of study**

- 1 Feb 2020 (Objectives 1, 2 and 4)
- Date of first positive evidence of infection from primary care record or SGSS testing data (Objective 3)

### **End of study:**

- To be determined based on when analysis will be done, using as much outcome data as is feasible.
- For analyses involving ONS death, end of study is the latest date at which death data is available until.
- For all other analyses, end of study is the latest date the outcome data are available to.
- Follow-up will for non-mortality outcomes will be censored at the earliest of date of death, date of leaving the TPP database, end of available outcome data.

## Descriptive

Absolute risks/rates of each outcome will be described separately by each ethnic group. Ethnicity stratified kaplan-meier plots or cumulative incidence plots will be generated for each outcome. All descriptive analyses will be reported firstly by ethnic group overall, secondly by ethnic group and sex, and thirdly, by ethnic group, sex, and age group. Some age categories may be collapsed together due to small numbers in each ethnic group-sex specific strata

## Univariable and multivariable models

Cox proportional hazards regression models will be fitted to each outcome to determine differences between the White British ethnic group and all other ethnic groups.

Timescale: Days since cohort entry (i.e. days since 1 Feb 2020)

Time at risk: from cohort entry until the first of death, the outcome of interest, or end of study

### **Model adjustment**

Model 1: Age, sex, region

Model 2: + All factors listed above except household composition

Model 3: all factors in model 2 plus household composition

## **Software and Reproducibility**

Data management will be performed using Python 3.8 and SQL, with analysis carried out using Stata 16 / Python 3.8. Code for data management and analysis as well as codelists archived online at <https://github.com/opensafely/ethnicity-covid-research>. Codelists are available for review and reuse at [codelists.opensafely.org](https://codelists.opensafely.org).

## **Issues to consider when interpreting results**

### **ICNARC data**

- ICNARC dataset includes people with either a) confirmed COVID b) clinical evidence of COVID despite negative test
- Admissibility to ICU is partially based on age. Younger people are preferentially admitted to ICU because they are likely to benefit more. As BAME groups are on average younger, they may be more likely to be admitted than white groups. Therefore increased risk of admission to ICU may not necessarily be a marker of more severe disease in BAME groups - and may confer differential protection for mortality or other downstream outcomes.

### **SGSS data**

- Many people early in the epidemic will have received a COVID-19 tests after being admitted to hospital, so we may be unable to examine 'risk of attending A&E' in these people. Later on the epidemic when primary care and community testing became more prevalent, this becomes less of an issue.
- Similarly, people may get their primary care COVID code after having been to hospital, so we need to treat dates and ordering of events carefully.
- UKB analysis shows that BAME groups twice as likely to test positive compared to white groups- suggests that
  - BAME groups need to be sicker before getting tested
  - Greater proportion of BAME groups hospitalized and subsequently tested may be frontline workers
  - Ethnic differences in testing rates/outcomes may change over time as testing criteria were updated - this will be explored in the exploratory analysis comparing pre-peak, peak and post-peak trends.

## Table Shells

Table 1. Description of baseline characteristics by ethnic group

	Ethnic Group								
	White British	White Irish	Indian	Pakistani	Bangladeshi	Caribbean	African	Mixed	Other
<b>Total</b>									
<b>Sex</b>									
Male									
Female									
<b>Age</b>									
18-<40									
...									
80+									
<b>IMD</b>									
Least deprived quintile									
...									
Most deprived quintile									
<b>Comorbidities</b>									
Asthma									
T1DM									
T2DM									
CHD									
Respiratory disease									
Etc ...									
<b>Household size</b>									
1-2									
3-4									
5+									

Table 2. Absolute rates of each outcome by ethnic group

Outcome	White British	White Irish	Indian	Pakistani	Bangladeshi	African	Caribbean
Primary care code	rate/100,000						
Tested							
Test positive							
...							
Death							

Table 3. Age-sex specific absolute rates of each outcome by ethnic group

All ages	Male							Female						
Outcome	White British	White Irish	Indian	Pakistani	Bangladeshi	African	Caribbean	White British	White Irish	Indian	Pakistani	Bangladeshi	African	Caribbean
Primary care code	rate/100,000													
Tested														
Test positive														
...														
Death														
18- <40	Male							Female						
Outcome	White British	White Irish	Indian	Pakistani	Bangladeshi	African	Caribbean	White British	White Irish	Indian	Pakistani	Bangladeshi	African	Caribbean
Primary care code	rate/100,000													
Tested														
Test positive														
...														
Death														
40- <50	Male							Female						
Outcome	White British	White Irish	Indian	Pakistani	Bangladeshi	African	Caribbean	White British	White Irish	Indian	Pakistani	Bangladeshi	African	Caribbean
Primary care code	rate/100,000													

Tested														
Test positive														
...														
Death														
50- <60	Male							Female						
Outcome	White British	White Irish	Indian	Pakistani	Bangladeshi	African	Caribbean	White British	White Irish	Indian	Pakistani	Bangladeshi	African	Caribbean
Primary care code	rate/100,000													
Tested														
Test positive														
...														
Death														
60 - <70	Male							Female						
Outcome	White British	White Irish	Indian	Pakistani	Bangladeshi	African	Caribbean	White British	White Irish	Indian	Pakistani	Bangladeshi	African	Caribbean
Primary care code	rate/100,000													
Tested														
Test positive														
...														
Death														
70+	Male							Female						
Outcome	White British	White Irish	Indian	Pakistani	Bangladeshi	African	Caribbean	White British	White Irish	Indian	Pakistani	Bangladeshi	African	Caribbean
Primary care code	rate/100,000													
Tested														
Test positive														
...														

Table 4a. Risk of each severe COVID-19 outcomes by ethnic group in the general population

	HR (95%CI)						
	White British (reference)	White Irish	Indian	Pakistani	Bangladeshi	Caribbean	African
A&E attendance	1						
ICU admission	1						
Ventilation	1						
In-hospital death	1						
Out of hospital death	1						

\* model 1: Adjusted for age, sex, region Model 2: Adjusted for all nature paper variables, model 3: additionally adjusted for household composition

Table 4b. Risk of each infection by ethnic group in the general population

	HR (95%CI)						
	White British (reference)	White Irish	Indian	Pakistani	Bangladeshi	Caribbean	African
Being coded with suspected COVID-19	1	Model1 HR, 95%CI Model2 HR, 95%CI Model3 HR, 95%CI					
Being coded with confirmed COVID-19							
Being tested for COVID (SGSS)	1						
Testing positive for COVID (SGSS)	1						

\* model 1: Adjusted for age, sex, region Model 2: Adjusted for all nature paper variables, model 3: additionally adjusted for household composition

Table4c. Risk of each severe COVID-19 outcomes by ethnic group in those with evidence of COVID-19 infection

	HR (95%CI)						
	White British (reference)	White Irish	Indian	Pakistani	Bangladeshi	Caribbean	African
A&E attendance	1	Model1 HR, 95%CI Model2 HR, 95%CI Model3 HR, 95%CI					
ICU admission	1						
Ventilation							
In-hospital death	1						
Out of hospital death	1						



\* model 1: Adjusted for age, sex, region Model 2: Adjusted for all nature paper variables, model 3: additionally adjusted for household composition

Table 5. Risk of each outcome by number of people within each household by ethnic group

	White British			White Irish			Indian		
HH size	1-2 (ref)	3-4	5+	1-2 (ref)	2-4	5+	1-2 (ref)	3-4	5+
	HR 95%CI p	HR 95%CI p	HR 95%CI p	HR 95%CI p	HR 95%CI p	HR 95%CI p	HR 95%CI p	HR 95%CI p	HR 95%CI p
Suspected covid	1			1			1		
...	1			1			1		
Covid death	1			1			1		

\* model 1: Adjusted for age, sex, region Model 2: Adjusted for all nature paper variables

Table 6. Risk of each outcome by diabetes type and diabetic control at baseline within each ethnic group

	Type 1 diabetes		Type 2 diabetes	
	HbA1c <=58 mmol/mol	HbA1c >58 mmol/mol	HbA1c <=58 mmol/mol	HbA1c >58 mmol/mol
Suspected COVID	HR, 95%CI			
...				
COVID death				

\* model 1: Adjusted for age, sex, region Model 2: Adjusted for all variables except diabetes and HbA1c

\*Baseline HbA1c defined as latest HbA1c value in the 12 months prior to Feb 1st, 2020

Table 6. Risk of each outcome by hypertension and blood pressure control at baseline within each ethnic group

	No hypertension diagnosed				Hypertension diagnosis ever			
	Normal BP/ no tx	Normal BP-treated	Raised BP – no tx	Raised BP-treated	Normal BP/ no tx	Normal BP-treated	Raised BP – no tx	Raised BP-treated
Suspected COVID	HR, 95%CI							
...								
COVID death								

\* model 1: Adjusted for age, sex, region Model 2: Adjusted for all variables except BP and hypertension

\*Baseline BP defined as latest SBP/DBP value in the 12 months prior to Feb 1st, 2020

# Administrative

## Ethics

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 has 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 via a virtual private network (VPN) connection, restricted to a small group of researchers, their specific machine and IP address; the researchers hold contracts with NHS England and only access the platform to initiate database queries and statistical models; all database activity is logged; only aggregate statistical outputs leave the platform environment following best practice for anonymisation of results such as statistical disclosure control for low cell counts.<sup>54</sup> 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.<sup>55</sup> Taken together, these provide the legal bases to link patient datasets on the OpenSAFELY platform. This study was approved by the Health Research Authority (REC reference 20/LO/0651) and by the LSHTM Ethics Board (reference 21863).

## References

1. Pan D, Sze S, Minhas JS, Bangash MN, Pareek N, Divall P, et al. The impact of ethnicity on clinical outcomes in COVID-19: A systematic review. *EClinicalMedicine*. 2020. doi:10.1016/j.eclinm.2020.100404
2. Hull SA, Williams C, Ashworth M, Carvalho C, Boomla K. Suspected COVID-19 in primary care: how GP records contribute to understanding differences in prevalence by ethnicity. *medRxiv*. 2020. doi:10.1101/2020.05.23.20101741
3. The OpenSAFELY Collaborative, Williamson E, Walker AJ, Bhaskaran K, Bacon S, Bates C, et al. OpenSAFELY: factors associated with COVID-19-related hospital death in the linked electronic health records of 17 million adult NHS patients. [cited 7 May 2020]. doi:10.1101/2020.05.06.20092999
4. Aldridge RW, Lewer D, Katikireddi SV, Mathur R, Pathak N, Burns R, et al. Black, Asian and Minority Ethnic groups in England are at increased risk of death from COVID-19: indirect standardisation of NHS mortality data. *Wellcome Open Res*. 2020;5: 88. doi:10.12688/wellcomeopenres.15922.1
5. Baumer T, Phillips E, Dhadda A, Szakmany T, Group GC-19. Insights into the Epidemiology of the First Wave of COVID-19 ICU Admissions in South Wales – the Interplay between Ethnicity and Deprivation. *Preprints*. 2020. doi:http://dx.doi.org/10.20944/preprints202006.0029.v1
6. Apea VJ, Wan YI, Dhairyawan R, Puthuchearu ZA, Pearse RM, Orkin CM, et al. Ethnicity and outcomes in patients hospitalised with COVID-19 infection in East London: an observational cohort study. *medRxiv*. 2020. doi:10.1101/2020.06.10.20127621
7. Kirby T. Evidence mounts on the disproportionate effect of COVID-19 on ethnic minorities. *Lancet Respir Med*. 2020. doi:10.1016/S2213-2600(20)30228-9
8. Khunti K, Singh AK, Pareek M, Hanif W. Is ethnicity linked to incidence or outcomes of covid-19? *BMJ*. 2020. doi:10.1136/bmj.m1548
9. Lassale C, Gaye B, Hamer M, Gale CR, Batty GD. Ethnic disparities in hospitalisation for COVID-19 in England: The role of socioeconomic factors, mental health, and inflammatory and pro-inflammatory factors in a community-based cohort study. *Brain Behav Immun*. 2020. doi:10.1016/j.bbi.2020.05.074
10. Hawkins D. Differential occupational risk for COVID-19 and other infection exposure according to race and ethnicity. *Am J Ind Med*. 2020. doi:10.1002/ajim.23145
11. Rimmer A. Covid-19: Disproportionate impact on ethnic minority healthcare workers will be explored by government. *BMJ*. 2020. doi:10.1136/bmj.m1562
12. Gupta R, Hussain A, Misra A. Diabetes and COVID-19: evidence, current status and unanswered research questions. *European Journal of Clinical Nutrition*. 2020. doi:10.1038/s41430-020-0652-1
13. Sheldon TA, Parker H, T.A. S, H. P. Race and ethnicity in health research. *J Public Health Med*. 1992/06/01. 1992;14: 104–110. Available: <http://ovidsp.ovid.com/ovidweb.cgi?T=JS&CSC=Y&NEWS=N&PAGE=fulltext&D=emed2&AN=19921>

97604

14. Mathur R, Grundy E, Smeeth L. Availability and use of UK based ethnicity data for health research. Natl Cent Res Methods Work Pap Ser. London: n/a; 2013 Mar. Report No.: 01/13. Available: [http://eprints.ncrm.ac.uk/3040/1/Mathur-\\_Availability\\_and\\_use\\_of\\_UK\\_based\\_ethnicity\\_data\\_for\\_health\\_res\\_1.pdf](http://eprints.ncrm.ac.uk/3040/1/Mathur-_Availability_and_use_of_UK_based_ethnicity_data_for_health_res_1.pdf)
15. Garner S, Bhattacharyya G. Poverty , ethnicity and place. 2011.