

## Short report: Ethnicity and COVID-19 death in the early part of the COVID-19 second wave in England: an analysis of OpenSAFELY data from 1<sup>st</sup> September to 9<sup>th</sup> November 2020

Krishnan Bhaskaran<sup>\*1</sup>, Rohini Mathur<sup>\*1</sup>, Christopher T Rentsch<sup>1</sup>, Caroline E Morton<sup>2</sup>, William J Hulme<sup>2</sup>, Anna Schultze<sup>1</sup>, Brian MacKenna<sup>2</sup>, Rosalind M Eggo<sup>1</sup>, Angel YS Wong<sup>1</sup>, Elizabeth J Williamson<sup>1</sup>, Harriet Forbes<sup>1</sup>, Kevin Wing<sup>1</sup>, Helen I McDonald<sup>1</sup>, Chris Bates<sup>3</sup>, Seb Bacon<sup>2</sup>, Alex J Walker<sup>2</sup>, David Evans<sup>2</sup>, Peter Inglesby<sup>2</sup>, Amir Mehrkar<sup>2</sup>, Helen J Curtis<sup>2</sup>, Nicholas J DeVito<sup>2</sup>, Richard Croker<sup>2</sup>, Henry Drysdale<sup>2</sup>, Jonathan Cockburn<sup>3</sup>, John Parry<sup>3</sup>, Frank Hester<sup>3</sup>, Sam Harper<sup>3</sup>, Ian J Douglas<sup>1</sup>, Laurie Tomlinson<sup>1</sup>, Stephen JW Evans<sup>1</sup>, Richard Grieve<sup>1</sup>, Liam Smeeth<sup>+1</sup>, Ben Goldacre<sup>+2</sup>

*\*contributed equally; + joint senior authors*

*1 London School of Hygiene and Tropical Medicine, WC1E 7HT*

*2 The DataLab, Nuffield Department of Primary Care Health Sciences, University of Oxford, OX26GG*

*3 TPP, TPP House, 129 Low Lane, Horsforth, Leeds, LS18 5PX*

### Summary

Black and minority ethnic groups were at raised risk of dying from COVID-19 during the first few months of the COVID-19 epidemic in England. We aimed to investigate whether ethnic inequalities in COVID-19 deaths were similar in the more recent “second wave” of the epidemic. Working on behalf of NHS England, we used primary care and linked ONS mortality data within the OpenSAFELY platform. All adults in the database at 1st September 2020 and with at least 1 year of prior follow-up and a record of ethnicity were included. The outcome was COVID-19-related death (death with COVID-19 listed as a cause of death on the death certificate). Follow-up was to 9th November 2020. Hazard ratios for ethnicity were calculated using Cox regression models adjusted for age and sex, and then further adjusted for deprivation. 13,223,154 people were included. During the study period, people of South Asian ethnicity were at higher risk of death due to COVID-19 than white people after adjusting for age and sex (HR = 3.47, 95% CI 2.99-4.03); the association attenuated somewhat on further adjustment for index of multiple deprivation (HR = 2.86, 2.46-3.33, Table 2). In contrast with the first wave of the epidemic, we found little evidence of a raised risk in black or other ethnic groups compared to white (HR for black vs white = 1.28, 0.87-1.88 adjusted for age and sex; and 1.01, 0.69-1.49 further adjusted for deprivation). Our findings suggest that ethnic inequalities in the risk of dying COVID-19-related death have changed between the first and early second wave of the epidemic in England.

**Background and aims:** As reported by OpenSAFELY<sup>1,2</sup> and others,<sup>3</sup> black and minority ethnic groups were at raised risk of dying from COVID-19 during the first few months of the COVID-19 epidemic in England. We aimed to investigate whether ethnic inequalities in COVID-19 deaths were similar in the more recent “second wave” of the epidemic.

**Methods:** The analysis was conducted within the OpenSAFELY platform on behalf of NHS England, using routinely collected electronic primary care data from practices using TPP SystmOne software, linked to ONS death registrations. The start date of the study was 1<sup>st</sup> September 2020, chosen as being approximately the date of minimum COVID-19 mortality between the first and second waves of the epidemic in England. All adults under follow-up in a TPP general practice on 1<sup>st</sup> September 2020, with at least one year of follow-up prior to this date, and with a record indicating their ethnicity were included. Ethnicity codes were classified into 5 main categories; a 16-level categorisation was used in a secondary analysis. COVID-19 mortality was defined as a record for death in linked ONS data (complete to 9<sup>th</sup> November 2020) with the COVID related ICD-10 codes U071 or U072 listed as a cause of death anywhere on the death certificate. Following our previous methodology,<sup>2</sup> Cox regression models were used to estimate the association between ethnicity and COVID-19 mortality, first adjusted for age (as a 4-knot cubic spline) and sex only, and then additionally adjusted for index of multiple deprivation, an area-based measure of socioeconomic status derived from the patient’s postcode. Analyses were stratified by the Sustainability and Transformation Partnership (STP, an NHS administrative region) of the patient’s general practice to account for geographical variation in infection rates across the country.

**Results:** Of 17,695,074 people registered in TPP on 1<sup>st</sup> September 2020 and with a year’s prior follow-up available, 13,223,154 (75%) had a record of ethnicity and were included (Table 1). People of South Asian ethnicity were at higher risk of death due to COVID-19 than white people during the study period after adjusting for age and sex (HR = 3.47, 95% CI 2.99-4.03); the association attenuated somewhat on further adjustment for index of multiple deprivation (HR = 2.86, 2.46-3.33, Table 2). Analysis using a finer categorisation of ethnicity suggested that the association was especially pronounced for Bangladeshi and Pakistani ethnicities. In contrast with results from the first wave of the epidemic, there was little evidence during the study period of raised risk in Black or Other ethnic groups compared to white (Figure 1). Further analysis stratified by age and sex showed similar patterns (Tables A1-A3). A further exploratory analysis suggested a largely consistent pattern across English regions (Table A4, Figure A1).

**Interpretation:** Ethnic inequalities in the risk of dying COVID-19-related death appear to have changed between the first and early second wave of the epidemic in England. In the period between September and early November 2020, South Asian ethnic groups (in particular those of Bangladeshi and Pakistani ethnicity) were at markedly increased risk compared to people of white ethnicity, while people of Black and other ethnicities had similar risks to white. This suggests that mitigation strategies may need to be targeted towards South Asian communities.

Any underlying ethnic differences in the tendency to experience more severe disease once infected would be expected to remain relatively constant over time. Here we have observed marked changes in ethnic differences

in outcomes within a short time period, strongly suggesting that ethnic inequalities in COVID-19 outcomes are driven by risk of infection as opposed to innate susceptibility to severe disease.

Ongoing monitoring of ethnic inequalities in COVID-19 outcomes, including by geographic region, is needed to detect any further changes in these patterns.

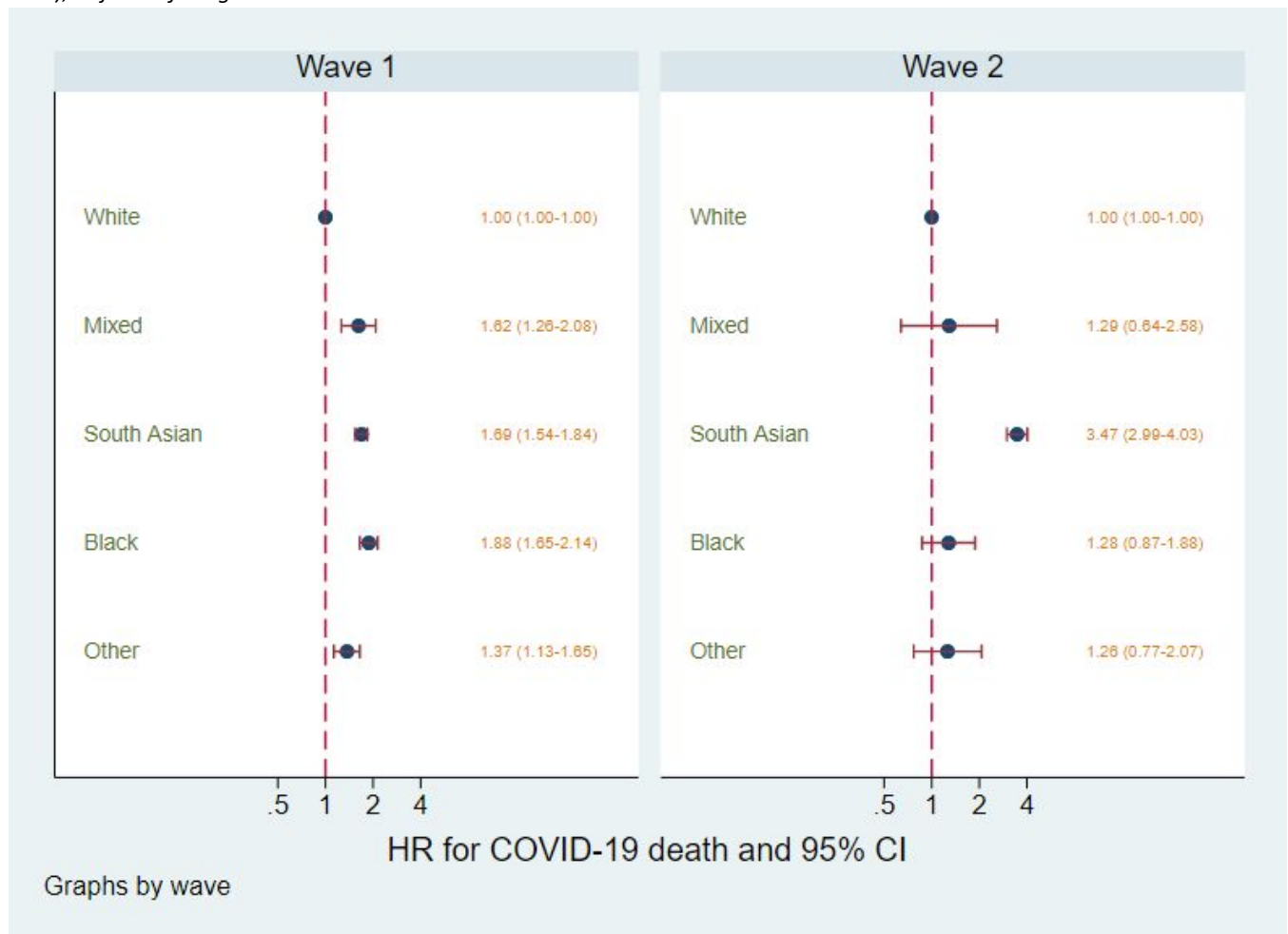
Table 1: Characteristics of the study population

|  | Ethnicity          |                 |                   |                 |                 |
|--|--------------------|-----------------|-------------------|-----------------|-----------------|
|  | White              | Mixed           | South Asian       | Black           | Other           |
| N                                      | 11,295,498 (100.0) | 179,253 (100.0) | 1,059,377 (100.0) | 355,195 (100.0) | 333,831 (100.0) |
| Age                                    |                    |                 |                   |                 |                 |
| 18-39                                  | 3,421,561 (30.3)   | 97,323 (54.3)   | 494,537 (46.7)    | 147,586 (41.6)  | 178,499 (53.5)  |
| 40-49                                  | 1,838,647 (16.3)   | 36,669 (20.5)   | 252,690 (23.9)    | 84,268 (23.7)   | 72,072 (21.6)   |
| 50-59                                  | 2,084,805 (18.5)   | 25,606 (14.3)   | 143,614 (13.6)    | 70,052 (19.7)   | 42,485 (12.7)   |
| 60-69                                  | 1,720,122 (15.2)   | 11,796 (6.6)    | 96,704 (9.1)      | 30,687 (8.6)    | 23,887 (7.2)    |
| 70-79                                  | 1,460,546 (12.9)   | 5,202 (2.9)     | 46,763 (4.4)      | 12,899 (3.6)    | 11,703 (3.5)    |
| 80+                                    | 769,817 (6.8)      | 2,657 (1.5)     | 25,069 (2.4)      | 9,703 (2.7)     | 5,185 (1.6)     |
| Median (IQR)                           | 51 (36-66)         | 38 (28-50)      | 41 (32-52)        | 43 (32-54)      | 38 (29-49)      |
| Sex                                    |                    |                 |                   |                 |                 |
| Female                                 | 5,872,445 (52.0)   | 92,162 (51.4)   | 508,560 (48.0)    | 177,044 (49.8)  | 163,852 (49.1)  |
| Male                                   | 5,423,053 (48.0)   | 87,091 (48.6)   | 550,817 (52.0)    | 178,151 (50.2)  | 169,979 (50.9)  |
| Index of multiple deprivation quintile |                    |                 |                   |                 |                 |
| 1 (least deprived)                     | 2,423,497 (21.5)   | 23,885 (13.3)   | 93,346 (8.8)      | 24,089 (6.8)    | 45,053 (13.5)   |
| 2                                      | 2,378,336 (21.1)   | 28,929 (16.1)   | 121,274 (11.4)    | 36,626 (10.3)   | 57,590 (17.3)   |
| 3                                      | 2,336,657 (20.7)   | 34,831 (19.4)   | 201,001 (19.0)    | 59,057 (16.6)   | 65,490 (19.6)   |
| 4                                      | 2,174,386 (19.3)   | 42,262 (23.6)   | 304,523 (28.7)    | 94,714 (26.7)   | 81,802 (24.5)   |
| 5                                      | 1,982,622 (17.6)   | 49,346 (27.5)   | 339,233 (32.0)    | 140,709 (39.6)  | 83,896 (25.1)   |

Table 2: Association between ethnicity and COVID-19 death 1<sup>st</sup> Sept - 9<sup>th</sup> Nov 2020

|  | Ethnicity  |  |  |  |  |
|--|--|--|--|--|--|
|  | White  | Mixed  | South Asian  | Black  | Other  |
| COVID-19 deaths  | 1771   | 8  | 224  | 27   | 16   |
| Person-years (millions)  | 2,131,944  | 33,853   | 200,061  | 67,078   | 63,054   |
| Age-sex adjusted HR  | 1.00 (REF)   | 1.29 (0.64-2.58)   | 3.47 (2.99-4.03)   | 1.28 (0.87-1.88)   | 1.26 (0.77-2.07)   |
| + deprivation adjusted   | 1.00 (REF)   | 1.13 (0.56-2.26)   | 2.86 (2.46-3.33)   | 1.01 (0.69-1.49)   | 1.14 (0.70-1.88)   |
| Finer ethnic categories  |  |  |  |  |  |
| Age-sex adjusted HRs   | <u>British/Mixed</u><br>1.00 (REF)<br><u>British</u><br>0.84 (0.48-1.49)<br><u>Irish</u><br>1.09 (0.88-1.35) | <u>White+Black</u><br>0.85 (0.21-3.40)<br><u>White+Black African</u><br>2.05 (0.51-8.22)<br><u>White + Asian</u><br>0.93 (0.13-6.64)<br><u>Other mixed</u><br>1.62 (0.52-5.03) | <u>Indian/British Indian</u><br>2.70 (2.12-3.43)<br><u>Pakistani/British</u><br>4.43 (3.62-5.43)<br><u>Pakistani</u><br>5.96 (3.63-9.79)<br><u>Bangladeshi/British</u><br>2.49 (1.62-3.81)<br><u>Bangladeshi</u><br>2.49 (1.62-3.81) | <u>Caribbean</u><br>1.27 (0.79-2.07)<br><u>African</u><br>0.99 (0.41-2.38)<br><u>Other Black</u><br>1.68 (0.70-4.06) | <u>Chinese</u><br>1.07 (0.40-2.85)<br><u>Other</u><br>1.32 (0.75-2.34) |
| PREVIOUS RESULTS FOR COMPARISON<br>Study period 1st Feb to 6th May<br>(Williamson et al <sup>2</sup> ) |  |  |  |  |  |
| Age-sex adjusted HRs   | 1.00 (REF)   | 1.62 (1.26-2.08)   | 1.69 (1.54-1.84)   | 1.88 (1.65-2.14)   | 1.37 (1.13-1.65)   |

Figure 1: Association between ethnicity and COVID-19 death in wave 1 (1st Feb - 6th May)<sup>2</sup> versus wave 2 (1<sup>st</sup> Sept - 9<sup>th</sup> Nov 2020), adjusted for age and sex



## **References**

1. Mathur R, Rentsch CT, Morton C, Hulme WJ, Schultze A, MacKenna B, Eggo RM, Bhaskaran K, Wong AY, Williamson EJ, Forbes H, Wing K, McDonald HI, Bates C, Bacon S, Walker AJ, Evans D, Inglesby P, Mehrkar A, Curtis HJ, DeVito NJ, Croker R, Drysdale H, Cockburn J, Parry J, Hester F, Harper S, Douglas IJ, Tomlinson L, Evans S, Grieve R, Harrison D, Rowan K, Khunti K, Chaturvedi N, Smeeth L, Goldacre B. Ethnic differences in COVID-19 infection, hospitalisation, and mortality: an OpenSAFELY analysis of 17 million adults in England. *medRxiv* 2020: 2020.09.22.20198754. doi: 10.1101/2020.09.22.20198754.
2. Williamson EJ, Walker AJ, Bhaskaran K, Bacon S, Bates C, Morton CE, Curtis HJ, Mehrkar A, Evans D, Inglesby P. OpenSAFELY: factors associated with COVID-19 death in 17 million patients. *Nature* 2020: 1-11. doi: 10.1038/s41586-020-2521-4.
3. Public Health England. Disparities in the risk and outcomes of COVID-19. 2020.  
[https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/908434/Disparities\\_in\\_the\\_risk\\_and\\_outcomes\\_of\\_COVID\\_August\\_2020\\_update.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/908434/Disparities_in_the_risk_and_outcomes_of_COVID_August_2020_update.pdf) (accessed 18th December 2020).

## Funding

This work was supported by the Medical Research Council MR/V015737/1. TPP provided technical expertise and infrastructure within their data centre pro bono in the context of a national emergency.

RM holds a fellowship funded by the Wellcome Trust. BG's work on better use of data in healthcare more broadly is currently funded in part by: NIHR Oxford Biomedical Research Centre, NIHR Applied Research Collaboration Oxford and Thames Valley, the Mohn-Westlake Foundation, NHS England, and the Health Foundation; all DataLab staff are supported by BG's grants on this work. LS reports grants from Wellcome, MRC, NIHR, UKRI, British Council, GSK, British Heart Foundation, and Diabetes UK outside this work. AS is employed by LSHTM on a fellowship sponsored by GSK. KB holds a Sir Henry Dale fellowship jointly funded by Wellcome and the Royal Society. HIM is funded by the National Institute for Health Research (NIHR) Health Protection Research Unit in Immunisation, a partnership between Public Health England and LSHTM. AYSW holds a fellowship from BHF. EW holds grants from MRC. RG holds grants from NIHR and MRC. ID holds grants from NIHR and GSK. HF holds a UKRI fellowship. RE is funded by HDR-UK and the MRC.

The views expressed are those of the authors and not necessarily those of the NIHR, NHS England, Public Health England or the Department of Health and Social Care. Funders had no role in the study design, collection, analysis, and interpretation of data; in the writing of the report; and in the decision to submit the article for publication.

## Information governance and ethical approval

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; 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. 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. Taken 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.

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

## SUPPLEMENTARY ANALYSIS

Table A1: HRs for the association between ethnicity and COVID-19 death stratified by sex, during the period 1<sup>st</sup> Sept to 9<sup>th</sup> Nov 2020

|             | Men              | Women            |
|-------------|------------------|------------------|
| White       | 1.00 (ref)       | 1.00 (ref)       |
| Mixed       | 1.08 (0.41-2.90) | 1.59 (0.59-4.25) |
| South Asian | 4.02 (3.37-4.80) | 2.62 (2.03-3.40) |
| Black       | 1.02 (0.59-1.78) | 1.66 (0.98-2.83) |
| Other       | 1.31 (0.70-2.45) | 1.19 (0.53-2.65) |

From an age and sex adjusted model with interaction between ethnicity and sex. P-interaction = 0.04

Table A2: HRs for the association between ethnicity (in fine categories\*) and COVID-19 death stratified by sex, during the period 1<sup>st</sup> Sept to 9<sup>th</sup> Nov 2020

|                           | Men               | Women             |
|---------------------------|-------------------|-------------------|
| <b>White groups</b>       |                   |                   |
| White British             | 1.00 (ref)        | 1.00 (ref)        |
| Other White               | 1.10 (0.85-1.42)  | 1.00 (0.74-1.36)  |
|                           |                   |                   |
| <b>Mixed</b>              | 1.08 (0.40-2.89)  | 1.57 (0.59-4.21)  |
|                           |                   |                   |
| <b>South Asian groups</b> |                   |                   |
| Indian                    | 3.33 (2.53-4.37)  | 1.72 (1.10-2.71)  |
| Pakistani                 | 4.94 (3.86-6.33)  | 3.68 (2.61-5.20)  |
| Bangladeshi               | 7.25 (4.09-12.85) | 3.91 (1.46-10.48) |
| Other Asian               | 2.74 (1.64-4.59)  | 2.09 (0.99-4.42)  |
|                           |                   |                   |
| <b>Black groups</b>       |                   |                   |
| Black Caribbean           | 1.27 (0.68-2.38)  | 1.29 (0.61-2.72)  |
| Other Black               | 0.61 (0.20-1.90)  | 2.26 (1.07-4.77)  |
|                           |                   |                   |
| <b>Other</b>              | 1.30 (0.70-2.44)  | 1.17 (0.52-2.62)  |

From an age and sex adjusted model with interaction between ethnicity and sex. \*All Mixed groups were combined, and Black African/Other Black were combined, due to small numbers of events. P-interaction = 0.15

Table A3: HRs for the association between ethnicity and COVID-19 death stratified by age, during the period 1<sup>st</sup> Sept to 9<sup>th</sup> Nov 2020

|             | Age <70 years    | Age ≥70 years    |
|-------------|------------------|------------------|
| White       | 1.00 (ref)       | 1.00 (ref)       |
| Mixed       | 1.60 (0.51-5.02) | 1.19 (0.49-2.86) |
| South Asian | 4.65 (3.54-6.10) | 3.25 (2.73-3.87) |
| Black       | 1.66 (0.82-3.36) | 1.21 (0.76-1.91) |
| Other       | 0.88 (0.28-2.76) | 1.45 (0.84-2.51) |

From an age and sex adjusted model (age fitted as a spline as per original model), with interaction added between ethnicity and binary age group (<70/≥70). P-interaction = 0.18

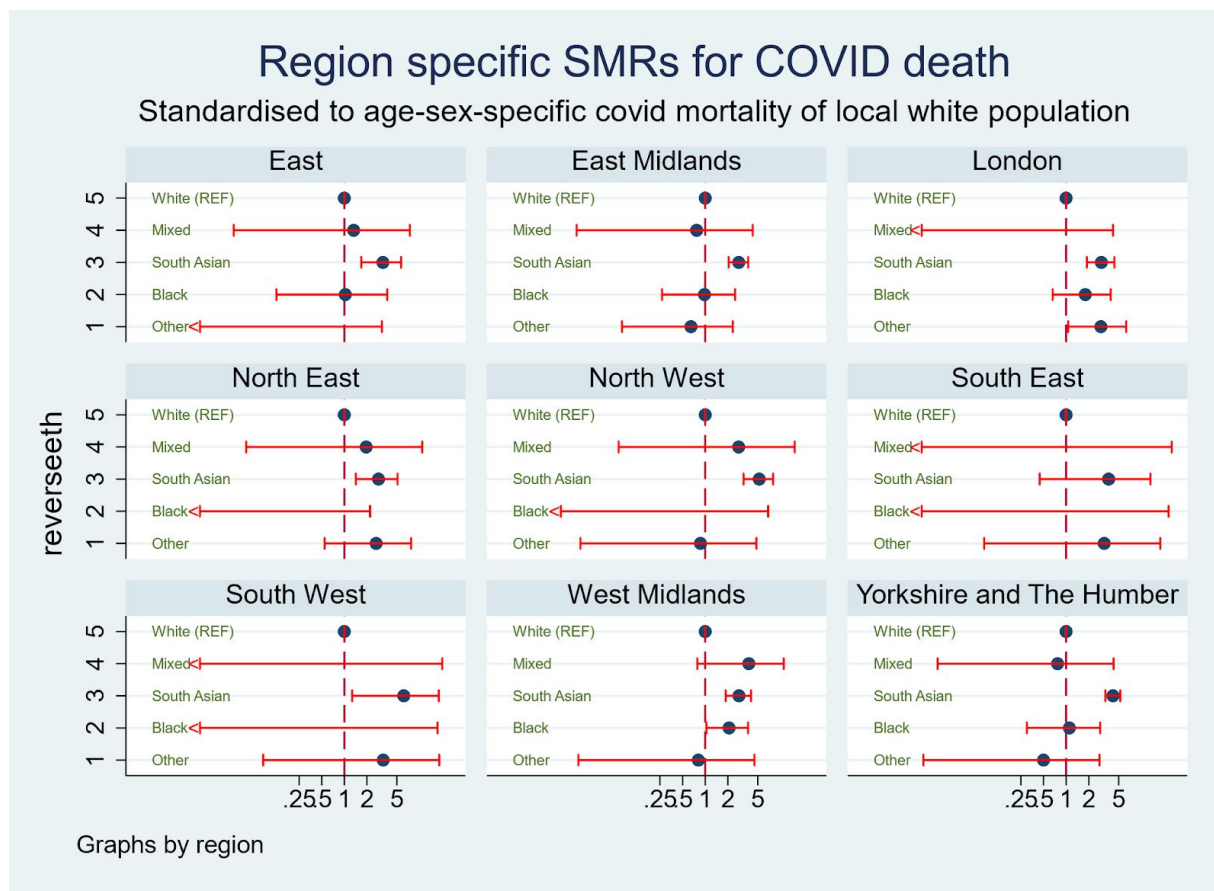
Table A4: Region-specific observed and expected\* deaths by ethnic group

|                    |            | White          | Mixed        | S Asian       | Black        | Other        |
|--------------------|------------|----------------|--------------|---------------|--------------|--------------|
| East               | Population | <b>2695831</b> | <b>42128</b> | <b>169390</b> | <b>84801</b> | <b>61624</b> |
|                    | Observed   | 188            | <=5          | 12            | <=5          | 0            |
|                    | Expected   | -              | 0.75         | 3.68          | 1.94         | 1.17         |
| E Midlands         | Population | <b>2041696</b> | <b>29024</b> | <b>203169</b> | <b>58240</b> | <b>50336</b> |
|                    | Observed   | 357            | <=5          | 46            | <=5          | <=5          |
|                    | Expected   | -              | 1.30         | 16.47         | 4.09         | 3.10         |
| London             | Population | <b>572729</b>  | <b>35563</b> | <b>250599</b> | <b>88759</b> | <b>92085</b> |
|                    | Observed   | 24             | 0            | 24            | 6            | 6            |
|                    | Expected   | -              | 0.88         | 8.15          | 3.33         | 2.07         |
| North East         | Population | <b>565126</b>  | <b>6366</b>  | <b>28299</b>  | <b>11041</b> | <b>13750</b> |
|                    | Observed   | 216            | <=5          | 11            | 0            | <=5          |
|                    | Expected   | -              | 0.51         | 3.86          | 1.68         | 1.13         |
| North West         | Population | <b>1113418</b> | <b>6428</b>  | <b>39869</b>  | <b>9145</b>  | <b>16323</b> |
|                    | Observed   | 354            | <=5          | 21            | 0            | <=5          |
|                    | Expected   | -              | 0.36         | 4.02          | 0.54         | 1.16         |
| South East         | Population | <b>716772</b>  | <b>11697</b> | <b>32274</b>  | <b>13276</b> | <b>22861</b> |
|                    | Observed   | 42             | 0            | <=5           | 0            | <=5          |
|                    | Expected   | -              | 0.15         | 0.54          | 0.16         | 0.31         |
| South West         | Population | <b>1588697</b> | <b>14890</b> | <b>37453</b>  | <b>12999</b> | <b>23906</b> |
|                    | Observed   | 81             | 0            | <=5           | 0            | <=5          |
|                    | Expected   | -              | 0.18         | 0.49          | 0.21         | 0.30         |
| W Midlands         | Population | <b>379409</b>  | <b>13013</b> | <b>104530</b> | <b>40173</b> | <b>18479</b> |
|                    | Observed   | 90             | <=5          | 28            | 11           | <=5          |
|                    | Expected   | -              | 0.79         | 9.96          | 5.30         | 1.23         |
| Yorkshire & Humber | Population | <b>1620066</b> | <b>20028</b> | <b>193702</b> | <b>36544</b> | <b>34428</b> |
|                    | Observed   | 418            | <=5          | 77            | <=5          | <=5          |
|                    | Expected   | -              | 1.30         | 18.32         | 3.62         | 2.00         |

Non-zero counts of <=5 are redacted. Observed = observed COVID-19 deaths. \*Expected = expected COVID-19 deaths if the age, sex and region-specific COVID-19 death rates for the local White population applied (for this analysis, age was grouped as 18-39, 40-49, 50-59, 60-69, 70-79, 80+).



Figure A1: Region-specific Standardised Mortality Ratios (SMR) for COVID-19 death



SMRs calculated for each region using White (in that region) as the reference; SMRs are Observed/Expected where Expected = expected COVID-19 deaths if the age, sex and region-specific COVID-19 death rates for the local White population applied (for this analysis, age was grouped as 18-39, 40-49, 50-59, 60-69, 70-79, 80+).