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

Rohini Mathur*, Christopher T Rentsch*, Caroline E Morton*, William J Hulme, Anna Schultze, Brian MacKenna, Rosalind M Eggo, Krishnan Bhaskaran, Angel Y S Wong, Elizabeth J Williamson, Harriet Forbes, Kevin Wing, Helen I McDonald, Chris Bates, Seb Bacon, Alex J Walker, David Evans, Peter Inglesby, Amir Mehrkar, Helen J Curtis, Nicholas J DeVito, Richard Croker, Henry Drysdale, Jonathan Cockburn, John Parry, Frank Hester, Sam Harper, Ian J Douglas, Laurie Tomlinson, Stephen J W Evans, Richard Grieve, David Harrison, Kathy Rowan, Kamlesh Khunti, Nishi Chaturvedi, Liam Smeeth†, Ben Goldacre†, for the OpenSAFELY Collaborative

Summary

Background COVID-19 has disproportionately affected minority ethnic populations in the UK. Our aim was to quantify ethnic differences in SARS-CoV-2 infection and COVID-19 outcomes during the first and second waves of the COVID-19 pandemic in England.

Methods We conducted an observational cohort study of adults (aged ≥18 years) registered with primary care practices in England for whom electronic health records were available through the OpenSAFELY platform, and who had at least 1 year of continuous registration at the start of each study period (Feb 1 to Aug 3, 2020 [wave 1], and Sept 1 to Dec 31, 2020 [wave 2]). Individual-level primary care data were linked to data from other sources on the outcomes of interest: SARS-CoV-2 testing and positive test results and COVID-19-related hospital admissions, intensive care unit (ICU) admissions, and death. The exposure was self-reported ethnicity as captured on the primary care record, grouped into five high-level census categories (White, South Asian, Black, other, and mixed) and 16 subcategories across these five categories, as well as an unknown ethnicity category. We used multivariable Cox regression to examine ethnic differences in the outcomes of interest. Models were adjusted for age, sex, deprivation, clinical factors and comorbidities, and household size, with stratification by geographical region.

Findings Of 17 288 532 adults included in the study (excluding care home residents), 10 877 978 (62·9%) were White, 1 025 319 (5·9%) were South Asian, 340 912 (2·0%) were Black, 170 484 (1·0%) were of mixed ethnicity, 320 788 (1·9%) were of other ethnicity, and 4 553 051 (26·3%) were of unknown ethnicity. In wave 1, the likelihood of being tested for SARS-CoV-2 infection was slightly higher in the South Asian group (adjusted hazard ratio 1·08 [95% CI 1·07–1·09]), Black group (1·08 [1·06–1·09]), and mixed ethnicity group (1·04 [1·02–1·05]) and was decreased in the other ethnicity group (0·77 [0·76–0·78]) relative to the White group. The risk of testing positive for SARS-CoV-2 infection was higher in the South Asian group (1·99 [1·94–2·04]), Black group (1·69 [1·62–1·77]), mixed ethnicity group (1·49 [1·39–1·59]), and other ethnicity group (1·20 [1·14–1·28]). Compared with the White group, the four remaining high-level ethnic groups had an increased risk of COVID-19-related hospitalisation (South Asian group 1·48 [1·41–1·55], Black group 1·78 [1·67–1·90], mixed ethnicity group 1·63 [1·45–1·83], other ethnicity group 1·54 [1·41–1·69]), COVID-19-related ICU admission (2·18 [1·92–2·48], 3·12 [2·65–3·67], 2·96 [2·26–3·87], 3·18 [2·58–3·93]), and death (1·26 [1·15–1·37], 1·51 [1·31–1·71], 1·41 [1·11–1·81], 1·22 [1·00–1·48]). In wave 2, the risks of hospitalisation, ICU admission, and death relative to the White group were increased in the South Asian group but attenuated for the Black group compared with these risks in wave 1. Disaggregation into 16 ethnicity groups showed important heterogeneity within the five broader categories.

Interpretation Some minority ethnic populations in England have excess risks of testing positive for SARS-CoV-2 and of adverse COVID-19 outcomes compared with the White population, even after accounting for differences in sociodemographic, clinical, and household characteristics. Causes are likely to be multifactorial, and delineating the exact mechanisms is crucial. Tackling ethnic inequalities will require action across many fronts, including reducing structural inequalities, addressing barriers to equitable care, and improving uptake of testing and vaccination.

Funding Medical Research Council.

Department of
Non-Communicable Disease Epidemiology (R Mathur PhD, C T Rentsch PhD, A Schultze PhD, Prof K Bhaskaran PhD, A Y S Wong PhD, H Forbes PhD, Prof I J Douglas PhD, L Tomlinson PhD, Prof L Smeeth FMedSci), Department of Medical Statistics (Prof E J Williamson PhD, Prof S J W Evans MSc), Department of Infectious Disease Epidemiology (R M Eggo PhD, H I McDonald PhD), and Department of Health Services Research and Policy (Prof R Grieve PhD), London School of Hygiene & Tropical Medicine, London, UK; The DataLab, Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford, UK (C E Morton MRCGP, W J Hulme PhD, B MacKenna MPharm, K Wing PhD, S Bacon BA, A J Walker PhD, D Evans MPhil, P Inglesby MPhil, A Mehrkar MB, H J Curtis DPhil, N J DeVito MPH, R Croker MSc, H Drysdale MRCS, B Goldacre MRCPsych); National Institute for Health Research Health Protection Research Unit in Immunisation, London,

UK (H I McDonald); **TPP, Leeds,
UK** (C Bates PhD,
J Cockburn BSc, J Parry MRCGP,
F Hester BSc, S Harper MSci);
**Intensive Care National Audit
and Research Centre, London,
UK** (D Harrison PhD,
K Rowan PhD); **Diabetes
Research Centre, University of
Leicester, Leicester, UK**
(Prof K Khunti FMedSci); **Medical
Research Council Unit for
Lifelong Health and Ageing,
University College London,
London, UK**
(Prof N Chaturvedi MD)

Correspondence to:
Dr Rohini Mathur, Department of
Non-Communicable Disease
Epidemiology, London School of
Hygiene & Tropical Medicine,
London WC1E 7HT, UK
rohini.mathur@lshtm.ac.uk

Research in context

Evidence before this study

We searched PubMed for population-based studies examining
the association between ethnicity and COVID-19. Keywords
included (ethnic* OR race) AND (COVID OR coronavirus OR
SARS-CoV-2) AND (UK or England) AND (risk OR rate OR odds).
Results were filtered to human studies published from
2019 onwards with abstracts available. We identified six studies
examining ethnic differences in SARS-CoV-2 infection and
COVID-19 outcomes in population-based samples. Five studies
from the UK Biobank reported an increased risk of SARS-CoV-2
infection and COVID-19-related hospitalisation in Black and
South Asian groups compared with White groups. As the UK
Biobank cohort is known to be healthier, less deprived, and less
ethnically diverse than the UK population, these findings are
not wholly generalisable to the wider UK population. Our
previous study using the OpenSAFELY platform showed an
increased risk of COVID-19-related death in minority ethnic
groups, but did not examine the role of household size or
examine ethnic differences in SARS-CoV-2 infection and
COVID-19 outcomes earlier in the care pathway.

Added value of this study

This is the largest study in the UK to examine ethnic inequalities
in testing positive for SARS-CoV-2 and in COVID-19-related
outcomes in a cohort covering 40% of the population in England.

Additionally, it is the only population-representative study to
date that accounts for household size in addition to
sociodemographic characteristics and clinical comorbidities.
By examining ethnicity according to both high-level and detailed
ethnic groupings, we have shown important ethnic differences in
the risk of testing positive for SARS-CoV-2 and the risks of
COVID-19-related hospital admission, intensive care unit
admissions, and death. We showed that multiple factors
contribute to ethnic inequalities in COVID-19 and the importance
of these factors varies by ethnic group. Compared with wave 1,
the risks of COVID-19-related hospitalisation and death in wave 2
were increased for South Asian groups and reduced in all other
ethnic minority groups relative to the White group.

Implications of all the available evidence

The risks of SARS-CoV-2 infection and severe COVID-19
outcomes are disproportionately increased in minority ethnic
groups, both in the UK and internationally. Reducing ethnic
inequalities in COVID-19 risks requires action on social
determinants including addressing disadvantage and
discrimination, reducing risk of infection and transmission,
improving quality of and access to quality clinical care, and
improving management of pre-existing clinical conditions.
The appropriate balance of these actions needs tailoring for
different ethnic groups.

Introduction

The risks of SARS-CoV-2 infection and COVID-19 disease
have been reported to be disproportionately increased in
minority ethnic groups compared with White groups[1–4] in
the UK and in other countries, including the USA[5] and
Brazil.[6] It is hypothesised that these differences might be
driven by factors such as living in deprived areas; working
in high-exposure or front-line occupations; living in
large, multigenerational households; a higher burden of
underlying conditions; discrimination; and poor access
to health or community services.[7–10]

In the UK, the collection of ethnicity data is con-
sidered essential for identifying and reducing ethnic
inequalities.[11,12] Although there is no single universally
accepted definition of ethnicity, it serves as an important
social construct and surrogate marker for shared
exposures or risks for people with similar social, biological,
and cultural characteristics.[13–15]

To date, many studies on COVID-19 have reported
findings according to broad ethnic categories—such as
White, South Asian, and Black—rather than considering
disaggregated groupings. Furthermore, most evidence has
been derived from populations with severe disease
requiring hospitalisation, making it difficult to extrapolate
findings to the general population.[16–21] Finally, although
previous studies have accounted for health status,
socioeconomic deprivation, or household composition,
none yet have considered these factors in conjunction.[22,23]

The aim of this study was to estimate the effect of
ethnicity on being tested and testing positive for
SARS-CoV-2, and on COVID-19-related hospitalisation,
intensive care unit (ICU) admission, and mortality,
recognising the potential role of sociodemographic,
clinical, and household-related factors.

Methods

Study design and population

We did a population-based, observational cohort study
using the OpenSAFELY platform, for which National
Health Service (NHS) England is the data controller.
OpenSAFELY holds electronic health record data for
24 million people registered with primary care practices
using TPP software, representing around 40% of the
population of England.

Individual-level primary care data were linked to
SARS-CoV-2 testing data from the Second Generation
Surveillance System, COVID-19-related hospital admis-
sions from the Secondary Uses Service, COVID-19-related
ICU admissions from the Intensive Care National Audit &
Research Centre,[24] and mortality data from the Office for
National Statistics (ONS). The study population comprised
adults aged 18 years and older who were registered with a
primary care practice on Feb 1, 2020. The study periods
were from Feb 1 to Aug 3, 2020 (for wave 1), and from
Sept 1 to Dec 31, 2020 (for wave 2). A minimum of
12 months of continuous registration before the start date

of each wave was required for inclusion, to ensure that baseline factors were adequately captured. Individuals residing in care homes were excluded from the main analyses because we hypothesised that the role of sociodemographic, clinical, and household characteristics would be systematically different for care home residents than for the general population.

This study was approved by the Health Research Authority (REC reference 20/LO/0651) and by the London School of Hygiene & Tropical Medicine's Ethics Board (reference 21863). Patient data were pseudonymised for analysis and linkage using industry standard cryptographic hashing techniques, and all pseudonymised datasets transmitted for linkage onto OpenSAFELY were encrypted. Only aggregate statistical outputs could leave the TPP 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 2016.

Exposures

The primary exposure was self-reported ethnicity as captured on the primary care record, collapsed into the five high-level and 16 detailed census categories of White (White British, White Irish, other White), South Asian (Indian, Pakistani, Bangladeshi, other South Asian), Black (African, Caribbean, other Black), other (Chinese, all others), and mixed (White and Asian, White and African, White and Caribbean, other mixed). An unknown ethnicity category was also included. Comparisons were reported for the five high-level ethnic groups with the White group as the reference category, and for the 16 disaggregated groups, with the White British group as the reference category.

Outcomes

Outcomes of interest included receiving a PCR test for SARS-CoV-2, testing positive for SARS-CoV-2, and COVID-19-related hospital admission, ICU admission, and death (defined as the presence of ICD-10 codes U07.1 [confirmed COVID-19] or U07.2 [suspected COVID-19] anywhere on the death certificate). Testing outcomes were obtained from the UK's pillar 1 (NHS and Public Health England laboratories) and pillar 2 (commercial partners) testing strategies and included results from PCR swab tests used to identify symptomatic individuals.[25,26]

Covariates

Demographic characteristics included age, sex, deprivation, household size, number of primary care consultations in the past 12 months, and geographical region (defined by the sustainability and transformation partnership [STP], an NHS administrative area). Deprivation was defined using quintiles of the Index of

Multiple Deprivation, an area-level composite measure of seven domains: income, employment, education, skills and training, health and disability, crime, and barriers to housing services and living environment.[27] Household size (categorised as 1–2, 3–5, 6–10, or ≥11 people) was determined using the number of individuals of all ages in OpenSAFELY residing at the same address on Feb 1, 2020.

Clinical covariates were identified using the Read clinical classification system[28] and included body-mass index (BMI), glycated haemoglobin (HbA_{1c}), blood pressure, and smoking status. BMI in kg/m^2 was grouped into six categories using the WHO classification, with adjustments for South Asian ethnicity: underweight ($<18 \cdot 5$ kg/m^2), normal weight ($18 \cdot 5$–$24 \cdot 9$ kg/m^2 [or $18 \cdot 5$–$23 \cdot 5$ kg/m^2 if South Asian]), overweight ($25 \cdot 0$–$29 \cdot 9$ kg/m^2 [$23 \cdot 6$–$27 \cdot 5$ kg/m^2]); obese I ($30 \cdot 0$–$34 \cdot 9$ kg/m^2 [$27 \cdot 5$–$32 \cdot 4$ kg/m^2]); obese II ($35 \cdot 0$–$39 \cdot 9$ kg/m^2 [$32 \cdot 5$–$37 \cdot 4$ kg/m^2]); and obese III (≥40 kg/m^2 [$\geq37 \cdot 5$ kg/m^2]). HbA_{1c} was grouped into five categories ($<6 \cdot 5\%$, $6 \cdot 5$–$7 \cdot 4\%$, $7 \cdot 5$–$7 \cdot 9\%$, $8 \cdot 0$–$8 \cdot 9\%$, and $\geq9 \cdot 0\%$). Blood pressure was grouped into four categories: normal ($<120/<80$); elevated (120 to $130/<80$); high, stage I (131 to $<140/80$ to <90); and high, stage II ($\geq140/\geq90$). Smoking status was grouped into current, former, and never smokers. Those with missing smoking status were categorised as never smokers. Those with missing BMI, HbA_{1c}, or blood pressure data were grouped into a separate category of "unknown".

Clinical comorbidities were considered present at baseline if they were recorded at any time before Feb 1, 2020 (for wave 1) or Sept 1, 2020 (for wave 2). Comorbidities included the following: hypertension, asthma, chronic respiratory disease, chronic heart disease, type 1 and type 2 diabetes, cancer, chronic liver disease, stroke, dementia, other chronic neurological diseases, chronic kidney disease (defined as an estimated glomerular filtration rate <60 mL/min per $1 \cdot 73$ m^2), end-stage renal failure, common autoimmune diseases (rheumatoid arthritis, systemic lupus erythematosus, or psoriasis), and immunosuppression (HIV infection, sickle cell disease, organ transplant, or asplenia). All codelists are available for review and reuse.[29]

Statistical analysis

Sociodemographic and clinical characteristics at baseline were summarised using descriptive statistics, stratified by ethnic group. Follow-up began on Feb 1, 2020, for wave 1 and on Sept 1, 2020, for wave 2, and ended at the earliest occurrence of the outcome of interest, death, deregistration from a primary care practice, or the censoring date for the dataset capturing the outcome of interest (from July 30 to Aug 3, 2020, [for wave 1] or to Dec 31, 2020 [for wave 2]).

Multivariable Cox proportional hazards regression was used to estimate ethnic differences in the cause-specific hazard of each outcome in the whole denominator
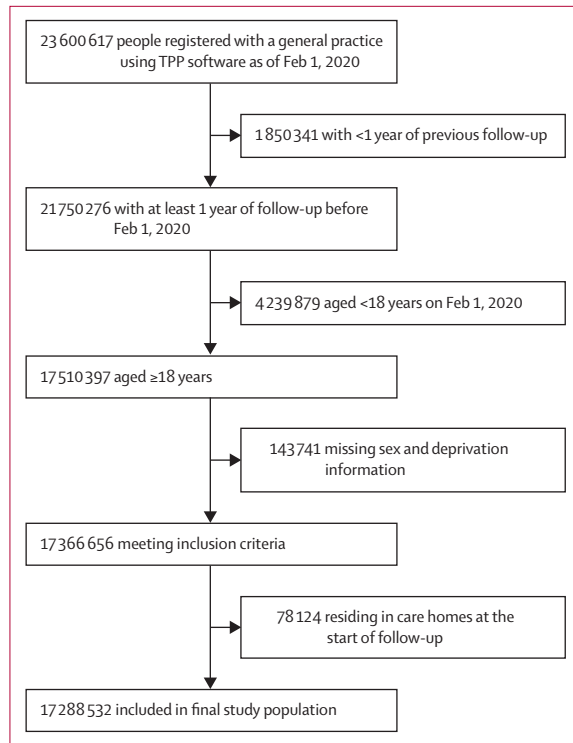
Figure 1: **Study population flowchart**

23 600 617 people registered with a general practice using TPP software as of Feb 1, 2020

1 850 341 with <1 year of previous follow-up

21 750 276 with at least 1 year of follow-up before Feb 1, 2020

4 239 879 aged <18 years on Feb 1, 2020

17 510 397 aged ≥18 years

143 741 missing sex and deprivation information

17 366 656 meeting inclusion criteria

78 124 residing in care homes at the start of follow-up

17 288 532 included in final study population

See **Online** for appendix

population.[30] All analyses were adjusted for age (using restricted cubic splines), sex, deprivation quintile, diagnosed comorbidities, BMI, HbA$_{1c}$, blood pressure, number of primary care consultations in the preceding 12 months, and household size. To investigate the extent to which age-sex-adjusted ethnicity-associated differences could further be explained by deprivation, comorbidities, and household size, we sequentially adjusted for age and sex in the first model, adding deprivation in the second model, comorbidities, clinical factors, and primary care consultations in the third model, and household size in the fourth model. All models were stratified by STP to account for clustering by geographical region. All analyses were done separately for wave 1 and wave 2. Analyses of COVID-19-related hospital admissions for wave 1 were added during the revision stage of this Article as data on hospitalisations became available in OpenSAFELY after the initial submission; therefore, cohort sizes slightly differ for this outcome compared with other outcomes.

In secondary analyses, we estimated ethnic differences in the risk of non-COVID-19 death (defined as any death without a COVID-19 diagnosis code anywhere on the death certificate). Additionally, we used logistic regression adjusting for all covariates to examine ethnic differences in the odds of testing positive among those tested for SARS-CoV-2. We also estimated differences between ethnic groups in all outcomes for care home residents, adjusting for all covariates except for household size.

In sensitivity analyses, we used multiple imputation to account for missing ethnicity data, examining differences between ethnic groups in the risk of death where COVID-19 was the underlying cause (rather than any cause) and exploring the effect of regional variation on ethnic differences in all outcomes. Proportional hazards assumptions were assessed by testing for a zero slope in the scaled Schoenfeld residuals and graphical inspection of Kaplan-Meier plots.

Data management was done with Python 3.8 and SQL, and analysis was done using Stata 16.

Role of the funding source
The funders of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

Results
From a total of 23 600 617 individuals in the OpenSAFELY database on Feb 1, 2020, 17 288 532 adults aged 18 years or older who met the selection criteria were included in the study (figure 1). The ethnic breakdown of the study population was 10 877 978 (62·9%) people of White ethnicity, 1 025 319 (5·9%) of South Asian ethnicity, 340 912 (2·0%) of Black ethnicity, 320 788 (1·9%) of other ethnicity, 170 484 (1·0%) of mixed ethnicity, and 4 553 051 (26·3%) of unknown ethnicity. Compared with the White population, minority ethnic groups were younger and over-represented in deprived neighbourhoods, large households, and diabetic populations (table 1; appendix p 3).

Between Feb 1 and Aug 3, 2020 (wave 1), 1 216 801 (7·0%) people in the study population received a test for SARS-CoV-2 infection, and 71 246 (0·4%) tested positive (table 2). The ethnic breakdown of individuals who received a test was similar to that of the general population, although test recipients were slightly older with more comorbid chronic conditions than the general population (appendix p 4). After accounting for all measured explanatory variables (age, sex, deprivation, comorbidities, clinical factors, primary care consultations in the preceding 12 months, and household size, with stratification by STP region), South Asian, Black, and mixed ethnicity groups were more likely to be tested (South Asian group adjusted hazard ratio [HR] 1·08 [95% CI 1·07–1·09], Black group 1·08 [1·06–1·09], mixed ethnicity group 1·04 [1·02–1·05]) and to test positive for SARS-CoV-2 (1·99 [1·94–2·04], 1·69 [1·62–1·77], 1·49 [1·39–1·59]) compared with the White ethnic group (figure 2). Across the 16 subcategories of ethnicity, risks of testing positive were similar to those of the respective high-level categories, except for the Chinese group, for whom risks of being tested and testing positive (0·49 [0·42–0·58]) were lower than for the White British group. When restricted to the population who had ever received a test (as opposed to the whole denominator population used in the primary analysis), patterns by

ethnic group remained unchanged, except for the Chinese group, who had an equivalent risk of testing positive (odds ratio [OR] 1·13 [95% CI 0·95–1·34], adjusted for all explanatory variables; appendix p 15).

Between Feb 1 and Aug 3, 2020, 32 473 (0·2%) individuals in the study population were admitted to hospital for COVID-19, 3096 (<0·1%) were admitted to ICU for COVID-19, and there were 11 649 (0·1%) COVID-19-related deaths (table 2). After accounting for all measured explanatory factors, risk of hospitalisation was increased, relative to the White reference groups, in all minority ethnic groups (South Asian group adjusted HR 1·48 [95% CI 1·41–1·55], Black group 1·78 [1·67–1·90], mixed ethnicity group 1·63 [1·45–1·83], other ethnicity group 1·54 [1·41–1·69]), including all minority ethnic subcategories except the Chinese group (0·97 [0·77–1·23]; figure 3A). Risk of ICU admission was around 2–3 times in the four broad minority ethnic groups (South Asian group 2·18 [1·92–2·48], Black group 3·12 [2·65–3·67], mixed ethnicity group 2·96 [2·26–3·87], other ethnicity group 3·18 [2·58–3·93]) relative to the White reference group, and around 2–5 times higher among South Asian, Black, mixed, and other ethnic subcategories relative to the White British group (figure 3B). Risk of COVID-19-related death was increased by 22–51% in the four broad minority ethnic groups relative to the White group (South Asian group 1·26 [1·15–1·37], Black group 1·51 [1·31–1·71], mixed ethnicity group 1·41 [1·11–1·81], other ethnicity group 1·22 [1·00–1·48]; figure 4).

After accounting for age and sex, further adjustment had little effect on the likelihood of being tested for COVID-19. In South Asian groups, adjustment for clinical characteristics led to the largest reduction in HRs for testing positive for SARS-CoV-2 and hospitalisation and ICU admission due to COVID-19, and adjustment for deprivation and household size made equivalent reductions in HRs for COVID-19-related mortality. In all other minority ethnic groups, adjustment for social deprivation led to the largest reduction in HRs for all outcomes after accounting for age and sex (table 2; appendix p 5).

Between Sept 1 and Dec 31, 2020 (wave 2), 2 647 756 (15·3%) individuals in the study population received a test for SARS-CoV-2, 506 773 (2·9%) tested positive, 18 885 (0·1%) were admitted to hospital for COVID-19, 3351 (<0·1%) had a COVID-19 related ICU admission, and there were 7366 (<0·1%) COVID-19-related deaths. In contrast to wave 1, all four of the broad minority ethnic groups were less likely to be tested than the White group, and this pattern was also seen in all minority ethnic subcategories relative to the White British subcategory, with the exception of the Pakistani group (figure 2). The South Asian group (and each of its subgroups) remained at higher risk of testing positive (adjusted HR 1·32 [95% CI 1·31–1·33]) than the White group, and the risks of COVID-19-related hospitalisation (1·89 [1·79–2·00]), ICU admission (2·68 [2·39–3·01]), and death (1·87 [1·68–2·07]) were relatively greater in magnitude in wave 2 than in wave 1 (figures 2–4). In contrast to wave 1, the Black group (and all Black subgroups) was less likely than the White group to test positive (0·85 [0·84–0·87]; figure 2), although the risk of testing positive remained higher among those ever tested (OR 1·03 [95% CI 1·02–1·06] adjusted for all

	Overall (n=17 288 532)	White (n=10 877 978)	South Asian (n=1 025 319)	Black (n=340 912)	Mixed (n=170 484)	Other (n=320 788)	Unknown (n=4 553 051)
Age, years	49·6 (18·6)	51·2 (18·2)	43·0 (15·4)	43·8 (15·5)	40·0 (14·9)	40·2 (15·0)	48·8 (20·0)
Sex							
Female	8 638 780 (50·0%)	5 650 496 (51·9%)	490 641 (47·9%)	169 832 (49·8%)	87 678 (51·4%)	157 832 (49·2%)	2 082 301 (45·7%)
Male	8 649 752 (50·0%)	5 227 482 (48·1%)	534 678 (52·1%)	171 080 (50·2%)	82 806 (48·6%)	162 956 (50·8%)	2 470 750 (54·3%)
Index of Multiple Deprivation							
1 (most affluent)	3 495 006 (20·2%)	2 294 962 (21·1%)	87 798 (8·6%)	22 716 (6·7%)	22 259 (13·1%)	43 209 (13·5%)	1 024 062 (22·5%)
2	3 481 686 (20·1%)	2 292 018 (21·1%)	114 552 (11·2%)	34 646 (10·2%)	27 255 (16·0%)	54 638 (17·0%)	958 577 (21·1%)
3	3 487 259 (20·2%)	2 220 520 (20·4%)	189 485 (18·5%)	54 731 (16·1%)	32 256 (18·9%)	60 843 (19·0%)	929 424 (20·4%)
4	3 482 514 (20·1%)	2 113 382 (19·4%)	295 273 (28·8%)	90 963 (26·7%)	40 543 (23·8%)	79 661 (24·8%)	862 692 (18·9%)
5 (most deprived)	3 342 067 (19·3%)	1 957 096 (18·0%)	338 211 (33·0%)	137 856 (40·4%)	48 171 (28·3%)	82 437 (25·7%)	778 296 (17·1%)
Number of people in household							
1–2	7 533 408 (43·6%)	5 169 700 (47·5%)	209 599 (20·4%)	118 164 (34·7%)	62 417 (36·6%)	118 594 (37·0%)	1 854 934 (40·7%)
3–5	6 166 295 (35·7%)	3 762 179 (34·6%)	430 932 (42·0%)	135 543 (39·8%)	69 564 (40·8%)	117 587 (36·7%)	1 650 490 (36·3%)
6–10	992 375 (5·7%)	443 780 (4·1%)	221 312 (21·6%)	50 678 (14·9%)	18 183 (10·7%)	39 841 (12·4%)	218 581 (4·8%)
≥11	169 045 (1·0%)	75 056 (0·7%)	41 039 (4·0%)	6486 (1·9%)	3475 (2·0%)	14 228 (4·4%)	28 761 (0·6%)
Unknown	2 427 409 (14·0%)	1 427 263 (13·1%)	122 437 (11·9%)	30 041 (8·8%)	16 845 (9·9%)	30 538 (9·5%)	800 285 (17·6%)
Number of general practitioner consultations in the 12 months before baseline	3·0 (0·0–8·0)	4·0 (1·0–8·0)	3·0 (0·0–8·0)	3·0 (0·0–7·0)	2·0 (0·0–7·0)	1·0 (0·0–5·0)	2·0 (0·0–7·0)

(Table 1 continues on next page)

	Overall (n=17 288 532)	White (n=10 877 978)	South Asian (n=1 025 319)	Black (n=340 912)	Mixed (n=170 484)	Other (n=320 788)	Unknown (n=4 553 051)
(Continued from previous page)							
Smoking status							
Never*	8 653 213 (50·1%)	4 882 899 (44·9%)	751 034 (73·2%)	232 984 (68·3%)	96 363 (56·5%)	208 450 (65·0%)	2 481 483 (54·5%)
Former	5 683 564 (32·9%)	4 037 473 (37·1%)	159 103 (15·5%)	63 290 (18·6%)	39 798 (23·3%)	58 768 (18·3%)	1 325 132 (29·1%)
Current	2 951 755 (17·1%)	1 957 606 (18·0%)	115 182 (11·2%)	44 638 (13·1%)	34 323 (20·1%)	53 570 (16·7%)	746 436 (16·4%)
BMI, kg/m²	27·4 (5·7)	27·5 (5·7)	26·6 (5·1)	28·0 (5·7)	26·8 (5·7)	25·2 (5·1)	27·7 (5·9)
BMI category, with adjustment for South Asian populations							
Underweight	294 735 (1·7%)	180 046 (1·7%)	26 124 (2·5%)	5012 (1·5%)	3800 (2·2%)	11 844 (3·7%)	67 909 (1·5%)
Normal	4 571 011 (26·4%)	3 234 076 (29·7%)	176 170 (17·2%)	81 236 (23·8%)	53 668 (31·5%)	120 216 (37·5%)	905 645 (19·9%)
Overweight	4 666 217 (27·0%)	3 218 869 (29·6%)	313 478 (30·6%)	97 270 (28·5%)	42 901 (25·2%)	71 463 (22·3%)	922 236 (20·3%)
Obese I	2 457 922 (14·2%)	1 633 030 (15·0%)	212 808 (20·8%)	54 909 (16·1%)	20 662 (12·1%)	27 228 (8·5%)	509 285 (11·2%)
Obese II	949 367 (5·5%)	631 273 (5·8%)	71 428 (7·0%)	21 383 (6·3%)	7825 (4·6%)	8355 (2·6%)	209 103 (4·6%)
Obese III	474 090 (2·7%)	318 199 (2·9%)	28 812 (2·8%)	10 282 (3·0%)	4076 (2·4%)	3505 (1·1%)	109 216 (2·4%)
Unknown	3 875 190 (22·4%)	1 662 485 (15·3%)	196 499 (19·2%)	70 820 (20·8%)	37 552 (22·0%)	78 177 (24·4%)	1 829 657 (40·2%)
HbA$_{1c}$, %	5·9 (1·0)	5·8 (1·0)	6·1 (1·2)	6·0 (1·2)	5·9 (1·1)	5·8 (1·0)	5·9 (1·0)
HbA$_{1c}$, mmol/mol	41·0 (92·7)	40·4 (62·2)	44·3 (218·2)	42·8 (138·9)	40·5 (12·0)	40·2 (10·7)	41·4 (101·9)
HbA$_{1c}$ category							
<6·5%	6 706 373 (38·8%)	4 546 662 (41·8%)	424 221 (41·4%)	127 134 (37·3%)	54 728 (32·1%)	92 552 (28·9%)	1 461 076 (32·1%)
6·5–7·4%	582 059 (3·4%)	350 237 (3·2%)	64 385 (6·3%)	15 294 (4·5%)	4500 (2·6%)	7680 (2·4%)	139 963 (3·1%)
7·5–7·9%	155 580 (0·9%)	94 158 (0·9%)	17 117 (1·7%)	3376 (1·0%)	1144 (0·7%)	1825 (0·6%)	37 960 (0·8%)
8·0–8·9%	168 963 (1·0%)	103 062 (0·9%)	18 312 (1·8%)	3558 (1·0%)	1226 (0·7%)	1965 (0·6%)	40 840 (0·9%)
≥9·0%	190 305 (1·1%)	115 910 (1·1%)	21 060 (2·1%)	5415 (1·6%)	1686 (1·0%)	2215 (0·7%)	44 019 (1·0%)
Unknown	9 485 252 (54·9%)	5 667 949 (52·1%)	480 224 (46·8%)	186 135 (54·6%)	107 200 (62·9%)	214 551 (66·9%)	2 829 193 (62·1%)
Blood pressure (mm Hg)							
Systolic	128·0 (15·7)	128·3 (15·5)	123·8 (15·4)	127·1 (16·0)	124·0 (15·3)	122·5 (15·4)	129·0 (15·9)
Diastolic	76·8 (9·9)	76·8 (9·8)	76·5 (9·7)	78·0 (10·3)	76·6 (10·1)	75·8 (10·0)	76·7 (10·0)
Blood pressure category							
Normal	2 653 823 (15·4%)	1 719 408 (15·8%)	211 095 (20·6%)	56 786 (16·7%)	32 711 (19·2%)	57 475 (17·9%)	576 348 (12·7%)
Elevated	1 818 544 (10·5%)	1 225 760 (11·3%)	99 291 (9·7%)	32 093 (9·4%)	15 713 (9·2%)	23 598 (7·4%)	422 089 (9·3%)
High, stage I	4 270 604 (24·7%)	2 880 135 (26·5%)	238 962 (23·3%)	80 605 (23·6%)	35 154 (20·6%)	53 458 (16·7%)	982 290 (21·6%)
High, stage II	3 010 178 (17·4%)	2 025 395 (18·6%)	120 401 (11·7%)	54 017 (15·8%)	19 008 (11·1%)	26 433 (8·2%)	764 924 (16·8%)
Unknown	5 535 383 (32·0%)	3 027 280 (27·8%)	355 570 (34·7%)	117 411 (34·4%)	67 898 (39·8%)	159 824 (49·8%)	1 807 400 (39·7%)
Comorbidities							
Type 1 diabetes	88 294 (0·5%)	59 940 (0·6%)	2508 (0·2%)	1363 (0·4%)	637 (0·4%)	637 (0·2%)	23 209 (0·5%)
Type 2 diabetes	1 234 858 (7·1%)	747 798 (6·9%)	135 741 (13·2%)	31 590 (9·3%)	9603 (5·6%)	15 527 (4·8%)	294 599 (6·5%)
Diagnosed hypertension	3 703 816 (21·4%)	2 437 571 (22·4%)	173 237 (16·9%)	69 524 (20·4%)	20 712 (12·1%)	31 706 (9·9%)	971 066 (21·3%)
Chronic heart disease	1 193 155 (6·9%)	803 572 (7·4%)	53 903 (5·3%)	11 092 (3·3%)	4731 (2·8%)	8136 (2·5%)	311 721 (6·8%)
Stroke	368 707 (2·1%)	249 742 (2·3%)	12 429 (1·2%)	4373 (1·3%)	1495 (0·9%)	2161 (0·7%)	98 507 (2·2%)
Chronic kidney disease†	978 300 (5·7%)	647 085 (5·9%)	27 979 (2·7%)	15 042 (4·4%)	3787 (2·2%)	4782 (1·5%)	279 625 (6·1%)
End-stage renal failure	25 348 (0·1%)	14 615 (0·1%)	2410 (0·2%)	878 (0·3%)	219 (0·1%)	352 (0·1%)	6874 (0·2%)
Cancer	979 433 (5·7%)	684 824 (6·3%)	19 868 (1·9%)	9249 (2·7%)	3718 (2·2%)	6640 (2·1%)	255 134 (5·6%)
Autoimmune disease	889 832 (5·1%)	615 132 (5·7%)	37 348 (3·6%)	6912 (2·0%)	4988 (2·9%)	7504 (2·3%)	217 948 (4·8%)
Immunosuppression	93 162 (0·5%)	54 143 (0·5%)	4602 (0·4%)	10 087 (3·0%)	2118 (1·2%)	1425 (0·4%)	20 787 (0·5%)
Chronic liver disease	104 781 (0·6%)	69 901 (0·6%)	5677 (0·6%)	3035 (0·9%)	894 (0·5%)	2445 (0·8%)	22 829 (0·5%)
Dementia	34 169 (0·2%)	22 855 (0·2%)	1199 (0·1%)	538 (0·2%)	106 (0·1%)	195 (0·1%)	9276 (0·2%)
Neurological disease	169 483 (1·0%)	116 300 (1·1%)	5803 (0·6%)	1951 (0·6%)	888 (0·5%)	1163 (0·4%)	43 378 (1·0%)
Asthma	2 663 321 (15·4%)	1 790 975 (16·5%)	125 745 (12·3%)	37 351 (11·0%)	25 776 (15·1%)	22 406 (7·0%)	661 068 (14·5%)
Chronic respiratory disease	718 047 (4·2%)	527 265 (4·8%)	17 435 (1·7%)	5003 (1·5%)	2274 (1·3%)	3520 (1·1%)	162 550 (3·6%)

Data are mean (SD), n (%), or median (IQR). BMI=body-mass index. HbA$_{1c}$=glycated haemoglobin. *Including those with missing data. †Defined as estimated glomerular filtration rate <60 mL/min per 1·73 m².

Table 1: **Baseline characteristics by ethnic group**

	Denominator	Events	Hazard ratio (95% CI)				
			Crude	Age-sex adjusted	Plus deprivation	Plus comorbidities and clinical factors*	Plus household size
Tested for SARS-CoV-2 (n=1 216 801)							
White	10 877 978	793 181 (7·3%)	1 (ref)	1 (ref)	1 (ref)	1 (ref)	1 (ref)
South Asian	1 025 319	82 647 (8·1%)	1·08 (1·08–1·09)	1·11 (1·10–1·12)	1·09 (1·08–1·09)	1·09 (1·08–1·09)	1·08 (1·07–1·09)
Black	340 912	25 305 (7·4%)	1·04 (1·03–1·05)	1·08 (1·06–1·09)	1·04 (1·03–1·06)	1·08 (1·06–1·09)	1·08 (1·06–1·09)
Mixed	170 484	12 126 (7·1%)	1·00 (0·98–1·01)	1·03 (1·01–1·05)	1·01 (1·00–1·03)	1·04 (1·02–1·06)	1·04 (1·02–1·05)
Other	320 788	15 824 (4·9%)	0·68 (0·67–0·69)	0·71 (0·69–0·72)	0·70 (0·69–0·71)	0·78 (0·77–0·79)	0·77 (0·76–0·78)
Unknown	4 553 051	287 718 (6·3%)	0·86 (0·86–0·87)	0·88 (0·88–0·89)	0·89 (0·88–0·89)	0·97 (0·97–0·98)	0·97 (0·97–0·98)
Tested positive for SARS-CoV-2 (n=71 246)							
White	10 877 978	41 180 (0·4%)	1 (ref)	1 (ref)	1 (ref)	1 (ref)	1 (ref)
South Asian	1 025 319	9679 (0·9%)	2·38 (2·32–2·43)	2·64 (2·57–2·70)	2·45 (2·39–2·51)	2·16 (2·10–2·21)	1·99 (1·94–2·04)
Black	340 912	2286 (0·7%)	1·82 (1·74–1·90)	2·04 (1·95–2·13)	1·86 (1·78–1·94)	1·74 (1·67–1·82)	1·69 (1·62–1·77)
Mixed	170 484	840 (0·5%)	1·37 (1·28–1·46)	1·59 (1·48–1·70)	1·52 (1·42–1·63)	1·51 (1·41–1·62)	1·49 (1·39–1·59)
Other	320 788	1213 (0·4%)	1·06 (1·00–1·12)	1·22 (1·15–1·29)	1·18 (1·11–1·25)	1·25 (1·18–1·33)	1·20 (1·14–1·28)
Unknown	4 553 051	16 048 (0·4%)	0·97 (0·95–0·99)	0·99 (0·98–1·01)	1·00 (0·98–1·02)	1·06 (1·04–1·08)	1·06 (1·04–1·08)
COVID-19-related hospital admission† (n=32 473)							
White	11 110 312	20 504 (0·2%)	1 (ref)	1 (ref)	1 (ref)	1 (ref)	1 (ref)
South Asian	1 026 551	2836 (0·3%)	1·26 (1·21–1·31)	2·04 (1·96–2·13)	1·83 (1·76–1·91)	1·59 (1·52–1·66)	1·48 (1·41–1·55)
Black	342 561	1051 (0·3%)	1·39 (1·30–1·48)	2·18 (2·05–2·33)	1·89 (1·77–2·01)	1·81 (1·70–1·94)	1·78 (1·67–1·90)
Mixed	171 933	302 (0·2%)	0·86 (0·77–0·97)	1·78 (1·59–1·99)	1·64 (1·47–1·84)	1·65 (1·47–1·85)	1·63 (1·45–1·83)
Other	323 529	504 (0·2%)	0·77 (0·71–0·85)	1·52 (1·39–1·66)	1·43 (1·31–1·57)	1·59 (1·45–1·74)	1·54 (1·41–1·69)
Unknown	4 408 386	7276 (0·2%)	0·91 (0·88–0·93)	0·96 (0·94–0·99)	0·97 (0·94–0·99)	1·06 (1·03–1·09)	1·06 (1·03–1·09)
COVID-19-related intensive care unit admission (n=3096)							
White‡	10 877 978	1700 (<0·1%)	1 (ref)	1 (ref)	1 (ref)	1 (ref)	1 (ref)
South Asian	1 025 319	410 (<0·1%)	2·38 (2·12–2·67)	3·30 (2·93–3·71)	3·05 (2·71–3·44)	2·34 (2·07–2·64)	2·18 (1·92–2·48)
Black	340 912	186 (0·1%)	3·08 (2·63–3·60)	3·91 (3·34–4·58)	3·52 (3·00–4·13)	3·21 (2·73–3·77)	3·12 (2·65–3·67)
Mixed	170 484	56 (<0·1%)	1·97 (1·51–2·57)	3·19 (2·44–4·17)	3·01 (2·30–3·94)	3·02 (2·31–3·95)	2·96 (2·26–3·87)
Other	320 788	97 (<0·1%)	1·83 (1·48–2·25)	2·92 (2·37–3·60)	2·80 (2·27–3·45)	3·28 (2·66–4·06)	3·18 (2·58–3·93)
Unknown	4 553 051	662 (<0·1%)	0·95 (0·87–1·04)	1·03 (0·94–1·12)	1·03 (0·94–1·13)	1·09 (0·99–1·19)	1·08 (0·99–1·19)
COVID-19-related death (n=11 649)							
White	10 877 978	7514 (0·1%)	1 (ref)	1 (ref)	1 (ref)	1 (ref)	1 (ref)
South Asian	1 025 319	734 (0·1%)	0·89 (0·83–0·97)	1·85 (1·71–2·01)	1·64 (1·52–1·78)	1·47 (1·35–1·60)	1·26 (1·15–1·37)
Black	340 912	268 (0·1%)	0·96 (0·85–1·09)	1·88 (1·66–2·13)	1·61 (1·42–1·82)	1·50 (1·32–1·70)	1·51 (1·33–1·71)
Mixed	170 484	65 (<0·1%)	0·51 (0·40–0·65)	1·58 (1·23–2·01)	1·44 (1·13–1·84)	1·43 (1·12–1·83)	1·41 (1·11–1·81)
Other	320 788	107 (<0·1%)	0·44 (0·37–0·54)	1·25 (1·03–1·51)	1·17 (0·97–1·42)	1·24 (1·02–1·50)	1·22 (1·00–1·48)
Unknown	4 553 051	2961 (0·1%)	0·95 (0·91–0·99)	0·92 (0·88–0·96)	0·93 (0·89–0·97)	1·00 (0·95–1·04)	1·01 (0·97–1·06)

Data are N, n (%), or hazard ratio (95% CI). *Including number of primary care consultations in the preceding 12 months. †Analyses of COVID-19-related hospital admissions for wave 1 were added during the revision stage of this Article as data on hospitalisations became available in OpenSAFELY after the initial submission; therefore, cohort sizes differ for this outcome compared with other outcomes. ‡For categories containing small numbers (≤5) within any subcategory, we have rounded all counts to the nearest 10, per data disclosure agreements.

Table 2: Associations between ethnicity in five categories and COVID-19 outcomes in wave 1, with serial adjustment

explanatory variables; appendix p 15), rather than the general denominator population. Risks of COVID-19-related hospitalisation (adjusted HR 1·23 [95% CI 1·11–1·37]) and ICU admission (1·67 [1·37–2·05]) remained higher for the Black group than for the White group in wave 2, but these risks were attenuated in magnitude compared with wave 1 (figure 3). However, in contrast to the pattern seen in the broader Black group, the risk of COVID-19-related hospitalisation in the Caribbean subgroup did not differ from that of the White British

group (0·98 [0·81–1·17]). Additionally, the excess risk of COVID-19 death was attenuated for the Black group in wave 2 compared with wave 1, so that no differences between the White and Black groups remained in wave 2 (0·92 [0·73–1·16]; figure 4).

71 920 non-COVID-19-related deaths occurred in the study population during wave 1. The risk of non-COVID-19-related death was 15–32% lower in all non-White ethnic groups (South Asian group adjusted HR 0·85 [95% CI 0·81–0·90], Black group 0·85 [0·78–0·92], mixed ethnicity

A Tested for SARS-CoV-2

	Wave 1			Wave 2		
	Denominator	Events	HR (95% CI)	Denominator	Events	HR (95% CI)
All White ethnicities*	10 877 978	793 181 (7·3%)	**1 (ref)**	11 323 083	1 733 272 (15·3%)	**1 (ref)**
White British*	9 461 409	713 121 (7·5%)	1 (ref)	9 845 315	1 567 742 (15·9%)	1 (ref)
White Irish	78 170	6160 (7·9%)	1·06 (1·04–1·09)	81 185	11 763 (14·5%)	0·94 (0·92–0·96)
Other White	1 338 399	73 900 (5·5%)	0·80 (0·80–0·81)	1 396 583	153 767 (11·0%)	0·67 (0·66–0·67)
All South Asian ethnicities	1 025 319	82 647 (8·1%)	**1·08 (1·07–1·09)**	1 061 336	166 856 (15·7%)	**0·96 (0·95–0·96)**
Indian	431 814	40 231 (9·3%)	1·15 (1·14–1·16)	447 812	66 056 (14·8%)	0·88 (0·87–0·88)
Pakistani	301 496	21 952 (7·3%)	0·94 (0·93–0·96)	309 214	58 882 (19·0%)	1·02 (1·02–1·03)
Bangladeshi	69 959	3924 (5·6%)	0·70 (0·67–0·72)	73 154	11 396 (15·6%)	0·81 (0·79–0·82)
Other South Asian	222 050	16 540 (7·4%)	1·06 (1·04–1·07)	231 156	30 522 (13·2%)	0·77 (0·76–0·77)
All Black ethnicities	340 912	25 305 (7·4%)	**1·08 (1·06–1·09)**	355 625	48 068 (13·5%)	**0·83 (0·82–0·84)**
Caribbean	82 561	5972 (7·2%)	0·97 (0·95–1·00)	84 219	10 749 (12·8%)	0·76 (0·75–0·78)
African	190 164	14 623 (7·7%)	1·09 (1·08–1·11)	199 012	27 417 (13·8%)	0·78 (0·77–0·79)
Other Black	68 187	4710 (6·9%)	0·96 (0·94–0·99)	72 394	9902 (13·7%)	0·77 (0·75–0·78)
All mixed ethnicities	170 484	12 126 (7·1%)	**1·04 (1·02–1·05)**	179 557	26 679 (14·9%)	**0·90 (0·88–0·91)**
White and Caribbean	39 208	2838 (7·2%)	0·96 (0·92–0·99)	41 228	6868 (16·7%)	0·90 (0·88–0·92)
White and African	33 194	2619 (7·9%)	1·15 (1·10–1·19)	34 991	5134 (14·7%)	0·83 (0·81–0·86)
White and Asian	34 264	2608 (7·6%)	1·07 (1·03–1·12)	36 257	5548 (15·3%)	0·87 (0·85–0·90)
Other mixed ethnicity	63 818	4061 (6·4%)	0·91 (0·89–0·94)	67 081	9129 (13·6%)	0·78 (0·76–0·80)
All other ethnicities	320 788	15 824 (4·9%)	**0·77 (0·76–0·78)**	334 962	32 904 (9·8%)	**0·64 (0·64–0·65)**
Chinese	103 892	2899 (2·8%)	0·45 (0·43–0·46)	105 658	6444 (6·1%)	0·38 (0·38–0·39)
All other ethnic groups	216 896	12 925 (6·0%)	0·87 (0·86–0·89)	229 304	26 460 (11·5%)	0·69 (0·68–0·70)
Unknown ethnicity	4 553 051	287 718 (6·3%)	0·97 (0·97–0·98)	4 381 803	639 977 (14·6%)	0·99 (0·99–0·99)

B Tested positive for SARS-CoV-2

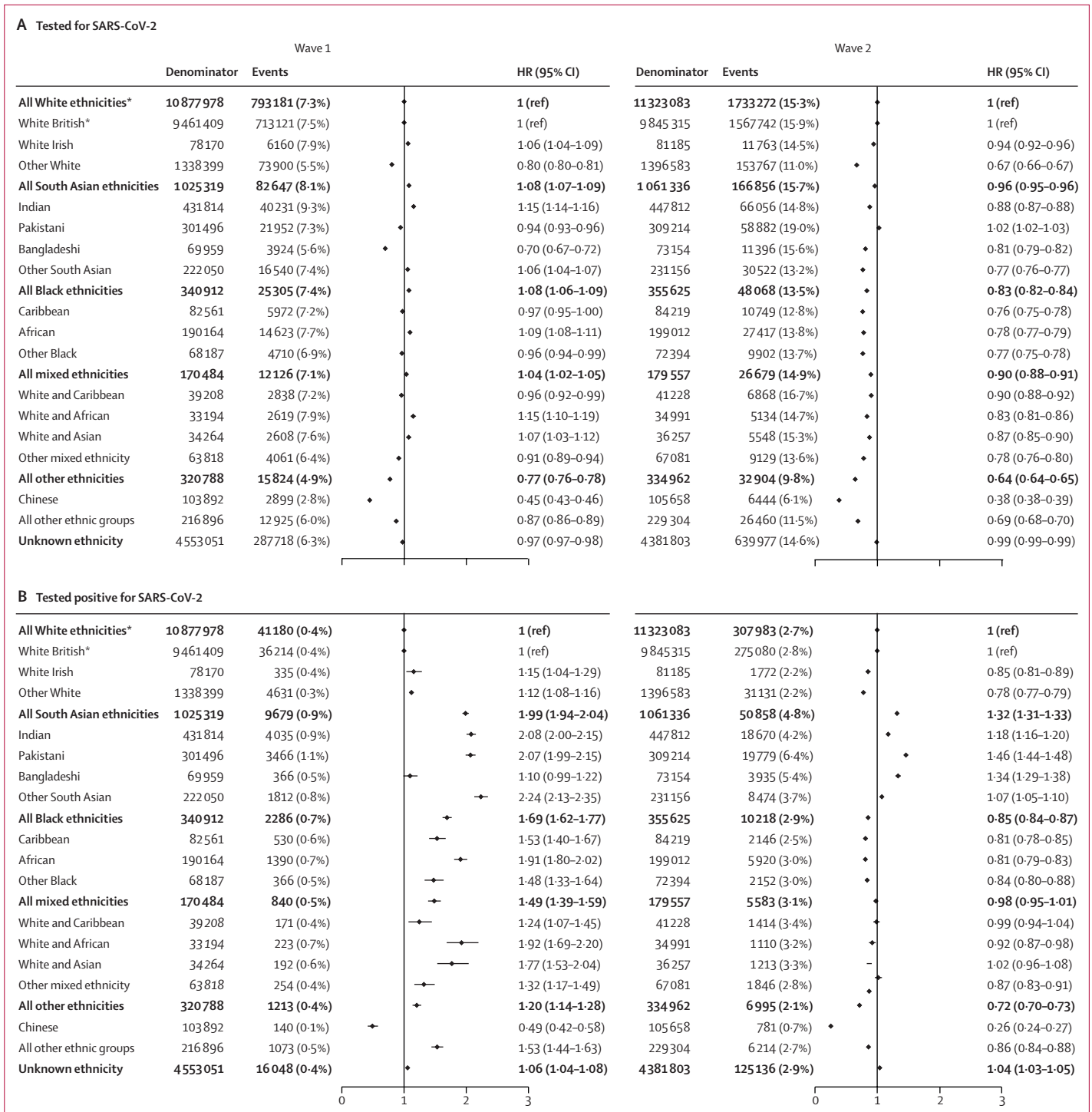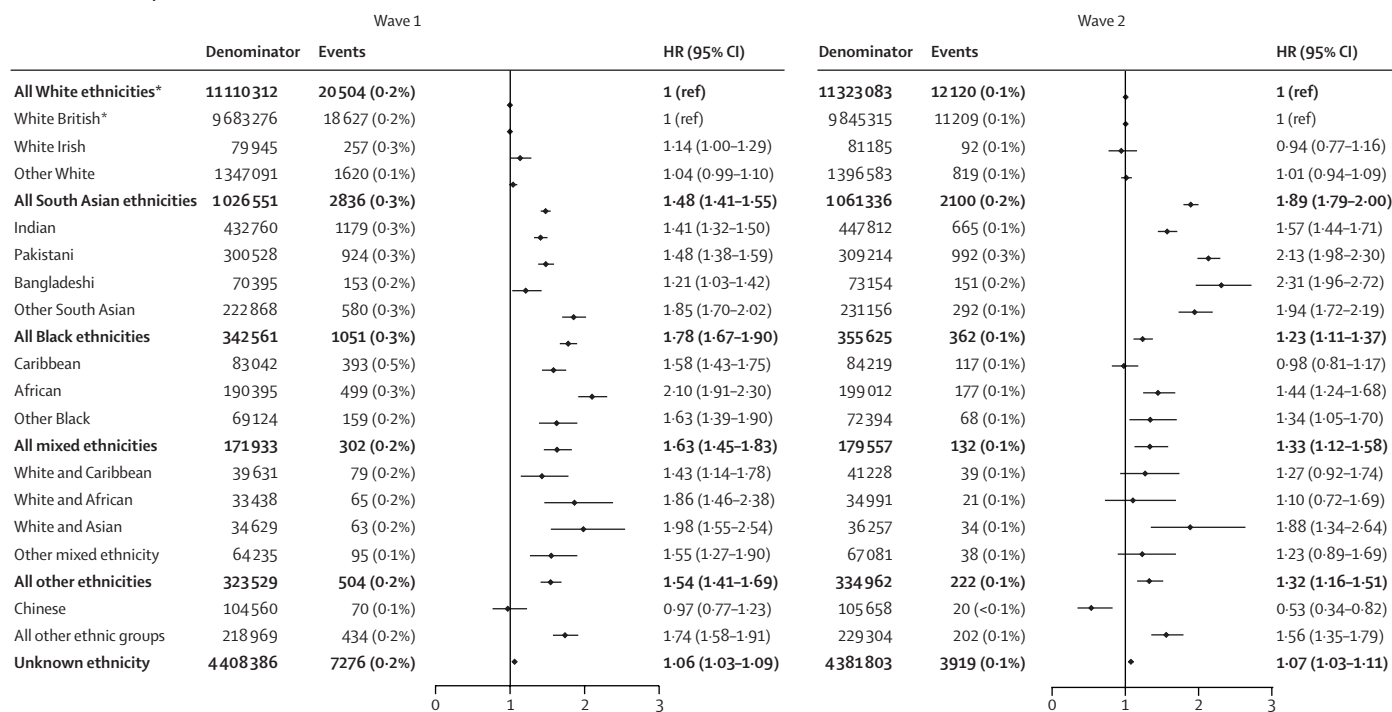|	Wave 1			Wave 2		
	Denominator	Events	HR (95% CI)	Denominator	Events	HR (95% CI)
All White ethnicities*	10 877 978	41 180 (0·4%)	**1 (ref)**	11 323 083	307 983 (2·7%)	**1 (ref)**
White British*	9 461 409	36 214 (0·4%)	1 (ref)	9 845 315	275 080 (2·8%)	1 (ref)
White Irish	78 170	335 (0·4%)	1·15 (1·04–1·29)	81 185	1772 (2·2%)	0·85 (0·81–0·89)
Other White	1 338 399	4631 (0·3%)	1·12 (1·08–1·16)	1 396 583	31 131 (2·2%)	0·78 (0·77–0·79)
All South Asian ethnicities	1 025 319	9679 (0·9%)	**1·99 (1·94–2·04)**	1 061 336	50 858 (4·8%)	**1·32 (1·31–1·33)**
Indian	431 814	4035 (0·9%)	2·08 (2·00–2·15)	447 812	18 670 (4·2%)	1·18 (1·16–1·20)
Pakistani	301 496	3466 (1·1%)	2·07 (1·99–2·15)	309 214	19 779 (6·4%)	1·46 (1·44–1·48)
Bangladeshi	69 959	366 (0·5%)	1·10 (0·99–1·22)	73 154	3935 (5·4%)	1·34 (1·29–1·38)
Other South Asian	222 050	1812 (0·8%)	2·24 (2·13–2·35)	231 156	8474 (3·7%)	1·07 (1·05–1·10)
All Black ethnicities	340 912	2286 (0·7%)	**1·69 (1·62–1·77)**	355 625	10 218 (2·9%)	**0·85 (0·84–0·87)**
Caribbean	82 561	530 (0·6%)	1·53 (1·40–1·67)	84 219	2146 (2·5%)	0·81 (0·78–0·85)
African	190 164	1390 (0·7%)	1·91 (1·80–2·02)	199 012	5920 (3·0%)	0·81 (0·79–0·83)
Other Black	68 187	366 (0·5%)	1·48 (1·33–1·64)	72 394	2152 (3·0%)	0·84 (0·80–0·88)
All mixed ethnicities	170 484	840 (0·5%)	**1·49 (1·39–1·59)**	179 557	5583 (3·1%)	**0·98 (0·95–1·01)**
White and Caribbean	39 208	171 (0·4%)	1·24 (1·07–1·45)	41 228	1414 (3·4%)	0·99 (0·94–1·04)
White and African	33 194	223 (0·7%)	1·92 (1·69–2·20)	34 991	1110 (3·2%)	0·92 (0·87–0·98)
White and Asian	34 264	192 (0·6%)	1·77 (1·53–2·04)	36 257	1213 (3·3%)	1·02 (0·96–1·08)
Other mixed ethnicity	63 818	254 (0·4%)	1·32 (1·17–1·49)	67 081	1846 (2·8%)	0·87 (0·83–0·91)
All other ethnicities	320 788	1213 (0·4%)	**1·20 (1·14–1·28)**	334 962	6995 (2·1%)	**0·72 (0·70–0·73)**
Chinese	103 892	140 (0·1%)	0·49 (0·42–0·58)	105 658	781 (0·7%)	0·26 (0·24–0·27)
All other ethnic groups	216 896	1073 (0·5%)	1·53 (1·44–1·63)	229 304	6214 (2·7%)	0·86 (0·84–0·88)
Unknown ethnicity	4 553 051	16 048 (0·4%)	1·06 (1·04–1·08)	4 381 803	125 136 (2·9%)	1·04 (1·03–1·05)

Figure 2: **Ethnic differences in the risks of being tested for SARS-CoV-2 infection (A) and testing positive for SARS-CoV-2 infection (B)**
Models adjust for age, sex, deprivation quintile, all prespecified clinical comorbidities, body-mass index category, glycated haemoglobin category, systolic and diastolic blood pressure category, number of primary care consultations in the previous 12 months, household size, and stratification by sustainability and transformation partnership region. HR=hazard ratio. *All White ethnicities is the reference category for comparison of ethnicity in the five broad categories, and White British is the reference category for comparison of ethnicity in the 16 subcategories.

group 0·81 [0·70–0·93]; other ethnicity group 0·68 [0·61–0·77]) than in the White group after adjustment for all measured explanatory variables (appendix p 5). In wave 2, the risk of non-COVID-19-related death remained lower for the South Asian, Black, and other ethnicity groups than for the White group (appendix p 10).
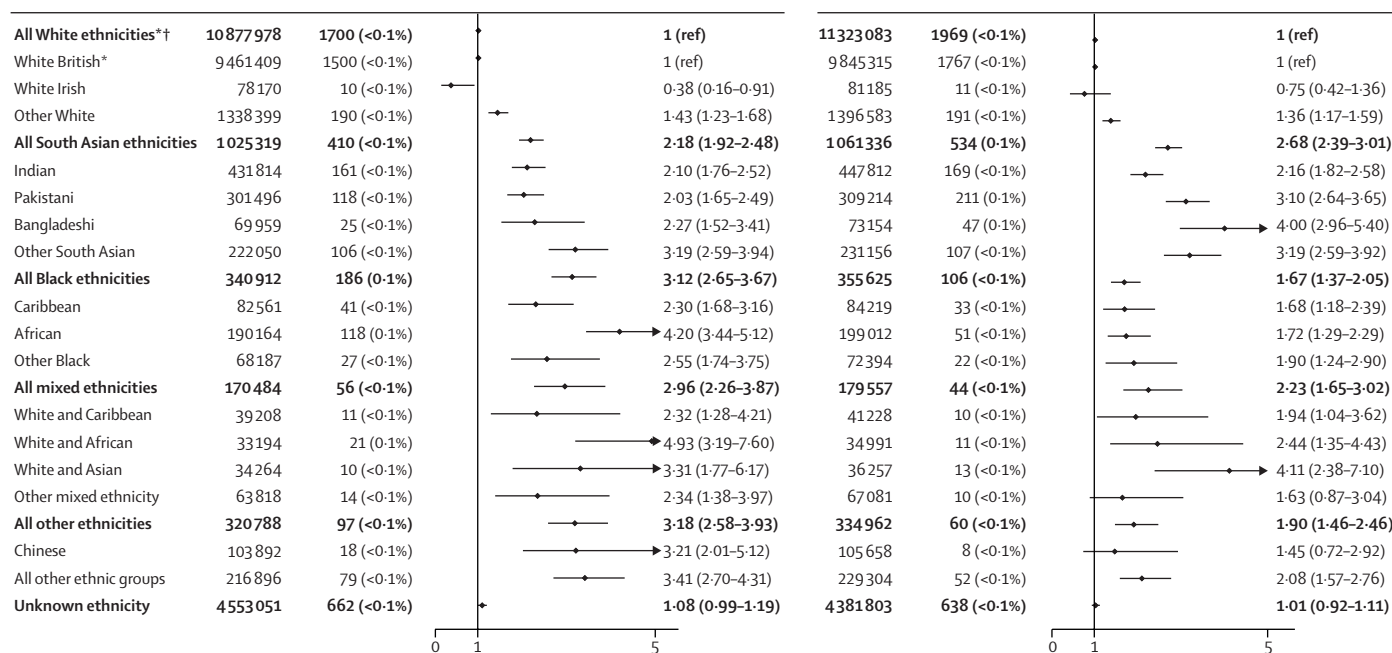
A Admitted to hospital for COVID-19

	Wave 1				Wave 2		
	Denominator	Events		HR (95% CI)	Denominator	Events	HR (95% CI)
All White ethnicities*	11 110 312	20 504 (0·2%)		1 (ref)	11 323 083	12 120 (0·1%)	1 (ref)
White British*	9 683 276	18 627 (0·2%)		1 (ref)	9 845 315	11 209 (0·1%)	1 (ref)
White Irish	79 945	257 (0·3%)		1·14 (1·00–1·29)	81 185	92 (0·1%)	0·94 (0·77–1·16)
Other White	1 347 091	1620 (0·1%)		1·04 (0·99–1·10)	1 396 583	819 (0·1%)	1·01 (0·94–1·09)
All South Asian ethnicities	1 026 551	2836 (0·3%)		1·48 (1·41–1·55)	1 061 336	2100 (0·2%)	1·89 (1·79–2·00)
Indian	432 760	1179 (0·3%)		1·41 (1·32–1·50)	447 812	665 (0·1%)	1·57 (1·44–1·71)
Pakistani	300 528	924 (0·3%)		1·48 (1·38–1·59)	309 214	992 (0·3%)	2·13 (1·98–2·30)
Bangladeshi	70 395	153 (0·2%)		1·21 (1·03–1·42)	73 154	151 (0·2%)	2·31 (1·96–2·72)
Other South Asian	222 868	580 (0·3%)		1·85 (1·70–2·02)	231 156	292 (0·1%)	1·94 (1·72–2·19)
All Black ethnicities	342 561	1051 (0·3%)		1·78 (1·67–1·90)	355 625	362 (0·1%)	1·23 (1·11–1·37)
Caribbean	83 042	393 (0·5%)		1·58 (1·43–1·75)	84 219	117 (0·1%)	0·98 (0·81–1·17)
African	190 395	499 (0·3%)		2·10 (1·91–2·30)	199 012	177 (0·1%)	1·44 (1·24–1·68)
Other Black	69 124	159 (0·2%)		1·63 (1·39–1·90)	72 394	68 (0·1%)	1·34 (1·05–1·70)
All mixed ethnicities	171 933	302 (0·2%)		1·63 (1·45–1·83)	179 557	132 (0·1%)	1·33 (1·12–1·58)
White and Caribbean	39 631	79 (0·2%)		1·43 (1·14–1·78)	41 228	39 (0·1%)	1·27 (0·92–1·74)
White and African	33 438	65 (0·2%)		1·86 (1·46–2·38)	34 991	21 (0·1%)	1·10 (0·72–1·69)
White and Asian	34 629	63 (0·2%)		1·98 (1·55–2·54)	36 257	34 (0·1%)	1·88 (1·34–2·64)
Other mixed ethnicity	64 235	95 (0·1%)		1·55 (1·27–1·90)	67 081	38 (0·1%)	1·23 (0·89–1·69)
All other ethnicities	323 529	504 (0·2%)		1·54 (1·41–1·69)	334 962	222 (0·1%)	1·32 (1·16–1·51)
Chinese	104 560	70 (0·1%)		0·97 (0·77–1·23)	105 658	20 (<0·1%)	0·53 (0·34–0·82)
All other ethnic groups	218 969	434 (0·2%)		1·74 (1·58–1·91)	229 304	202 (0·1%)	1·56 (1·35–1·79)
Unknown ethnicity	4 408 386	7276 (0·2%)		1·06 (1·03–1·09)	4 381 803	3919 (0·1%)	1·07 (1·03–1·11)

(HR axis: 0, 1, 2, 3)

B Admitted to ICU for COVID-19

	Wave 1				Wave 2		
	Denominator	Events		HR (95% CI)	Denominator	Events	HR (95% CI)
All White ethnicities*†	10 877 978	1700 (<0·1%)		1 (ref)	11 323 083	1969 (<0·1%)	1 (ref)
White British*	9 461 409	1500 (<0·1%)		1 (ref)	9 845 315	1767 (<0·1%)	1 (ref)
White Irish	78 170	10 (<0·1%)		0·38 (0·16–0·91)	81 185	11 (<0·1%)	0·75 (0·42–1·36)
Other White	1 338 399	190 (<0·1%)		1·43 (1·23–1·68)	1 396 583	191 (<0·1%)	1·36 (1·17–1·59)
All South Asian ethnicities	1 025 319	410 (<0·1%)		2·18 (1·92–2·48)	1 061 336	534 (0·1%)	2·68 (2·39–3·01)
Indian	431 814	161 (<0·1%)		2·10 (1·76–2·52)	447 812	169 (<0·1%)	2·16 (1·82–2·58)
Pakistani	301 496	118 (<0·1%)		2·03 (1·65–2·49)	309 214	211 (0·1%)	3·10 (2·64–3·65)
Bangladeshi	69 959	25 (<0·1%)		2·27 (1·52–3·41)	73 154	47 (0·1%)	4·00 (2·96–5·40)
Other South Asian	222 050	106 (<0·1%)		3·19 (2·59–3·94)	231 156	107 (<0·1%)	3·19 (2·59–3·92)
All Black ethnicities	340 912	186 (0·1%)		3·12 (2·65–3·67)	355 625	106 (<0·1%)	1·67 (1·37–2·05)
Caribbean	82 561	41 (<0·1%)		2·30 (1·68–3·16)	84 219	33 (<0·1%)	1·68 (1·18–2·39)
African	190 164	118 (0·1%)		4·20 (3·44–5·12)	199 012	51 (<0·1%)	1·72 (1·29–2·29)
Other Black	68 187	27 (<0·1%)		2·55 (1·74–3·75)	72 394	22 (<0·1%)	1·90 (1·24–2·90)
All mixed ethnicities	170 484	56 (<0·1%)		2·96 (2·26–3·87)	179 557	44 (<0·1%)	2·23 (1·65–3·02)
White and Caribbean	39 208	11 (<0·1%)		2·32 (1·28–4·21)	41 228	10 (<0·1%)	1·94 (1·04–3·62)
White and African	33 194	21 (0·1%)		4·93 (3·19–7·60)	34 991	11 (<0·1%)	2·44 (1·35–4·43)
White and Asian	34 264	10 (<0·1%)		3·31 (1·77–6·17)	36 257	13 (<0·1%)	4·11 (2·38–7·10)
Other mixed ethnicity	63 818	14 (<0·1%)		2·34 (1·38–3·97)	67 081	10 (<0·1%)	1·63 (0·87–3·04)
All other ethnicities	320 788	97 (<0·1%)		3·18 (2·58–3·93)	334 962	60 (<0·1%)	1·90 (1·46–2·46)
Chinese	103 892	18 (<0·1%)		3·21 (2·01–5·12)	105 658	8 (<0·1%)	1·45 (0·72–2·92)
All other ethnic groups	216 896	79 (<0·1%)		3·41 (2·70–4·31)	229 304	52 (<0·1%)	2·08 (1·57–2·76)
Unknown ethnicity	4 553 051	662 (<0·1%)		1·08 (0·99–1·19)	4 381 803	638 (<0·1%)	1·01 (0·92–1·11)

(HR axis: 0, 1, 5)

***Figure 3*: Ethnic differences in the risk of COVID-19-related hospital admission (A) and COVID-19-related ICU admission (B)**
Models adjust for age, sex, deprivation quintile, all prespecified clinical comorbidities, body-mass index category, glycated haemoglobin category, systolic and diastolic blood pressure category, number of primary care consultations in the previous 12 months, household size, and stratification by sustainability and transformation partnership region. Analyses of COVID-19-related hospital admissions for wave 1 were added during the revision stage of this Article as data on hospitalisations became available in OpenSAFELY after the initial submission; therefore, cohort sizes differ for this outcome compared with other outcomes. HR=hazard ratio. ICU=intensive care unit. *All White ethnicities is the reference category for comparison of ethnicity in the five broad categories, and White British is the reference category for comparison of ethnicity in the 16 subcategories. †For categories containing small numbers (≤5) within any subcategory, we have rounded all counts to the nearest 10, per data disclosure agreements.

	Wave 1			Wave 2		
	Denominator	Events	HR (95% CI)	Denominator	Events	HR (95% CI)
All White ethnicities*	10877978	7514 (0.1%)	**1 (ref)**	11323083	4874 (<0.1%)	**1 (ref)**
White British*	9461409	6992 (0.1%)	1 (ref)	9845315	4599 (<0.1%)	1 (ref)
White Irish	78170	86 (0.1%)	0.92 (0.74–1.14)	81185	32 (<0.1%)	0.69 (0.49–0.98)
Other White	1338399	436 (<0.1%)	0.91 (0.82–1.00)	1396583	243 (<0.1%)	0.96 (0.84–1.09)
All South Asian ethnicities	1025319	734 (0.1%)	**1.26 (1.15–1.37)**	1061336	532 (0.1%)	**1.87 (1.68–2.07)**
Indian	431814	335 (0.1%)	1.29 (1.14–1.45)	447812	205 (<0.1%)	1.84 (1.58–2.15)
Pakistani	301496	210 (0.1%)	1.03 (0.89–1.19)	309214	240 (0.1%)	1.94 (1.67–2.24)
Bangladeshi	69959	49 (0.1%)	1.29 (0.97–1.71)	73154	39 (0.1%)	2.26 (1.64–3.13)
Other South Asian	222050	140 (0.1%)	1.57 (1.32–1.87)	231156	48 (<0.1%)	1.41 (1.05–1.88)
All Black ethnicities	340912	268 (0.1%)	**1.51 (1.33–1.71)**	355625	75 (<0.1%)	**0.92 (0.73–1.16)**
Caribbean	82561	137 (0.2%)	1.33 (1.12–1.58)	84219	44 (0.1%)	0.85 (0.63–1.15)
African	190164	91 (<0.1%)	1.77 (1.43–2.19)	199012	19 (<0.1%)	0.95 (0.61–1.50)
Other Black	68187	40 (0.1%)	1.61 (1.18–2.20)	72394	12 (<0.1%)	1.08 (0.61–1.90)
All mixed ethnicities	170484	65 (<0.1%)	**1.41 (1.11–1.81)**	179225	26 (<0.1%)	**1.24 (0.85–1.83)**
White and Caribbean	39208	28 (0.1%)	1.62 (1.12–2.36)	41228	7 (<0.1%)	0.80 (0.38–1.69)
White and African	33194	9 (<0.1%)	1.26 (0.65–2.42)	34991	6 (<0.1%)	1.93 (0.86–4.30)
White and Asian	34264	9 (<0.1%)	1.21 (0.63–2.33)	36257	5 (<0.1%)	1.51 (0.63–3.64)
Other mixed ethnicity	63818	19 (<0.1%)	1.31 (0.84–2.06)	67081	8 (<0.1%)	1.35 (0.67–2.70)
All other ethnicities	320788	107 (<0.1%)	**1.22 (1.00–1.48)**	334962	35 (<0.1%)	**0.92 (0.66–1.29)**
Chinese	103892	20 (<0.1%)	1.05 (0.67–1.63)	105658	6 (<0.1%)	0.70 (0.31–1.55)
All other ethnic groups	216896	87 (<0.1%)	1.26 (1.01–1.56)	229304	29 (<0.1%)	0.97 (0.67–1.40)
Unknown ethnicity	4553051	2961 (0.1%)	**1.01 (0.97–1.06)**	4381803	1824 (<0.1%)	**1.17 (1.11–1.24)**

Figure 4: **Ethnic differences in the risk of COVID-19-related death**

Models adjust for age, sex, deprivation quintile, all prespecified clinical comorbidities, body-mass index category, glycated haemoglobin category, systolic and diastolic blood pressure category, number of primary care consultations in the previous 12 months, household size, and stratification by sustainability and transformation partnership region. For categories containing small numbers (≤5) within any subcategory, we have rounded all counts to the nearest 10, per data disclosure agreements. HR=hazard ratio. *All White ethnicities is the reference category for comparison of ethnicity in the five broad categories, and White British is the reference category for comparison of ethnicity in the 16 subcategories.

In wave 1, among the 78124 care home residents, 46065 (59.0%) individuals were tested for SARS-CoV-2, 6330 (8.1%) tested positive, 2517 (3.2%) were admitted to hospital, and 3877 (5.0%) died from COVID-19. Although no ethnic differences in being tested for or testing positive for SARS-CoV-2 infection were apparent, people of Black ethnicity (adjusted HR 1.43 [95% CI 1.02–2.00]) and other ethnicity (1.73 [1.19–2.50]) were more likely to die from COVID-19 than people of White ethnicity, after adjustment for all measured explanatory variables except household size. In wave 2, no ethnic differences among care home populations were evident for any outcome except COVID-19 death, which was raised for South Asian groups (1.81 [1.07–3.05]; appendix p 16). Because of small numbers, we were unable to explore ethnic differences in ICU admissions or differences according to ethnicity in 16 categories among care home residents.

Using multiple imputation to account for unknown ethnicity did not materially change any of the associations observed in the complete case analysis (appendix p 17), nor did restricting the definition of COVID-19-related death to underlying cause only (appendix p 18) or removing adjustment for STP region (appendix p 19). We detected no evidence of deviations from the proportional hazards assumption (appendix p 20).

Discussion

In a population-based cohort study of 17 million adults in England we found that, although ethnic differences in testing were small, minority ethnic groups were at increased risk of testing positive for SARS-CoV-2 and of COVID-19-related hospitalisation, ICU admission, and death. Disaggregation into detailed ethnic categories revealed important within-group heterogeneity, emphasising the importance of disaggregated reporting wherever possible. In wave 2, minority ethnic groups were less likely to be tested than White groups, and risks of severe COVID-19 outcomes (ie, hospitalisation, ICU admission, and death) increased for South Asian groups but were attenuated in all other ethnic groups relative to the White group compared with wave 1.

In the largest UK-based study to date, we captured high-quality clinical data across a range of health-care settings and linked individual-level COVID-19 datasets, which enabled us to generate timely insights into ethnic disparities at different severity levels of COVID-19, from being tested for infection to dying from the disease. We were able to report findings according to self-reported ethnicity in 16 categories, whereas other UK-based studies have aggregated ethnicity into higher-level groups because of small numbers. We also reported differences in outcomes using a general population-based sample,

which allowed us to overcome issues commonly faced by studies limited to individuals with SARS-CoV-2 infection or hospitalisation, wherein the populations under study might not represent the true general population at risk.[31]

Our inability to capture all potential explanatory factors of ethnic disparities in COVID-19 outcomes is likely to have affected our observed associations. For example, we were unable to account for ethnic differences in genetic ancestry,[32,33] occupation,[34] experiences of racism or structural discrimination,[9,35,36] and health-related behaviour.[37,38] Because of invalid address information, we were unable to estimate household size for 13% of our population. We might have underestimated household size for homes including people registered at non-TPP primary care practices and overestimated it for individuals living in large apartment blocks, or for people who had not updated their address after moving homes. In recognition of these limitations, we grouped household size into four levels rather than considering it as a continuous measure. Furthermore, it is possible that cause of death might have been misclassified on death certificates, and that the extent of this misclassification might have differed by time period and ethnicity. A limitation of SARS-CoV-2 testing data was the selective opportunity to be tested early in the pandemic, which was skewed towards health-care workers and people with severe or symptomatic disease, particularly during the first wave. Although OpenSAFELY is broadly representative of the English population, it includes data from a single software system that is known to have lower coverage in London than in other regions of the UK. However, our results mirror other studies done in the UK[1] and in the USA,[5,39] suggesting that potential mechanisms underpinning ethnic differences in COVID-19 might be common across countries with similar population structures. OpenSAFELY data are collected prospectively in real time by clinicians and practice staff, and are subject to the same strengths and biases as other UK-based electronic health record databases.

Despite these limitations, this study represents the most comprehensive examination of ethnic inequalities in England during the COVID-19 pandemic in 2020. Using the OpenSAFELY data analytics platform, we capitalised on the rapid real-time linkage of routine datasets in a highly secure environment to explore a range of urgent questions around the patterning of ethnic inequalities in the UK.

This study builds on previous research in several ways. First, we have confirmed ethnic differences in COVID-19-related mortality and provided novel data across a range of other outcomes (testing, hospitalisation, and ICU admission). Second, we have explored whether household size has an effect beyond sociodemographic and clinical characteristics. Finally, we have reported on both the general population and care home residents during the first and second waves of the pandemic in England.

Although some minority ethnic groups were less likely to be tested for SARS-CoV-2 in this study, all non-White groups were more likely to test positive, even when restricted to those ever tested. This finding might suggest that White populations are tested more frequently with mild or asymptomatic disease, or that minority ethnic groups get tested at more severe stages of the disease. Disparities in testing might relate to a lack of access to testing sites, poorer health literacy, lack of tailored and accessible health communications, or differences in testing-related behaviours.[40] Emerging evidence suggests that individuals might avoid seeking a test for fear of losing income or employment if required to quarantine after testing positive.[41] Given that minority ethnic groups are more likely to work in insecure jobs with poor workplace protections, and in essential or key-worker roles associated with higher risk of death from COVID-19,[42–44] it is likely that social and economic barriers to testing are greater in minority ethnic groups.

Our finding that minority ethnic groups have higher risks of COVID-19-related hospitalisation, ICU admission, and death after accounting for clinical comorbidities suggests that improving equity in clinical care and understanding potential interactions between COVID-19 and underlying conditions are essential for mitigating inequalities in the downstream effects of SARS-CoV-2 infection. The fact that inequalities worsened for South Asian groups in wave 2 compared with in wave 1 suggests that more aggressive and tailored interventions are needed to meet the needs in these communities.[45] However, our finding of relatively attenuated risks in all other minority ethnic groups is a potential positive finding; further investigation is warranted into which public health actions were most influential in mitigating health disparities for these groups.

Our finding that the magnitude of ethnic differences in testing positive in wave 1 were similar to those in COVID-19-related death suggests that ethnic differences in death might be mediated through exposure or susceptibility to infection, rather than through susceptibility to severe disease once infected. This hypothesis is supported by recent findings from the REACT-2 study, which found higher levels of SARS-CoV-2 antibodies in minority ethnic groups but no ethnic differences in the infection-to-mortality ratio.[46]

After accounting for sociodemographic and clinical factors, household size further explained differences in COVID-19 outcomes for South Asian groups. This finding is consistent with an ONS study that found that multigenerational living was causally associated with an increased risk of death due to COVID-19 in South Asian women, but not in any other ethnic groups.[47] According to data from the 2011 census, 21% of South Asian groups live in multigenerational households, in contrast to around 7% of White groups.[24,48] We hypothesise that household size and deprivation might proxy viral exposure by capturing aspects of occupational and community-level exposure. Although multigenerational living might increase the risks of exposure and

transmission (from children or working-age adults to older or vulnerable family members), such households and extended communities also offer valuable informal care networks and facilitate engagement with health and community services.[49] In light of emerging evidence that minority ethnic groups are less likely to take up the COVID-19 vaccine, co-designing culturally competent and non-stigmatising engagement strategies with these communities is increasingly important.[50,51]

National data from England and Scotland have shown that some minority ethnic groups have both better overall health and lower all-cause mortality than White groups.[52,53] We were able to confirm this pattern in our sensitivity analyses, and our findings of disparities in SARS-CoV-2 positivity and COVID-19-related outcomes, some of which have continued to widen over the course of the epidemic in the UK, are, therefore, particularly concerning.

Our findings mirror large studies in the USA, which have found that minority racial and ethnic communities have elevated risks of testing positive for SARS-CoV-2 infection and of COVID-19-related hospitalisation and death that differentially vary over time, even after accounting for sociodemographic characteristics and underlying health conditions.[5,39] These parallel findings suggest that mechanisms underpinning ethnic differences in COVID-19 outcomes in England might be common in other settings, and that learnings across settings should be shared.

Improving the quality and completeness of ethnicity data across health and administrative datasets is essential for building a complete picture of ethnic disparities.[54] Furthermore, although the recording of ethnicity on death certificates has been the norm in Scotland for the past decade, it is only now being considered for use in England.[55–57] Prioritising linkage between health, social, and employment data will be essential in building a complete picture of ethnic differences in COVID-19 risk and outcomes.

Minority ethnic groups in the UK have had disproportionately high levels of poor COVID-19 outcomes, with disparities increasing even within the course of the epidemic for some groups. Reducing ethnic inequalities will need action across a broad range of measures such as addressing the wider adverse effects of disadvantage and structural discrimination, reducing within-household and between-household transmission, and improving control of clinical conditions. The relative importance of each of these measures will differ by both ethnic group and stage of COVID-19 progression. Equality is difficult to achieve, but structural and persistent inequalities must be addressed in a civilised society.

References

1 Pan D, Sze S, Minhas JS, et al. The impact of ethnicity on clinical outcomes in COVID-19: a systematic review. *EClinicalMedicine* 2020; **23:** 100404.

2 Williamson EJ, Walker AJ, Bhaskaran K, et al. Factors associated with COVID-19-related death using OpenSAFELY. *Nature* 2020; **584:** 430–36.

3 Aldridge RW, Lewer D, Katikireddi SV, 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.

4 Public Health England. Beyond the data: understanding the impact of COVID-19 on BAME groups. June, 2020. https://assets. publishing.service.gov.uk/government/uploads/system/uploads/ attachment_data/file/892376/COVID_stakeholder_engagement_ synthesis_beyond_the_data.pdf (accessed Sept 2, 2020).

5 Rentsch CT, Kidwai-Khan F, Tate JP, et al. Patterns of COVID-19 testing and mortality by race and ethnicity among United States veterans: a nationwide cohort study. *PLoS Med* 2020; **17:** e1003379.

6 Baqui P, Bica I, Marra V, Ercole A, van der Schaar M. Ethnic and regional variations in hospital mortality from COVID-19 in Brazil: a cross-sectional observational study. *Lancet Glob Health* 2020; **8:** e1018–26.

7 Chen J, Krieger N. Revealing the unequal burden of COVID-19 by income, race/ethnicity, and household crowding: US county versus zip code analyses. *J Public Health Manag Pract* 2021; **27:** S43–56.

8 Kenway P, Holden J. Accounting for the variation in the confirmed COVID-19 caseload across England: an analysis of the role of multi-generation households, London and time. New Policy Institute, 2020. https://www.npi.org.uk/publications/housing-and-homelessness/accounting-variation-confirmed-covid-19-caseload-across-england-analysis-role-multi-generation-households-london-and-time/ (accessed Sept 15, 2020).

9 Razai MS, Kankam HKN, Majeed A, Esmail A, Williams DR. Mitigating ethnic disparities in COVID-19 and beyond. *BMJ* 2021; **372:** m4921.

10 Ethnicity sub-group of the Scientific Advisory Group for Emergencies. Drivers of the higher COVID-19 incidence, morbidity and mortality among minority ethnic groups. 2020. https://assets. publishing.service.gov.uk/government/uploads/system/uploads/ attachment_data/file/925135/S0778_Drivers_of_the_higher_ COVID-19_incidence__morbidity_and_mortality_among_minority_ ethnic_groups.pdf (accessed Nov 3, 2020).

11 Warden J. NHS to come under Race Relations Act. *BMJ* 1999; **318:** 625.

12 UK Government. Race Relations (Amendment) Act 2000. https://www.legislation.gov.uk/ukpga/2000/34/contents (accessed Sept 9, 2020).

13 Mathur R, Grundy E, Smeeth L. NCRM Working Paper Series: availability and use of UK based ethnicity data for health research. March 27, 2013. http://eprints.ncrm.ac.uk/3040/ (accessed June 5, 2020).

14 Lee C. "Race" and "ethnicity" in biomedical research: how do scientists construct and explain differences in health? *Soc Sci Med* 2009; **68:** 1183–90.

15 Bhopal RS. Terminology and classifications for migrant, ethnic, and racial groups: the centrality of census and population registers. In: Bhopal RS, ed. Migration, ethnicity, race, and health in multicultural societies. Oxford: Oxford University Press, 2013: 33–61.

16 Kabarriti R, Brodin NP, Maron MI, et al. Association of race and ethnicity with comorbidities and survival among patients with COVID-19 at an urban medical center in New York. *JAMA Netw Open* 2020; **3:** e2019795.

17 Yehia BR, Winegar A, Fogel R, et al. Association of race with mortality among patients hospitalized with coronavirus disease 2019 (COVID-19) at 92 US hospitals. *JAMA Netw Open* 2020; **3:** e2018039.

18 Muñoz-Price LS, Nattinger AB, Rivera F, et al. Racial disparities in incidence and outcomes among patients with COVID-19. *JAMA Netw Open* 2020; **3:** e2021892.

19 Harrison EM, Docherty AB, Barr B, et al. Ethnicity and outcomes from COVID-19: the ISARIC CCP-UK prospective observational cohort study of hospitalised patients. *SSRN* 2020; published online June 17. https://doi.org/10.2139/ssrn.3618215 (preprint).

20 Sapey E, Gallier S, Mainey C, et al. Ethnicity and risk of death in patients hospitalised for COVID-19 infection in the UK: an observational cohort study in an urban catchment area. *BMJ Open Respir Res* 2020; **7:** e000644.

21 Apea VJ, Wan YI, Dhairyawan R, et al. Ethnicity and outcomes in patients hospitalised with COVID-19 infection in East London: an observational cohort study. *medRxiv* 2020; published online June 12. https://doi.org/10.1101/2020.06.10.20127621 (preprint).

22 Lassale C, Gaye B, Hamer M, Gale CR, Batty GD. Ethnic disparities in hospitalization for COVID-19: a community-based cohort study in the UK. *medRxiv* 2020; published online May 26. https://doi.org/10.1101/2020.05.19.20106344 (preprint).

23 Raisi-Estabragh Z, McCracken C, Bethell MS, et al. Greater risk of severe COVID-19 in Black, Asian and Minority Ethnic populations is not explained by cardiometabolic, socioeconomic or behavioural factors, or by 25(OH)-vitamin D status: study of 1326 cases from the UK Biobank. *J Public Health (Oxf)* 2020; **42:** 451–60.

24 Smith CW. 2011 Census analysis: what does the 2011 Census tell us about concealed families living in multi-family households in England and Wales? Office for National Statistics. Feb 6, 2014. https://www.ons.gov.uk/peoplepopulationandcommunity/ birthsdeathsandmarriages/families/articles/2011censusanalys is/2014-02-06 (accessed Sept 15, 2020).

25 OpenSAFELY. COVID-19 test results. https://docs.opensafely.org/ en/latest/dataset-sgsscovid/ (accessed Jan 13, 2021).

26 UK Department of Health and Social Care. Coronavirus (COVID-19): scaling up our testing programmes. April 6, 2020. https://www.gov.uk/government/publications/coronavirus-covid-19-scaling-up-testing-programmes/coronavirus-covid-19-scaling-up-our-testing-programmes (accessed Jan 13, 2021).

27 Ministry of Housing, Communities & Local Government. English indices of deprivation 2015: Index of Multiple Deprivation. Sept 30, 2015. https://www.gov.uk/government/statistics/english-indices-of-deprivation-2015 (accessed June 5, 2020).

28 Chisholm J. The Read clinical classification. *BMJ* 1990; **300:** 1092.

29 OpenSAFELY. Codelists. https://codelists.opensafely.org/ (accessed Jan 24, 2021).

30 Lau B, Cole SR, Gange SJ. Competing risk regression models for epidemiologic data. *Am J Epidemiol* 2009; **170:** 244–56.

31 Griffith G, Morris TT, Tudball M, et al. Collider bias undermines our understanding of COVID-19 disease risk and severity. *Nat Commun* 2020: **11:** 5749.

32 Ortiz-Fernández L, Sawalha AH. Genetic variability in the expression of the SARS-CoV-2 host cell entry factors across populations. *Genes Immun* 2020; **21:** 269–72.

33 Maiti AK. The African-American population with a low allele frequency of SNP rs1990760 (T allele) in IFIH1 predicts less IFN-beta expression and potential vulnerability to COVID-19 infection. *Immunogenetics* 2020; **72:** 387–91.

34 Hawkins D. Differential occupational risk for COVID-19 and other infection exposure according to race and ethnicity. *Am J Ind Med* 2020; **63:** 817–20.

35 Laster Pirtle WN. Racial capitalism: a fundamental cause of novel coronavirus (COVID-19) pandemic inequities in the United States. *Health Educ Behav* 2020; **47:** 504–08.

36 Bhala N, Curry G, Martineau AR, Agyemang C, Bhopal R. Sharpening the global focus on ethnicity and race in the time of COVID-19. *Lancet* 2020; **395:** 1673–76.

37 Asimakopoulou K, Hoorens V, Speed E, et al. Comparative optimism about infection and recovery from COVID-19; implications for adherence with lockdown advice. *Health Expect* 2020; **23:** 1502–11.

38 Aleta A, Martín-Corral D, Pastore Y Piontti A, et al. Modelling the impact of testing, contact tracing and household quarantine on second waves of COVID-19. *Nat Hum Behav* 2020; **4:** 964–71.

39 Rubin-Miller L, Alban C, Artiga S, Sullivan S. COVID-19 racial disparities in testing, infection, hospitalization, and death: analysis of Epic patient data. Sept 16, 2020. https://www.kff.org/coronavirus-covid-19/issue-brief/covid-19-racial-disparities-testing-infection-hospitalization-death-analysis-epic-patient-data/ (accessed Sept 20, 2020).

40 Office for National Statistics. Ethnicity facts and figures: English language skills. Feb 7, 2018. https://www.ethnicity-facts-figures.service.gov.uk/uk-population-by-ethnicity/demographics/ english-language-skills/latest (accessed Sept 8, 2020).

41   Tapper J. Low-paid shun COVID tests because the cost of self-isolating is too high. Jan 16, 2021. http://www.theguardian.com/society/2021/jan/16/low-paid-shun-covid-tests-cost-of-self-isolating-too-high (accessed Jan 24, 2021).

42   Windsor-Shellard B, Butt A. Coronavirus (COVID-19) related deaths by occupation, England and Wales: deaths registered between 9 March and 25 May 2020. June 26, 2020. https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/causesofdeath/bulletins/coronaviruscovid19relateddeathsbyoccupationenglandandwales/deathsregisteredbetween9marchand25may2020 (accessed Sept 3, 2020).

43   Mahase E. COVID-19: ethnic minority doctors feel more pressured and less protected than white colleagues, survey finds. *BMJ* 2020; **369:** m2506.

44   Trades Union Congress. BME workers have been asked to "shoulder more risk" during pandemic, says TUC. Sept 15, 2020. https://www.tuc.org.uk/news/bme-workers-have-been-asked-shoulder-more-risk-during-pandemic-says-tuc (accessed Jan 24, 2021).

45   Iacobucci G. COVID-19: PHE review has failed minorities, leaders tell *BMJ. BMJ* 2020; **369:** m2264.

46   Ward H, Atchison C, Whitaker M, et al. Antibody prevalence for SARS-CoV-2 following the peak of the pandemic in England: REACT2 study in 100,000 adults. *medRxiv* 2020; published online Aug 21. https://doi.org/10.1101/2020.08.12.20173690 (preprint, version 2).

47   Nafilyan V, Islam N, Ayoubkhani D, et al. Ethnicity, household composition and COVID-19 mortality: a national linked data study. *medRxiv* 2020; published online Dec 6. https://doi.org/10.1101/2020.11.27.20238147 (preprint, version 3).

48   Ministry of Housing, Communities and Local Government. Ethnicity facts and figures: overcrowded households. Sept 9, 2020. https://www.ethnicity-facts-figures.service.gov.uk/housing/housing-conditions/overcrowded-households/latest (accessed Sept 15, 2020).

49   Bear L, James D, Simpson N, et al. A right to care: the social foundations of recovery from COVID-19. http://eprints.lse.ac.uk/107060/ (accessed Nov 4, 2020).

50   Scientific Pandemic Influenza Group on Behaviours, Environmental and Modelling Group. SPI-B/EMG: MHCLG housing impacts paper. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/923605/s0744-4a-mhclg-housing-impacts-summary-paper.pdf (accessed Oct 30, 2020).

51   Scientific Advisory Group for Emergencies. Factors influencing COVID-19 vaccine uptake among minority ethnic groups, 17 December 2020. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/952716/s0979-factors-influencing-vaccine-uptake-minority-ethnic-groups.pdf (accessed Jan 18, 2021).

52   Bhopal RS, Gruer L, Cezard G, et al. Mortality, ethnicity, and country of birth on a national scale, 2001-2013: a retrospective cohort (Scottish Health and Ethnicity Linkage Study). *PLoS Med* 2018; **15:** e1002515.

53   Scott AP, Timæus IM. Mortality differentials 1991-2005 by self-reported ethnicity: findings from the ONS Longitudinal Study. *J Epidemiol Community Health* 2013; **67:** 743–50.

54   Khunti K, Routen A, Banerjee A, Pareek M. The need for improved collection and coding of ethnicity in health research. *J Public Health (Oxf)* 2020; published online Dec 7. https://doi.org/10.1093/pubmed/fdaa198.

55   Morris S, Stevenson O. Coronavirus: record ethnicity on all death certificates to start building a clearer picture. May 11, 2020. http://theconversation.com/coronavirus-record-ethnicity-on-all-death-certificates-to-start-building-a-clearer-picture-137860 (accessed Aug 24, 2020).

56   Christie B. Scotland introduces record of ethnicity on death certificates. *BMJ* 2012; **344:** e475.

57   London Health Observatory. Missing record: the case for recording ethnicity at birth and death registrations. July, 2003. https://www.kent.ac.uk/chss/docs/Births_Deaths_Reg1.pdf (accessed Aug 24, 2020).