

Ethnicity and risk of death in patients hospitalised for COVID-19 infection in the UK: an observational cohort study in an urban catchment area

Elizabeth Sapey ,^{1,2} Suzy Gallier,³ Chris Mainey ,⁴ Peter Nightingale,⁵ David McNulty,⁴ Hannah Crothers,⁴ Felicity Evison,⁶ Katharine Reeves,⁴ Domenico Pagano,⁷ Alastair K Denniston,⁸ Krishnarajah Nirantharakumar,⁹ Peter Diggle,^{10,11} Simon Ball,¹² All clinicians and students at University Hospitals Birmingham NHS Foundation Trust

To cite: 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 Res* 2020;7:e000644. doi:10.1136/bmjresp-2020-000644

► Additional material is published online only. To view, please visit the journal online (<http://dx.doi.org/10.1136/bmjresp-2020-000644>).

Received 21 May 2020
Revised 12 August 2020
Accepted 13 August 2020

ABSTRACT

Background Studies suggest that certain black and Asian minority ethnic groups experience poorer outcomes from COVID-19, but these studies have not provided insight into potential reasons for this. We hypothesised that outcomes would be poorer for those of South Asian ethnicity hospitalised from a confirmed SARS-CoV-2 infection, once confounding factors, health-seeking behaviours and community demographics were considered, and that this might reflect a more aggressive disease course in these patients.

Methods Patients with confirmed SARS-CoV-2 infection requiring admission to University Hospitals Birmingham NHS Foundation Trust (UHB) in Birmingham, UK between 10 March 2020 and 17 April 2020 were included. Standardised admission ratio (SAR) and standardised mortality ratio (SMR) were calculated using observed COVID-19 admissions/deaths and 2011 census data. Adjusted HR for mortality was estimated using Cox proportional hazard model adjusting and propensity score matching.

Results All patients admitted to UHB with COVID-19 during the study period were included (2217 in total). 58% were male, 69.5% were white and the majority (80.2%) had comorbidities. 18.5% were of South Asian ethnicity, and these patients were more likely to be younger and have no comorbidities, but twice the prevalence of diabetes than white patients. SAR and SMR suggested more admissions and deaths in South Asian patients than would be predicted and they were more likely to present with severe disease despite no delay in presentation since symptom onset. South Asian ethnicity was associated with an increased risk of death, both by Cox regression (HR 1.4, 95% CI 1.2 to 1.8), after adjusting for age, sex, deprivation and comorbidities, and by propensity score matching, matching for the same factors but categorising ethnicity into South Asian or not (HR 1.3, 95% CI 1.0 to 1.6).

Conclusions Those of South Asian ethnicity appear at risk of worse COVID-19 outcomes. Further studies need to establish the underlying mechanistic pathways.

INTRODUCTION

COVID-19 was identified in January 2020¹ and given its designated name by the WHO

Key messages

- There were more admissions from South Asian patients to our hospital than would be expected based on our local population.
- These patients were admitted with a worse severity of COVID-19-related respiratory compromise without a significant delay in presentation and experience a higher level of mortality even when differences in age, sex, deprivation and key comorbidities were taken into account.
- South Asian ethnicity may form another ‘at risk’ population from COVID-19, and further studies are needed to identify any treatable factors to improve outcomes as well as to refine our understanding and communication around non-modifiable risk factors.

in February 2020.² Initial reports from China, Italy and the USA focused on risk factors which predisposed individuals to severe manifestations of infection such as viral pneumonia and adult respiratory distress syndrome (ARDS) requiring critical care support and death, including age, male sex and comorbidities.^{3–6}

A more recent report from the Centers for Disease Control and Prevention in the USA described an early sign of non-Hispanic, black people being disproportionately affected by COVID-19 hospitalisation.⁷ In the UK, Intensive Care National Audit and Research Centre reports^{8–9} described that a higher proportion of patients requiring critical care for COVID-19 were of Asian and black ethnicity compared with pre-COVID-19 historical data (2017–2019) for patients who required critical care for viral pneumonia. Since then, two studies, one from Office for National Statistics analysing over 10 000 deaths in UK¹⁰ and



© Author(s) (or their employer(s)) 2020. Re-use permitted under CC BY. Published by BMJ.

For numbered affiliations see end of article.

Correspondence to
Dr Elizabeth Sapey;
e.sapey@bham.ac.uk



another examining around 5000 deaths from primary care records,¹¹ concluded twofold to threefold high death rates in ethnic minority groups after accounting for important confounders. However, both studies reported mortality at a population level that could be a reflection of higher infection rates in these ethnic groups rather than a higher case fatality rate.

The West Midlands is experiencing a high incidence of COVID-19-associated hospitalisations. This is particularly marked in Birmingham, which has a higher than average percentage of minority ethnic groups, with the 2011 census reporting the following percentages: South Asian 23.4%, black 7.8%, mixed ethnicity 4.1%, others 1.8% and a lower than average white ethnic group at 63%.¹²

University Hospitals Birmingham NHS Foundation Trust (UHB) is one of the largest National Health Service (NHS) Trusts in England, providing direct acute services and specialist care across four hospital sites, including 2.2 million patient episodes per year, with 2750 beds and an expanded ITU capacity of up to 250 beds during the COVID-19 pandemic. UHB constitutes four acute hospital sites following organisational merger in 2018. At present the Queen Elizabeth Hospital Birmingham (QEHB) runs a fully electronic healthcare record (EHR) (PICS; Birmingham Systems, in place since 1999), while the other three sites currently run mixed electronic and paper healthcare records and a shared primary and secondary care record (Your Care Connected). UHB provides secondary care to a diverse population of 1.3 million in Birmingham and Solihull and provides a full range of tertiary services to the West Midlands region.

It was hypothesised that South Asian ethnicity would form a risk factor for the most severe respiratory manifestations of COVID-19 infection, even once age, sex, medical conditions and social deprivation were taken into account, and therefore:

- ▶ They have more admissions than would be expected given the proportion of different ethnic groups within the Birmingham Trust catchment area (based on the last census data).
- ▶ They have worse outcomes (death and/or admission to critical care) from hospitalised COVID-19 viral infection than white ethnic group, even once age, gender, deprivation and comorbidities were accounted for.
- ▶ They have more severe disease on presentation based on a severity score which could not be explained by duration of symptoms, compared with non-South Asian patients.

The study had the following aims:

- ▶ To identify all COVID-19 confirmed patients admitted to UHB hospital within a determined timeframe.
- ▶ To determine the expected and observed admission and death rates given the local population.
- ▶ To determine whether South Asian ethnicity was associated with poor outcomes following hospitalisation with confirmed COVID-19 infection, once potentially confounding factors were considered.

- ▶ To explore if disease presentation was more severe in patients of South Asian ethnicity and if there was any evidence of a delayed presentation to hospital.

METHODS

This retrospective cohort study, using prospectively collected data, was conducted in affiliation with PIONEER, the UK Health Data Research Hub in acute care.

Study population

All patients with a confirmed positive SARS-CoV-2 swab result between 09:00 on 10 March 2020 and 16:00 on 17 April 2020 and who were admitted to UHB at the time of or up to 2 weeks following their first positive SARS-CoV-2 swab test were included. COVID-19 cases were confirmed following a nasopharyngeal and oropharyngeal swab in all cases,¹³ which were processed in accordance with NHS guidance within UHB NHS laboratories.¹⁴ Mortality and (in those alive) patient admission status (discharged and alive, continued admission and alive) were assessed on 12 May 2020.

Data collection and variable definitions

Patient demographics and clinical data were collected from the EHR and from mandatory data sets within the Trust. Clinician-confirmed comorbidities were available from the EHR, the depth of which was enhanced by access to a summary primary care record (Your Care Connected) and further enriched with diagnostic codes derived from previous hospital episodes. The EHR encodes diagnoses using NHS Digital SNOMED CT browser¹⁵ alongside and mapped on to the International Classification of Diseases (ICD)-10 codes,¹⁶ allowing for the presentation and inclusion of historically entered ICD-10 codes. Comorbidities of interest were defined by those associated with poor outcomes from previous publications^{17 18} in order to determine the impact of multimorbidity.¹⁸ The most common clusters of diagnostic categories are listed in [table 1](#). A simple count of comorbidities was undertaken to determine the impact of multimorbidity, as described.¹⁸ English Indices of Deprivation scores were calculated using postcodes from the current data provided by the UK's Ministry of Housing, Communities and Local Government (2019) report.¹⁹ Seven main types of deprivation are considered in the Index of Multiple Deprivation 2019—income, employment, education, health, crime, access to housing and services, and living environment—and these are combined to form the overall measure of multiple deprivation.

Ethnicity was self-reported by the patient or their family members on admission to hospital. Where these data were missing, it was gathered from previous admissions and by reviewing primary and secondary medical records. If this was not available (as was the case in 91 patients), ethnicity was imputed from the modal ethnicity

Table 1 All patients with a confirmed positive SARS-CoV-2 swab during the study period

	All patients	White ethnicity	Asian ethnicity	Black ethnicity
n	2217	1540	410	134
Age in years, median (IQR)	73 (58–84)	77 (66–86)	61 (45–73)	62 (53–79)
Sex, n (%)				
Female	927 (41.8)	639 (41.5)	170 (41.5)	59 (44.0)
Male	1290 (58.2)	901 (58.5)	240 (58.5)	75 (56.0)
Self-reported ethnicity, n (%)				
White	1540 (69.5)	N/A	N/A	N/A
Mixed/multiple	18 (0.8)			
South Asian/South Asian British	410 (18.5)			
Black/African/Caribbean/black British	134 (6.0)			
Other ethnic group	67 (3.0)			
Preferred not to say	22 (1.0)			
Not known	26 (1.2)			
Comorbidity count, n (%)				
None	439 (19.8)	255 (16.6)	114 (27.8)	25 (18.7)
1–2	888 (40.1)	620 (40.3)	155 (37.8)	61 (45.5)
3 or more	890 (40.1)	665 (43.2)	141 (34.4)	48 (35.8)
Morbidities, n (%)				
Hypertension	864 (39.0)	649 (42.1)	126 (30.7)	50 (37.3)
Cerebrovascular disease	268 (12.1)	233 (15.1)	18 (4.4)	10 (7.5)
Atrial fibrillation	464 (20.9)	404 (26.2)	26 (6.3)	20 (14.9)
Ischaemic heart disease, angina, myocardial infarct	546 (24.6)	404 (26.2)	100 (24.4)	24 (17.9)
Diabetes (type 1 and 2)	752 (33.9)	434 (28.2)	197 (48.0)	72 (53.7)
Asthma	439 (19.8)	290 (18.8)	91 (22.2)	34 (25.4)
COPD	376 (17.0)	333 (21.6)	20 (4.9)	14 (10.5)
Interstitial lung disease	49 (2.2)	40 (2.6)	4 (1.0)	2 (1.5)
Chronic kidney disease	511 (23.0)	338 (21.9)	111 (27.1)	42 (31.3)
Any active malignancy	152 (6.9)	124 (8.1)	12 (2.9)	7 (5.2)
Dementia (all types)	326 (14.7)	283 (18.4)	22 (5.4)	10 (7.5)
Obesity	267 (12.0)	174 (11.3)	60 (14.6)	12 (9.0)
English Indices of Deprivation, n (%)				
1 (most deprived)	1003 (45.2)	554 (36.0)	276 (67.3)	89 (66.4)
2	416 (18.8)	320 (20.8)	50 (12.2)	26 (19.4)
3	311 (14.1)	263 (17.1)	29 (7.1)	11 (8.2)
4	230 (10.3)	195 (12.7)	18 (4.4)	4 (3.0)
5 (least deprived)	225 (10.1)	195 (12.7)	20 (4.9)	3 (2.2)
Missing, n (%)	32 (1.4)	13 (0.8)	17 (4.2)	1 (0.8)
Recovered and discharged, n (%)				
Recovered and discharged, n (%)	1052 (47.5)	679 (44.1)	225 (54.9)	76 (56.7)
Remain admitted at point of data census, n (%)				
Remain admitted at point of data census, n (%)	554 (25.0)	411 (26.7)	79 (19.3)	26 (19.4)
Died, n (%)				
Died, n (%)	611 (27.5)	450 (29.2)	106 (25.9)	32 (23.9)
Duration of symptoms prior to admission (in days), n where data were available (% of 567)				
Median (IQR)	7 (3–10)	6 (3–10)	7 (4–10)	7 (3.5–13.5)
Severity of COVID-19 on admission				

Continued



Table 1 Continued

	All patients	White ethnicity	Asian ethnicity	Black ethnicity
Documented, n	736	483	137	53
Mild, n (% of those recorded)	449 (61)	295 (61.1)	79 (57.7)	36 (67.9)
Moderate, n (% of those recorded)	185 (25.1)	134 (27.7)	24 (17.5)	10 (18.9)
Severe, n (% of those recorded)	102 (13.9)	54 (11.2)	34 (24.8)	7 (13.2)
Care escalation to ITU, n (%)	269 (12.1)	133 (8.6)	86 (21.0)	21 (15.7)
LOS in full days, median (IQR)				
LOS for total population	6 (2–12)	6 (2–12)	5 (2–9)	5 (3–11)
LOS for patients discharged	5 (2–11)	6 (2–13)	4.5 (2–8)	5 (3–9)
LOS for patients who died	6 (3–11)	6 (3–12)	5 (2–8)	6 (3–9)

Data are number (percentage) unless otherwise stated.

Ethnicity was self-reported or inferred (see the Methods section). Medical conditions were physician-confirmed and checked against admission and linked primary care notes. English Indices of Deprivation were calculated using postcode. Severity was determined by respiratory oxygen requirements (see the Methods section). Subgroup data are provided for those ethnicities which represented more than 5% of the whole population.

For English Indices of Deprivation, the quintiles were as follows: quintile 1=33.5–78.1; quintile 2=21.7–33.2; quintile 3=14.4–21.5; quintile 4=8.8–14.1; quintile 5=1.4–8.6.

COPD, chronic obstructive pulmonary disease; LOS, length of stay; N/A, not applicable.

of patients with the same surname in the EHR database where possible, as previously described,²⁰ but remained unavailable in 48 patients (see table 1 for missing data). Ethnicity was grouped as per national guidelines.²¹

Severity of COVID-19 on admission

Physician-determined severity of COVID-19 on first admission was categorised using a pragmatic and locally developed score which made use of baseline physiological assessments and oxygen requirements to identify those on admission to hospital who were in need of urgent critical care assessments for respiratory support, and is as follows:

- ▶ Patients were considered to have severe respiratory manifestations of COVID-19 infection if COVID-19 was suspected and the patient required inspired oxygen $\geq 50\%$ to maintain targeted oxygen saturations ($>93\%$ except in the presence of type 2 respiratory failure where the target saturations were 88%–92%) with respiratory pathology thought driven by COVID-19 illness.
- ▶ If not severe, patients were considered to have moderate severity respiratory manifestations of COVID-19 infection if COVID-19 was suspected and the patient required inspired oxygen of >4 L/min or inspired oxygen $>28\%$ to maintain target oxygen saturations.
- ▶ Patients were considered to have mild severity respiratory manifestations of COVID-19 infection if the patient had respiratory symptoms but did not meet the severe or moderate criteria as described above.

Baseline physiological assessments to determine severity of COVID-19 were considered to be those taken within 24 hours either side of the SARS-CoV-2 swab collection time,

of which the earliest available measurement was used. Since not all patients were admitted within 24 hours of their SARS-CoV-2 swab test, and since these assessments are only routinely recorded in the EHR system for QEHB patients, baseline severity scores were only available for a subset of patients (736 of 2217).

To determine if disease severity on admission reflected duration of illness, medical clerking notes were reviewed to determine the duration of symptoms prior to admission. This was available in only a subset of patients (567 of 2217).

Outcomes

The primary outcome was death while in hospital or post discharge until 12 May 2020. For those patients discharged from hospital, primary care records were checked and any patients admitted to hospital with COVID-19 and discharged who had died in the community within the censor period were noted. Those with an ongoing admission were censored on the study end date.

Statistics

Baseline characteristics

Baseline characteristics for the total population and ethnic communities are presented as mean (SD) or median (IQR) for continuous variables and as frequency (percentage) for categorical variables. Ethnic groups were compared by age, sex, comorbidity and severity on presentation.

Standardised admission rate and standardised mortality rate by ethnicity

Ethnicity data for the Birmingham and Solihull area from the 2011 census were used to estimate expected numbers of admissions and deaths for each ethnic category conditional on the observed numbers and sex-specific age distributions of COVID-19 admissions and deaths in UHB. The ratios of observed to expected numbers were calculated to provide standardised admission ratio (SAR) and standardised mortality ratio (SMR) for each ethnic category, and 95%, 99% and 99.9% CIs were obtained using the mid-p exact test.²²

Predictors of mortality

The overall effects of age, sex, ethnicity, comorbidities and admission severity on mortality were tested by univariable analysis. The effect of ethnicity on mortality was then considered adjusting for age, sex, comorbidity counts and deprivation in a multivariable analysis. Cox model was used to derive adjusted HR (aHR) for mortality, defined as death from any cause after COVID-19. Survival time was calculated as the time between the collection of a sample on clinical suspicion to the date of death or study end date and was used for Kaplan-Meier estimates. Multiple parameterisations were tested for patient age, including linear fit, square-root transformation, categorical groupings and natural cubic splines. Categorical variables were fitted for sex, ethnicity, deprivation score quintiles and number of clinically assessed comorbidities. The proportional hazards assumption was tested through correlation of the scaled Schoenfeld residuals with survival time, with hypothesis tests for independence.²³ Models were then tested using propensity score matching,²⁴ where age, sex, social deprivation index and comorbidities were matched and ethnicity was treated as a dichotomous variable. Models were fitted in R V.3.6.3 using the survival package and twang package in 'r' (www.r-project.org).²⁵ All p values are reported exactly and no corrections were made for multiple comparisons unless stated.

Patient and public involvement

Three hundred and two patients and public members were consulted as to the use of health data to improve the care for people with acute, unplanned illness. A group of patients recovering from COVID-19 specifically were asked and supported the use of routinely collected health data to investigate the relationship between poor outcomes and ethnicity. A working group of staff and patients from black and Asian minority ethnic groups discussed the results and how they should be disseminated.

RESULTS

The study analysed 2217 consecutive patients admitted to UHB with a swab-proven diagnosis of COVID-19. A modified Consolidated Standards of Reporting Trials diagram

(online supplementary figure S1) and a summary of the demographic and clinical characteristics on admission (table 1) are shown. Most patients (n=2132) had a Birmingham postcode. The majority of patients were male (1290 of 2217; 58.2%) and white (1540 of 2217; 69.5%). High levels of comorbidity were identified across all ethnic groups (with 40.1% of patients having three or more comorbidities).

Of note, up to 50% had missing morbidity data and 30% had missing ethnicity data using the secondary care records for the COVID-19-related admission alone, but >96% of ethnicity data and all available morbidity data were resolved through review of primary and secondary care records.

Comparison of baseline characteristics between ethnicities

South Asian patients, when compared with white patients, were younger (median age 61 years vs 77 years, $p<0.001$), had a higher prevalence of diabetes mellitus (48.1% vs 28.2%, $p<0.001$) but lower prevalence of dementia (5.4% vs 18.9%), chronic obstructive pulmonary disease (COPD) (4.9% vs 21.6%), atrial fibrillation (6.3% vs 26.2%) and cerebrovascular disease (4.4% vs 15.1%) ($p<0.0001$ for all). Similar to the South Asian patients, black ethnic population were younger (median age 62 years), had a higher prevalence of diabetes mellitus (53.7%) but lower prevalence of other key comorbidities than the white population. Two-thirds of the South Asian and black ethnicity came from the most deprived quintile of deprivation compared with 36% from the white patients.

Across the course of data collection, the proportion of patients presenting from different ethnic groups was relatively stable (online supplementary figure S2), suggesting no differential transmission related to ethnic group within the location and timeframe studied.

SAR in different ethnicities

Age and sex SAR (95% CI) for South Asian women was 74% higher (SAR 1.7, 95% CI 1.5 to 2.0) and for South Asian men 63% higher (SAR 1.6, 95% CI 1.4 to 1.9) than the standard population (see online supplementary table S1). In contrast white patients were less likely to be admitted in comparison with the standard population (white women: 0.8 (0.8–0.9); white men: 0.9 (0.8–0.9)). Admission rates were similar to expected rates of the standard population for black ethnicity.

Severity of COVID-19 at presentation in different ethnicities

In 736 patients admitted directly to the QEHB site, the severity of COVID-19 recorded on admission identified 185 (25.1%) classified as moderate and 102 (13.9%) classified as severe. In this data set a higher proportion of South Asians than whites were assessed to have severe disease on presentation (34 of 137 (24.8%) vs 54 of 483 (11.2%), $p<0.0001$) online supplementary table

Table 2 Comparison of demographics of patients who met the primary endpoint of death up to and including 12 May 2020

	Patients currently alive (admitted or discharged)	Patients discharged and alive	Patients still admitted/receiving inpatient care and alive	Patients who died	P value comparing all alive patients and patients who died
n	1448	1372	76	769	
Age, IQR and range	68 (52–80)	68 (52–80)	66 (54–76)	80 (71–887)	<0.001
Sex, n (%) (female)	637 (44.0)	616 (44.9)	21 (27.6)	290 (37.7)	0.004
Self-reported ethnicity, n (%)					<0.001
White	958 (66.2)	906 (66.0)	52 (68.4)	582 (75.7)	
Mixed/multiple	10 (0.7)	10 (0.7)	0 (0.0)	8 (1.0)	
South Asian/South Asian British	290 (20.0)	275 (20.0)	15 (19.7)	120 (15.6)	
Black/African/Caribbean/black British	94 (6.5)	92 (6.7)	2 (2.6)	40 (5.2)	
Other ethnic group	60 (4.1)	53 (3.9)	7 (9.2)	7 (0.9)	
Preferred not to say	16 (1.1)	16 (1.2)	0 (0.0)	6 (0.8)	
Unknown	20 (1.4)	20 (1.5)	0 (0.0)	6 (0.8)	
Comorbidity count, n (%)					<0.0001
None	377 (26.0)	362 (26.4)	15 (19.7)	62 (8.1)	
1 or 2	593 (41.0)	555 (40.5)	38 (50.0)	295 (38.4)	
3 or more	478 (33.0)	455 (33.2)	23 (30.3)	412 (53.6)*	
Morbidities					
Hypertension	490 (33.8)	457 (33.3)	33 (43.4)	374 (48.6)	<0.001
Cerebrovascular disease	128 (8.8)	125 (9.1)	3 (3.9)	140 (18.2)	<0.001
Atrial fibrillation	236 (16.3)	228 (16.6)	8 (10.5)	228 (29.6)	<0.001
Ischaemic heart disease, angina, myocardial infarct	286 (19.8)	276 (20.1)	10 (13.2)	260 (33.8)	<0.001
Diabetes (type 1 and 2)	459 (31.7)	431 (31.4)	28 (36.8)	293 (38.1)	0.002
Asthma	296 (20.4)	280 (20.4)	16 (21.1)	143 (18.6)	0.299
COPD	202 (14.0)	194 (14.1)	8 (10.5)	174 (22.6)	<0.001
Interstitial lung disease	24 (1.7)	23 (1.7)	1 (1.3)	25 (3.3)	0.015
Chronic kidney disease	267 (18.4)	253 (18.4)	14 (18.4)	244 (31.7)	<0.001
Any active malignancy	87 (6.0)	84 (6.1)	3 (3.9)	65 (8.5)	0.030
Dementia (all types)	131 (9.0)	129 (9.4)	2 (2.6)	195 (25.4)	<0.001
Obesity	176 (12.2)	163 (11.9)	13 (17.1)	91 (11.8)	0.825
English Indices of Deprivation, n (%)					0.848
1	667 (46.1)	633 (46.1)	34 (44.7)	336 (43.7)	
2	271 (18.7)	253 (18.4)	18 (23.7)	145 (18.9)	
3	200 (13.8)	191 (13.9)	9 (11.8)	111 (14.4)	
4	150 (10.4)	146 (10.6)	4 (5.3)	80 (10.4)	
5	139 (9.6)	128 (9.3)	11 (14.5)	86 (11.2)	
Missing	21 (1.5)	21 (1.5)	0 (0.0)	11 (1.4)	
Duration of symptoms prior to admission (in days), n where data were available (%)	245 (16.9)	223 (16.3)	22 (28.9)	110 (14.3)	
Median (IQR)	9 (2–16)	9 (3–16)	7 (5–12)	7 (3–15)	
Severity of COVID-19 on admission, n (%)	n=489	n=453	n=36	n=247	<0.001
Mild	348 (71.2)	338 (74.6)	10 (27.8)	101 (40.8)	
Moderate	91 (18.6)	81 (17.8)	10 (27.8)	94 (38.1)	
Severe	50 (10.2)	34 (7.5)	16 (44.4)	52 (21.1)†	

Data are number (percentage) unless otherwise stated.

Medical conditions were self-reported and checked against admission and linked primary care notes.

Groups are compared using χ^2 analysis, except for pregnancy where Fisher's exact test was used, and age distribution where Kruskal-Wallis was used.

Of note, p values compared all patients currently alive (in patients or discharged) versus those who had died.

For English Indices of Deprivation, the quintiles were as follows: quintile 1=33.5–78.1; quintile 2=21.7–33.2; quintile 3=14.4–21.5; quintile 4=8.8–14.1; quintile 5=1.4–8.6.

*Increased deaths with 3 or more ($p<0.0001$) comorbidities on post-hoc analysis.

†Increased deaths in moderate and severe cases versus mild cases on post-hoc analysis ($p<0.0001$) for comparisons with patients discharged alive and currently alive.

COPD, chronic obstructive pulmonary disease.

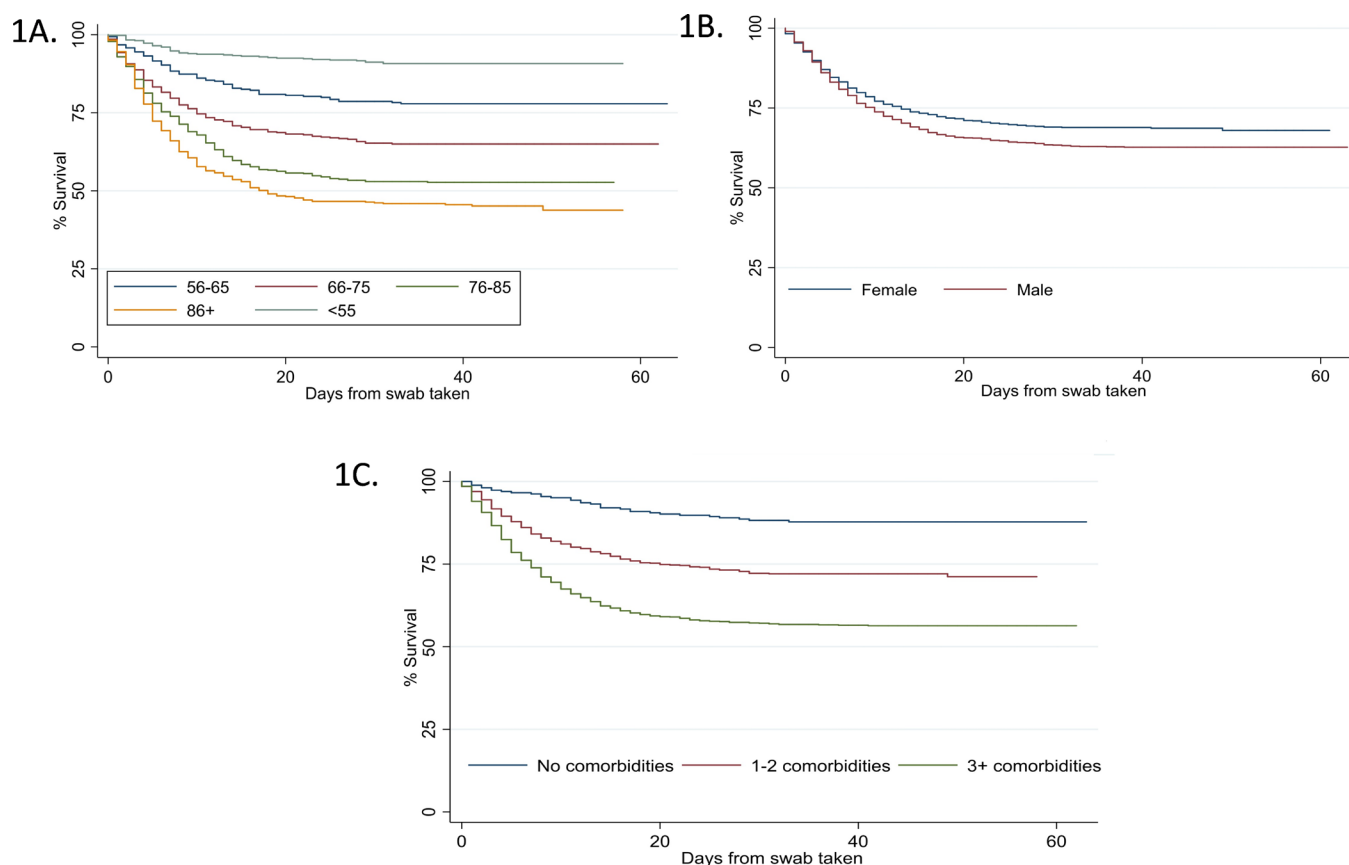


Figure 1 Kaplan-Meier estimates of survival for COVID-19-positive patients. Data compared survival status of patients by age (A, $p < 0.001$), sex (B, $p < 0.001$) and simple comorbidity counts as listed in table 1 (C, $p < 0.001$). Comparison done using the log-rank test.

S2. A higher proportion of South Asians than whites were admitted to intensive care unit (86 of 410 (21.0%) vs 133 of 1540 (8.6%), $p < 0.001$). There were no differences in the duration of symptoms prior to admission by ethnic group (median: South Asian=7 days, white=6 days, black=7 days, $p = 0.40$).

Characteristics of patients who died

On 12 May 2020, 769 of 2217 (34.6%) patients had died in hospital or following discharge. These patients were older, more likely to be male, white, have multiple comorbidities, and in the 736 patients admitted directly to the QEHB site, in whom data were available, more likely to have moderate or severe disease on admission, in comparison with all other groups of patients (table 2).

In those in whom data were available, people who died had a shorter duration of symptoms prior to admission compared with those who were still alive (median 7 days (3–36) vs 9 (5–36) days). In all in whom data were available, the duration of symptoms prior to admission did not relate to disease severity on admission ($p = 0.46$). The relationships between survival from diagnosis and age, gender and number of comorbidities are further illustrated in figure 1.

Ethnic group outcome analysis

SMR between different ethnic groups

In comparison with the expected number of deaths based on Birmingham and Solihull 2011 census data age and sex structure, there were significantly more South Asian women and men who died with a positive COVID-19 swab than would be expected: SMR (95% CI): Asian women 1.9 (1.4 to 2.6) and Asian men 1.7 (1.4 to 2.1). In contrast fewer white women and men died than would be expected: SMR (95% CI): white women 0.9 (0.8 to 1.0) and white men 0.9 (0.8 to 1.0) (see online supplementary table S1). For those of black ethnicity, death rates were not different from the expected rates in the standard population.

Survival curves for mortality in different ethnicities

An age-adjusted Kaplan-Meier showed that although there were no differences in age-adjusted survival in white and black ethnicities, patients from Asian ethnic groups were less likely to survive (see figure 2).

Multivariable analysis

In a multivariable Cox regression model adjusted for age, sex, comorbidity counts and deprivation, South Asian ethnicity (aHR 1.4, 95% CI 1.2 to 1.8) was associated with

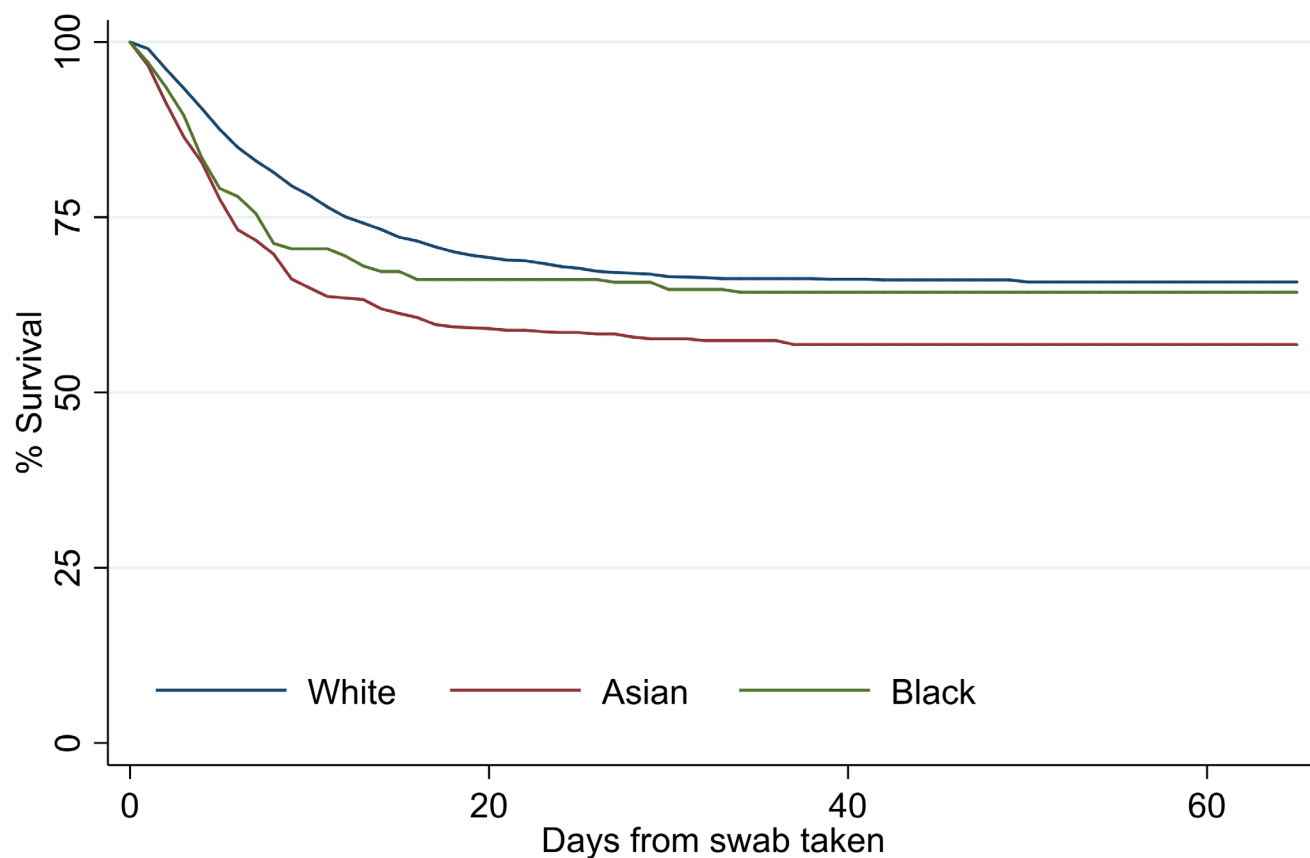


Figure 2 Age-adjusted Kaplan-Meier estimates of survival for different ethnic groups of COVID-19-positive patients. Data compared age-adjusted survival status of patients by ethnicity ($p < 0.001$ using the age ranges as listed in figure 1A). Comparison done using the log-rank test.

a significantly higher risk of death. Within the limits of the power of the study, there was absence of a significant difference in survival for black ethnicity compared with the white population (aHR 1.1, 95% CI 0.8 to 1.5). In addition to this we found age z-score (aHR 2.4, 95% CI 1.8 to 3.2) and comorbidities (1–2 comorbidities, aHR 1.7, 95% CI 1.3 to 2.2; 3 or more comorbidities, aHR 2.3, 95% CI 1.7 to 3.0) as significant predictors. We also found a significant effect of sex (aHR for male sex, 1.3, 95% CI 1.1 to 1.5), with the interaction between age and sex suggesting an amplified risk in men with increasing age (table 3). The main effects of age and multiple comorbidities were modified by an interaction (aHR for age interaction with 1–2 comorbidities, 0.9, 95% CI 0.7 to 1.2; age interaction with 3+ comorbidities, 0.7, 95% CI 0.5 to 0.9) that attenuated the relative impact of increasing comorbidity at advanced age. Of note, the HR of 1.4 (1.2–1.8) did not change (1.4, 1.2–1.8) if the interactions were removed from the model.

To further test the assumption that South Asian ethnicity was associated with worse outcomes even when

comorbidities, age and sex were considered, propensity score matching was conducted, matching for the same factors in the Cox regression but categorising ethnicity into South Asian or not. With propensity score matching the HR for risk of death was 1.3 (1.0–1.6) for South Asian patients compared with non-South Asian patients. When comorbidities were added independently (rather than as a count) using propensity score matching, South Asian ethnicity was still associated with a significantly higher risk of death (HR 1.4, 95% CI 1.1 to 1.7).

DISCUSSION

This is the first study to specifically describe the impact of South Asian ethnicity on the outcome of COVID-19 infection using highly characterised and accurate primary and secondary data from patients admitted to hospital in the UK. South Asians were significantly younger and twice likely to have diabetes than white patients, and when accounted for the age structure of the local population had a high admission and death rate. South Asians

**Table 3** Adjusted HR of risk factors for mortality

Predictors	Estimates	CI	P value
Age (z-score)	2.4	1.8 to 3.2	<0.001
White	Reference		
Asian	1.4	1.2 to 1.8	0.001
Black	1.1	0.8 to 1.5	0.536
Mixed	1.9	0.9 to 3.9	0.072
Unknown	0.8	0.4 to 1.9	0.676
Other	0.6	0.3 to 1.3	0.226
Not stated	1.2	0.6 to 2.8	0.603
Female	Reference		
Male	1.3	1.1 to 1.5	0.014
Deprivation quintile 1	Reference		
Deprivation quintile 2	0.9	0.7 to 1.0	0.126
Deprivation quintile 3	0.9	0.7 to 1.1	0.439
Deprivation quintile 4	0.8	0.7 to 1.1	0.183
Deprivation quintile 5	0.9	0.7 to 1.2	0.593
Deprivation quintile missing	1.0	0.5 to 1.8	0.930
No comorbidity	Reference		
Comorbidity group 1–2	1.7	1.3 to 2.2	<0.001
Comorbidity group 3+	2.3	1.7 to 3.0	<0.001
Age (z-score) × sex (male)	1.2	1.0 to 1.5	0.055
Age (z-score) × comorbidity group 1–2	0.9	0.7 to 1.2	0.529
Age (z-score) × comorbidity group 3+	0.7	0.5 to 0.9	0.015

Multivariable Cox regression model including age (z-score), ethnicity, sex and comorbidity count as covariates. Adjusted HRs along with their CI are presented.

were also more likely to present with severe symptoms, but with no difference in the duration of symptoms and more likely to be admitted to ITU. Importantly, after adjusting for age, deprivation and multiple comorbidities, the effect of South Asian ethnicity on mortality was 42% higher.

Our study is in line with two population-based studies^{10 11} where South Asians were found to be at increased risk of death, with similar effect size with the study¹¹ that considered similar covariates for adjusting (aHR 1.4 (1.2–1.8) vs 1.6 (1.4–1.8)). However, the findings of the two population-based studies could reflect infection rates rather than case fatality rates, and therefore differentiating this is important from a public health and research perspective. If increased deaths in South Asian patients reflected high infection rates, then our focus should be on looking at barriers and emphasising the need for adherence to current social distancing guidelines. If this reflects an increased susceptibility to poorer outcomes from SARS-CoV-2 infection, we need to urgently understand the reasons for severity of infection

and mitigate risk or develop targeted treatments. While no firm conclusion can be drawn from the current data set, our study potentially supports the latter, with South Asian patients more likely to be admitted, more likely to present with severe symptoms and have an increased risk of mortality. This supports the call by the government and research community for urgent research on the reasons underpinning these observations. Our study was not sufficiently powered to report on other ethnic groups, particularly those of black ethnicity, and therefore the findings among these ethnic groups should be interpreted cautiously.

The excess age-adjusted mortality in COVID-19 is not solely attributable to a range of cardiovascular and metabolic risk factors that are over-represented in this ethnic group. In sensitivity analysis the HR for South Asian ethnicity using both Cox regression and propensity score matching was stable when ‘number of co-morbidities’ was exchanged for the presence of specific comorbidities including diabetes mellitus and hypertension. To place this in context, the effect of South Asian on mortality is significantly less than the effect of one or more comorbidities (present in 80.2% of all admitted patients) and approximates to the effect of ageing 10 years in the white population. In this study we did not observe an independent signal related to higher deprivation levels²⁶ and poor outcomes in contrast to the population-based study,¹¹ suggesting deprivation is likely to be related to high infection rates and thereby high mortality, rather than high rates of severe infection leading to increased mortality.

It is notable that in the subanalysis of patients admitted to QEHB, where we were able to immediately integrate a COVID-19-specific assessment into our EHR, South Asian patients appear to present with more severe disease, but there was no difference in the duration of symptoms prior to admission, suggesting disease severity was not caused simply by delayed presentation to medical services or differences in health service utilisation (although this cannot be fully excluded, given the unknown burden of COVID-19 in the community). Indeed, when comparing those still alive and those who had died at the end of the study, the patients who had died had a shorter history of symptoms prior to admission, suggesting a different disease course.

A significantly higher rate of admission to ITU in the South Asian ethnic group could relate to this more severe disease at presentation. It may also relate to patient-level differences in joint decision-making regarding ITU treatment, in patients who have higher levels of specific comorbidities such as dementia and COPD, groups that are significantly over-represented in the white ethnic group, which was also significantly older.

The limitations in our overall analysis need to be considered, specifically that 5% of patients remain in hospital at the time of the data lock and more patients have been admitted, so our findings will evolve. Since the proportion of patients presenting from different ethnic groups



was stable across the course of data collection, any consequence for our main conclusion on the mortality risk in South Asians admitted to hospital is likely to be small. Data on the admission severity scoring did not include all patients, which is a limitation. This limitation reflects the real-world response within a global pandemic which includes designing a score to inform care escalation decisions and updating the UHB electronic health record to capture this information during the first wave of patient admissions. However, these data suggest further exploration of severity of disease at presentation is warranted.

It is also important to acknowledge that standardised admission and mortality ratios from Birmingham and Solihull use the most recent census data, but that these are from 2011. Estimates of the contemporary age structure do not however suggest a need to significantly qualify these findings. The UK has not undertaken widespread screening or diagnosis of patients in the community, and we are therefore unable to comment on the natural history of COVID-19 prior to admission to secondary care, irrespective of ethnic group. This testing regimen is likely to evolve with the development of capacity and methodology and will provide a more complete picture of COVID-19. A description of disease in the community will help build a clearer understanding of the apparent excess mortality following admission, for which there remain a number of possible explanations. A limitation of this study (and, arguably, of any observational study) is that it cannot exclude the possibility that another, unmeasured variable could account for the ethnicity effect described here. The assessment of comorbidities does not reflect the degree of severity of the condition, nor disease treatment or control, and the assessment of social deprivation might impact on chronic disorders. There is also the possibility that differences in work or home living arrangements might impact on potential transmission and this may be different across ethnic communities. However, a real strength has been the ability to study a highly curated and complete data set, without the inherent issues of significant undercoding seen with morbidity and ethnicity data when using a secondary care data set.

The biological basis of any difference in outcome can only be speculated on at present. There are reported differences in outcomes for non-white ethnic groups from ARDS even after adjusting for sex, age, disease severity, type of hospital and median household income.^{27 28} The worst clinical manifestations of COVID-19 appear to be associated with a cytokine storm syndrome. Here a hypercytokinaemia is seen,⁴ with predictors of mortality reflecting a virally induced inflammatory state which can be assessed using a scoring system including validated clinical laboratory tests.²⁹ Candidate genes associated with ARDS have been identified in bioinformatic analyses, with a strong predominance of inflammatory pathways, including reactive oxygen species, innate immunity-related inflammation and endothelial vascular signalling pathways.³⁰ Ethnicity may influence cytokine gene polymorphisms and inflammatory profiles following specific challenges,³¹ with some

ethnic groups more prone to a heightened inflammatory response. Of note, socioeconomic factors might also impact on inflammatory pathways and gene expression.³² These factors remain poorly understood, and were a priority for our patient and public involvement group who were consulted for this study, and there is an urgent need to understand the genomic and associated phenomic and socioeconomic characteristics of patients who are susceptible or resistant to the severe manifestations of COVID-19 to understand this further.

Although our study includes only one NHS Foundation Trust, it covers an ethnically diverse contiguous population of 1.3 million people for which it is the sole provider of adult acute secondary care across four hospital sites. This provides for continuity of data, clinical protocols and access to therapy. The immediate availability of access to an electronic representation of a primary care record to support the care of admitted patients also supports the integrity of data collection, the quality of which might otherwise be more limited.³³

Our findings, which describe and quantify the risk of COVID-19 in the South Asian population, are relevant to national policy and to understanding the underlying biological mechanisms in 'at risk' populations. Future studies will extend our observations and explore underlying epidemiology and biological mechanisms, to improve interventions based in the community, the emergency department, ward and ITU. Perhaps most importantly our findings inform the UK's national discussion on at 'at risk' groups and the ensuing fear arising from uncertainty.

Author affiliations

¹PIONEER HDR-UK Hub, University of Birmingham, Birmingham, UK

²Acute Medicine, University Hospitals Birmingham NHS Foundation Trust, Birmingham, UK

³PIONEER Technical Director, Health Informatics, University Hospitals Birmingham NHS Foundation Trust, Birmingham, UK

⁴Health Informatics, University Hospitals Birmingham NHS Foundation Trust, Birmingham, UK

⁵Statistics, University Hospitals Birmingham NHS Foundation Trust, Birmingham, UK

⁶Department of Informatics, University Hospitals Birmingham NHS Foundation Trust, Birmingham, UK

⁷Cardiothoracic Surgery, University Hospitals Birmingham, Birmingham, UK

⁸INSIGHT: HDRUK Health Data Research Hub for Eye Health, University Hospitals Birmingham NHS Foundation Trust, Birmingham, UK

⁹Public Health, University of Birmingham, Birmingham, UK

¹⁰Faculty of Health and Medicine, Lancaster University, Lancaster, UK

¹¹Epidemiology and Population Health, University of Liverpool, Liverpool, UK

¹²HDR-UK Midlands Physical Site, University Hospitals Birmingham, Birmingham, UK

Acknowledgements We would like to acknowledge the contribution of all staff, key workers, patients and the community who have supported our hospitals and the wider NHS at this time. This work uses data provided by patients and collected by the NHS as part of their care and support.

Collaborators The manuscript was prepared and submitted on behalf of all of the staff and students at UHB NHS Foundation Trust, with particular thanks to medical students Lylah Irshad, Maxim Harris and Theodore Nabavi, and health informatics A Kolesnyk, M Ahmed, A Liaqat, Tanya Pankhurst, Jamie Coleman, Chirag Dave, Khaled ElFandi, Rifat Rashid and Paul Cockwell.

Contributors ES designed the study, collated the data, performed some analyses and wrote the manuscript. SG collated the data, performed the analysis and wrote the manuscript. CM, DM, HC, FE and KR performed the statistical analysis. PN performed the statistical analysis and helped write the manuscript. DP assisted with the design of the study and manuscript preparation. PD assisted with analysis. KN and AKD assisted with analysis and manuscript writing. SB designed the study, oversaw data collection and helped write the manuscript. All authors amended the manuscript and approved the final version. Data were curated and analysed on behalf of all clinicians at UHB.

Funding This work was supported by PIONEER, the Health Data Research UK (HDR-UK) Health Data Research Hub in acute care. HDR-UK is an initiative funded by the UK Research and Innovation, Department of Health and Social Care (England) and the devolved administrations, and leading medical research charities.

Competing interests SB reports funding support from the HDR-UK. KR reports funding support from the NIHR. ES reports funding support from HDR-UK, MRC, Wellcome Trust, NIHR and British Lung Foundation. KN reports funding from MRC, Wellcome Trust, NIHR, Vifor and AstraZeneca. AKD reports funding from HDR-UK, Wellcome Trust and Fight for Sight.

Patient and public involvement Patients and/or the public were involved in the design, or conduct, or reporting, or dissemination plans of this research. Refer to the Methods section for further details.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request. To facilitate knowledge in this area, the anonymised participant data and a data dictionary defining each field will be available to others through application to PIONEER via the corresponding author.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution 4.0 Unported (CC BY 4.0) license, which permits others to copy, redistribute, remix, transform and build upon this work for any purpose, provided the original work is properly cited, a link to the licence is given, and indication of whether changes were made. See: <https://creativecommons.org/licenses/by/4.0/>.

ORCID iDs

Elizabeth Sapey <http://orcid.org/0000-0003-3454-5482>

Chris Mainey <http://orcid.org/0000-0002-3018-6171>

REFERENCES

- Phelan AL, Katz R, Gostin LO. The novel coronavirus originating in Wuhan, China: challenges for global health governance. *JAMA* 2020;323:709–10.
- World Health Organisation. Novel Coronavirus(2019-nCoV) Situation Report - 11, 2020. Available: https://www.who.int/docs/default-source/coronavirus/situation-reports/20200131-sitrep-11-nCoV.pdf?sfvrsn=de7c0f7_4
- Chen T, Wu D, Chen H, *et al*. Clinical characteristics of 113 deceased patients with coronavirus disease 2019: retrospective study. *BMJ* 2020;368:m1091.
- Huang Y, Yang R, Xu Y, *et al*. Clinical characteristics of 36 non-survivors with COVID-19 in Wuhan, China. *medRxiv* 2020. doi:10.1101/2020.06.17.20131433. [Epub ahead of print: 22 Apr 2020].
- Onder G, Rezza G, Brusaferro S. Case-Fatality rate and characteristics of patients dying in relation to COVID-19 in Italy. *JAMA* 2020. doi:10.1001/jama.2020.4683. [Epub ahead of print: 23 Mar 2020].
- Richardson S, Hirsch JS, Narasimhan M, *et al*. Presenting characteristics, comorbidities, and outcomes among 5700 patients hospitalized with COVID-19 in the new York City area. *JAMA* 2020. doi:10.1001/jama.2020.6775. [Epub ahead of print: 22 Apr 2020].
- Centers of Disease Control and Prevention. Non Hispanic black people disproportionately affected by COVID-19 hospitalizations in CDC data, 2020. Available: <https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/>
- Intensive Care National Audit and Research Centre. ICNARC report on COVID-19 in critical care 10th April, 2020. Available: <https://www.icnarc.org/Our-Audit/Audits/Cmp/Reports>
- Centre. ICNAR. ICNARC report on COVID-19 in critical care 17th April, 2020. Available: <https://www.icnarc.org/Our-Audit/Audits/Cmp/Reports>
- Office for National Statistics. Coronavirus (COVID-19) related deaths by ethnic group, England and Wales: 2 March 2020 to 10 April 2020, 2020. Available: <https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/articles/coronavirusrelateddeathsbyethnicgroupenglandandwales/2march2020to10april2020>
- OpenSafely Collaborative. OpenSAFELY: factors associated with COVID-19-related Hospital death in the linked electronic health records of 17 million adult NHS patients, 2020. Available: <https://opensafely.org/outputs/2020/05/covid-risk-factors/>
- Office for National Statistics. CT0430_2011 Census - Sex by age by country of birth by ethnic group - Birmingham LA and Solihull LA combined, 2011. Available: <https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/adhocs/004295ct04302011censussexbyagebycountryofbirthbyethnicgroupbirminghamlaandsolihulllacombined>
- Carver K, Jones N. Comparative accuracy of oropharyngeal and nasopharyngeal swabs for diagnosis of COVID-19. Oxford COVID-19 evidence service team centre for evidence-based medicine, Nuffield department of primary care health sciences University of Oxford, 2020. Available: <https://www.cebm.net/covid-19/comparative-accuracy-of-oropharyngeal-and-nasopharyngeal-swabs-for-diagnosis-of-covid-19/>
- NHS England. Guidance and standard operating procedure COVID-19 virus testing in NHS laboratories, 2020. Available: <https://www.england.nhs.uk/coronavirus/wp-content/uploads/sites/52/2020/03/guidance-and-sop-covid-19-virus-testing-in-nhs-laboratories-v1.pdf>
- NHS Digital. The NHS digital SNOMED CT Browser, 2020. Available: <https://termbrowser.nhs.uk/>
- Organisation. WH. Icd-10 online versions, 2019. Available: <https://www.who.int/classifications/icd/icdonlineversions/en/>
- Caramelo F, Ferreira N, Oliveiros B. Estimation of risk factors for COVID-19 mortality - preliminary results. *medRxiv* 2020.
- Guan W-jie, Liang W-hua, Zhao Y, *et al*. Comorbidity and its impact on 1590 patients with COVID-19 in China: a nationwide analysis. *Eur Respir J* 2020;55:2000547.
- Ministry of Housing Calg. The English indices of deprivation 2019 technical report, 2019. Available: https://dera.ioe.ac.uk/34259/1/loD2019_Technical_Report.pdf
- Gagne F. *Discriminant function analysis. Biochemical Ecotoxicology: principles and methods*, 2014.
- Office for National Statistics. Ethnic group, National identity and religion. Measuring equality: a guide for the collection and classification of ethnic group, National identity and religion data in the UK, 2016. Available: <https://www.ons.gov.uk/methodology/classificationsandstandards/measuringequality/ethnicgroupnationalidentityandreligion> [Accessed 20 Mar 2020].
- Dean AG, Sullivan KM, Soe MM. OpenEpi: open source epidemiological statistics for public health, 2013. Available: www.OpenEpi.com
- Schoenfeld D. Chi-squared goodness-of-fit tests for the proportional hazards regression model. *Biometrika* 1980;67:145–53.
- Austin PC. An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behav Res* 2011;46:399–424.
- Therneau TM. *Package survival*, 2020. Available: <https://cran.r-project.org/web/packages/survival/survival.pdf>
- Khunti K, Singh AK, Pareek M, *et al*. Is ethnicity linked to incidence or outcomes of covid-19? *BMJ* 2020;369:m1548.
- Erickson SE, Shlipak MG, Martin GS, *et al*. Racial and ethnic disparities in mortality from acute lung injury. *Crit Care Med* 2009;37:1–6.
- Bime C, Poongkunran C, Borgstrom M, *et al*. Racial differences in mortality from severe acute respiratory failure in the United States, 2008–2012. *Ann Am Thorac Soc* 2016;13:2184–9.
- Mehta P, McAuley DF, Brown M, *et al*. COVID-19: consider cytokine storm syndromes and immunosuppression. *Lancet* 2020;395:1033–4.
- Lynn H, Sun X, Casanova N, *et al*. Genomic and genetic approaches to deciphering acute respiratory distress syndrome risk and mortality. *Antioxid Redox Signal* 2019;31:1027–52.
- Hoffmann SC, Stanley EM, Cox ED, *et al*. Ethnicity greatly influences cytokine gene polymorphism distribution. *Am J Transplant* 2002;2:560–7.
- Thames AD, Irwin MR, Breen EC, *et al*. Experienced discrimination and racial differences in leukocyte gene expression. *Psychoneuroendocrinology* 2019;106:277–83.
- Iqbal G, Gumber A, Johnson MRD, *et al*. Improving ethnicity data collection for health statistics in the UK, 2009. Available: <https://diversityhealthcare.imedpub.com/improving-ethnicity-data-collection-for-health-statistics-in-the-uk.php?aid=2045>