Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.

Patient factors and temporal trends associated with COVID-19 in-hospital mortality in England: an observational study using administrative data

Annakan V Navaratnam, William K Gray, Jamie Day, Julia Wendon, Tim W R Briggs

Summary

Background Analysis of the effect of COVID-19 on the complete hospital population in England has been lacking. Our aim was to provide a comprehensive account of all hospitalised patients with COVID-19 in England during the early phase of the pandemic and to identify the factors that influenced mortality as the pandemic evolved.

Methods This was a retrospective exploratory analysis using the Hospital Episode Statistics administrative dataset. All patients aged 18 years or older in England who completed a hospital stay (were discharged alive or died) between March 1 and May 31, 2020, and had a diagnosis of COVID-19 on admission or during their stay were included. In-hospital death was the primary outcome of interest. Multilevel logistic regression was used to model the relationship between death and several covariates: age, sex, deprivation (Index of Multiple Deprivation), ethnicity, frailty (Hospital Frailty Risk Score), presence of comorbidities (Charlson Comorbidity Index items), and date of discharge (whether alive or deceased).

Findings 91 541 adult patients with COVID-19 were discharged during the study period, among which 28 200 (30·8%) in-hospital deaths occurred. The final multilevel logistic regression model accounted for age, deprivation score, and date of discharge as continuous variables, and sex, ethnicity, and Charlson Comorbidity Index items as categorical variables. In this model, significant predictors of in-hospital death included older age (modelled using restricted cubic splines), male sex (1·457 [1·408–1·509]), greater deprivation (1·002 [1·001–1·003]), Asian (1·211 [1·128–1·299]) or mixed ethnicity (1·317 [1·080–1·605]; *vs* White ethnicity), and most of the assessed comorbidities, including moderate or severe liver disease (5·433 [4·618–6·392]). Later date of discharge was associated with a lower odds of death (0·977 [0·976–0·978]); adjusted in-hospital mortality improved significantly in a broadly linear fashion, from 52·2% in the first week of March to 16·8% in the last week of May.

Interpretation Reductions in the adjusted probability of in-hospital mortality for COVID-19 patients over time might reflect the impact of changes in hospital strategy and clinical processes. The reasons for the observed improvements in mortality should be thoroughly investigated to inform the response to future outbreaks. The higher mortality rate reported for certain ethnic minority groups in community-based studies compared with our hospital-based analysis might partly reflect differential infection rates in those at greatest risk, propensity to become severely ill once infected, and health-seeking behaviours.

Funding None.

Copyright © 2021 Elsevier Ltd. All rights reserved.

Lancet Respir Med 2021; 9: 397–406

Published Online February 15, 2021 https://doi.org/10.1016/S2213-2600(20)30579-8

See Comment page 322

Getting It Right First Time programme, London, UK (A V Navaratnam FRCS, W K Gray PhD, J Day PhD, T W R Briggs FRCS); **University College London Hospitals NHS Foundation Trust, London, UK** (A V Navaratnam); **Kings College Hospital NHS Foundation Trust, London, UK** (J Wendon FRCP); **Royal National Orthopaedic Hospital NHS Trust, London, UK** (T W R Briggs)

Correspondence to: Mr Annakan Navaratnam, Royal National Ear, Nose and Throat Hospital, University College London Hospitals NHS Foundation Trust, London W1T 7PA, UK **annakan.navaratnam@nhs.net**

Introduction

The COVID-19 pandemic has had an extraordinary impact on health-care systems globally. As of early February, 2021, more than 100 million people had been infected with SARS-CoV-2, with over 2·2 million COVID-19-related deaths reported worldwide.[1]

Internationally, large COVID-19 cohort studies have shown that older age and male sex are strong non-modifiable risk factors that contribute to poor outcomes.[2,3] Furthermore, the presence of pre-existing comorbidities such as cardiovascular disease, hypertension, liver disease, diabetes, respiratory conditions, cancer, and severe obesity have been associated with a higher mortality risk.[4–6]

Analysis of patient primary care electronic health-care records in England suggests that, in addition to the aforementioned risk factors, individuals of mixed, Asian, or Black ethnicity are at greater risk of COVID-19-related mortality than White ethnic groups.[7] Furthermore, this increased risk was only partly attributable to a greater burden of comorbidities and greater deprivation in these groups.[7] The association of Black and Asian ethnicities with higher COVID-19 in-hospital mortality rates is supported by other studies done in the UK.[8,9] By contrast, a study from Louisiana, USA, suggested that, after adjusting for covariates—including age, sex, Charlson Comorbidity Index score, residence in a low-income area, obesity, and indicators for baseline vital signs and laboratory measures—hospitalised Black patients had no greater COVID-19 mortality risk than that of White patients.[10]

Several studies have used hospital patient datasets to identify trends in clinical characteristics and outcomes

Research in context

Evidence before this study

We searched PubMed on Aug 10, 2020, for articles that documented risk factors for COVID-19-related in-hospital death, using the search terms "SARS-CoV-2" OR "COVID-19" AND "mortality" AND "hospital". Of the 2393 papers identified, we found 244 publications that included original clinical data from patients admitted to hospital with COVID-19, primarily reported from China (35%) and the USA (22%). We identified 15 studies from England (including studies covering the UK). However, no study based in England covered the entire hospital population nationally, and most studies reported incomplete data, with a large group of patients in the cohorts still hospitalised at the end of the study period.

Added value of this study

To our knowledge, this is the first study that presents data for all completed hospital episodes involving patients with a diagnosis of COVID-19 in any large region or whole country. The study includes data for 91 541 patients in 500 hospitals across 164 National Health Service trusts in England between March 1 and May 31, 2020. The profile of hospital patients with COVID-19 changed over time, with those discharged during March, 2020, being younger, more ethnically diverse, and more likely to be male than those discharged in May, 2020. After adjusting for covariates, the probability of in-hospital death declined significantly over time in a broadly linear fashion. In our hospitalised cohort, the effect of ethnicity and deprivation on mortality was modest compared with that in previous reports based on community-level data, with patients of Asian or mixed ethnicity, but not Black ethnicity, having a significantly increased risk of in-hospital death compared with White patients.

Implications of all the available evidence

This is one of very few studies that identifies the importance of temporal trends on mortality in the characteristics of hospital patients with COVID-19. Our findings on ethnicity and deprivation are consistent with previous hospital-based studies from the UK and elsewhere. Comparing hospital-based and community-based studies sheds some light on why disparities in COVID-19 mortality might be evident at a community level, with differences in infection rates in those most at risk, differences in propensity to become ill once infected, and differences in health-seeking behaviour likely to be key reasons for higher death rates in some ethnic groups. The reasons for higher in-hospital mortality in patients of Asian or mixed ethnicity should be investigated further. It is likely that public health strategy and hospital management of patients changed rapidly in the early stages of the COVID-19 pandemic in England and helped to significantly reduce in-hospital mortality in a short period. The mechanisms used to achieve this should be thoroughly investigated to ensure that lessons learned can be used to inform practice across all hospitals in England and internationally.

of patients with COVID-19 in selected cohorts, but few have used nationally representative datasets.[11] Owing to the nature of health-care provision in many countries that have been severely affected by COVID-19, there is no single available repository of this information. Additionally, efforts to capture information directly from hospitals can be constrained by inconsistent reporting patterns, especially if individual hospitals are overwhelmed by surges of patients.

The National Health Service (NHS) Hospital Episodes Statistics (HES) database is an administrative dataset that contains a wide range of details regarding all NHS-funded hospital admissions in England. The Getting It Right First Time (GIRFT) programme uses this resource to understand clinical activity in NHS hospitals and to help hospitals to improve performance. The aim of this study was to use HES data to gain insight into the course of the COVID-19 pandemic as experienced by hospitals in England, and specifically to understand the temporal trends and the roles of patient demographics and disease factors as determinants of in-hospital death.

Methods

Study design and data collection

This was a retrospective exploratory analysis of HES data. HES data are collected by NHS Digital for all NHS-funded patients admitted to hospitals in England. The data are collected primarily to allow NHS trusts (which run secondary and tertiary care hospitals in England) to be reimbursed for providing hospital care. Data are entered by trained coders in each trust and data collection is mandatory. The data were complete and final for the study period (March 1 to May 31, 2020) at the time of extraction and analysis during August and September, 2020.

Consent from individuals involved in this study was not required, and ethical approval was not sought because the study did not directly involve human participants. The analysis and presentation of data follows current NHS Digital guidance for the use of HES data for research purposes and is anonymised to the level required by the ISB1523 anonymisation standard for publishing health and social care data.[12] The study was completed in accordance with the Helsinki Declaration as revised in 2013.

Timing, case ascertainment, and inclusion and exclusion criteria

We reviewed HES data for all completed episodes of hospital care that involved a diagnosis of COVID-19 in England with a discharge date between March 1 and May 31, 2020. The data collection period was defined in

terms of the discharge date rather than the admission date because our interest was in completed episodes of care in which the outcome (alive at discharge or died during hospital stay) was known. The term discharge is used here to refer to patients discharged alive and those discharged following death. Patients younger than 18 years were excluded. Cases of COVID-19 were identified using the International Classification of Diseases 10th edition (ICD-10) codes U071 and U072. U071 is assigned when the presence of COVID-19 has been confirmed by laboratory testing. U072 is assigned to a clinical or epidemiological diagnosis of COVID-19 in which laboratory confirmation is inconclusive or not available.

When a patient had multiple admissions during the study period, only the chronologically last admission was retained in order to ensure that all admissions were independent of one another at a patient level and to avoid biasing the mortality data, as it would be impossible for a patient to die during an admission that was not their final admission in the time period.

Outcome

The primary outcome was in-hospital death as recorded by the Office for National Statistics (ONS). An in-hospital death was recorded if the date of death was the same as or within 1 day either side of the date of hospital discharge recorded in HES. Although rare, some dates of discharge can differ from dates of death if formal discharge was on the day following death or if the death or discharge was close to midnight. Recorded deaths were checked against two independent sources of COVID-19 mortality data in England: NHS England's COVID-19 Patient Notification System (CPNS), which records daily deaths as close to the time that they occur as possible;[13] and discharge data recorded directly in HES by clinical coders at each hospital. Deaths in HES were identified if either of two codes were used: discharge destination (code 79, patient died) and discharge method (code 4, patient died).

Covariates

The following covariates were assessed: age, sex, ethnicity, deprivation, frailty, comorbidities, obesity, and temporal trends. Age was categorised as 18–39 years, 40–49 years, 50–59 years, 60–69 years, 70–79 years, and 80 years and older for exploratory analysis, and treated as continuous in the final multivariate model. Sex was categorised as male or female. Ethnicity was coded in broad categories to reflect those used by NHS Digital:[7] White, mixed, Asian or Asian British, Black or Black British, other, or not stated. In a secondary analysis, the Asian or Asian British category was subdivided into South Asian (Bangladeshi, Indian, and Pakistani) and other Asian. Deprivation was recorded using the Index of Multiple Deprivation for the lower-layer super output area of the patient's home address, with scores categorised into quintiles based on national averages.

Frailty was defined using Hospital Frailty Risk Score (HFRS) bands (none, mild, moderate, or severe).[14] The HFRS uses ICD-10 codes over the previous 2 years to identify frailty and is calculated post-hoc. The HFRS was preferred over the Charlson Comorbidity Index as a global measure of frailty because it was deemed to cover a greater range of potential causes of frailty. 14 comorbidities were used to construct the Charlson Comorbidity Index: peripheral vascular disease, congestive heart failure, acute myocardial infarction, cerebrovascular disease, dementia, chronic pulmonary disease, connective tissue disease or rheumatic disease, peptic ulcer, liver disease (mild and moderate or severe), diabetes (with and without chronic complications), paraplegia or hemiplegia, renal disease, cancer (primary and metastatic), and HIV/AIDS.[15] A comorbidity was deemed to be present if it was recorded in HES as a secondary diagnosis in the index admission or as a primary or secondary diagnosis in any admission during the previous year, in accordance with the recommendations of Quan and colleagues.[16] Obesity was recorded as present if the ICD-10 code E66 was used as a diagnostic code during the admission.

Temporal trends were categorised into day, week, or fortnight of discharge (starting from March 1, 2020) depending on the analysis. Fortnightly categorisation was used to avoid over-fitting during exploratory model building and for ease of interpretation of tabulated data. For the weekly categorisation, the first 2 weeks were combined because of the smaller number of discharges in these weeks and the last 2 weeks were combined because the final week (week 14) contained only 1 day (May 31). Daily data were used for the final multivariable model.

Data management and statistical analysis

Data were extracted onto a secure encrypted server controlled by NHS England and NHS Improvement. Analysis within this secure environment was done with standard statistical software: Microsoft Excel 2019, Stata 13, and Alteryx 2020.4.

In descriptive analysis, data were categorised as detailed above and summarised as frequency and percentage. Length of stay was non-normally distributed, with a right-skew, and summarised as median and IQR.

As an initial inferential exploration of the relationship between death and the covariates, a series of four multilevel (mixed-effects) logistic regression models were constructed with use of the melogit command in Stata. For ease of interpretation, all variables were categorised for this analysis. Two-level intercept-only models were constructed, allowing adjustment for clustering of patients within NHS hospital trusts.

This exploratory analysis then informed a model that considered age, day of discharge, and deprivation score as continuous variables, modelled using restricted cubic splines where non-linearity was evident. Linearity was identified where only the first linear spline was

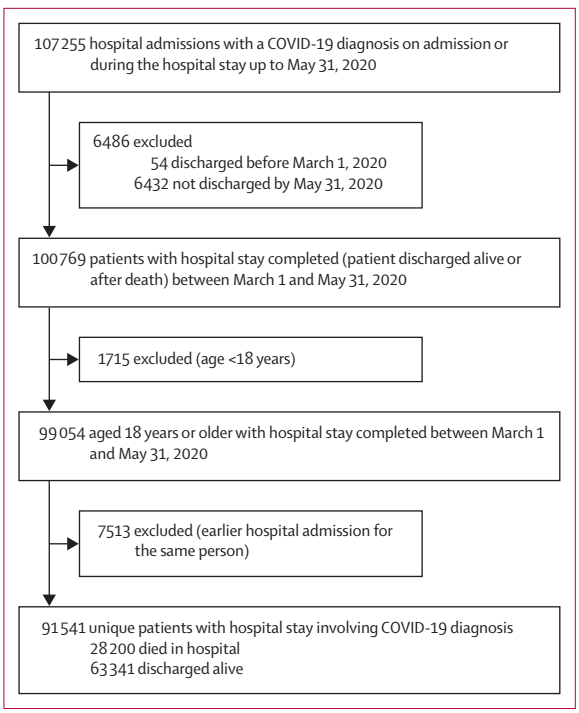

Figure 1: **Data extraction process**

significant. Only age was found to be non-linear and three knots were placed at equally spaced percentiles of the data and the optimal number of knots assessed using the Akaike Information Criterion (AIC). The relationship between age and the estimated probability of death was plotted using the adjustrcspline command. The marginal relationship between date of discharge and the probability of death was modelled using the margins command with fixed effects only.

A logistic model was preferred to a proportional hazards survival model as it was felt that the short length of stay, the potential for time from symptom onset to presentation to vary systematically with some of the covariates under consideration, differences in admission criteria and discharge practice between trusts, and the unsuitability of discharge date as a censoring point would mean that a survival model would add little in terms of model fit, be less robust, and might complicate interpretation.

Missing data were uncommon. Only ethnicity had a substantial number of missing values in cases in which patients were unwilling to state their ethnicity. No attempt was made to impute missing values. Where data were missing, the numbers involved are stated. The models were summarised in terms of odds ratios (ORs) and 95% CIs. An OR with a 95% CI not crossing 1 indicates statistical significance. For covariates modelled linearly, the OR represents the change in the odds of the outcome for a one unit change in the covariate.

Sensitivity analysis was done to assess the effect of including only patients for whom COVID-19 was confirmed by a test (ICD-10 code U071) and counting only deaths for which COVID-19 was the primary cause of death identified by the ONS (U071 and U072).

Role of the funding source

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. All authors had full access to all the data in the study and accept responsibility to submit for publication.

Results

The data extraction process (figure 1) yielded a dataset of 91 541 unique patients across 500 hospitals within 164 NHS trusts who had a diagnosis of COVID-19 either on admission or during their stay. 74 660 (81·6%) had COVID-19 confirmed by a test. There were 28 200 (30·8%) in-hospital deaths. Of the 26 240 in-hospital deaths with a cause listed, the cause of death involved COVID-19 in 23 153 (88·2%) patients. The number of deaths recorded was slightly higher than that reported by CPNS, through which 27 128 deaths were recorded, but was very similar to that recorded directly in HES using discharge destination and method codes (n=28 089).

The demographic, socioeconomic, and clinical characteristics of the dataset are summarised in table 1 together with the number of in-hospital deaths. The number of admissions increased substantially with age and deprivation, and was higher for men than for women. The ratio of White:Asian:Black patients was 13·0:1·4:1·0 in our study, compared with a ratio of 26:2:1 for the population of England and Wales at the 2011 census.[17] The most common comorbidities in the study were diabetes (25 382 [27·7%] patients), chronic pulmonary disease (23 719 [25·9%]), renal disease (16 713 [18·3%]), dementia (14 019 [15·3%]), and congestive heart failure (12 706 [13·9%]).

Crude in-hospital death rates (table 1) increased substantially with age and frailty, were higher in men than in women, and higher in the White ethnic group compared with other ethnic groups. There was no clear trend across deprivation quintiles. The comorbidities with the highest death rates as a proportion of cases were moderate or severe liver disease, congestive heart failure, dementia, and renal disease, although most of the listed comorbidities when present conferred a higher unadjusted death rate than in the wider study population (with mild liver disease, HIV/AIDS, and obesity being exceptions).

Fortnightly trends in the demographic, socioeconomic, and frailty profiles of the study population are shown table 2. Over the study period, older age bands comprised an increasing proportion of discharges (whether alive or deceased): comparing the first fortnight (March 1–14) with the last one (May 24–31), the proportion of patients who were in the older two age bands (aged ≥70 years) more than doubled. This trend was mirrored in the frailty

	Patients	In-hospital deaths
Age, years		
18–39	6984	204 (2·9%)
40–49	7172	514 (7·2%)
50–59	12157	1725 (14·2%)
60–69	13866	3549 (25·6%)
70–79	19466	7287 (37·4%)
≥80	31896	14921 (46·8%)
Sex		
Male	50668	17093 (33·7%)
Female	40697	11059 (27·2%)
Missing data	176	48 (27·3%)
Deprivation quintile		
1 (most deprived)	22956	6784 (29·6%)
2	20180	6166 (30·6%)
3	17158	5501 (32·1%)
4	15441	4963 (32·1%)
5 (least deprived)	13938	4392 (31·5%)
Missing	1868	394 (21·1%)
Ethnicity		
White	64615	21351 (33·0%)
Asian or Asian British	7117	1802 (25·3%)
Black or Black British	4983	1266 (25·4%)
Mixed	746	165 (22·1%)
Other ethnic groups	3291	689 (20·9%)
Missing data	10789	2933 (27·2%)
Hospital Frailty Risk Score		
None	34658	4244 (12·2%)
Mild	7213	2055 (28·5%)
Moderate	21137	8955 (42·4%)
Severe	28533	12946 (45·4%)

(Table 1 continues in next column)

	Patients	In-hospital deaths
(Continued from previous column)		
Charlson Comorbidity Index items*		
Peripheral vascular disease	4609	2053 (44·5%)
Congestive heart failure	12706	6319 (49·7%)
Acute myocardial infarction	8152	3542 (43·4%)
Cerebrovascular disease	8262	3538 (42·8%)
Dementia	14019	6765 (48·3%)
Chronic pulmonary disease	23719	8088 (34·1%)
Connective tissue disease or rheumatic disease	2698	989 (36·7%)
Peptic ulcer	572	203 (35·5%)
Liver disease	3720	1206 (32·4%)
Mild liver disease	2808	709 (25·2%)
Moderate or severe liver disease	912	497 (54·5%)
Diabetes	25382	9376 (36·9%)
Diabetes without chronic complications	22687	8317 (36·7%)
Diabetes with chronic complications	2695	1059 (39·3%)
Paraplegia and hemiplegia	2160	875 (40·5%)
Renal disease	16713	7691 (46·0%)
Cancer	7820	3442 (44·0%)
Primary cancer	5201	2356 (45·3%)
Metastatic carcinoma	2619	1086 (41·5%)
HIV/AIDS	149	26 (17·4%)
Obesity	7920	2155 (27·2%)

Data are n or n (%) of patients in that category. *Only those with each disease are listed; there were no missing data; individual patients can appear in multiple disease categories.

Table 1: **Demographic, socioeconomic, and clinical profile of patients (n=91541) and in-hospital deaths (n=28200)**

profile of patients. The proportion of admitted patients who were female also increased over time, but there was no obvious temporal trend in deprivation. An increasing proportion of patients of White ethnicity was observed over time, with the proportions of Black patients and other ethnic groups more than halving.

The unadjusted numbers of discharges and associated death rates per week are summarised in the appendix (p 3). The number of weekly discharges peaked at 15839 in week 6 (April 5–11) and the death rate peaked at 38·2% in week 5 (March 29–April 4) before declining to an average of 22·2% in the last 3 full weeks of data (May 10–31).

Data on length of stay over the study period are shown in the appendix (p 4). In those who died in hospital and those who survived to discharge, median length of stay increased steadily across the study period (March 1–14 to May 24–31), from 2 (IQR 0–5) to 10 (4–22) days in those who survived to discharge and from 4·5 (3–10) to 11 (5–20) days in those who died during their hospital stay. Across the whole study period, the median length of stay was 7 days (IQR 3–14).

After stratifying the unadjusted mortality by age and sex (figure 2), the relatively low mortality in early March (appendix p 3) is much less apparent. In men and women, the week-on-week improvement in mortality was most obvious in the oldest two age bands, with mortality in patients aged 80 years and older more than halving.

To further investigate the relationship between demographic, socioeconomic, and clinical characteristics and in-hospital death, a series of four exploratory, multilevel, logistic regression models were constructed with each variable categorised (appendix p 1). In all four models, there was a significant and consistent trend towards greater odds of death at older age, although the effect was attenuated by the inclusion of frailty (model 3) or comorbidities (model 4) as predictor variables. Male sex, greater deprivation, and Asian and mixed ethnicities were also associated with significantly greater odds of death in all four models, and the effect was largely unaffected by the addition of fortnight of discharge, frailty, or comorbidity variables to the models. Frailty and all comorbidities investigated, except mild liver disease and peptic ulcer, were significant predictors of in-hospital

See Online for appendix

	March 1–14	March 15–28	March 29–April 11	April 12–25	April 26–May 9	May 10–23	May 24–31
Age, years	n=291	n=6296	n=26 884	n=25 215	n=17 111	n=11 485	n=4259
18–39	77 (26·5%)	843 (13·4%)	2098 (7·8%)	1488 (5·9%)	1267 (7·4%)	900 (7·8%)	311 (7·3%)
40–49	40 (13·7%)	610 (9·7%)	2323 (8·6%)	1916 (7·6%)	1245 (7·3%)	738 (6·4%)	300 (7·0%)
50–59	45 (15·5%)	914 (14·5%)	3920 (14·6%)	3476 (13·8%)	2038 (11·9%)	1294 (11·3%)	470 (11·0%)
60–69	38 (13·1%)	913 (14·5%)	4324 (16·1%)	3956 (15·7%)	2430 (14·2%)	1634 (14·2%)	571 (13·4%)
70–79	40 (13·7%)	1241 (19·7%)	5876 (21·9%)	5343 (21·2%)	3679 (21·5%)	2411 (21·0%)	876 (20·6%)
≥80	51 (17·5%)	1775 (28·2%)	8343 (31·0%)	9036 (35·8%)	6452 (37·7%)	4508 (39·3%)	1731 (40·6%)
Sex*	n=291	n=6294	n=26 857	n=25 161	n=17 050	n=11 460	n=4252
Male	171 (58·8%)	3539 (56·2%)	15 661 (58·3%)	14 137 (56·2%)	9060 (53·1%)	5885 (51·4%)	2215 (52·1%)
Female	120 (41·2%)	2755 (43·8%)	11 196 (41·7%)	11 024 (43·8%)	7990 (46·9%)	5575 (48·6%)	2037 (47·9%)
Deprivation quintile*	n=283	n=6117	n=26 288	n=24 766	n=16 783	n=11 253	n=4183
1 (most deprived)	59 (20·8%)	1502 (24·6%)	6786 (25·8%)	6459 (26·1%)	4246 (25·3%)	2810 (25·0%)	1094 (26·2%)
2	58 (20·5%)	1420 (23·2%)	6073 (23·1%)	5590 (22·6%)	3699 (22·0%)	2456 (21·8%)	884 (21·1%)
3	59 (20·8%)	1181 (19·3%)	4841 (18·4%)	4836 (19·5%)	3248 (19·4%)	2208 (19·6%)	785 (18·8%)
4	42 (14·8%)	1046 (17·1%)	4435 (16·9%)	4216 (17·0%)	2915 (17·4%)	2019 (17·9%)	768 (18·4%)
5 (least deprived)	65 (23·0%)	968 (15·8%)	4153 (15·8%)	3665 (14·8%)	2675 (15·9%)	1760 (15·6%)	652 (15·6%)
Ethnicity*	n=230	n=5487	n=23 605	n=22 139	n=15 186	n=10 304	n=3801
White	166 (72·2%)	3890 (70·9%)	17 713 (75·0%)	17 804 (80·4%)	12 820 (84·4%)	8916 (86·5%)	3306 (87·0%)
Asian or Asian British	25 (10·9%)	641 (11·7%)	2562 (10·9%)	1907 (8·6%)	1068 (7·0%)	673 (6·5%)	241 (6·3%)
Black or Black British	19 (8·3%)	591 (10·8%)	1942 (8·2%)	1328 (6·0%)	636 (4·2%)	344 (3·3%)	123 (3·2%)
Mixed	2 (0·9%)	56 (1·0%)	234 (1·0%)	198 (0·9%)	149 (1·0%)	80 (0·8%)	27 (0·7%)
Other ethnic groups	18 (7·8%)	309 (5·6%)	1154 (4·9%)	902 (4·1%)	513 (3·4%)	291 (2·8%)	104 (2·7%)
Hospital Frailty Risk Score	n=291	n=6296	n=26 884	n=25 215	n=17 111	n=11 485	n=4259
None	188 (64·6%)	2964 (47·1%)	11 136 (41·4%)	9203 (36·5%)	5921 (34·6%)	3862 (33·6%)	1384 (32·5%)
Mild	17 (5·8%)	676 (10·7%)	2869 (10·7%)	1873 (7·4%)	972 (5·7%)	583 (5·1%)	223 (5·2%)
Moderate	45 (15·5%)	1391 (22·1%)	6388 (23·8%)	6107 (24·2%)	3816 (22·3%)	2441 (21·3%)	949 (22·3%)
Severe	41 (14·1%)	1265 (20·1%)	6491 (24·1%)	8032 (31·9%)	6402 (37·4%)	4599 (40·0%)	1703 (40·0%)

Data are n (%). *Missing data: sex 176; deprivation quintile 1868; ethnicity 10 789.

Table 2: **Patient demographic, socioeconomic, and frailty profiles by fortnight throughout the study period**

death. Moderate or severe liver disease had a particularly high OR (5·345 [95% CI 4·542–6·289]).

The model accounting for age, sex, fortnight of discharge, deprivation, ethnicity, and Charlson Comorbidity Index items (model 4) had the lowest AIC value and this model was taken forward, with age, deprivation score, and day of discharge treated as continuous rather than categorical variables. The relationships of the covariates with in-hospital death in this model were very similar to those found during exploratory model building: male sex, greater deprivation, most comorbidities, and mixed ethnicity (*vs* White) were associated with significantly greater odds of death (table 3; appendix p 5). Asian ethnicity was also associated with greater odds of death, although this differed between South Asian ethnicity (OR 1·246 [95% CI 1·152–1·348]; n=5117) and other Asian ethnicity (1·108 [0·973–1·262]); n=2000). Age had a shallow sigmoid relationship with death rate, with odds of death increasing with older age (figure 3). Later date of discharge was associated with lower odds of in-hospital mortality, with the adjusted probability falling from an average of 52·2% during the first 7 days of data collection to an average of 16·8% during the last 7 days of data collection.

In sensitivity analyses, we also analysed deaths over time in which COVID-19 was the primary cause (as opposed to all deaths; appendix p 3), as well as discharges and deaths for patients with test-confirmed COVID-19 over time (as opposed to all discharges and deaths; appendix p 6). In addition, we assessed predictors of in-hospital death (based on the final multilevel logistic regression model) among patients with test-confirmed COVID-19 only (appendix p 2). All comparisons showed very similar trends and associations to those seen in the main analyses.

Discussion

To the best of our knowledge, this is the most complete analysis of COVID-19-related hospital activity in any large region or whole country to date. We used an administrative dataset to describe the characteristics of all people in hospital in England with a diagnosis of COVID-19 during the 3 months with greatest hospital activity in the early phase of the pandemic.

One of the main findings of our study was that, after adjusting for demographics, socioeconomic factors, and comorbidity, the probability of in-hospital death fell by more than half over the 3 month period in a broadly linear fashion. Therefore, our study does not provide evidence that high death rates early in the pandemic were caused by increasing overall COVID-19 patient numbers straining services, since adjusted in-hospital death rates were falling as case numbers were increasing in the first half of April, 2020. The fall in mortality was most apparent in older age groups. Temporal effects on in-hospital death rates during the COVID-19 pandemic have received little attention. However, the trend towards reduced morality rates over time has been shown in a cohort of patients with COVID-19 in critical care in England.[18] The Intensive Care National Audit and Research Centre (ICNARC) collects and analyses case-mix and outcome data for individual patients for all adult general intensive care units, as well as other critical care units, in England, Wales, and Northern Ireland. They found that, after adjusting for patient characteristics, 28-day COVID-19-related mortality decreased from 43·5% in the pre-peak period (before March 29, 2020) to 34·3% in the post-peak period (April 16 to May 21, 2020).

Mutations in SARS-CoV-2[19] and changes in host human leukocyte antigen susceptibility[20] over time might have affected the clinical presentation, and subsequent risk of death, of COVID-19 cases. However, it seems likely that non-patient-related factors will have been key drivers for the observed trend.[4] Such factors are likely to include increased individual hospital strain due to COVID-19 patient surges early in the pandemic, and changes in hospital practice with regard to admission and treatment decisions as new knowledge became available. Increasing length of stay over time provides some evidence to support both these postulations.

From March through to May, the proportions of older, frail, female, and White patients increased steadily. These demographic changes in case-mix with time are likely to be driven, to an extent, by the profile of those infected in the community, with people of working age and from minority ethnic backgrounds more likely to be infected in the early stages of the pandemic, before the full UK lockdown on March 23.[21] High infection rates early in the pandemic in demographically younger and more ethnically diverse areas, such as London, will have played a part. As infections spread out geographically and lockdown measures started to have an effect, the demographic profile of hospital patients with COVID-19 appears to have shifted.

It has been postulated that there might be an increased risk of infectivity in minority ethnic groups because of their over-representation in frontline occupations and higher incidences of multigenerational households.[22–24] Differences in access to health care has been proposed as an explanation for higher COVID-19 mortality seen in some ethnic groups in the USA[25] and Brazil.[26] Additionally, public health messaging regarding prevention, early diagnosis, and treatment of COVID-19 might have been less effective in certain ethnic minority groups, leading to later presentation.[24]

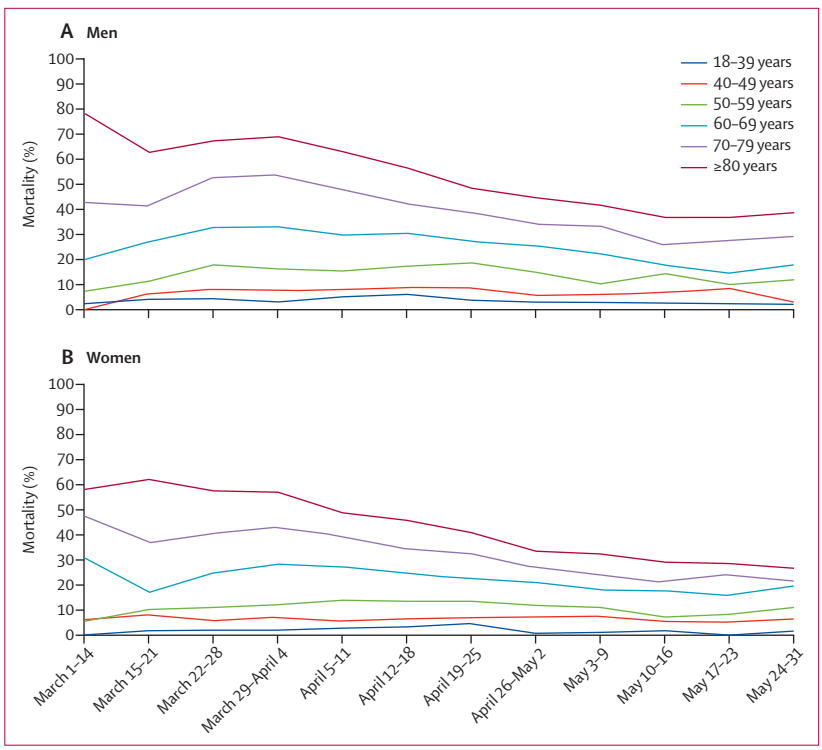

Figure 2: **Weekly mortality by age between March 1 and May 31, 2020**
Data are presented separately for men (A) and women (B).

In our hospital-based study, patients of Black and White ethnicities had similar odds of death, and South Asian patients a significantly higher odds of death than White patients. Other hospital-based studies from the USA and UK support these findings.[10,27,28] A UK study that followed hospital patients until late May noted that South Asian patients had a modestly, but significantly, increased risk of death (hazard ratio 1·19 [95% CI 1·05–1·36]) relative to White patients—very similar to our findings.[27] By contrast, studies of community-based cohorts tend to report significantly higher mortality for patients of Black ethnicity and greater effects of Asian or mixed ethnicities on mortality than we report.[7–9,29] The OpenSAFELY group reported hazard ratios (relative to the White ethnic group) for COVID-19-related mortality of 1·48 (95% CI 1·29–1·69) for Black ethnic groups, 1·45 (1·32–1·58) for South Asian groups, and 1·43 (1·11–1·84) for mixed ethnic groups.[7]

However, community-based and hospital-based studies provide different insights. Hospital-based studies look at the risk of death for those who are most severely ill and attend (and are admitted to) hospital. For this denominator population, increased mortality is due to

	OR (95% CI)
Sex	
Female	1 (reference)
Male	1·457 (1·408–1·509)*
Deprivation score	1·002 (1·001–1·003)*
Ethnicity	
White	1 (reference)
Asian	1·211 (1·128–1·299)*
Black	1·015 (0·935–1·103)
Mixed	1·317 (1·080–1·605)*
Other	0·989 (0·893–1·096)
Date of discharge (alive or deceased)	0·977 (0·976–0·978)*
Charlson Comorbidity Index items†	
Peripheral vascular disease	1·227 (1·146–1·315)*
Congestive heart failure	1·606 (1·535–1·681)*
Acute myocardial infarction	1·073 (1·016–1·132)*
Cerebrovascular disease	1·176 (1·110–1·245)*
Dementia	1·496 (1·431–1·565)*
Chronic pulmonary disease	1·089 (1·049–1·130)*
Connective tissue disease or rheumatic disease	1·242 (1·134–1·360)*
Peptic ulcer	1·056 (0·865–1·289)
Mild liver disease	1·017 (0·917–1·127)
Moderate or severe liver disease	5·433 (4·618–6·392)*
Diabetes without chronic complications	1·159 (1·115–1·204)*
Diabetes with chronic complicaions	1·295 (1·179–1·422)*
Paraplegia and hemiplegia	1·154 (1·093–1·219)*
Renal disease	1·158 (1·134–1·182)*
Primary cancer	1·542 (1·445–1·646)*
Metastatic carcinoma	2·053 (1·870–2·255)*
Obesity	1·476 (1·383–1·575)*

Models are based on data for 79 124 patients with no missing data. Age is modelled in years using restricted cubic splines and so the model output cannot be summarised as an OR. The relationship between age and in-hospital mortality is depicted graphically in figure 3. Deprivation was modelled from the Index of Multiple Deprivation score as a linear term, and date of discharge in days as a linear term. A stable OR for the comorbidity HIV/AIDS could not be calculated due to small numbers. OR=odds ratio. *95% CI indicates a significant difference compared with the reference category. †The reference category is patients without the specified comorbidity; for items relating to liver disease, diabetes, or cancer, three mutually exclusive categories (including the reference category) were used.

Table 3: **Multilevel logistic regression model of variables, modelled as linear or categorical, predicting in-hospital mortality**

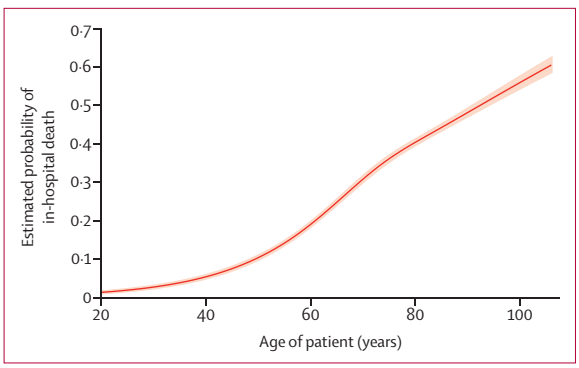

Figure 3: **Estimated probability of in-hospital mortality by age**
Estimates are adjusted for sex, deprivation, ethnicity, date of discharge, and comorbidities. The shaded area represents the 95% CI.

factors not adjusted for by the model used (eg, genetics, illness severity on admission, comorbidity not accounted for in modelling, hospital admission, and treatment protocols at the time of admission). In community-based studies, differences in infection rates in those most at risk will be a key driver for the death rate in an ethnic group, with propensity to become severely ill once infected and health-seeking behaviours acting as potentially mediating factors. Comparing the findings from both types of study provides additional insight and suggests that such factors might be responsible for the differences reported. Access to services for some ethnic groups could also be an issue. A recent case-control study for patients admitted to two hospitals in London was able to look at differences in outcomes for community dwellers and hospital patients, with patients of Black and mixed (but not Asian) ethnicities having higher hospital admission rates than community-dwelling controls, but Asian patients having a higher mortality if admitted.[28]

Although greater deprivation was significantly associated with higher mortality in our study, this association was modest and much weaker than that described in community-based cohort studies in England.[7] Age-standardised comparisons using ONS data for England and Wales between March 1 and June 30, 2020, found a two-fold difference in COVID-19-related mortality between the most deprived areas (139·6 deaths per 100 000 people) and the least deprived areas (63·4 deaths per 100 000).[29] A similar argument to that used above for ethnicity can be applied to explain these contrasting findings.

Of the comorbidities considered, moderate or severe liver disease, metastatic carcinoma, and congestive heart failure had particularly large ORs. Increased risk of death has been reported previously for hospitalised patients in the UK with COVID-19 and cardiac, pulmonary, or renal diseases, as well as obesity, cancer, or dementia.[30] Cardiovascular disease, diabetes, malignancy, and respiratory disease have been reported as risk factors for higher COVID-19 death rates internationally,[31] and in China,[32] the USA,[33] and Italy.[4]

HES data provide a complete record of all COVID-19-related hospital activity in England. As such, this dataset represents a record of the pandemic as experienced by hospitals in England during this time. Furthermore, the 3-month study period allowed us to look at temporal trends in hospital admissions and mortality. By including all patients hospitalised in England, the risk of collider bias through selective patient inclusion criteria is minimised in our study compared with other hospital-based studies.[34] However, we emphasise that our findings should not be directly extrapolated beyond the hospital setting. Comparison of our findings with community-based studies should only be done where differences in the denominator population are acknowledged.

There are inherent limitations in using HES data because of the reliance on hospitals completing administrative information accurately. Some inconsistencies in data entry have been noted for specific procedures, mostly with regard to differences in coding practice across trusts. Although HES data are entered by trained coders who are independent of clinicians, they rely on patients' notes for information and only if this is recorded accurately will HES data be reliable. Furthermore, HES does not capture detailed data on admission criteria within each trust, and thresholds for admission and treatment escalation are likely to vary across trusts. Nevertheless, by including trust as a random variable in our models, we have adjusted for clustering of deaths. HES provides only limited clinical information and no record of how acutely unwell a patient was during their stay. Comorbidities (including obesity) will be under-reported in HES because they are recorded only if relevant to a specific hospital admission. As such, they will capture only cases of severe or relevant comorbidity. Some conditions (eg, hypertension) are not well recorded in HES, and body-mass index data are not collected. Such factors could influence the risk of in-hospital death reported for certain groups of patients.

In capturing all activity for patients diagnosed with COVID-19, we will have picked up some admissions in which COVID-19 was not the reason for admission, a major complication during the stay, or a factor in a death. This limitation would be more of a problem later in our data collection period, in which subsequent admissions unrelated to COVID-19 might still contain the code for COVID-19. Use of the two ICD-10 codes by coders could also have changed over time, and testing capacity was lower early in the pandemic. However, our analysis of ONS cause-of-death data, and triangulation of our death data with that reported by CPNS, suggests that our data align closely with other available sources and are not a substantial over-estimate. Our decision to include all recorded COVID-19-related hospitalisations and deaths ensured that our data are as complete a record of the pandemic as possible. Our study includes all in-hospital COVID-19-related deaths and as such does not consider deaths that occurred in the community or after discharge. By excluding patients still in hospital at the end of the study period, some activity has been missed, although the numbers involved are quite small (figure 1). Finally, we acknowledge that some data were missing in our dataset, particularly for ethnicity. However, the extent of missing data was modest compared with that of other available data sources.

Our data shed light on the nature of the relationship between several covariates and risk of COVID-19-related death. The relationships of male sex and greater age with the probability of death are broadly consistent with previous research. Comparing our findings to community-based studies provides further insight into the role of ethnicity and deprivation in predicting outcomes from COVID-19.

Temporal analysis in this study showed a large reduction in mortality with time. The reasons for this reduction should be further investigated. Learning from individual hospitals with the greatest levels of improvement, or consistently good performance throughout, could help to support others to improve. Further examination of these temporal trends—with consideration of changes in public health strategy, central directives from arm's-length bodies, alterations in clinical processes, and availability of local health-care resources—will be vital to inform preparations and adjustments for ongoing and future case surges.

Contributors

This study was designed and organised by AVN, WKG, JD, and TWRB. WKG and JD accessed and verified the underlying data (Hospital Episodes Statistics). Data cleaning and analysis was done by WKG, supported by JD. Writing of the first draft was done by AVN and WKG. All authors critically reviewed the manuscript and agreed to submission of the final draft.

Declaration of interests

We declare no competing interests.

Data sharing

This report does not contain patient identifiable data. Consent from individuals involved in this study was not required. Requests for any underlying data cannot be granted by the authors because the data were acquired from data under licence/data sharing agreement from NHS Digital, for which conditions of use (and further use) apply. Individuals and organisations wishing to access HES data can make a request directly to NHS Digital.

Acknowledgments

We thank NHS Digital for permission to use their data in this report. The Getting It Right First Time (GIRFT) programme is providing a framework for examining contemporary clinical practice in unprecedented detail and breadth. We also thank all staff within individual NHS trusts who collected and entered the data used in this study; the GIRFT clinical leads for advice (Philip Dyer, Adrian Hopper, Michael Jones, Cliff Mann, Chris Moutlon, Anna Batchelor, Michael Swart, Christopher Snowden, Martin Allen, Partha Kar, and Gerry Rayman); and GIRFT clinical fellows (Ini Adelaja, Pratusha Babu, Shona MacKinnon, and Sam Bartlett-Pestell).

References

1 WHO. WHO coronavirus disease (COVID-19) dashboard. https://covid19.who.int/ (accessed Oct 4, 2020).
2 Guan WJ, Ni ZY, Hu Y, et al. Clinical characteristics of coronavirus disease 2019 in China. *N Engl J Med* 2020; **382:** 1708–20.
3 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; **323:** 2052–59.
4 Grasselli G, Greco M, Zanella A, et al. Risk factors associated with mortality among patients with COVID-19 in intensive care units in Lombardy, Italy. *JAMA Intern Med* 2020; **180:** 1345–55.
5 Liu J, Zhang S, Wu Z, et al. Clinical outcomes of COVID-19 in Wuhan, China: a large cohort study. *Ann Intensive Care* 2020; **10:** 99.
6 Cummings MJ, Baldwin MR, Abrams D, et al. Epidemiology, clinical course, and outcomes of critically ill adults with COVID-19 in New York City: a prospective cohort study. *Lancet* 2020; **395:** 1763–70.
7 Williamson EJ, Walker AJ, Bhaskaran K, et al. OpenSAFELY: factors associated with COVID-19 death in 17 million patients. *Nature* 2020; **584:** 430–36.
8 Perez-Guzman PN, Daunt A, Mukherjee S, et al. Clinical characteristics and predictors of outcomes of hospitalized patients with COVID-19 in a multi-ethnic London NHS Trust: a retrospective cohort study. *Clin Infect Dis* 2020; published online Aug 7. https://doi.org/10.1093/cid/ciaa1091.

9 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.

10 Price-Haywood EG, Burton J, Fort D, Seoane L. Hospitalization and mortality among Black patients and White patients with COVID-19. *N Engl J Med* 2020; **382:** 2534–43.

11 Karagiannidis C, Mostert C, Hentschker C, et al. Case characteristics, resource use, and outcomes of 10 021 patients with COVID-19 admitted to 920 German hospitals: an observational study. *Lancet Respir Med* 2020; **8:** 853–62.

12 NHS Digital, Information Standard Board for Health and Social Care. Anonymisation standard for publishing health and social care data specification (process standard). London, 2013. https://digital.nhs.uk/data-and-information/information-standards/information-standards-and-data-collections-including-extractions/publications-and-notifications/standards-and-collections/isb1523-anonymisation-standard-for-publishing-health-and-social-care-data (accessed Feb 2, 2021).

13 NHS England. COVID-19 daily deaths. https://www.england.nhs.uk/statistics/statistical-work-areas/covid-19-daily-deaths/ (accessed Aug 16, 2020).

14 Gilbert T, Neuburger J, Kraindler J, et al. Development and validation of a Hospital Frailty Risk Score focusing on older people in acute care settings using electronic hospital records: an observational study. *Lancet* 2018; **391:** 1775–82.

15 Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis* 1987; **40:** 373–83.

16 Quan H, Li B, Couris CM, et al. Updating and validating the Charlson comorbidity index and score for risk adjustment in hospital discharge abstracts using data from 6 countries. *Am J Epidemiol* 2011; **173:** 676–82.

17 UK Government. Ethnicity facts and figures—population of England and Wales. https://www.ethnicity-facts-figures.service.gov.uk/uk-population-by-ethnicity/national-and-regional-populations/population-of-england-and-wales/latest (accessed Aug 16, 2020).

18 Doidge JC, Gould DW, Ferrando-Vivas P, et al. Trends in intensive care for patients with COVID-19 in England, Wales and Northern Ireland. *Am J Respir Crit Care Med* 2020; published online Dec 11. https://doi.org/10.1164/rccm.202008-3212OC.

19 Korber B, Fischer WM, Gnanakaran S, et al. Tracking changes in SARS-CoV-2 spike: evidence that D614G increases infectivity of the COVID-19 virus. *Cell* 2020; **182:** 812–27.e19.

20 Nguyen A, David JK, Maden SK, et al. Human leukocyte antigen susceptibility map for severe acute respiratory syndrome coronavirus 2. *J Virol* 2020; **94:** e00510-20.

21 Niedzwiedz CL, O'Donnell CA, Jani BD, et al. Ethnic and socioeconomic differences in SARS-CoV-2 infection: prospective cohort study using UK Biobank. *BMC Med* 2020; **18:** 160.

22 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.

23 Ho FK, Celis-Morales CA, Gray SR, et al. Modifiable and non-modifiable risk factors for COVID-19, and comparison to risk factors for influenza and pneumonia: results from a UK Biobank prospective cohort study. *BMJ Open* 2020; **10:** e040402.

24 Khunti K, Singh AK, Pareek M, Hanif W. Is ethnicity linked to incidence or outcomes of covid-19? *BMJ* 2020; **369:** m1548.

25 Yancy CW. COVID-19 and African Americans. *JAMA* 2020; **323:** 1891–92.

26 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.

27 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).

28 Zakeri R, Bendayan R, Ashworth M, et al. A case-control and cohort study to determine the relationship between ethnic background and severe COVID-19. *EClinicalMedicine* 2020; **28:** 100574.

29 Office for National Statistics. Deaths involving COVID-19 by local area and socioeconomic deprivation: deaths occurring between 1 March and 30 June 2020. https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/bulletins/deathsinvolvingcovid19bylocalareasanddeprivation/deathsoccurringbetween1marchand30june2020 (accessed Aug 18, 2020).

30 Docherty AB, Harrison EM, Green CA, et al. Features of 20 133 UK patients in hospital with covid-19 using the ISARIC WHO Clinical Characterisation Protocol: prospective observational cohort study. *BMJ* 2020; **369:** m1985.

31 Zhou Y, Yang Q, Chi J, et al. Comorbidities and the risk of severe or fatal outcomes associated with coronavirus disease 2019: a systematic review and meta-analysis. *Int J Infect Dis* 2020; **99:** 47–56.

32 Deng G, Yin M, Chen X, Zeng F. Clinical determinants for fatality of 44,672 patients with COVID-19. *Crit Care* 2020; **24:** 179.

33 Gupta S, Hayek SS, Wang W, et al. Factors associated with death in critically ill patients with coronavirus disease 2019 in the US. *JAMA Intern Med* 2020; **180:** 1436–46.

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