

# BMJ Open Risk of COVID-19 hospital admission and COVID-19 mortality during the first COVID-19 wave with a special emphasis on ethnic minorities: an observational study of a single, deprived, multiethnic UK health economy

Baldev M Singh ,<sup>1,2</sup> James Bateman ,<sup>1</sup> Ananth Viswanath,<sup>1</sup> Vijay Klaire,<sup>1</sup> Sultan Mahmud,<sup>1,3</sup> Alan Nevill ,<sup>4</sup> Simon J Dunmore <sup>2</sup>

**To cite:** Singh BM, Bateman J, Viswanath A, *et al.* Risk of COVID-19 hospital admission and COVID-19 mortality during the first COVID-19 wave with a special emphasis on ethnic minorities: an observational study of a single, deprived, multiethnic UK health economy. *BMJ Open* 2021;**11**:e046556. doi:10.1136/bmjopen-2020-046556

► Prepublication history for this paper is available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2020-046556>).

Received 02 November 2020  
Revised 20 January 2021  
Accepted 21 January 2021



© Author(s) (or their employer(s)) 2021. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

## Correspondence to

Professor Baldev M Singh;  
baldev.singh@nhs.net

## ABSTRACT

**Objectives** The objective of this study was to describe variations in COVID-19 outcomes in relation to local risks within a well-defined but diverse single-city area.

**Design** Observational study of COVID-19 outcomes using quality-assured integrated data from a single UK hospital contextualised to its feeder population and associated factors (comorbidities, ethnicity, age, deprivation).

**Setting/participants** Single-city hospital with a feeder population of 228 632 adults in Wolverhampton.

**Main outcome measures** Hospital admissions (defined as COVID-19 admissions (CA) or non-COVID-19 admissions (NCA)) and mortality (defined as COVID-19 deaths or non-COVID-19 deaths).

**Results** Of the 5558 patients admitted, 686 died (556 in hospital); 930 were CA, of which 270 were hospital COVID-19 deaths, 47 non-COVID-19 deaths and 36 deaths after discharge; of the 4628 NCA, there were 239 in-hospital deaths (2 COVID-19) and 94 deaths after discharge. Of the 223 074 adults not admitted, 407 died. Age, gender, multimorbidity and black ethnicity (OR 2.1 (95% CI 1.5 to 3.2),  $p < 0.001$ , compared with white ethnicity, absolute excess risk of  $< 1/1000$ ) were associated with CA and mortality. The South Asian cohort had lower CA and NCA, lower mortality compared with the white group (CA, 0.5 (0.3 to 0.8),  $p < 0.01$ ; NCA, 0.4 (0.3 to 0.6),  $p < 0.001$ ) and community deaths (0.5 (0.3 to 0.7),  $p < 0.001$ ). Despite many common risk factors for CA and NCA, ethnic groups had different admission rates and within-group differing association of risk factors. Deprivation impacted only the white ethnicity, in the oldest age bracket and in a lesser (not most) deprived quintile.

**Conclusions** Wolverhampton's results, reflecting high ethnic diversity and deprivation, are similar to other studies of black ethnicity, age and comorbidity risk in COVID-19 but strikingly different in South Asians and for deprivation. Sequentially considering population and then hospital-based NCA and CA outcomes, we present a complete single health economy picture. Risk factors

## Strengths and limitations of this study

- In contrast to the majority of other studies of factors related to COVID-19 morbidity and mortality, we used data from both a single-city hospital and its feeder population.
- Our observational study used a high-quality and complete data set from the local population and the hospital serving it to examine the association of purported risk factors with severity and mortality.
- Our study method enables assessment of the importance of evaluating such risks in the local, and not just national, population setting, taking into account the local variations in patient backgrounds.
- This nuanced approach factors in regional variation in elements such as ethnicity and deprivation by being specifically linked to the source population.
- Although limiting our study to the local population makes our findings less generalisable, it nevertheless allows evaluation of the importance of demographic and geographical variation.

may differ within ethnic groups; our data may be more representative of communities with high Black, Asian and minority ethnic populations, highlighting the need for locally focused public health strategies. We emphasise the need for a more comprehensible and nuanced conveyance of risk.

## INTRODUCTION

In understanding the natural history of disease, fundamental to healthcare, the COVID-19 pandemic highlights issues within data repositories. Constructing multiple-source data sets has complexity in case definition, data acquisition, integration, quality,

completeness, coding accuracy and the clinical meaning of analysis outcomes.<sup>1-4</sup> Emphasising this challenge, national UK data were initially collated via the Patient Notification System, requiring a positive swab test up until 28 April 2020 but revised to include clinical definitions given an estimated false-negative testing rate of up to 29%.<sup>5-8</sup> Well-established primary care databases may have significant inaccuracy and do not include hospital secondary care information.<sup>9</sup> A large UK primary care epidemiological study also used national COVID-19-positive (SARS-CoV-2) swab cases for case definition.<sup>2</sup> Conversely, secondary care case series and international registry studies for specific diseases are not linked to primary care data sets.<sup>10-14</sup> Numerous studies have described the risk factors associated with COVID-19 mortality, which have been from primary care, secondary care and meta-analyses.<sup>15 16</sup> Although these studies describe risk factors for severity, admission and mortality with COVID-19 infection, they typically use either large secondary care or primary care data sources, without amalgamating these data. Therefore, important caveats exist in using and interpreting such data and drawing clinically important conclusions regarding the adverse associations of ethnicity with outcomes.<sup>2 17</sup>

Our objective, therefore, was to establish a tightly governed comprehensive, multisource, integrated, quality-assured local structured clinical data set, used for purposes of direct care; to define cohorts at risk; to systematically improve clinical coding and mortality-recording accuracy; to enable an informed understanding of factors influencing hospital activity, including admissions; and especially to describe variations in COVID-19 outcomes in relation to local risks within a well-defined but diverse single-city area. This approach should ultimately inform public health initiatives. We present a proof-of-principle study to evaluate the utility of this approach in relation to a single UK city-wide health district, reporting our findings regarding population-wide factors that may have an association with two key COVID-19 outcomes, hospital admission and mortality, over the first 12 weeks of the pandemic in this city.

## METHODS

### General method

The time frame spanned from 1 March 2020 to 24 May 2020.

Data were integrated into an SQL database from primary care, from community and hospital clinical and pathology systems for all people residing in Wolverhampton or registered to Wolverhampton practices, and from immediately adjacent districts with emergency admission to New Cross Hospital (NXH). Only those alive at the start point were included, and subsequently death and date of death were tracked. The final total population aged >18 years was 228 632, of whom 1063 were resident but not Wolverhampton General Practitioner (GP) registered, 1521 were registered but not city resident and

1026 were neither resident nor registered from immediately surrounding areas with an emergency admission to NXH, such that 99.5% of the cohort were registered and/or resident constituting 88% of COVID-19 admissions (CA) and 91% of COVID-19 deaths. The Index of Multiple Deprivation (IMD) was allocated according to postcode. Unavailable smoking status (15%) was reallocated to 'non-smoker' (recognising that this is an assumption which may potentially introduce slight bias). Missing body mass index (BMI) (22%) was replaced by age-related (5-year band) mean value in the cohort. Ethnicity data from all sources were reviewed, and only unambiguous data were accepted and recoded into Caucasian (white), South Asian, black, mixed ethnicity and Chinese, with 7.5% remaining 'unknown'. Comorbidities were accrued and cross-checked from primary care and hospital coding to include asthma, chronic obstructive pulmonary disease (COPD), diabetes, hypertension, coronary artery disease, stroke and peripheral arterial disease, chronic heart failure, atrial fibrillation, chronic kidney disease (CKD), cancer, dementia, depression, other mental health disorders, epilepsy, learning difficulties, osteoarthritis, rheumatoid arthritis (RA), and recorded nursing home residency and palliative care status. Non-elective admissions over the preceding 12 months were ascertained. During admission, the COVID-19 clinical status was recorded by the Infection Diseases team or the clinical team in daily updates as 'COVID-19 definite', 'COVID-19 probable' or 'not COVID-19'. Formal end point coding was in duplicate, with a rolling triangulation audit in place comparing the clinical diagnosis, the coded diagnosis and the COVID-19 pathology status for coding accuracy. Mortality and cause of death were certified in our Medical Examiner System and also were continuously cross-checked against the coded status. COVID-19 coding and death certification arbitration were supported by the accountable senior responsible consultant (AV). Further validation against the National Strategic Tracing Service captured deaths outside hospital.

### Statistical analysis

This analysis was undertaken in SPSS V.26. Factors analysis of all variables considered confounding effects and redundancy, yielding a nine-component rotated solution explaining 48% of the variance: deprivation and ethnicity were strongly coassociated in a single component, whereas the two principal outcome measures of hospital admission and mortality were in another distinct component. We adopted a multinomial regression analysis approach. This allowed the association of those independent factors with the dependent categorical variable, yielding ORs with their 95% CIs and statistical significance. The analysis was undertaken sequentially to ensure an a priori justification for further analysis. Statistical tests are described in the text and their results considered significant at  $p < 0.05$ .

### Patient and public involvement

None (not applicable to this type of study).

**Table 1** Demographic and clinical features and mortality outcomes of a whole adult population (n=228 632) categorised according to their hospital admission status during 12 weeks of the UK COVID-19 pandemic

	Whole population	Not admitted	Non-COVID-19 admission	COVID-19 admission	
Number (% of total)	228 632	223 074 (97.6)	4628 (2)	930 (0.4)	
Age (years)	50.0±18.8	47.6±18.5	63.1±20.5	71.4±16.5	p<0.001
Gender (male, %)	114 866 (50.2)	50.3	48.4	55.2	p<0.001
Ethnicity (% in category)					
White	142 781 (62.5)	62.1	77.7	73.5	p<0.001
South Asian	44 229 (19.3)	19.6	10.4	8.9	
Black	17 858 (7.8)	7.9	4.6	12.2	
Mixed	5809 (2.5)	2.6	1.3	0.6	
Chinese	806 (0.4)	0.4	0.1	0.2	
Unknown	17 149 (7.5)	7.5	5.8	4.5	
Index of Multiple Deprivation	32.8±15.9	32.9±15.9	30.4±14.4	30.7±14.4	p<0.001
Smoking status (% never)	148 046 (64.8)	65	73	81	p<0.001
Body mass index (kg/m <sup>2</sup> )	27.3 (±5.7)	27.3 (±5.7)	27.8 (±5.1)	27.9 (±4.6)	p<0.001
Prior admission (any 1 year (≥3))	15 119 (7%) (0.6%)	6% (0.4%)	31% (7%)	35% (8%)	p<0.001
Any comorbidity (≥3)	110 564 (48%) (12%)	48% (11%)	77% (41%)	88% (58%)	p<0.001
Palliative care registered (%)	1530 (0.7)	0.5	4.9	13.9	p<0.001
Nursing home resident (%)	2130 (0.9)	0.8	4.7	9.7	p<0.001
Hypertension (%)	47 830 (20.9)	20.2	47.6	64.2	p<0.001
Depression (%)	39 153 (17)	17.0	22.4	18.7	p<0.001
Asthma (%)	30 335 (13.3)	13.2	16.1	16.3	p<0.001
Diabetes (%)	20 529 (9.0)	8.6	21.4	33.8	p<0.001
Ischaemic heart disease (%)	10 965 (4.8)	4.4	19.9	23.2	p<0.001
Chronic kidney disease (%)	10 265 (4.5)	4.1	17.6	29.0	p<0.001
Cancer (%)	9796 (4.3)	4.0	16.0	17.8	p<0.001
Atrial fibrillation (%)	6771 (3.0)	2.6	4.7	8.7	p<0.001
Chronic obstructive pulmonary disease (%)	6762 (3.0)	2.7	12.3	15.1	p<0.001
Cerebrovascular disease (%)	5359 (2.3)	2.2	8.8	13.5	p<0.001
Cardiac failure (%)	4940 (2.2)	1.9	13.1	20.8	p<0.001
Osteoarthritis (%)	4267 (1.9)	1.7	8.0	11.2	p<0.001
Epilepsy (%)	4035 (1.8)	1.7	3.8	4.5	p<0.001
Rheumatoid arthritis (%)	3284 (1.4)	1.4	4.2	6.2	p<0.001
Mental health disorder (%)	2967 (1.3)	1.3	1.7	1.7	p<0.05
Peripheral vascular disease (%)	2724 (1.2)	1.1	4.7	8.7	p<0.001
Dementia (%)	2304 (1.0)	0.9	4.0	7.8	p<0.001
Learning difficulties (%)	1597 (0.7)	0.7	1.0	1.2	p<0.05
Vital status (n, % died)	1093 (0.5)	407, 0.2	333, 7.2*	353, 38†	p<0.001

Data are presented as mean±SD or as percentages. Between-groups analysis is by analysis of variance or by  $\chi^2$  for scale or categorical variables, respectively. Comorbidities are listed in descending order of frequency.

\*237 non-COVID-19 hospital deaths, 94 postdischarge deaths in community and 2 hospital COVID-19 deaths.

†270 hospital COVID-19 deaths, 47 non-COVID-19 hospital deaths and 36 postdischarge deaths in community.

**Table 2** COVID-19 admission and death by ethnic category showing numbers, absolute rates per 1000 population, and excess risk and ORs (95% CI) versus the white group as comparator

	White	South Asian	Black	Other or unknown
Total numbers	142 781 (63%)	44 229 (19%)	17 858 (8%)	23 764 (10%)
COVID-19 admission	684 (74%)	83 (9%)	113 (12%)	50 (5%)
COVID-19 admission/1000	4.8	1.9	6.3	2.1
COVID-19 admission excess risk/1000	Comparator	-2.9	1.5	-2.7
COVID-19 admission OR	Comparator	0.39 (0.31 to 0.48), p<0.001	1.31 (1.07 to 1.59), p<0.001	0.43 (0.33 to 0.58), p<0.001
COVID-19 death	190 (70%)	32 (12%)	39 (14%)	11 (4%)
COVID-19 death/1000	1.3	0.7	2.2	0.5
COVID-19 death excess risk/1000	Comparator	-0.6	0.9	-0.9
COVID-19 death OR	Comparator	0.54 (0.37 to 0.79), p<0.01	1.64 (1.16 to 2.31), p<0.01	0.35 (0.19 to 0.64), p<0.01
COVID-19 death/COVID-19 admission/1000	0.3	0.4	0.3	0.2
COVID-19 death in COVID-19 admission OR	Comparator	1.55 (0.96 to 2.51), p=0.075, ns	1.19 (0.78 to 1.83), p=0.423, ns	0.63 (0.32 to 1.25), p=0.187, ns

'Chinese', 'mixed' and 'unknown' categories showed no significant associations as individual categories in this analysis or when merged as 'other or unknown'.

## RESULTS

### Hospital admissions

The population characteristics are shown in [table 1](#), grouped according to admission status (no admission (NA), non-COVID-19 admission (NCA) and CA) together with their mortality rates.

Compared with NA, there was an increased association of all variables with NCA and CA, including age, the number of comorbidities, most individual comorbidities, surrogate measures of dependency (of being on a palliative care register or being a nursing home resident) and history of emergency admissions. Male gender, BMI, IMD and smoking status were significantly different between the three categories. Ethnic minority groupings were significantly different between admission types, with the South Asian population prevalence in CA being 46% of that in the comparator NA population, whereas the black population appeared to have a 56% excess. [Table 2](#) gives further numerical details.

The three hospital admission categories (NA, NCA, CA) were taken as the response variable and submitted to multinomial regression ([table 3](#)). The complete model was highly significant ( $\chi^2=8869.1$ ,  $p<0.001$ ). Male gender was more prevalent in CA.

The age distribution ([figure 1A](#)) differed significantly for CA and NCA versus NA, and the two admission groups differed significantly from each other, reflecting the higher mean age in CA. The pattern for deprivation ([figure 1B](#)) showed the peak admission rates to be in the second least deprived quintile, with the most deprived quintile not being significantly different from the least deprived quintile, whereas the two admission groups did not differ significantly from each other in this regard.

There was a decreased relative risk for admission in either group with current or previous smoking. Both admission groupings had a significantly increased history of prior emergency admissions, established multimorbidity, being a nursing home resident or being in a palliative phase of care, with these latter two characteristics in significantly higher prevalence in CA compared with NCA. Both groups shared individual comorbidities in higher risk, but with some differential effect for diabetes, hypertension, atrial fibrillation and peripheral vascular disease which were increased in CA.

The South Asian ethnic group was less likely to have a CA or NCA (60% and 50% crude percentage reduced risk, respectively) compared with the white ethnic reference category, whereas the black ethnic group shared the significant propensity not to have an NCA but had a markedly increased relative risk (70%) for CA. Ethnicity-related outcomes were examined specifically among those with CA, by comparing those admitted with those not admitted within their ethnic category in separate binary regression analyses (all  $\chi^2>252.4$ ,  $p<0.001$ ) ([table 4](#)).

Age, gender, prior emergency admissions, palliative phase, comorbidity and nursing home residence were significant associations in two or more of the ethnic groups. Of note, patterns of significantly associated individual comorbidities were different between the ethnic groups: black—hypertension, atrial fibrillation and cardiac failure; South Asian—diabetes, peripheral vascular disease and atrial fibrillation; white—specific association with COPD, CKD and RA. Deprivation had a significant impact only in the white group. The inter-relationship of age, deprivation with ethnicity and the impact of white ethnicity in the oldest quintile in lesser deprived categories can be seen



**Table 3** Multinomial regression for the association of factors with COVID-19-related or non-COVID-19-related emergency hospital admissions (HA) compared with the reference category of those not admitted (n=223 074)

	CHA	NCHA	CHA vs NCHA
Number (% of population)	930 (0.4%)	4628 (2%)	
Gender (male)	1.5 (1.3 to 1.7), p<0.001	ns, p=0.48	1.3 (1.1 to 1.5), p<0.001
Age category Q2 (30–40)	2.7 (1.5 to 5), p<0.01	1.2 (1 to 1.3), p<0.05	2.2 (1.2 to 4), p<0.05
Age category Q3 (41–51)	5.1 (2.9 to 9), p<0.001	1.4 (1.2 to 1.6), p<0.001	3.2 (1.8 to 5.8), p<0.001
Age category Q4 (52–65)	7.8 (4.6 to 13.4), p<0.001	1.7 (1.5 to 1.9), p<0.001	3.7 (2.1 to 6.5), p<0.001
Age category Q5 (66–113)	16.6 (9.7 to 28.3), p<0.001	2.9 (2.5 to 3.3), p<0.001	4.7 (2.7 to 8.1), p<0.001
IMD category Q2 (16.5–27.7)	1.8 (1.4 to 2.2), p<0.001	2 (1.8 to 2.2), p<0.001	ns
IMD category Q3 (27.8–39.0)	1.7 (1.4 to 2.1), p<0.001	1.5 (1.4 to 1.7), p<0.001	ns
IMD category Q4 (39.3–45.7)	1.6 (1.3 to 2), p<0.001	1.4 (1.3 to 1.6), p<0.001	ns
IMD category Q5 (45.7–71.8)	ns, p=0.517	ns, p=0.574	ns
Ethnicity South Asian	0.4 (0.3 to 0.5), p<0.001	0.5 (0.4 to 0.5), p<0.001	1 (0.7 to 1.2), ns, 0.735
Ethnicity black	1.7 (1.3 to 2.1), p<0.001	0.6 (0.5 to 0.7), p<0.001	3.1 (2.4 to 4), p<0.001
Smoking current or prior	0.3 (0.2 to 0.3), p<0.001	0.5 (0.4 to 0.5), p<0.001	0.7 (0.5 to 0.8), p<0.001
Prior emergency admissions (1 year)	1.6 (1.5 to 1.7), p<0.001	1.8 (1.7 to 1.8), p<0.001	0.9 (0.9 to 1), p<0.05
Palliative care registered	4.1 (3.2 to 5.1), p<0.001	1.7 (1.4 to 2), p<0.001	2.4 (1.9 to 3.1), p<0.001
Nursing home resident	1.7 (1.3 to 2.3), p<0.001	1.2 (1 to 1.5), ns, 0.058	1.7 (1.3 to 2.2), p<0.001
Comorbidities ≥3	1.5 (1.2 to 1.9), p<0.01	1.4 (1.3 to 1.6), p<0.001	ns
Chronic obstructive pulmonary disease	1.9 (1.5 to 2.3), p<0.001	1.8 (1.6 to 2), p<0.001	ns
Peripheral vascular disease	1.8 (1.4 to 2.3), p<0.001	1.2 (1 to 1.4), p<0.05	1.5 (1.1 to 2), p<0.01
Atrial fibrillation	1.7 (1.4 to 2.1), p<0.001	1.4 (1.2 to 1.5), p<0.001	1.3 (1.1 to 1.6), p<0.01
Diabetes	1.6 (1.4 to 1.9), p<0.001	1.3 (1.2 to 1.4), p<0.001	1.4 (1.2 to 1.6), p<0.001
Cardiac failure	1.6 (1.3 to 1.9), p<0.001	1.4 (1.2 to 1.6), p<0.001	ns
Rheumatoid arthritis	1.5 (1.1 to 2), p<0.01	ns, p=0.134	ns
Epilepsy	1.4 (1 to 2), p<0.05	1.4 (1.2 to 1.6), p<0.001	ns
Chronic kidney disease	1.3 (1.1 to 1.6), p<0.01	1.2 (1 to 1.3), p<0.01	ns
Hypertension	1.2 (1 to 1.5), p<0.05	ns, p=0.196	1.2 (1 to 1.4), p<0.05
Cancer	1.2 (1 to 1.5), p<0.05	1.6 (1.5 to 1.8), p<0.001	0.8 (0.7 to 1), p<0.05
Cerebrovascular disease	ns, p=0.101	1.2 (1 to 1.3), p<0.05	ns
Ischaemic heart disease	ns, p=0.859	1.5 (1.3 to 1.6), p<0.001	0.7 (0.6 to 0.9), p<0.01
Depression	ns, p=0.34	1.1 (1 to 1.2), p<0.01	ns
Dementia	ns, p=0.059	0.8 (0.7 to 1), p<0.05	ns

Data are OR with 95% CIs. For age and IMD as categorical ordinal variables (data ranges shown), the comparators were the youngest and least deprived quintiles, respectively. Comorbidity associations are listed in descending OR order for the CA group. The comparison of CA vs NCA was by binary logistic regression. Variables not listed (table 1) were excluded stepwise as not significant. CA, COVID-19 admissions; CHA, COVID-19 hospital admissions; IMD, Index of Multiple Deprivation; NCA, non-COVID-19 admissions; NCHA, non-COVID-19 hospital admissions.

in figure 1C. In a simplified model of admission type and ethnicity ( $\chi^2=4542.9$ ,  $p<0.001$ ) with only age and deprivation entered categorically together with their interaction ( $\chi^2=412.7$ ,  $p<0.001$ ), the ORs for CA compared with the white group were as follows: black, 2.08 (1.70 to 2.57) ( $p<0.001$ ) and South Asian, 0.56 (0.44 to 0.70) ( $p<0.001$ ), with both groups still less likely to have NCA ( $p<0.001$ ).

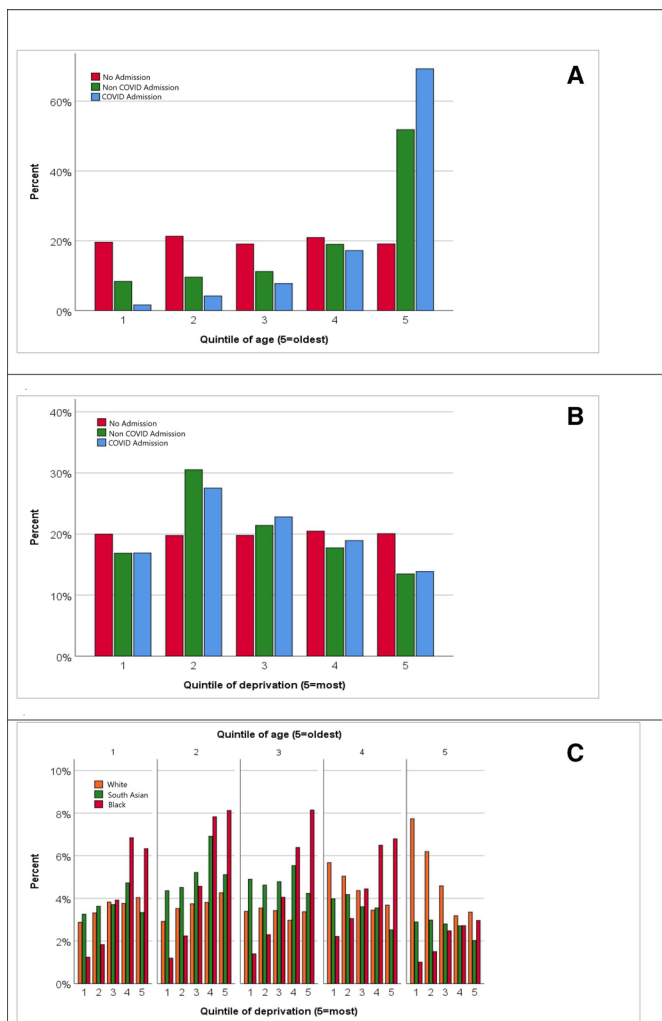
#### Absolute risk of CA within ethnic groups

The absolute risk from COVID-19 hospital admission was 4.8/1000 population, and table 2 shows this broken

down by ethnic grouping giving numbers, percentages, absolute risk and excess risk with ORs compared with the white group, with the South Asian group showing a lower and the black group a higher absolute risk as reflected in ORs.

#### Mortality outcomes

COVID-19 hospital deaths, non-COVID-19 hospital deaths and deaths in the community (CHD, NCHD, DIC) were analysed in stepwise backward multinomial regression ( $\chi^2=5548.3$ ,  $p<0.001$ ) (table 5).



**Figure 1** Age and deprivation in relation to hospital admission and in whole population. The association of age (A) and deprivation (B) with hospital admission type. (C) Inter-relationship of age, deprivation and ethnicity in the whole population (n=228 632) (other/unknown ethnic groups not shown).

Male gender was significantly positively associated with mortality in all three categories. Increasing age was a significant factor, but there was no significant difference in age quintile distribution ( $\chi^2=12.168$ ,  $p=0.144$ , ns) with 89%, 84% and 86% in the oldest quintile in the CHD, NCHD and DIC groups, respectively. For deprivation, for CHD and NCHD, the pattern mirrored that of hospital admission with significantly increased mortality rates in the lesser deprived quintiles but not in the highest quintile, whereas for DIC, a significant effect showing an increased mortality rate was only seen in the most deprived quintile. All categories shared a propensity for greater prior emergency admissions, multimorbidity and being in a palliative phase of care, while being a nursing home resident was associated with death in the community rather than HD. Individual morbidities varied in their associations, noting that diabetes and CKD were in increased association with mortality only in the CHD group. The black ethnic minority had significantly higher and the South

Asian group had significantly lower COVID-19 hospital mortality rate, proportionately mirroring admission rates. Directly comparing CHD with NCHD confirmed a significantly increased association with black ethnicity (OR 4.6 (2 to 10.2),  $p<0.003$ ), diabetes (OR 1.5 (1 to 2.3),  $p<0.005$ ) and CKD (OR 1.6 (1.1 to 2.3),  $p<0.004$ ) and an even greater negative association with current or previous smoking (OR 0.1 (0 to 0.3),  $p<0.002$ ).

### Absolute risk of COVID-19 death by ethnic group

Specifically for COVID-19 death, [table 2](#) shows numbers, percentages, absolute risk and excess risk with unadjusted ORs for the ethnic minorities compared with the white group, and [figure 2A](#) shows the distribution of mortality outcome by ethnic category ( $\chi^2=126.1$ ,  $p<0.001$ ). The absolute risk of COVID-19 death was 1.32, 0.73 and 2.2 per 1000 population in the white, South Asian and black ethnic groups, and the excess risk was  $-0.61$  (negative) and 0.85 deaths per 1000 population in the South Asian and black versus white group, respectively. Compared with the white population, the unadjusted OR (95% CI) for COVID-19 death for the black and Asian groups was 1.6 (1.2 to 2.3) and 0.5 (0.4 to 0.8), respectively (both  $p<0.01$ ). The ethnic groups differed significantly in age (white,  $50\pm 20$ ; South Asian,  $45\pm 16$ ; black,  $45\pm 17$  years;  $F=1868.9$ ,  $p<0.001$ ), and age was the dominant factor associated with hospital admission and death ([tables 1 and 3–5](#)). To avoid any potential misrepresentation of mortality outcomes by statistical age adjustment, the absolute effects were considered for the oldest quintile only where 84% of all COVID-19 deaths occurred, in which case the ORs were as follows: black, 3.9 (2.7 to 5.6) ( $p<0.001$ ) and South Asian, 0.9 (0.6 to 1.4) ( $p=0.72$ , ns) ([figure 2B](#)).

### COVID-19 hospital admission and COVID-19 mortality

By introducing hospital admission status, the COVID-19 mortality ORs—black, 1.3 (0.9 to 2.0) ( $p=0.206$ , ns) and South Asian, 1.5 (0.9 to 2.3) ( $p=0.098$ , ns)—were similar, indicating similar in-hospital mortality in contrast to the whole population effect. To negate this potential effect of prior propensity for acquisition of serious COVID-19 infection, and focusing on the black and South Asian minorities compared with the white majority, a narrower assessment of those who were admitted with COVID-19 and had a COVID-19 death was made. Among 930 CA, excluding those with a CA but with non-COVID-19 death (n=83 (9%)), COVID-19 death occurred in 270 patients (32%) (white, 189; South Asian, 32; black, 38; other, 11). The ORs for the association of ethnicity with COVID-19 mortality were as follows: black, 1.2 (0.7 to 1.8) ( $p=0.423$ , ns) and South Asian, 1.6 (1.0 to 2.5) ( $p=0.075$ , ns) ( $\chi^2=5.92$ ,  $p=0.115$ , ns) ([figure 2C](#)). Using the full model with all independent variables, including age, which remained significantly different between ethnic groups ( $F=13.23$ ,  $p<0.001$ ), the significantly associated variables were age, gender, smoking status, BMI, palliative phase of life, multimorbidity and the individual comorbidities of

**Table 4** Outcomes by individual ethnic category for those specifically with COVID-19 hospital admissions compared those not admitted within individual ethnic grouping as examined in binary regression analysis

	Black	South Asian	White
Numbers admitted/not admitted	113 (0.6%)/17 530	83 (0.2%)/43 663	684 (0.5%)/138 500
Gender (male)	ns	2 (1.2 to 3.2), p<0.01	1.6 (1.4 to 1.9), p<0.001
Age category Q2 (30–40)	2.3 (0.6 to 8.6), ns, p=0.225	1.8 (0.5 to 6.8), ns, p=0.417	3 (1.3 to 6.8), p<0.01
Age category Q3 (41–51)	3.1 (0.9 to 11), ns, p=0.083	2.6 (0.7 to 9.3), ns, p=0.148	4.9 (2.2 to 10.5), p<0.001
Age category Q4 (52–65)	3.8 (1.1 to 13.4), p<0.05	3.1 (0.9 to 10.8), ns, p=0.082	9.7 (4.7 to 20), p<0.001
Age category Q5 (66–113)	10 (2.8 to 36), p<0.001	4.5 (1.3 to 16), p<0.05	21.2 (10.4 to 43.3), p<0.001
IMD category Q2 (16.5–27.7)	ns	ns	1.7 (1.4 to 2.1), p<0.001
IMD category Q3 (27.8–39.0)	ns	ns	1.6 (1.3 to 2.1), p<0.001
IMD category Q4 (39.3–45.7)	ns	ns	1.5 (1.2 to 2), p<0.01
IMD category Q5 (45.7–71.8)	ns	ns	ns
Smoking current or prior	0.3 (0.2 to 0.6), p<0.001	0.2 (0.1 to 0.6), p<0.01	0.3 (0.2 to 0.3), p<0.001
Prior emergency admissions (1 year)	ns	1.5 (1.2 to 1.8), p<0.001	1.5 (1.4 to 1.6), p<0.001
Palliative care registered	3.8 (1.8 to 8), p<0.001	11 (5.5 to 21.7), p<0.001	3.8 (2.9 to 5), p<0.001
Nursing home resident	ns	4.5 (1.4 to 14.2), p<0.05	1.6 (1.1 to 2.2), p<0.01
Comorbidities ≥3	2.6 (1.5 to 4.4), p<0.001	ns	1.7 (1.3 to 2.1), p<0.001
Atrial fibrillation	2 (1 to 3.7), p<0.05	2.8 (1.4 to 5.6), p<0.01	1.6 (1.3 to 2), p<0.001
Cardiac failure	2.2 (1.2 to 3.9), p<0.05	ns	1.7 (1.3 to 2.1), p<0.001
Chronic kidney disease	ns	ns	1.4 (1.1 to 1.7), p<0.01
Chronic obstructive pulmonary disease	ns	ns	1.9 (1.5 to 2.3), p<0.001
Diabetes	ns	4.4 (2.6 to 7.5), p<0.001	1.5 (1.3 to 1.9), p<0.001
Hypertension	2.2 (1.2 to 4.1), p<0.01	ns	ns
Peripheral vascular disease	ns	3.7 (1.6 to 8.8), p<0.01	1.9 (1.4 to 2.5), p<0.001
Rheumatoid arthritis	ns	ns	1.6 (1.1 to 2.1), p<0.01

Values are OR with 95% CIs.  
IMD, Index of Multiple Deprivation.

cardiac failure, CKD and peripheral vascular disease but not ethnic grouping or deprivation score (table 6).

Finally, table 2 shows the absolute risks of COVID-19 death in COVID-19 hospital admission and ORs, which are consistent with the findings of the modelled data.

## DISCUSSION

### Principal findings

Over and above known general associations with hospital admission and mortality, our study suggests a complex association of deprivation and points to heterogeneity of the impact of ethnicity, both of which may vary by locality. We highlight the need for local health economies to have robust, accurate and integrated clinical data to assess and inform local decision-making, and, in particular, at a time of heightened anxiety, we raise a concern about the conveyance of risk to local communities. The crucial differences in relationship to other studies are as follows:

### General associations

Uncontroversially, factors associated with NCA or COVID-19 hospital admission and death included age, gender, prior emergency admissions, palliative phase of life, nursing home residence and multimorbidity with specific comorbidities associated with CA or death and with ethnic status. Within the limitations of our study, we have found smokers as an under-represented group in CA and mortality. Although a number of hypotheses have been proposed to account for a possible protective effect, this remains an area under further evaluation.<sup>18</sup> It is suggested that any association of smoking with better COVID-19 outcomes, observed in some other studies,<sup>2 19</sup> may be questioned when taken in the context of this being common to NCA and death during this period.

### CA versus NCA

The significant differences were age, gender and degree of comorbidity complexity (palliative care, nursing home),

**Table 5** Multinomial regression of mortality outcomes over the 12-week period for those who died (n=1093) either in the community (n=537) or non-COVID-19-related (n=284) and COVID-19-related (n=272) hospital deaths (HD) (see table 1) compared with those who were alive (n=227 539)

	COVID-19 HD	Non-COVID-19 HD	Community death
Gender (male)	2 (1.5 to 2.6), p<0.001	1.5 (1.2 to 1.9), p<0.01	1.7 (1.4 to 2.1), p<0.001
Age category Q2 (30–40)	ns, p=0.997	>100, p<0.001	ns, p=0.85
Age category Q3 (41–51)	2.5 (0.5 to 14), ns, p=0.283	>100, p<0.001	6.6 (1.9 to 22.6), p<0.01
Age category Q4 (52–65)	10.5 (2.4 to 44.8), p<0.01	>100, p<0.001	13.9 (4.3 to 44.9), p<0.001
Age category Q5 (66–113)	39.9 (9.6 to 165.5), p<0.001	>100, p<0.001	41.3 (13 to 130.9), p<0.001
IMD category Q2 (16.5–27.7)	ns, p=0.163	1.8 (1.2 to 2.6), p<0.01	ns, p=0.708
IMD category Q3 (27.8–39.0)	1.5 (1 to 2.2), p<0.05	2.2 (1.5 to 3.1), p<0.001	ns, p=0.054
IMD category Q4 (39.3–45.7)	ns, p=0.319	1.8 (1.2 to 2.7), p<0.01	ns, p=0.384
IMD category Q5 (45.7–71.8)	ns, p=0.713	ns, p=0.455	1.6 (1.2 to 2.1), p<0.01
Ethnicity South Asian	0.5 (0.3 to 0.8), p<0.01	0.4 (0.3 to 0.6), p<0.001	0.5 (0.3 to 0.7), p<0.001
Ethnicity black	2.1 (1.5 to 3.2), p<0.001	0.5 (0.3 to 1), p<0.05	ns, p=0.841
Smoking current or prior	0 (0 to 0), p<0.001	0.1 (0.1 to 0.2), p<0.001	0.1 (0.1 to 0.1), p<0.001
Body mass index	1 (0.9 to 1), p<0.01	1 (1 to 1), ns, p=0.12	1 (1 to 1), p<0.05
Prior emergency admissions (1 year)	1.2 (1.1 to 1.3), p<0.001	1.3 (1.2 to 1.4), p<0.001	1.3 (1.2 to 1.4), p<0.001
Palliative care registered	9.5 (6.8 to 13.2), p<0.001	5.9 (4.2 to 8.4), p<0.001	5.9 (4.7 to 7.5), p<0.001
Nursing home resident	ns, p=0.547	ns, p=0.634	3.5 (2.7 to 4.6), p<0.001
Comorbidities ≥3	3 (2 to 4.4), p<0.001	2.3 (1.6 to 3.3), p<0.001	1.7 (1.3 to 2.3), p<0.001
Peripheral vascular disease	2.6 (1.8 to 3.9), p<0.001	1.9 (1.3 to 2.9), p<0.01	ns, p=0.232
Chronic obstructive pulmonary disease	2.1 (1.4 to 3), p<0.001	3.8 (2.8 to 5.2), p<0.001	2.2 (1.7 to 2.9), p<0.001
Cardiac failure	2 (1.5 to 2.8), p<0.001	1.7 (1.2 to 2.3), p<0.01	1.8 (1.4 to 2.3), p<0.001
Chronic kidney disease	1.6 (1.2 to 2.1), p<0.01	ns, p=0.095	ns, p=0.328
Diabetes	1.4 (1.1 to 1.9), p<0.05	ns, p=0.465	ns, p=0.426
Atrial fibrillation	1.4 (1 to 1.9), p<0.05	1.9 (1.4 to 2.6), p<0.001	ns, p=0.602
Cancer	ns, p=0.338	1.5 (1.1 to 2), p<0.01	2.1 (1.7 to 2.6), p<0.001
Dementia	ns, p=0.878	ns, p=0.505	2 (1.5 to 2.6), p<0.001
Asthma	ns, p=0.337	0.7 (0.5 to 1), p<0.05	ns, p=0.376

Results are OR with 95% CIs. For age and Index of Multiple Deprivation categories, the comparators were the youngest and least deprived quintile, respectively. Variables not listed from table 1 were excluded stepwise (backwards) as not significant. IMD, Index of Multiple Deprivation.

but as it is likely that patterns of emergency admissions differed at this time, comparisons of COVID-19 with non-COVID-19 hospital admission may have little relevance to COVID-19 outcomes, noteworthy for studies that have reported on COVID-19 hospital admission alone.<sup>12 13 20</sup>

### Deprivation

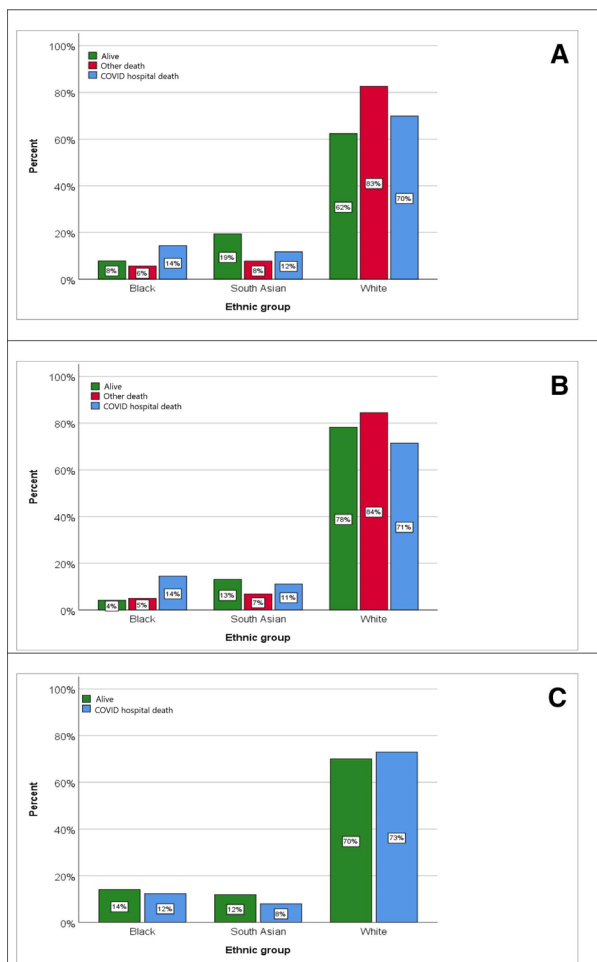
For hospital NCA and CA and death, the pattern was for excess in lesser deprived quintiles in the white ethnic population but not within ethnic minority groups where deprivation was not a significant factor. This contrasts with other studies<sup>25</sup>: in some, deprivation was not a significantly associated factor in fully adjusted models,<sup>21</sup> whereas other UK studies,<sup>20</sup> and most overseas studies, have not considered this.<sup>12 13</sup> Following the H1N1 pandemic influenza of 2009, many studies indicated effects of deprivation, including a rural–urban divide impact,<sup>22</sup> as is seen in this

pandemic.<sup>15</sup> Our findings within a health economy (ie, based on a local population) with significant deprivation call for the need to explore this association within larger studies, specifically within urban areas.

### Ethnicity

We note that a recent meta-analysis shows heterogeneity in the association of ethnicity with COVID-19 mortality.<sup>23</sup> In a large population study reporting adverse ORs for all ethnic groups, their crude unadjusted data showed significantly increased risk in the black ( $\chi^2=17.464$ , p<0.001) but not in the South Asian group ( $\chi^2=3.238$ , p=0.072).<sup>2</sup> This was also shown in another population-level study.<sup>15</sup> In the largest reported hospital admission series, both ethnic groups had significantly lower unadjusted mortality rates,<sup>17</sup> whereas in their modelled data no effect was seen among the black population. In a study





**Figure 2** Mortality by ethnicity. Crude mortality by ethnic grouping as percentages (A,  $\chi^2=184.4$ ,  $p<0.001$ ), within the oldest quintile (B,  $\chi^2=92.2$ ,  $p<0.001$ ) or restricted to those with a COVID-19 admission excluding those with a non-COVID death (C,  $\chi^2=5.92$ ,  $p=0.115$ , ns) (other/unknown ethnic categories are not shown but were included in the analysis).

from New York, there was no adverse ethnicity signal,<sup>12 13</sup> and early reported adverse ethnicity outcomes in the 2009 UK influenza pandemic<sup>24</sup> did not withstand subsequent review.<sup>25</sup> We find that our black population had significantly higher and the South Asian population lower crude and adjusted CA rates compared with the white population, also observing that both ethnic subgroups had lower non-COVID-19 hospital admissions, further contextualising the strong effect in the black cohort. In both groups, their crude and adjusted patterns of COVID-19 mortality mirrored that of COVID-19 hospital admission, but from the numerical base of CA, there was no significant difference between the black and South Asian groups compared with the white group, highlighting pitfalls of examining effects in isolation. Our data in the black population are broadly in keeping with some studies showing excess COVID-19 hospitalisation and mortality, but the South Asian group’s lower absolute and adjusted rate of admission and death from COVID-19 are strikingly different. Given the variation in findings to date, we do not consider

**Table 6** Multinomial regression among those with a COVID-19 admission restricted to white, black and South Asian ethnic groups (n=797) comparing those who have died due to COVID-19 (n=259) with those who were alive at 12 weeks

	OR COVID-19 death vs alive
Age	1.05 (1.03 to 1.07), $p<0.001$
Gender (male)	1.91 (1.3 to 2.83), $p<0.01$
Smoking current or prior	0.01 (0 to 0.04), $p<0.001$
Body mass index (kg/m <sup>2</sup> )	0.92 (0.86 to 0.98), $p<0.01$
Palliative care registered	7.83 (4.21 to 14.56), $p<0.001$
Comorbidities $\geq 3$	1.64 (1 to 2.67), $p<0.05$
Cardiac failure	1.82 (1.14 to 2.92), $p<0.05$
Chronic kidney disease	1.69 (1.09 to 2.63), $p<0.05$
Peripheral vascular disease	2.27 (1.12 to 4.59), $p<0.05$

Results are OR with 95% CIs. Variables not listed from table 1 were excluded stepwise (backwards) as not significant.

this an ‘unexpected finding’ and hypothesise that many local population factors are at play, including population density, family size, housing, duration of immigration, country of birth (including UK born) and occupation, and the precise ethnic group within the ‘South Asian’ population may well be of importance. A recent updated analysis by the UK Office for National Statistics has emphasised that ‘ethnic differences in mortality involving COVID-19 are most strongly associated with demographic and socio-economic factors, such as place of residence and occupational exposures, and cannot be explained by pre-existing health conditions’, which conclusion is consistent with our locally dictated findings.<sup>26</sup> Otherwise within Black, Asian and minority ethnic groups, we find specific individual comorbidities vary in their association with COVID-19 risks: in South Asians, these are diabetes, peripheral vascular disease and atrial fibrillation; in the Black population, hypertension, atrial fibrillation and cardiac failure; in white ethnicity, most comorbidities but in particular COPD, chronic kidney disease and RA.

### Strengths and weaknesses

Combining Wolverhampton’s health data enabled us to evaluate our local population’s heterogeneous demographic factors and their associations with community or hospital NCA and CA and mortality, by uniquely approaching these outcomes simultaneously. This local nuance complements larger studies, informing appraisal of risk from an urban and multiethnic and deprived setting, highlighting concerns of extrapolation from larger data sets to UK localities. An example of a particular strength of the data quality was the cross-check ascertainment of CA, without sole reliance on COVID-19 testing, permitting specific categorisation of deaths (COVID-19, non-COVID-19 and postdischarge) rather than less accurately into global mortality.



### Limitations of the study

This is a 12-week evaluation spanning the pandemic's upsurge and peak; the population and event number were comparatively small; the cause of death in the community was unknown, and it is likely that people died away from the hospital undiagnosed with COVID-19. A further weakness of the study is that there were some missing data, but this was very limited in magnitude and only affected three variables: BMI, smoking and ethnicity. We are confident that these were dealt with appropriately; for BMI, it is as described in Methods; for smoking, we coded all unknown smoking as non-smokers on the very likely assumption that the vastly greater majority were non-smokers, whereas missing ethnicity was coded as 'unknown' and analysed as such. Given the degree of completeness rather than incompleteness of our data, we consider our approach approximates to a complete case analysis, arising from significant effort on multisource data accrual, integration and quality. We thus do not feel that multiple imputation should be applied to replace missing data, since we do not feel this can possibly improve precision. In so doing, we are thus also avoiding the greater and well-recognised potential to introduce bias from poorly fitting imputation models.<sup>27</sup> We consider this to be a strength of the article. One further consideration is that while being aligned to the population at 99.5% concordance, hospital data were not totally drawn from the city population, which varied by GP registration, residency or admission from immediately surrounding areas, and a small proportion of admissions were non-resident or non-registered, so this is not strictly an epidemiological study but an observational study comparing defined cohorts in tiers of analysis (eg, COVID-19 death among CA) where this caveat does not apply.<sup>28</sup>

### Implications for clinicians and policymakers

We show that a variety of recognised factors were associated with COVID-19 death, as with non-COVID-19 death. At our local level, CA and death were not strongly associated with worsening deprivation, with a novel potential different relationship in the white population.

Higher absolute and adjusted CA and mortality occurred in the black population, whereas they were reduced in Wolverhampton's South Asian community. We point out the non-significant association between in-hospital COVID-19 case fatality and ethnicity, raising the probability that COVID-19 mortality relates to differential risks of exposure, susceptibility and disease contraction before hospital admission, let alone the possible avoidance of hospital admission. Two important considerations are the potential excessive use of multiple factors and the disruption of the perspective from a population's base through hospital admissions to COVID-19-specific admission, leading to widely varying conclusions and highlighting the difficulties of using observational data and the potential for collider bias.<sup>29</sup> We support the case for more localised population-based studies of both hospital admission and subsequent death, such as ours, in which

the denominator and numerator populations can be clearly linked and are fully and transparently ascertained and characterised. To avoid associations in the data being due to the way in which data are sampled, local health economies should be mandated to link hospital and primary care data across their population level down to the unanonymised individual level; they should, in preparation for future epidemics, have data quality mechanisms in place to ensure accuracy in their demographics, the accrual of important missing data and the triangulation of key outcomes to minimise false-positive and false-negative results. This includes the need to have a robust, systematic, accurate and timely approach to the recording of death whether in the community or hospital setting. A defined data set and its capture in routine clinical systems seem apposite.<sup>30</sup> Accepting that variation in findings in different population subsets is both inevitable and valid, we would suggest the need for the public health and research community to accommodate uncertainty in emergent evidence, learning from the experience of previous viral pandemics. This includes the need to have a robust, systematic, accurate and timely approach to the recording of death whether in the community or hospital setting.<sup>31</sup>

### The conveyance of risk

Public health messages are vital to convey, but population-adjusted risk rates may confuse, adversely impacting behaviours such that, it is feared, hospital admission patterns may change unfavourably. Absolute, absolute excess, relative, unadjusted and adjusted risk is complex to communicate even for healthcare professionals, making them susceptible to reasoning errors and misinterpretation of probabilities,<sup>32</sup> and individuals, with erstwhile health risk, should know about the magnitude of risk in a way that can be conceptualised.<sup>33,34</sup> For our black population, the full modelled OR for COVID-19 mortality was 2.1, with an absolute risk of 2.2/1000 people or an excess risk of 0.9/1000. A black person in Wolverhampton ought to be informed that 'twice as likely to die of COVID-19' compared with the white community can also mean 'a 1 in 1000 excess risk'.

### Future research

Crucially, we therefore argue that in reporting future research of this kind, during the current pandemic and beyond, there is an ethical obligation for the standardisation of the conveyance of risk in a manner that spans the absolute to the relative so that is easily comprehensible to the individuals and populations at risk and others, including health professionals, politicians and the media. These are all matters in which the editorial and peer review mechanism of our medical journals have a vital role.

### Author affiliations

<sup>1</sup>New Cross Hospital, Royal Wolverhampton Hospitals NHS Trust, Wolverhampton, UK  
<sup>2</sup>School of Medicine & Clinical Practice, Faculty of Science & Engineering, University of Wolverhampton, Wolverhampton, UK

<sup>3</sup>Faculty of Health, Education & Life Sciences, Birmingham City University, Birmingham, UK

<sup>4</sup>Faculty of Education Health & Wellbeing, Walsall Campus, University of Wolverhampton, Wolverhampton, UK

**Contributors** BMS was accountable senior author; BMS, JB, SJD and AV helped with data analysis and manuscript writing; SJD, BMS and JB helped with preparation for submission; AV, BMS and VK ensured database quality and helped with data integration, and data quality and integration. SM helped with reading drafts as lay expert. AN provided statistical advice. All authors contributed intellectual content during the drafting and revision of the work and approved the final version.

**Funding** The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

**Competing interests** None declared.

**Patient consent for publication** Not required.

**Ethics approval** This was not sought nor deemed necessary as this work represents a continuous quality improvement programme of the informatics component of service changes required between various local National Health Service (NHS) organisations for integrated working stipulated during the COVID-19 emergency. Data governance was in line with Trust policy and with the COVID-19 emergency directive of NHS England.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data availability statement** Data are available upon reasonable request. Anonymised data will be shared on reasonable request to the corresponding author Prof B M Singh (baldev.singh@nhs.net)

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

#### ORCID iDs

Baldev M Singh <http://orcid.org/0000-0003-3460-6759>

James Bateman <http://orcid.org/0000-0003-1952-2821>

Alan Nevill <http://orcid.org/0000-0003-0506-3652>

Simon J Dunmore <http://orcid.org/0000-0001-7227-5597>

#### REFERENCES

- Xu B, Kraemer MUG, Open COVID-19 Data Curation Group. Open access epidemiological data from the COVID-19 outbreak. *Lancet Infect Dis* 2020;20:534.
- Williamson EJ, Walker AJ, Bhaskaran K, *et al*. Factors associated with COVID-19-related death using OpenSAFELY. *Nature* 2020;584:430–6.
- Mehra MR, Ruschitzka F, Patel AN. Retraction-Hydroxychloroquine or chloroquine with or without a macrolide for treatment of COVID-19: a multinational registry analysis. *Lancet* 2020;395:1820.
- Rubin EJ, Baden LR, Morrissey S, *et al*. Medical journals and the 2019-nCoV outbreak. *N Engl J Med* 2020;382:866.
- Public Health England. Disparities in the risk and outcomes of COVID-19. Available: [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/892085/disparities\\_review.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/892085/disparities_review.pdf) [Accessed 20 Jul 2020].
- Office for national statistics. Available: <https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths> [Accessed 20 Jul 2020].
- Willett K. COVID-19 patient notification system (CPNS). Available: <https://www.england.nhs.uk/coronavirus/wp-content/uploads/sites/52/2020/04/C0389-update-to-cpns-reporting-letter-27-april-2020.pdf> [Accessed 20 Jul 2020].
- Woloshin S, Patel N, Kesselheim AS. False Negative Tests for SARS-CoV-2 Infection - Challenges and Implications. *N Engl J Med* 2020;383:e38.
- Bradley SH, Lawrence NR, Carder P. Using primary care data for health research in England - an overview. *Future Healthc J* 2018;5:207–12.
- Zhou F, Yu T, Du R. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study [published correction appears in *Lancet*. 2020 Mar 28;395(10229):1038] [published correction appears in *Lancet*. 2020 Mar 28;395(10229):1038]. *Lancet* 2020;395:1054–62.
- Goyal P, Choi JJ, Pinheiro LC, *et al*. Clinical characteristics of Covid-19 in New York City. *N Engl J Med* 2020;382:2372–4.
- Argenziano MG, Bruce SL, Slater CL, *et al*. Characterization and clinical course of 1000 patients with coronavirus disease 2019 in New York: retrospective case series. *BMJ* 2020;369:m1996.
- Petrilli CM, Jones SA, Yang J, *et al*. Factors associated with hospital admission and critical illness among 5279 people with coronavirus disease 2019 in New York City: prospective cohort study. *BMJ* 2020;2:m1966.
- Gianfrancesco M, Hyrich KL, Al-Adely S, *et al*. Characteristics associated with hospitalisation for COVID-19 in people with rheumatic disease: data from the COVID-19 global rheumatology alliance physician-reported registry. *Ann Rheum Dis* 2020;79:859–66.
- de Lusignan S, Dorward J, Correa A, *et al*. Risk factors for SARS-CoV-2 among patients in the Oxford Royal College of general practitioners research and surveillance centre primary care network: a cross-sectional study. *Lancet Infect Dis* 2020;20:1034–42.
- Izcovich A, Ragusa MA, Tortosa F, *et al*. Prognostic factors for severity and mortality in patients infected with COVID-19: a systematic review. *PLoS One* 2020;15:e0241955.
- Harrison EM, Docherty AB, Barr B. Ethnicity and outcomes from COVID-19: the ISARIC CCP-UK prospective observational cohort study of hospitalised patients (5/31/2020). Available: <https://ssrn.com/abstract=3618215>
- Usman MS, Siddiqi TJ, Khan MS, *et al*. Is there a smoker's paradox in COVID-19? *BMJ Evid Based Med* 2020. doi:10.1136/bmjebm-2020-111492. [Epub ahead of print: 11 Aug 2020].
- Farsalinos K, Barbouni A, Niaura R. Systematic review of the prevalence of current smoking among hospitalized COVID-19 patients in China: could nicotine be a therapeutic option? *Intern Emerg Med* 2020;15:845–52.
- 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.
- 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.
- Rutter PD, Mytton OT, Mak M, *et al*. Socio-Economic disparities in mortality due to pandemic influenza in England. *Int J Public Health* 2012;57:745–50.
- Pan D, Sze S, Minhas JS, *et al*. The impact of ethnicity on clinical outcomes in COVID-19: a systematic review. *EClinicalMedicine* 2020;23:100404.
- Nguyen-Van-Tam JS, Openshaw PJM, Hashim A, *et al*. Risk factors for hospitalisation and poor outcome with pandemic A/H1N1 influenza: United Kingdom first wave (May–September 2009). *Thorax* 2010;65:645–51.
- Nyland GA, McKENZIE BC, Myles PR, *et al*. Effect of ethnicity on care pathway and outcomes in patients hospitalized with influenza A(H1N1)pdm09 in the UK. *Epidemiol Infect* 2015;143:1129–38.
- Office for national statistics. Available: <https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/articles/updatingethniccontrastsindeathsinvolvingthecoronaviruscovid19englandandwales/deathsoccurring2marchto28july2020#main-points> [Accessed 22 Oct 20].
- Lee KJ, Carlin JB. Recovery of information from multiple imputation: a simulation study. *Emerg Themes Epidemiol* 2012;9.
- Pearce N. Classification of epidemiological study designs. *Int J Epidemiol* 2012;41:393–7.
- 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.
- Cross M. Show us the data: why clinical outcomes matter. *BMJ* 2012;344:e66.
- Luce T. COVID 19: death certification in England and Wales. *BMJ* 2020;369:m1571.
- Grant RL. Converting an odds ratio to a range of plausible relative risks for better communication of research findings. *BMJ* 2014;348:f7450.
- Paling J. Strategies to help patients understand risks. *BMJ* 2003;327:745–8.
- Ahmed H, Naik G, Willoughby H, *et al*. Communicating risk. *BMJ* 2012;344:e3996.