



OPEN

The interaction of ethnicity and deprivation on COVID-19 mortality risk: a retrospective ecological study

Kausik Chaudhuri¹, Anindita Chakrabarti¹, Jose Martin Lima¹, Joht Singh Chandan² & Siddhartha Bandyopadhyay³✉

Black, Asian and Minority Ethnic (BAME) populations are at an increased risk of developing COVID-19 and consequentially more severe outcomes compared to White populations. The aim of this study was to quantify how much of the disproportionate disease burden can be attributed to ethnicity and deprivation as well as its interaction. An ecological study was conducted using data derived from the Office for National Statistics data at a Local Authority District (LAD) level in England between 1st March and 17th April 2020. The primary analysis examined how age adjusted COVID-19 mortality depends on ethnicity, deprivation, and the interaction between the two using linear regression. The secondary analysis using spatial regression methods allowed for the quantification of the extent of LAD spillover effect of COVID-19 mortality. We find that in LADs with the highest deprivation quartile, where there is a 1 percentage point increase in “Black-African (regression coefficient 2.86; 95% CI 1.08–4.64)”, “Black-Caribbean (9.66; 95% CI 5.25–14.06)” and “Bangladeshi (1.95; 95% CI 1.14–2.76)” communities, there is a significantly higher age-adjusted COVID-19 mortality compared to respective control populations. In addition, the spatial regression results indicated positive significant correlation between the age-adjusted mortality in one LAD and the age-adjusted mortality in a neighbouring LAD, suggesting a spillover effect. Our results suggest targeted public health measures to support those who are deprived and belong to BAME communities as well as to encourage restricted movement between different localities to limit disease propagation.

The COVID-19 pandemic, caused by the Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-CoV-2) was originally declared a public health emergency on January 30th 2020¹, and has since led to the deaths of over 700,000 individuals globally². The first recorded case in England was on the 29th January 2020, following which England has become one of most globally affected countries in terms of case numbers and mortality^{2,3}.

The United Kingdom (UK) is an ethnically diverse country, and as the pandemic has progressed we find that the risk of both COVID-19 susceptibility and mortality varies with ethnicity. Medical records from primary and secondary care as well as death certificates examined by the Office for National Statistics (ONS), Public Health England (PHE) and other academics make it clear that those from Black, Asian and Minority Ethnic (BAME) communities are at the highest risk of both COVID-19 susceptibility and associated mortality^{4–6}. These findings have been replicated internationally and raise concern for urgent public health action to identify the cause for such a correlation^{7,8}.

However, an important point to note is that within the UK and abroad, those from BAME communities are more likely to live in deprived areas which may in turn increase their risk of health inequalities when compared to the reference White population in the UK⁹. Existing data taken by the ONS and PHE examined the impact of age, gender and deprivation separately on COVID-19 mortality rates but did not examine the interaction between ethnicity and deprivation and how this may affect such risk^{4,5}. More recent studies examining the impact of ethnicity on COVID-19 using intensive care unit data and hospital data clearly raise the need for examining the

¹Economics Division, Leeds University Business School, Leeds LS2 9JT, UK. ²Institute of Applied Health Research, University of Birmingham, Birmingham B152TT, UK. ³Department of Economics, Birmingham Business School and Centre for Crime Justice and Policing, University of Birmingham, Birmingham B15 2TT, UK. ✉email: s.bandyopadhyay@bham.ac.uk

Variable	Overall ^b Mean (SD)	Quartile 1 Mean (SD)	Quartile 2 Mean (SD)	Quartile 3 Mean (SD)	Quartile 4 Mean (SD)
Age-adjusted Mortality (per 100,000 people)	35.501 (24.858)	27.727 (11.256)	32.057 (23.603)	35.009 (25.656)	47.210 (32.245)
IMD (Average score)	19.641 (7.986)	10.538 (1.887)	15.680 (1.483)	21.703 (2.030)	30.643 (4.579)
Black-African ^a	1.311 (2.420)	0.451 (0.341)	1.086 (1.608)	1.551 (2.598)	2.155 (3.552)
Black-Caribbean ^a	0.739 (1.518)	0.270 (0.272)	0.557 (0.855)	0.880 (1.746)	1.249 (2.218)
Black-Other ^a	0.340 (0.756)	0.101 (0.077)	0.244 (0.395)	0.416 (0.827)	0.601 (1.041)
Chinese ^a	0.600 (0.585)	0.503 (0.327)	0.624 (0.680)	0.544 (0.571)	0.730 (0.674)
Bangladeshi ^a	0.592 (2.051)	0.183 (0.230)	0.345 (0.579)	0.446 (0.793)	1.393 (3.935)
Indian ^a	2.051 (3.638)	1.510 (2.287)	2.002 (3.836)	1.991 (3.301)	2.702 (4.677)
Pakistani ^a	1.389 (2.894)	0.589 (1.179)	0.736 (1.638)	1.291 (2.718)	2.939 (4.341)
Asian-Other ^a	1.252 (1.671)	0.961 (1.066)	1.393 (2.129)	1.330 (1.814)	1.323 (1.482)
Mixed ^a	1.893 (1.382)	1.566 (0.665)	1.741 (1.150)	1.979 (1.599)	2.288 (1.759)
Arab ^a	0.294 (0.616)	0.158 (0.217)	0.221 (0.313)	0.393 (1.025)	0.404 (0.538)
Other Ethnic Group ^a	0.462 (0.708)	0.263 (0.230)	0.382 (0.536)	0.529 (0.839)	0.674 (0.941)
White ^a	89.076 (12.996)	93.445 (5.099)	90.669 (11.717)	88.649 (13.683)	83.542 (16.687)
Observations	316	79	79	79	79

Table 1. Characteristics of LADs: summary statistics. Quartile 1–4 denotes the quartiles based on index of multiple deprivation. Quartile 1 is least deprived and Quartile 4 is the most deprived. ^aDenotes in terms of percentage of total population in the LAD. ^bOverall indicates for the whole of England.

association with socio-economic characteristics in more detail and also for the urgent need to identify cultural risk factors which may be susceptible for intervention^{10,11}.

When considering the effect between ethnicity and deprivation, it must also be important to analyse “neighbourhood effects,” which as seen even in the Global Burdens of Disease work have a substantial effect on population health^{12,13}. Neighbourhood effect is defined as the impact from living in a particular locality¹⁴, in this case we can think of it as a LAD. A large literature exists that suggest that one’s neighbourhood can affect individual outcomes, including health¹⁵. Spillovers refer to effects percolating from one LAD to a contiguous one. Therefore, for an effective analysis examining the interaction between ethnicity, deprivation and COVID-19 risk we must consider the impact of the neighbourhood effect as well as spillover effects between neighbouring regions. Not considering these spillover effects would potentially underestimate the full effect of interaction between ethnicity, deprivation and COVID mortality. Although some work has started to examine spatial distribution of COVID-19 even within the UK, to our knowledge no studies have examined spatial distribution of infection rates between Local Authorities Districts (LAD) in the UK which would give an indication of how COVID-19 is being transmitted due to social interaction and local travel^{16–18}.

Therefore, the first aim of this study was to examine the interaction between ethnicity and deprivation when associating ethnicity to COVID-19 mortality. In addition, we aimed to examine spillover effects between neighbouring LADs in England.

Methods

Study design and participants. In this study we used open source LAD (a subnational division of England used for the purposes of local government) level data from England to undertake a cross sectional analysis of our variables of interest. There are 317 LADs in England, and we used data from 315 of them excluding the City of London and the Isle of Sicily which were dropped due to missing observations. The primary focus of this study was to analyse the disparity in COVID-19 deaths for individuals belonging to BAME groups vis-à-vis the White population (while neither of these categories are homogenous, we use sub-groups as used by the Census)¹⁹. Secondly, we assessed the importance of the index of multiple deprivation (IMD), in explaining the age adjusted mortality across the different local areas in England after we control for ethnicity. Finally, we quantified the extent of a spillover effect of COVID-19 disease burden that helped us to assess the impact of social-distancing measures introduced by the government.

Data sources and definitions. The data are taken from different sources. Our dependent variable is the age standardised death rate data for all persons at LADs in England is taken from the ONS across 315 LADs from 1 March to 17 April 2020²⁰. This is defined as the age adjusted mortality per 100,000 of the population of the LAD. We took into account LAD boundary changes which occurred in April 2019, for example Bournemouth, Poole and Christchurch combined as one LAD. The average score for the index of multiple deprivation, ethnicity data and educational attainment used were taken from the Office for National Statistics^{19,21}. Data on density of population expressed as total population per hectare was taken from Census 2011, Table PHP01. A-level educational qualification (proportion with GCE A level or equivalent—aged 16–64) were taken from <https://www.nomisweb.co.uk/query/construct/components/stdSearch.asp?menuopt=7&subcomp=131>.

We classified ethnicity using the broad headings used in the 2011 Census²². Our dependent variable and independent variables are described in Table 1.

The dependent variable is the age adjusted COVID-19 death rate per 100,000 in the population. The 'Quartile Dummy_{*i*} (*j* = 1, 2, 3, 4)' presents the quartile dummies for the index of multiple socio-economic deprivation. We use 'Quartile 1 Dummy' representing the bottom 25% (least deprived) as the base. In addition, two other explanatory variables were incorporated to capture population per square kilometre within each LAD (i.e. *pop - density*) and education status within a LAD (*edu-status*) measured in terms of proportion of individuals within LAD who have completed A-level. Each ethnicity variable represented the % of people with that ethnicity in that LAD. Our categorisation of the ethnicity variable departs from other previous studies by disaggregating within broad ethnic groups¹⁰. Instead of using Black as a whole, we divide 'Black' into 'Black-African', 'Black-Caribbean' and 'Black-Others'. For the Asian ethnicity, we use the following categories: 'Bangladeshi', 'Indian', 'Pakistani', 'Arab', 'Chinese' and 'Asian-Other'. The 'Mixed' category is a combination of 'Mixed White & Asian', 'Mixed White & Black African', 'Mixed White and Black Caribbean' and 'Mixed Other'. Along with the above, we also have 'Other Ethnic group' and the excluded group is the 'White' community which hence serves as the reference category.

Statistical analysis. Two types of statistical analysis were performed.

A: An OLS (ordinary least squares) regression model was estimated to see the impact of ethnicity, particularly the interaction of ethnicity with economic deprivation in explaining the mortality differential across the LADs while controlling for population density and education status that may also affect mortality.

We note that the sum of the ethnic population percentages in each LAD equals 100. Therefore, we have excluded the 'White' community, which acts as the reference category, to eliminate the collinearity problem. The estimated equation takes the following form:

$$\begin{aligned}
 \text{COVID}_i = & \alpha + \sum_{j=2}^4 \beta_j \text{QuartileDummy}_i + \beta_5 \text{Bangladeshi}_i + \beta_6 \text{Indian}_i \\
 & + \beta_7 \text{Pakistani}_i + \beta_8 \text{Arab}_i + \beta_9 \text{Chinese}_i + \beta_{10} \text{Asian - Other}_i + \beta_{11} \text{Black - African}_i \\
 & + \beta_{12} \text{Black - Caribbean}_i + \beta_{13} \text{Black - Other}_i + \beta_{14} \text{Mixed}_i + \beta_{15} \text{Other - Ethnic}_i \\
 & + \sum_{j=2}^4 \gamma_j \text{Quartile}_j * \text{Bangladeshi}_i + \sum_{j=2}^4 \varnothing_j \text{Quartile}_j * \text{Indian}_i + \sum_{j=2}^4 \mu_j \text{Quartile}_j * \text{Pakistani}_i \\
 & + \sum_{j=2}^4 \tau_j \text{Quartile}_j * \text{Arab}_i + \sum_{j=2}^4 \vartheta_j \text{Quartile}_j * \text{Chinese}_i + \sum_{j=2}^4 \eta_j \text{Quartile}_j * \text{Asian - Other}_i \\
 & + \sum_{j=2}^4 \delta_j \text{Quartile}_j * \text{Black - African}_i + \sum_{j=2}^4 \theta_j \text{Quartile}_j * \text{Black - Caribbean}_i \\
 & + \sum_{j=2}^4 \lambda_j \text{Quartile}_j * \text{Black - Other}_i + \sum_{j=2}^4 \psi_j \text{Quartile}_j * \text{Mixed}_i \\
 & + \sum_{j=2}^4 \omega_j \text{Quartile}_j * \text{Other - Ethnic}_i + \vartheta \text{pop - density}_i + \Psi \text{edu - status}_i + u_i
 \end{aligned} \tag{1}$$

where the dependent variable is the age adjusted COVID-19 death rate per 100,000 in the population in LAD *i*. The independent variables are as described above. The term *u* represents the error term.

To justify the linear specification and rule out a non-linear relationship, we performed a specification test where the predicted value and its square from the estimated OLS regression were used as explanatory variables in a separate regression with the same dependent variable. It turned out that the squared of the predicted value had a *p*-value of 0.416, implying that our OLS model was correctly specified. Therefore, we excluded the possibility of a non-linear relationship. To test for presence of multicollinearity, we ran a model without the interaction terms. The average variance inflation factor was 5.60 which is lower than the tolerance value of 10.

The model performance in terms of goodness of fit statistics (R^2) is 0.782. We also performed an information matrix test for the regression model and an orthogonal decomposition into tests for heteroskedasticity, skewness, and kurtosis. The *p* value for test of heteroskedasticity, skewness, and kurtosis was 0.474, 0.991 and 0.150 respectively justifying our specified model.

As a robustness check, given the dependent variables is censored between lower-limit (minimum) and upper-limit (maximum), we ran a censored regression model. The results obtained from the censored regression model were very similar to the one reported in the paper from the linear regression.

B: Spatial analysis. Figure 1 displays the age-adjusted mortality rates along with index of multiple deprivations across 315 LADs in England, comprising an area with 54,706,877 inhabitants and 173,673 inhabitants per LAD on average. The colour scheme depicted in a colour gradient showed lowest to highest values of age-adjusted mortality rates (in shades of blue) and index of multiple deprivation (in shades of green).

The COVID-19 age standardised mortality rate is clustered in the London metropolitan area and its surroundings, the Midlands, the Liverpool-Manchester area, the south of the Cumbria region, and in the North East (especially Durham and Newcastle), and lower down in the South, the South West, the Yorkshire and Humber region, the East, and in the East Midlands. This picture contrasts with the existing disparities in IMD across

Covid19 age standardised mortality rate, April 2020

Index of Multiple Deprivation-Average Score, 2019

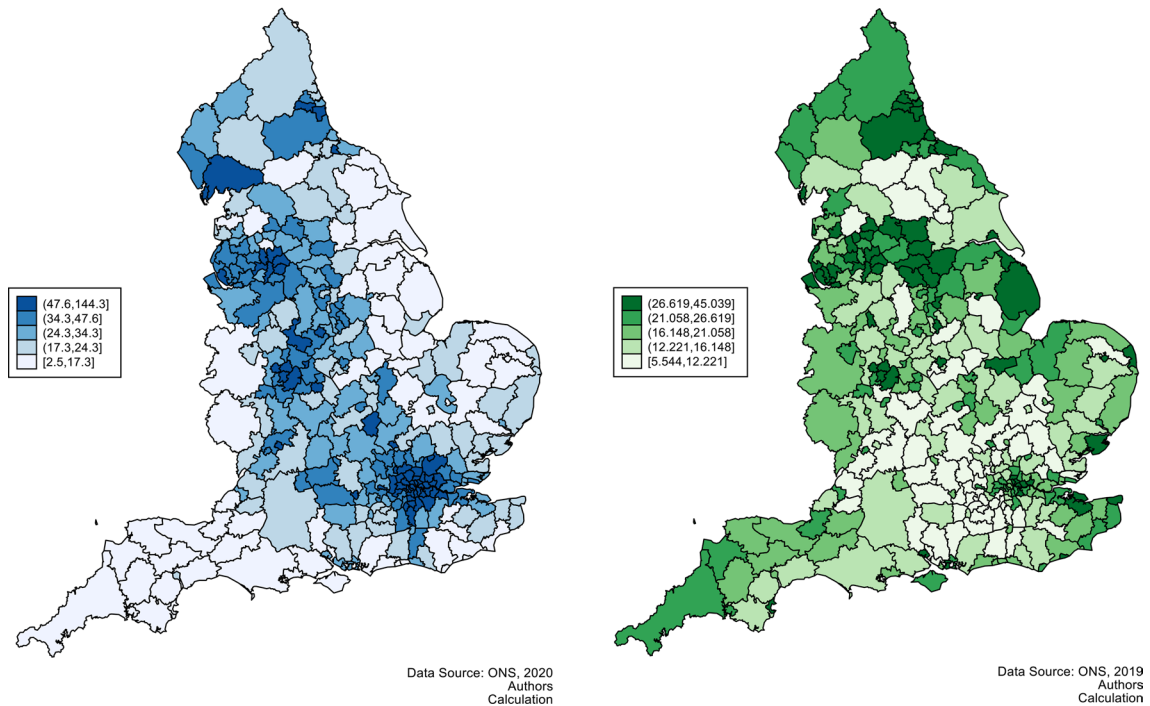


Figure 1. Age adjusted mortality rate and index of multiple deprivation.

England, where the South and the East present high levels of IMD, suggesting that the mortality rate may not be completely explained by deprivation at the local level. Other factors, like ethnic backgrounds and derived social habits, could also help to explain the mortality rate.

To investigate spillover effect from neighbouring LADs (keeping in mind the greater likelihood of social interaction, even during the lockdown, across neighbouring LADs), we undertook a spatial regression analysis using IMD and percentages of ethnic population as the explanatory variables. We allowed for spatial dependence in the age standardised mortality rate (especially given the North–South divide in England). In particular, we estimated the following equation:

$$\begin{aligned}
 \text{COVID}_i = & \alpha + \phi W\text{Covid}_i + \sum_{j=2}^4 \beta_j \text{QuartileDummy}_i + \beta_5 \text{Bangladeshi}_i + \beta_6 \text{Indian}_i \\
 & + \beta_7 \text{Pakistani}_i + \beta_8 \text{Arabi}_i + \beta_9 \text{Chinese}_i + \beta_{10} \text{Asian - Other}_i + \beta_{11} \text{Black - African}_i \\
 & + \beta_{12} \text{Black - Caribbean}_i + \beta_{13} \text{Black - Other}_i + \beta_{14} \text{Mixed}_i + \beta_{15} \text{Other - ethnic}_i \\
 & + \sum_{j=2}^4 \gamma_j \text{Quartile}_j * \text{Bangladeshi}_i + \sum_{j=2}^4 \varnothing_j \text{Quartile}_j * \text{Indian}_i + \sum_{j=2}^4 \mu_j \text{Quartile}_j * \text{Pakistani}_i \\
 & + \sum_{j=2}^4 \tau_j \text{Quartile}_j * \text{Arabi}_i + \sum_{j=2}^4 \vartheta_j \text{Quartile}_j * \text{Chinese}_i + \sum_{j=2}^4 \eta_j \text{Quartile}_j * \text{Asian - Other}_i \\
 & + \sum_{j=2}^4 \delta_j \text{Quartile}_j * \text{Black - African}_i + \sum_{j=2}^4 \theta_j \text{Quartile}_j * \text{Black - Carribean}_i \\
 & + \sum_{j=2}^4 \lambda_j \text{Quartile}_j * \text{Black - Other}_i + \sum_{j=2}^4 \psi_j \text{Quartile}_j * \text{Mixed}_i \\
 & + \sum_{j=2}^4 \omega_j \text{Quartile}_j * \text{Other - ethnic}_i + \vartheta \text{pop - density}_i + \Psi \text{edu - status}_i + e_i
 \end{aligned}
 \tag{2}$$

where W denotes a spatial contiguity matrix, and e the i.i.d disturbance term (Drukker et al.²³).

We also estimated a spatial regression model without the interaction terms. The estimated coefficient on the spatial lag of age-adjusted mortality is 0.407 (CI 0.292–0.522, $p < 0.001$), indicating positive correlation between the age-adjusted mortality in one LAD and the age-adjusted mortality in a neighbouring LAD. The total effect remained qualitatively the same as the model with interaction terms as reported in the paper.

Variable	Coefficient	(Confidence interval), <i>p</i> value
Quartile dummy 1	–	–
Quartile dummy 2	– 2.108	(– 16.570 to 12.353), <i>p</i> = 0.774
Quartile dummy 3	5.977	(– 7.203 to 19.156), <i>p</i> = 0.373
Quartile dummy 4	16.270	(2.355 to 30.185), <i>p</i> = 0.022
White	–	–
Black-African in 2nd quartile dummy	4.438	(– 0.096 to 8.972), <i>p</i> = 0.055
Black-African in 3rd quartile dummy	2.925	(1.102 to 4.748), <i>p</i> = 0.002
Black-African in 4th quartile dummy	2.861	(1.080 to 4.642), <i>p</i> = 0.002
Black-Caribbean in 2nd quartile dummy	12.166	(4.209 to 20.123), <i>p</i> = 0.003
Black-Caribbean in 3rd quartile dummy	7.276	(2.224 to 12.328), <i>p</i> = 0.005
Black-Caribbean in 4th quartile dummy	9.655	(5.248 to 14.061), <i>p</i> < 0.001
Black-Other in 2nd quartile Dummy	– 33.252	(– 49.193 to – 17.311), <i>p</i> < 0.001
Black-Other in 3rd quartile dummy	– 6.861	(– 19.842 to 6.121), <i>p</i> = 0.299
Black-Other in 4th quartile dummy	– 7.882	(– 20.924 to 5.160), <i>p</i> = 0.235
Bangladeshi in 2nd quartile dummy	– 16.183	(– 22.999 to – 9.366), <i>p</i> < 0.001
Bangladeshi in 3rd quartile dummy	– 3.645	(– 7.277 to – 0.013), <i>p</i> = 0.049
Bangladeshi in 4th quartile dummy	1.952	(1.144 to 2.760), <i>p</i> < 0.001
Indian in 2nd quartile dummy	5.122	(3.114 to 7.129), <i>p</i> < 0.001
Indian in 3rd quartile dummy	1.604	(0.037 to 3.171), <i>p</i> = 0.045
Indian in 4th quartile dummy	0.225	(– 0.990 to 1.441), <i>p</i> = 0.715
Pakistani in 2nd quartile dummy	1.225	(– 0.899 to 3.348), <i>p</i> = 0.257
Pakistani in 3rd quartile dummy	0.230	(– 1.190 to 1.650), <i>p</i> = 0.750
Pakistani in 4th quartile dummy	0.487	(– 0.366 to 1.339), <i>p</i> = 0.262
Chinese in 2nd quartile dummy	0.034	(– 8.790 to 8.858), <i>p</i> = 0.994
Chinese in 3rd quartile dummy	6.242	(– 6.134 to 18.619), <i>p</i> = 0.322
Chinese in 4th quartile dummy	– 4.708	(– 12.532 to 3.116), <i>p</i> = 0.237
Asian-Other in 2nd quartile dummy	– 1.922	(– 3.731 to – 0.112), <i>p</i> = 0.038
Asian-Other in 3rd quartile dummy	– 0.362	(– 3.461 to 2.738), <i>p</i> = 0.818
Asian-other in 4th quartile dummy	– 3.181	(– 11.553 to 5.192), <i>p</i> = 0.455
Arab in 2nd quartile dummy	– 37.071	(– 76.589 to 2.446), <i>p</i> = 0.066
Arab in 3rd quartile dummy	– 0.096	(– 3.206 to 3.014), <i>p</i> = 0.952
Arab in 4th quartile dummy	16.539	(– 0.659 to 33.736), <i>p</i> = 0.059
Mixed in 2nd quartile dummy	3.408	(– 5.701 to 12.517), <i>p</i> = 0.462
Mixed in 3rd quartile dummy	– 4.008	(– 9.237 to 1.220), <i>p</i> = 0.132
Mixed in 4th quartile dummy	– 6.836	(– 11.473 to – 2.199), <i>p</i> = 0.004
Other ethnic group in 2nd quartile dummy	23.201	(10.627 to 35.775), <i>p</i> < 0.001
Other ethnic group in 3rd quartile dummy	3.163	(– 3.582 to 9.908), <i>p</i> = 0.357
Other ethnic group in 4th quartile dummy	6.422	(2.266 to 10.579), <i>p</i> = 0.003
Proportion of people with A-level	– 30.625	(– 75.139 to 13.888), <i>p</i> = 0.177
Population density	0.275	(0.111 to 0.439), <i>p</i> = 0.001
No. of observations	315	

Table 2. Age adjusted mortality rates and association between ethnicity and index of multiple deprivation, linear regression model: coefficients (95% confidence interval, *P* value). White ethnicity is being used as the base category.

Results

Table 1 captures the variation in means across some key variables viz. age standardized COVID-19 mortality, the index of multiple deprivation (IMD) and the dis-aggregated ethnicity profiles. The overall average age-standardized mortality was 35.501 per 100,000 people with a standard deviation of 24.858. However, the average age-standardized mortality for the most deprived areas (Quartile 4) was 47.21 per 100,000 people compared to 27.727 per 100,000 people for the least deprived areas (Quartile 1). The percentage of ‘Black-African’, ‘Black-Caribbean’, ‘Black-Other’, ‘Bangladeshi’ and ‘Pakistani’ population increased across the quartiles of Index of multiple deprivation whereas for the ‘White’, the trend reverses. Percentage of ‘Indian’, ‘Chinese’ and ‘Asian-Other’ population increases for higher quartiles (except for Quartile 3).

Ethnicity	Median versus 25th percentile	75th percentile versus 25th percentile	75th percentile versus median
Deprivation Quartile 2			
Black-African	2.294 (− 0.255 to 4.842), $p=0.078$	8.158 (− 0.906 to 17.221), $p=0.078$	5.864 (− 0.651 to 12.379), $p=0.078$
Black-Caribbean	2.096 (− 0.120 to 4.312), $p=0.064$	8.111 (− 0.464 to 16.687), $p=0.064$	6.015 (− 0.344 to 12.374), $p=0.064$
Bangladeshi	− 1.995 (− 2.914 to − 1.075), $p<0.001$	− 8.184 (− 11.955 to − 4.413), $p<0.001$	− 6.189 (− 9.042 to − 3.337), $p<0.001$
Indian	1.819 (0.682 to 2.955), $p=0.002$	8.191 (3.072 to 13.311), $p=0.002$	6.373 (2.390 to 10.355), $p=0.002$
Pakistani	0.713 (0.098 to 1.329), $p=0.023$	4.443 (0.608 to 8.278), $p=0.023$	3.730 (0.510 to 6.949), $p=0.023$
Deprivation Quartile 3			
Black-African	2.419 (0.027 to 4.810), $p=0.047$	8.603 (0.097 to 17.109), $p=0.047$	6.184 (0.070 to 12.299), $p=0.047$
Black-Caribbean	3.092 (1.105 to 5.080), $p=0.002$	11.966 (4.274 to 19.658), $p=0.002$	8.874 (3.170 to 14.578), $p=0.002$
Bangladeshi	− 1.220 (− 1.962 to − 0.477), $p=0.001$	− 5.004 (− 8.049 to − 1.959), $p=0.001$	− 3.784 (− 6.087 to − 1.481), $p=0.001$
Indian	1.344 (0.130 to 2.557), $p=0.030$	6.052 (0.588 to 11.516), $p=0.030$	4.708 (0.457 to 8.959), $p=0.030$
Pakistani	0.045 (− 0.379 to 0.469), $p=0.836$	0.279 (− 2.363 to 2.922), $p=0.836$	0.235 (− 1.984 to 2.453), $p=0.836$
Deprivation Quartile 4			
Black-African	1.795 (− 0.597 to 4.187), $p=0.141$	6.385 (− 2.122 to 14.893), $p=0.141$	4.590 (− 1.526 to 10.706), $p=0.141$
Black-Caribbean	3.462 (1.577 to 5.348), $p<0.001$	13.398 (6.102 to 20.695), $p<0.001$	9.936 (4.525 to 15.347), $p<0.001$
Bangladeshi	− 0.557 (− 1.098 to − 0.017), $p=0.043$	− 2.287 (− 4.505 to − 0.069), $p=0.043$	− 1.730 (− 3.407 to − 0.052), $p=0.043$
Indian	0.629 (− 0.272 to 1.529), $p=0.171$	2.833 (− 1.223 to 6.889), $p=0.171$	2.204 (− 0.952 to 5.359), $p=0.171$
Pakistani	0.142 (− 0.214 to 0.498), $p=0.434$	0.884 (− 1.332 to 3.100), $p=0.434$	0.742 (− 1.118 to 2.603), $p=0.434$

Table 3. Marginal effects of ethnicity on age adjusted mortality, spatial regression model: coefficients (95% confidence interval, P value).

Regression results. The results are reported in Table 2. The coefficient associated with each ethnicity interacted with deprivation quartile in Table 2 reflects the aggregate effect of being in the particular deprivation quartile and belonging to the corresponding ethnicity.

It is clear that even after controlling for ethnicity, deprivation matters: most deprived LADs also exhibited significantly larger mortality compared to the affluent quartile (see ‘Quartile Dummy 4’ in Table 2). Interaction of economic deprivation with dis-aggregated ethnicity confirmed considerable heterogeneity within the BAME community. For example, a 1 percentage point increase in ‘Black-African’ population in the poorest LADs (Quartile 4) increased mortality rate by 2.861 per 100,000 population (CI 1.080–4.642, $p=0.002$) and the corresponding increase for the ‘Black-Caribbean’ population in the fourth Quartile is 9.655 (CI 5.248–14.061, $p<0.001$). The results were not as clear for the ‘Bangladeshi’ community, whereby the second and third Quartile effect was negative. Whereas a 1% point increase in ‘Bangladeshi’ community in the poorest LADs (Quartile 4) significantly increased COVID-19 mortality by 1.952 per 100,000 population (CI 1.144–2.760, $p<0.001$) compared to the ‘White’ community. Similar positive significant impact was also established for the ‘Indian’ in the second and in the third Quartiles. ‘Other ethnic group’ across all the three deprivation quartiles record a significantly higher age-adjusted mortality rate compared to the ‘White’ group. A unit increase in population density (see *pop-density* in Table 2) increased mortality by 0.275 per 100,000 population (CI 0.111–0.439, $p=0.001$).

Spatial regression results. The spatial regression results based on dis-aggregated ethnicity are reported in Table 3. The estimated coefficient on the spatial lag of age-adjusted mortality ($\hat{\phi}$) was 0.462 (CI 0.355–0.569, $p<0.001$), indicating positive correlation between the age-adjusted mortality in one LAD and the age-adjusted mortality in a neighbouring LAD.

Instead of regression coefficients, we focused on the obtained marginal effects for select ethnicity profile. The results are reported in Table 3 for five ethnicity groups: ‘Black-African’, ‘Black-Caribbean’, ‘Bangladeshi’, ‘Indian’ and ‘Pakistani’. Table 3 looks at the obtained results in a more disaggregated manner. Here we conducted comparison of three scenarios for each of the deprivation quartile: keeping the ethnicity at the 25th percentile, at the median and at the 75th percentile within each deprivation category. The results show that if the ‘Black-Caribbean’ population stood at the 75th percentile in comparison to the 25th percentile, the mortality rate would increase by 8.111 (CI − 0.464–16.687, $p=0.064$), 11.966 (CI 4.274–19.658, $p=0.002$), and 13.398 (CI 6.102–20.695, $p<0.001$) for the ‘Quartile dummy’ 2, 3 and 4 respectively. The ‘Black-Caribbean’ showed an increase of 2.096 (CI − 0.120–4.312, $p=0.064$), 3.092 (CI 1.105–5.080, $p=0.002$) and 3.462 (CI 1.577–5.348, $p<0.001$) when we compared the median in comparison to the 25th percentile. The other statistically significant increase was observed for the ‘Black-African’, and ‘Indian’ for the second and third Quartile of deprivation and for ‘Pakistani’ only for the second Quartile of deprivation. Contrary, to previous literature, ‘Bangladeshi’ had a significantly lower age-adjusted mortality with deterioration in deprivation status. For example, if the ‘Bangladeshi’ population stood at the 75th percentile in comparison to the 25th percentile, the mortality rate would decrease by 8.184 (CI − 11.955 to − 4.413, $p<0.001$), 5.004 (CI − 8.049 to − 1.959, $p<0.001$) and 2.287 (CI − 4.505 to − 0.069, $p=0.043$) for the ‘Quartile dummy’ 2, 3 and 4 respectively.

Discussion

We found that COVID-19 mortality disproportionately affects the local areas with an over-representation of individuals who are relatively socio-economically deprived and also belong to ethnic minority particularly the 'Black-Caribbean' and 'Black-African'. Linear regression estimates showed that considering 'South-Asian' community as one homogeneous entity can be misleading because there was considerable discrepancy among the ethnic subgroups ('Bangladeshi', 'Indian' and 'Pakistani') across the different deprivation quartiles. Impact of living in a particular LAD (neighbourhood effect) is also confirmed by positive significant impact of population density within each LAD on the age adjusted COVID-19 mortality. The spatial regression results suggest a strong spillover effect in the disease burden. The spatial regression estimates also corroborates the higher risks of particularly the 'Black-Caribbean'—for each deprivation quartile, age adjusted mortality rate became significantly higher if the percentage of 'Black Caribbean' increased in that LAD. A similar pattern was also seen for the 'Black-African', 'Indian' in second and third Quartile, and for 'Pakistani' in the second Quartile. Amongst the 'South-Asian' population, the 'Bangladeshi' behaves differently.

Pareek et al.⁷ have rightly called for an urgent public health approach to understand the differential effect of ethnicity and understand the interplay of ethnicity with several factors. Apart from the inherent genetic disposition, our analysis sheds light on deprivation e.g. in inequality in access to resources for the higher prevalence of mortality in certain segments of the population. Our analysis findings largely correlate with existing evidence suggesting that economic deprivation with dis-aggregated ethnicity exhibits considerable heterogeneity within the BAME community. Consistent with ONS⁴ and Apea et al.²⁴ individuals from 'Black-African' and 'Black-Caribbean' community and who are also economically more deprived record a significantly higher age-adjusted COVID-19 mortality compared to the 'White'.

However, our results are particularly interesting when we disaggregate by ethnic group. Our results are in contrast to previous literature¹⁰ whereby no consistent pattern emerges when we compare the combined effect of ethnicity and economic deprivation across the different 'South Asian' community profile. Impact for the 'Bangladeshi' is negative and significant contrary to previous findings¹⁰. While we cannot fully explain this, many of these areas may have other compensating factors such as higher social capital and other LAD specific factors that we cannot observe in the data. An empirical analysis of how such differences in behavioural norms and choices across ethnicities have impacted their risk of contracting the deadly virus is left for future investigation. Despite some similarities with pre-existing literature, our results suggest a differential public health approach would be warranted to capture the nuances within the different ethnic minority groups.

The most important limitation with this type of study is the possibility of ecological fallacy, whereby the effect seen at a LAD level may not be generalisable to individuals in the region. Another important limitation is that in this study we were unable to adjust for important co-variables such as occupation and comorbidity status all of which may be risk factors for COVID-19⁶ although we were able to account for population density and educational attainment. It is important to highlight that ethnic groups do have differing occupational profiles which may have an impact on their COVID-19 risk. Of the total employed Black population predominant share belonged to 'Caring Personal Service Occupations' (15.4%), 'Elementary Administration and Service Occupations' (14.3%) and in 'Health Professionals' (9%). In comparison for the employed White, there was a more uniform distribution across different occupations (for example while 6.9% were working in 'Caring Personal Service Occupations' a very large proportion (8.0%) were also 'Corporate Managers and Directors')²⁵. The largest share of employed Asian is in 'Elementary Administration and Service Occupations' (50.5%). If we look at NHS per se, the latest data reveals that "Asian people made up a higher percentage of medical staff (at 29.7%) than non-medical staff (at 8.0%)"²⁶. These occupational roles may put BAME communities at a greater risk of COVID-19 and may make it more difficult to practice social distancing²⁷. Additionally, it is clear that BAME communities have a higher level of general comorbidity which may pose as an independent risk for COVID-19²⁸.

In conclusion, our results clearly show that deprivation and its interaction with ethnicity play an important role in explaining COVID-19 mortality. The presence of spatial effects and spillover suggest family structures and social networks play an important role too. Social interactions between people across neighbouring regions can also spread the disease.

Data availability

The data used in this study was open source provided by the Local Authority. Any requests for the extracts used during this study period can be made to the corresponding author.

Received: 2 September 2020; Accepted: 18 May 2021

Published online: 02 June 2021

References

1. World Health Organization. Statement on the second meeting of the International Health Regulations. Emergency Committee regarding the outbreak of novel coronavirus (2019-nCoV). [https://www.who.int/news-room/detail/30-01-2020-statement-on-the-second-meeting-of-the-international-health-regulations-\(2005\)-emergency-committee-regarding-the-outbreak-of-novel-coronavirus-\(2019-ncov\)](https://www.who.int/news-room/detail/30-01-2020-statement-on-the-second-meeting-of-the-international-health-regulations-(2005)-emergency-committee-regarding-the-outbreak-of-novel-coronavirus-(2019-ncov)). Accessed 8 June 2020.
2. John Hopkins University. COVID-19 Map—Johns Hopkins Coronavirus Resource Center. 2020. <https://coronavirus.jhu.edu/map.html>. Accessed 8 June 2020.
3. Embury-Dennis T. Coronavirus: A timeline of how Britain went from 'low risk' to an unprecedented national shutdown. *Indep.* 2020. <https://www.independent.co.uk/news/uk/home-news/coronavirus-uk-timeline-deaths-cases-covid-19-nhs-social-distancing-a9416331.html>. Accessed 22 Aug 2020.
4. Office for National Statistics. Coronavirus (COVID-19) related deaths by ethnic group, England and Wales—Office for National Statistics. *Off Natl Stat* 1–10 (2020).

5. Public Health England. Disparities in the risk and outcomes of COVID-19. 2020. www.facebook.com/PublicHealthEngland. Accessed 18 June 2020.
6. Williamson, E. J., Walker, A. J., Bhaskaran, K., Bacon, S., Bates, C., Morton, C. E. *et al.* OpenSAFELY: factors associated with COVID-19 death in 17 million patients. *Nature* 1–11 (2020).
7. Pareek, M. *et al.* Ethnicity and COVID-19: an urgent public health research priority. *Lancet* **395**, 1421–1422 (2020).
8. Pan, D. *et al.* The impact of ethnicity on clinical outcomes in COVID-19: A systematic review. *EClinicalMedicine* **23**, 100404 (2020).
9. Office for National Statistics. People living in deprived neighbourhoods—GOV.UK Ethnicity facts and figures. <https://www.ethnicity-facts-figures.service.gov.uk/uk-population-by-ethnicity/demographics/people-living-in-deprived-neighbourhoods/latest>. Accessed 22 Aug 2020.
10. ISARIC. Ethnicity and Outcomes from COVID-19: The ISARIC CCP-UK Prospective Observational Cohort Study of Hospitalised Patients. ISARIC. <https://isaric.tghn.org/articles/ethnicity-and-outcomes-covid-19-isaric-ccp-uk-prospective-observational-cohort-study-hospitalised-patients/>. Accessed 22 Aug 2020.
11. Alaaahmed, M., Qian, Z., Rashbass, J., Pinto, K. G., Bengler, J., van der Schaar, M. Ethnicity and outcomes of coronavirus (COVID-19) patients in England. <https://digital.nhs.uk/coronavirus/coronavirus-data-services-updates/ethnicity-and-outcomes-of-covid-19-patients-in-england>. Accessed 22 Aug 2020.
12. Shouls, S., Congdon, P. & Curtis, S. Modelling inequality in reported long term illness in the UK: Combining individual and area characteristics. *J. Epidemiol. Community Health* **50**, 366–376 (1996).
13. Newton, J. N. *et al.* Changes in health in England, with analysis by English regions and areas of deprivation, 1990–2013: A systematic analysis for the Global Burden of Disease Study 2013. *Lancet* **386**, 2257–2274 (2015).
14. Wilson, W. J. *The truly disadvantaged: the inner city, the underclass, and public policy* (Chicago University Press, 1989).
15. Flowerdew, R., Manley, D. J. & Sabel, C. E. Neighbourhood effects on health: Does it matter where you draw the boundaries?. *Soc. Sci. Med.* **66**, 1241–1255 (2008).
16. Krisztin T, Piribauer P, Wögerer M. The spatial econometrics of the coronavirus pandemic. *Lett. Spat. Resour. Sci.* **1** (2020).
17. Almagro, M. & Orane, H. A. The differential impact of COVID-19 across demographic groups: Evidence from NYC. *SSRN Electron. J.* <https://doi.org/10.2139/ssrn.3573619> (2020).
18. Kulu H, Dorey P. Infection rates from Covid-19 in Great Britain by geographical units: a model-based estimation from mortality data. *Health Place.* <https://doi.org/10.31235/osf.io/84f3e> (2020).
19. Office for National Statistics. Regional ethnic diversity—GOV.UK Ethnicity facts and figures. GOV.UK **1** (2019).
20. Office for National Statistics. Deaths involving COVID-19 by local area and socioeconomic deprivation. *Off Natl Stat* 1–23 (2020).
21. Office for National Statistics. English Indices of Deprivation 2019. <https://www.gov.uk/government/statistics/english-indices-of-deprivation-2019>. Accessed 22 Aug 2020.
22. Office for National Statistics. 2011 Census. local authorities in the United Kingdom. <https://www.ons.gov.uk/census/2011census>. Accessed 22 Aug 2020.
23. Drukker, D. M. *et al.* Maximum likelihood and generalized spatial two-stage least-squares estimators for a spatial-autoregressive model with spatial-autoregressive disturbances. *Stata J* **13**, 221–241 (2013).
24. Apea, V. J., Wan, Y. I., Dhairyawan, R., Puthuchery, Z. A., Pearse, R. M., Orkin, C. M. *et al.* Ethnicity and outcomes in patients hospitalised with COVID-19 infection in East London: an observational cohort study. *medRxiv* 2020.
25. Office for National Statistics. Annual Population Survey—Nomis—Official Labour Market Statistics. <https://www.nomisweb.co.uk/articles/1167.aspx>. Accessed 22 Aug 2020.
26. NHS Digital. NHS workforce ethnicity facts and figures. Gov.Uk. <https://www.ethnicity-facts-figures.service.gov.uk/workforce-and-business/workforce-diversity/nhs-workforce/latest#title>. Accessed 22 Aug 2020.
27. Bavel, J. J. V. *et al.* Using social and behavioural science to support COVID-19 pandemic response. *Nat. Hum. Behav.* **4**, 460–471 (2020).
28. Khunti, K., Singh, A. K., Pareek, M. & Hanif, W. Is ethnicity linked to incidence or outcomes of Covid-19?. *BMJ* **369**, m1548 (2020).

Acknowledgements

We thank Juste Abramovaite and Eddie Kane for comments on a draft version of this paper and Karen Nokes for research assistance.

Author contributions

K.C. and J.M.L. led the quantitative analysis, A.C. and S.B. wrote the first draft and J.S.C. provide expertise on the public health implications of the research. All authors reviewed and contributed to the final draft.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-91076-8>.

Correspondence and requests for materials should be addressed to S.B.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021