medRxiv preprint doi: https://doi.org/10.1101/2020.11.27.20238147; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Ethnicity, Household Composition and COVID-19 Mortality: A National Linked Data Study

Vahé Nafilyan[a,b], Nazrul Islam[c], Daniel Ayoubkhani[a], Clare Gilles[d], Srinivasa Vittal Katikireddi[e], Rohini Mathur[b], Karen Tingay[a], Miqdad Asaria[f], Ann John[g], Peter Goldblatt[h], Amitava Banerjee[i], Kamlesh Kunti[d]

[a] Office for National Statistics, Newport, UK
[b] London School of Hygiene and Tropical Medicine, London, UK
[c] Nuffield Department of Population Health, Big Data Institute, University of Oxford, Oxford, UK
[d] Diabetes Research Centre, University of Leicester, Leicester, UK
[e] University of Glasgow, Glasgow, UK
[f] London School of Economics and Political Sciences, London, UK
[g] Swansea University, Swansea, UK
[h] UCL Institute of Health Equity, University College London, London, UK
[i] Institute of Health Informatics, University College London, London, UK

Abstract

Background: Ethnic minorities have experienced disproportionate COVID-19 mortality rates. We estimated associations between household composition and COVID-19 mortality in older adults (≥65 years) using a newly linked census-based dataset, and investigated whether living in a multi-generational household explained some of the elevated COVID-19 mortality amongst ethnic minority groups.

Methods: Using retrospective data from the 2011 Census linked to Hospital Episode Statistics (2017-2019) and death registration data (up to 27th July 2020), we followed adults aged 65 years or over living in private households in England from 2 March 2020 until 27 July 2020 (n=10,078,568). We estimated hazard ratios (HRs) for COVID-19 death for people living in a multi-generational household compared with people living with another older adult, adjusting for geographical factors, socio-economic characteristics and pre-pandemic health. We conducted a causal mediation analysis to estimate the proportion of ethnic inequalities explained by living in a multi-generational household.

Results: Living in a multi-generational household was associated with an increased risk of COVID-19 death. After adjusting for confounding factors, the HRs for living in a multi-generational household with dependent children were 1.13 [95% confidence interval 1.01-1.27] and 1.17 [1.01-1.35] for older males and females. The HRs for living in a multi-generational household without dependent children were 1.03 [0.97 - 1.09] for older males and 1.22 [1.12 - 1.32] for older females. Living in a multi-generational household explained between 10% and 15% of the elevated risk of COVID-19 death among older females from South Asian background, but very little for South Asian males or people in other ethnic minority groups.

Conclusion: Older adults living with younger people are at increased risk of COVID-19 mortality, and this is a notable contributing factor to the excess risk experienced by older South Asian females compared to White females. Relevant public health interventions should be directed at communities where such multi-generational households are highly prevalent.

Funding: This research was funded by the Office for National Statistics.

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.27.20238147; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Introduction

People of ethnic minority background in the UK and the USA have been disproportionately affected by the Coronavirus disease 2019 (COVID-19) [1, 2, 3, 4, 5] compared to the White population, particularly Black and South Asian groups. Whilst several studies have investigated whether adjusting for socio-demographic and economic factors and medical history reduces the estimated difference in risk of mortality and hospitalisation [6, 7, 8] , the reasons for the differences in the risk of experiencing harms from COVID-19 are still being explored.

One important driver of these ethnic inequalities may be differences in household structure between ethnic groups. Household composition varies substantially between ethnic groups, with some ethnic minority populations more likely to live in large, multi-generational households [9]. While living in multi-generational households is associated with increased social capital [10], which could have beneficial health effects [11], it may also increase the risk of potential viral transmission [12, 13]. For older people, who are at greater risk of experiencing severe complications if infected, residing with younger people may represent an increase in exposure to infection, which could lead to an increased risk of hospitalisation and mortality from COVID-19. To the best of our knowledge, no study has yet examined whether the difference in household composition partly explains the elevated risk of COVID-19 mortality in ethnic minority groups.

In this study, we examined the relationship between household composition and COVID-19 mortality risk amongst older adults (≥65 years) in England, with a focus on multi-generational households (older adults living with younger adults or dependent children). We then investigated how the propensity to live in a multi-generation household varies across ethnic groups, and whether this heterogeneity contributes to the raised risk of COVID-19 mortality amongst ethnic minority groups compared to the White population.

Methods

Data

This retrospective cohort study was based on the 2011 Census of England linked to mortality registration data and Hospital Episodes Statistics (2017 – 2019). The study population included all usual residents of England aged 65 years or over in 2020, who had been enumerated in private households at the time of the 2011 Census (27 March 2011), had not moved to a care home by 2019 (identified by linking to the NHS Patient Register) and were still alive on 2 March 2020. We further excluded individuals who entered the UK in the year before Census due to their higher propensity to leave the UK prior to the study period, and those aged over 100 years at the time of the Census. Our study population consisted of 10,078,568 individuals aged 65 years or over at 2 March 2020 (See Supplementary Table 1 details on the number of participants at each stage of the sample selection).

To adjust for out-migration, we applied weights reflecting the probability of having remained in the country until March 2020 after being enumerated in March 2011, based on data from the NHS Patient Register and the International Passenger Survey (IPS). Further information on the data has already been published [6]. All the variables used in the analysis, including their definitions and sources, are detailed in Table 1.

Outcome and exposure

Deaths involving COVID-19 included those with an underlying cause, or any mention, of International Statistical Classification of Diseases and Related Health Problems 10th Revision (ICD-10) codes U07.1

medRxiv preprint doi: https://doi.org/10.1101/2020.11.27.20238147; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

(COVID-19, virus identified) or U07.2 (COVID-19, virus not identified). We analysed deaths that occurred between 2 March 2020 and 28 July 2020, registered by 24 August 2020, which corresponds to the deaths that occurred during the first COVID-19 wave.

Household composition in 2020 was derived based the household composition at the time of the Census. We excluded people who died between 27 March 2011 and 1 March 2020 or had moved to a care home by 2019. To mitigate measurement error, we removed people aged 10 to 24 at the time of the Census because they are more likely to have moved out in 2020. We defined a multi-generational household to be a household in which someone aged 65 years or over on 2 March in 2020 co-resided with at least one other adult aged more than 20 years younger, or with at least one child. Our household composition variable classified households in five categories: Single; Two older adults; Multi-generational household without dependent children; Multi-generational household with dependent children; three or more older adults. As sensitivity analyses, we removed people aged 10 to 19 instead of 10 to 24).We also defined a multi-generational household to be one in which someone aged 65 years or over in 2020 co-resided with at least one other adult aged more than 15 years (instead of 20 years) younger.

In the mediation analysis, the exposure was self-reported ethnic affiliation based on a nine-group classification (see Table 1). The two mediators were binary variables for living in a multi-generational household with or without children.

Covariates

Demographic factors, geographical variables, socio-economic characteristics and measures of pre-pandemic health are listed in Table 1. These covariates were generally considered to be confounders of the relationship between household composition and COVID-19 mortality risk, and mediators of the ethnicity-mortality relationship (See Figure 1).

Table 1. Variables used for the statistical modelling

Variable	Coding	Source
Demographic factors		
Single year of age	Second-order polynomial	2011 Census
Ethnicity	Black Caribbean, Black African, Bangladeshi, Pakistani, Indian, Chinese, White, Mixed, Other	2011 Census
Household composition	Single; Two older adults; Multi-generational household without dependent children; Multi-generational household with dependent children; three or more older adults.	2011 Census
Geographical variables		
Government office region	Dummy variables representing local authority districts	NHS Patient register 2019
Urban-rural classification	Urban major conurbation, Urban minor conurbation, Urban city and town, Urban city & town in a sparse, Setting, Rural town and fringe, Rural village, Rural hamlet & isolated dwellings, Rural town & fringe in a sparse, setting, Rural	2011 Census

medRxiv preprint doi: https://doi.org/10.1101/2020.11.27.20238147; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

	village in a sparse setting Rural hamlet & isolated dwellings, in a sparse setting	
Population density of Lower Super Output Area (LSOA)	Second-order polynomial, allowing for a different slope beyond the 99th percentile of the distribution to account for extreme values	
Socio-economic variables		
Index of Multiple Deprivation (IMD), Welsh Index of Multiple Deprivation (WIMD) [14, 15]	Dummy variables representing deciles of deprivation	Department for Communities and Local Government/ Welsh Government
Household deprivation (see table note)	Not deprived, deprived in one dimension, deprived in two dimensions, deprived in three dimensions, deprived in four dimensions	2011 Census
Household tenure	Own outright, own with mortgage, social rented, private rented, other	2011 Census
Approximated social grade	Higher & intermediate managerial, administrative, professional occupations; Supervisory, clerical & junior managerial, administrative, professional occupations; Skilled manual occupations; Semi-skilled & unskilled manual occupations; Unemployed and lowest grade occupations	2011 Census
Level of highest qualification	Degree, A-level or equivalent, GCSE or equivalent, no qualification	2011 Census
Overcrowding	Overcrowded household (occupancy rating < -1)	2011 Census
Type of accommodation	Detached house, semi-detached house, terraced house, flat, other	2011 Census
Health-related variables		
Self-reported health status	Very good, good, fair, poor, very poor	2011 Census
Self-reported limiting long term heath problem or disability	Not limited, daily activity limited a lot, daily activity limited a little	2011 Census
Pre-existing conditions	Dummies for cancer, cardiovascular diseases, digestive disorders, mental health conditions, metabolic disorders, musculoskeletal disorders, neurological disorders, renal disorders, respiratory disorders	HES (2017-2019)
Number of hospital admissions	0, 1, 2-3, 4-5, 6-9, 10+	HES (2017-2019)
Total days spent in hospital	0, 1, 2-4, 5-9, 10-19, 20-39, 40-69, 70+	HES (2017-2019)

Note: Household deprivation is defined according to four dimensions: employment (at least one household member is unemployed or long-term sick, excluding full-time students); education (no household members have at least Level 2 education, and no one aged 16-18 years is a full-time student); health and disability (at least one household member reported their health as being 'bad'/'very bad' or has a long-term health problem); and housing (the household's accommodation is overcrowded, with an occupancy rating -1 or less, or is in a shared dwelling, or has no central heating).

medRxiv preprint doi: https://doi.org/10.1101/2020.11.27.20238147; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

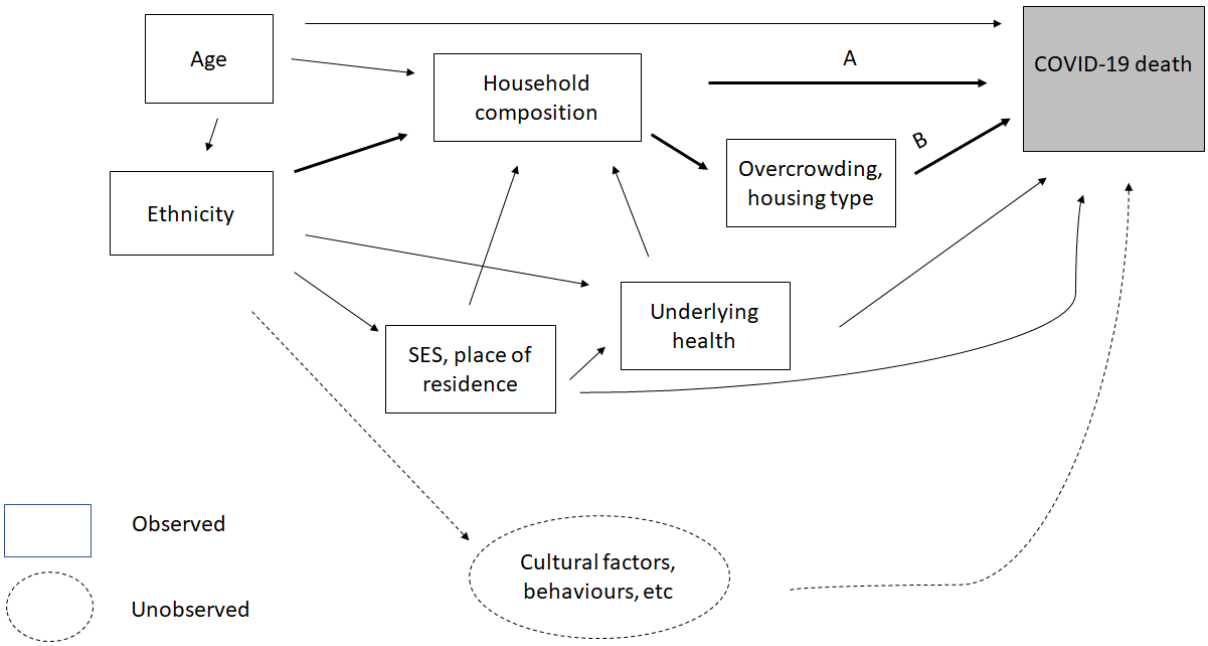
For people aged over 75 years at time of the 2011 Census, approximated social grade was imputed based on household tenure.

Statistical analyses

We calculated age-standardised mortality rates (ASMRs) stratified by household composition and sex, separately for COVID-19 related deaths and other deaths. The ASMRs were standardised to the 2013 European Standard Population and can be interpreted as deaths per 100,000 of the population at risk during the analysis period.

We estimated Cox proportional hazard models to assess whether the risk of COVID-19-related death varies by household type (using living with one other older adult with as the reference category) after adjusting for the geographical factors, socio-economic characteristics and measures of health listed in Table 1. These factors could confound the relationship between household composition and COVID-19-related mortality, as shown in Figure 1. We estimated separate models for males and females, as the risk of death involving COVID19 differs markedly by sex [7]. When fitting the Cox models, we included all individuals who died during the analysis period and a weighted random sample of those who did not (5% of White people, and 20% amongst ethnic minority groups), and applied case weights to reflect the original population totals.

Figure 1 Directed Acyclic Graphs summarising the relationship between ethnicity, household composition and COVID-19 mortality



Note: When analysing whether household composition directly affects the risk of COVID-19 death, our effect of interest is *A*. In the mediation analysis, where we estimate the proportion of the ethnic disparity in COVID that is explained by living in a multi-generational household, the effects of interest are *A+B*

We conducted a causal mediation analysis [16] to estimate the proportion of excess risk in ethnic minority groups which is attributable to living in a multigenerational household . As a measure of inequality in COVID-19 mortality, ethnicity-specific odds ratios for COVID-19 mortality were estimated using logistic regression models, fitted to males and females separately and adjusting solely for age in the baseline model. The proportion of the difference in COVID-19 mortality rates between ethnic groups mediated by living in a multi-generational household was then estimated as

medRxiv preprint doi: https://doi.org/10.1101/2020.11.27.20238147; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

the Average Causal Mediated Effect (ACME) as a proportion of the age-adjusted difference in the probability of COVID-19 mortality, using a non-parametric approach [17] (see the Technical Appendix for more details). The mediator models and the full outcome model were adjusted for geographical factors (region, population density, urban/rural classification), socio-economic characteristics (IMD decile, educational attainment, social grade, household tenancy) and health (self-reported health and disability from the Census, pre-existing conditions based on hospital contacts) but not for overcrowding or housing type, as these are likely to be consequences of living in a multi-generational household rather than confounding factors (See Figure 1). Confidence intervals were obtained via bootstrapping, using 500 replications. All statistical analyses were performed using R version 3.5.

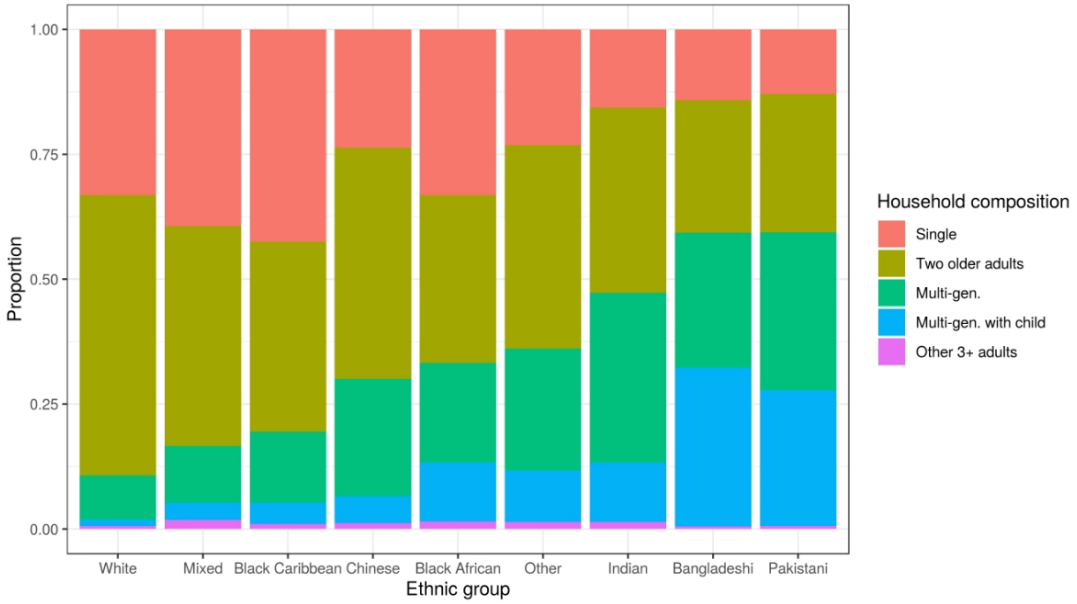
Results

Characteristics of the study population

Characteristics of the study population are reported in Supplementary Table 2 in the Appendix. In our study population of 10,078,568 individuals in England aged 65 years or over who were not in a care home in 2019 and were still alive on 2 March 2020, just over half (53.9%) were female, the mean age was 75.2 years, and 93.9% reported being from a White ethnic background (Table 2). Over the outcome period (2 March 2020 to 28 July 2020), 27,989 (0.28%) died of COVID-19, and 123,551 (1.2%) died of other causes.

Compared with older adults living with one other older adult (*n* = 5,538,963), people living by themselves (*n* = 3,287,395) had a higher mean age, were more likely to be female, and tended to be more deprived. Older people living in a multi-generational household without dependent children (*n* = 987,306) and with dependent children (n = 199,112) were on average younger and were more likely to be from an ethnic minority group, live in London and large urban conurbations, and tended to be more deprived than older people living with another older adult.

Figure 2 Household composition by ethnic group for people in England aged ≥65 years



medRxiv preprint doi: https://doi.org/10.1101/2020.11.27.20238147; this version posted December 2, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Note: Linked 2011 Census and mortality registration data for people in England aged ≥65 years, excluding those living in a care home in 2019. The number of adults in the household was calculated as the number of people aged ≥25 years who lived in the household at the time of the Census, minus those who died between 27 March 2011 and 1 March 2020.

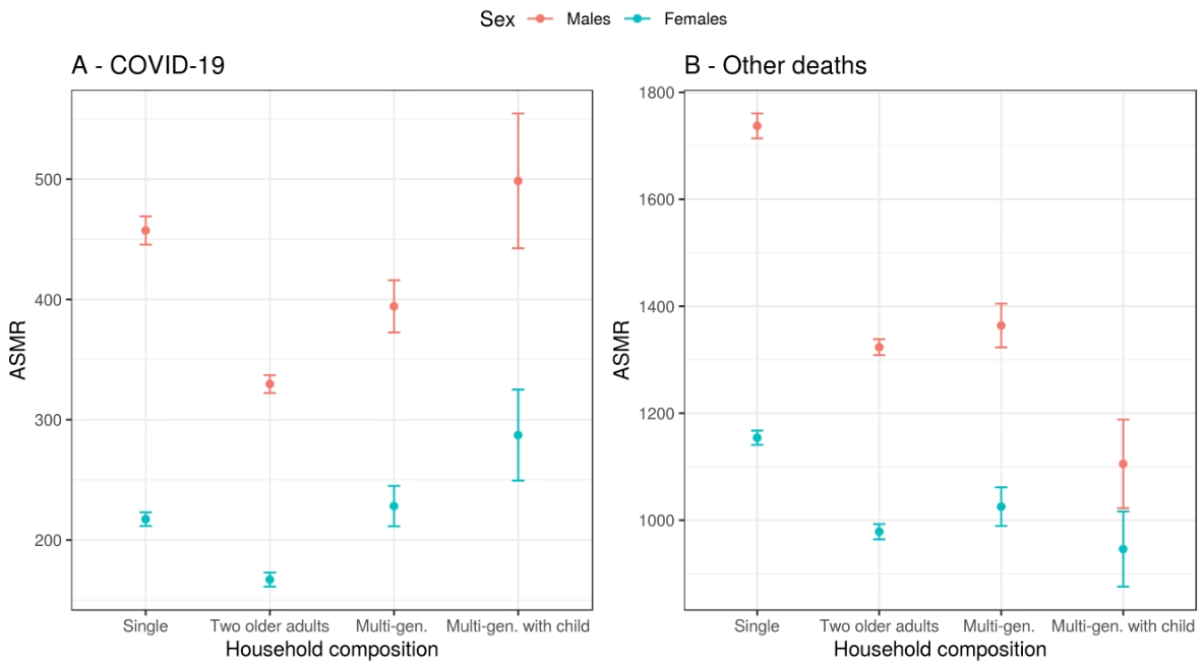
Figure 2 shows that household composition varied substantially between ethnic groups. Among older people, just over 10% of those of White background lived in a multi-generational household, compared to over half of Bangladeshi or Pakistani background (58.7% and 58.8% respectively) and 45.8% of Indian background. The patterns were similar for males and females, although a larger proportion of females live by themselves (See Supplementary Figure 1).

Household composition and death involving COVID-19

Older people living by themselves were more likely to have died from COVID-19 over the study period than those living with another adult. For males, the ASMRs were 457 [95% confidence interval (CI): 446-469] and 330 [322-337] per 100,000 of the population for those living by themselves and those living with another adult, respectively (Figure 2, Panel A). For females, the ASMRs were 217 [212-223] and 167 [161-173] per 100,000, respectively. A similar pattern is observed for deaths from other causes (Figure 2, Panel B).

There was a positive association between the risk of COVID-19 death and living a in multi-generational household. Both older males and females living a multi-generational household without school-age children were more likely to die from COVID -19 than older people living with another older adult (ASMR 394 [373-416] per 100,000 for males, 228 [211-245] for females), with the risk of death being greater still if there were children in the household (ASMR 499 [443-555] per 100,000 for males, 287 [249-325] per 100,000 for females). The risk of COVID-19 mortality was higher in males than that in females across all the household compositions. There was no clear relationship between living in a multi-generational household and mortality from other causes.

Figure 3 Age-standardised mortality rates per 100,000 adults aged 65 years or over, stratified by sex and household composition



medRxiv preprint doi: https://doi.org/10.1101/2020.11.27.20238147; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Note: Deaths occurring between 2 March 2020 and 28 July 2020. 95% Confidence intervals are reported. Mortality rates are standardised to the 2013 European Standard Population.

Adjusting for individual- and household-level characteristics (including age, geographical factors, socio-economic characteristics and measures of pre-pandemic health) reduced the estimated differences in COVID-19 mortality rates between older adults living in different types households (Table 2). However, even after adjusting for these characteristics, living in a multi-generational household, especially with children, remained associated with an increased risk of COVID-19-related death. Compared to living with another older adult aged 65 years or above, the rate of COVID-19-related death was 1.22 [1.12 - 1.32] and 1.17 [1.01 - 1.35] times greater for older females living in a multi-generational household without and with children, respectively. For older males, after adjusting for individual and household characteristics, there was no evidence of an association between living in a multi-generational household without children and the risk of COVID-19 mortality (HR: 1.03 [0.97-1.09]). However, living in a multi-generational household with children was associated with a 1.13 [1.01-1.27] times greater risk of COVID-19 related death. The rate of COVID-19 related death was also 1.19 [1.15-1.23] times greater for older males living alone than for those living with another older adult, and 1.13 [1.08 -1.18] times greater for older females. The results were similar in the sensitivity analyses using different definitions of household composition (See Supplementary Table 3).

Table 2 . Hazard ratios for COVID-19 related death for older adults (aged ≥65 years) in England, compared to living in a household with one other older adult, stratified by sex

	Men		Women	
	Age-adjusted	Fully-adjusted	Age-adjusted	Fully-adjusted
Single	1.405 [1.358 - 1.453]	1.186 [1.145 - 1.228]	1.288 [1.234 - 1.344]	1.115 [1.066 - 1.165]
Multi-generational household without children	1.247 [1.181 - 1.318]	1.027 [0.971 - 1.087]	1.425 [1.319 - 1.539]	1.215 [1.122 - 1.317]
Multi-generational household with children	1.629 [1.464-1.812]	1.128 [1.007 - 1.265]	1.852 [1.622 - 2.114]	1.169 [1.010 - 1.352]
3+ older adults	0.852 [0.695 - 1.046]	0.900 [0.733 - 1.106]	0.724 [0.556 - 0.944]	0.824 [0.632 - 1.075]
Observations (unweighted)	432,894		489,192	
Concordance	0.7432	0.8284	0.7588	0.8551

Note: Hazard ratios compared to living in a household with one other older adult. Fully adjusted Cox regression models include geographical factors (region, population density, urban/rural classification), ethnicity, socio-economic characteristics (IMD decile, household deprivation, educational attainment, social grade, household tenancy), health (self-reported health and disability from the Census, pre-existing conditions based on hospital contacts, number of hospital admissions, total days spent in hospital), a measure for overcrowding, and property type.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.27.20238147; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Living in a multi-generational household as a mediator for the disparity in COVID-19 death between ethnic groups

Among people aged 65 years or over, those from all ethnic minority groups except Chinese were at greater risk of COVID-19 related death than those from the White population. The odds of COVID-19 death were 3.1 [2.7-3.6] times greater for older Black African males and 2.2 [1.8-2.7] times greater for Black African females than for males and females from White ethnic group. The odds of death were also notably greater for people of Bangladeshi, Black Caribbean or Pakistani ethnic background than the White population, with odds ratios of 2.8 [2.3-3.4], 2.6 [2.4-2.9], and 2.2 [1.9-2.5], respectively, for males and 2.5 [1.9-3.3], 2.0 [1.7-2.3] and 2.2 [1.8-2.7], respectively, for females.

Living in a multi-generational household did not explain much of the difference in COVID-19 mortality rates amongst older males. However, it did explain a substantial part of the difference between older females of South Asian background and White older females. For older females of Pakistani background, living in a multi-generational household accounted for 13.8% [3.4%-23.9%] of the difference in COVID-19 mortality rates compared to White females; for females of Bangladeshi or Indian background, it accounted for 12.3% [7.3%-19.5%] and 12.2% [5.7% - 23.0%] of the difference in mortality rates, respectively (See Supplementary Table 4 in appendix for full results). The results were similar in the sensitivity analyses using different definitions of household composition (See Supplementary Table 5).

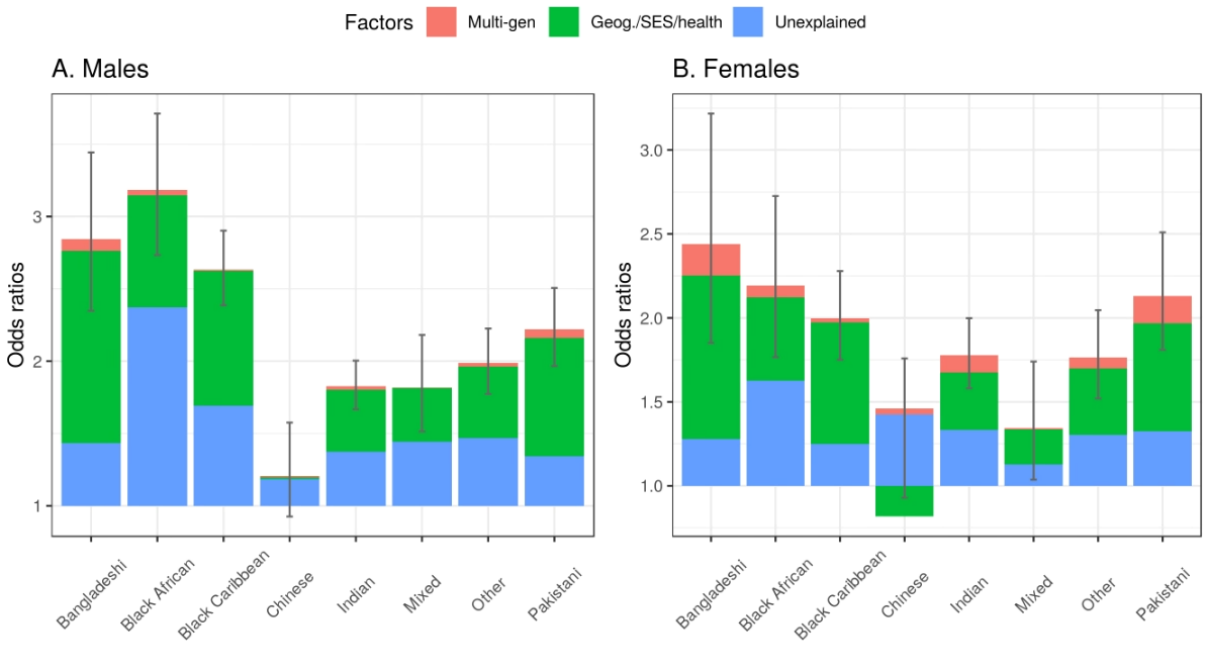
Figure 4 shows the decomposition of the age-adjusted odds ratios of COVID-19 death for ethnic minority groups compared to those of white ethnic group into three parts: (i) the part explained by living in a multi-generational household (in red); (ii) the part explained by other individual and household characteristics, such as geographical factors, socio-economic factors and pre-pandemic health (in green); and (iii) a residual component that is not explained by our model (in blue).

Among people aged 65 years or over, those from all ethnic minority groups except Chinese were at greater risk of COVID-19 related death than those from the White population. The odds of COVID-19 death were 3.1 [2.7-3.6] times greater for older Black African males and 2.2 [1.8-2.7] times greater for Black African females than for males and females from White ethnic group. The odds of death were also notably greater for people of Bangladeshi, Black Caribbean or Pakistani ethnic background than the White population, with odds ratios of 2.8 [2.3-3.4], 2.6 [2.4-2.9], and 2.2 [1.9-2.5], respectively, for males and 2.5 [1.9-3.3], 2.0 [1.7-2.3] and 2.2 [1.8-2.7], respectively, for females.

Living in a multi-generational household did not explain much of the difference in COVID-19 mortality rates amongst older males. However, it did explain a substantial part of the difference between older females of South Asian background and White older females. For older females of Pakistani background, living in a multi-generational household accounted for 13.8% [3.4%-23.9%] of the difference in COVID-19 mortality rates compared to White females; for females of Bangladeshi or Indian background, it accounted for 12.3% [7.3%-19.5%] and 12.2% [5.7% - 23.0%] of the difference in mortality rates, respectively (See Supplementary Table 4 in appendix for full results). The results were similar in the sensitivity analyses using different definitions of household composition (See Supplementary Table 5).

medRxiv preprint doi: https://doi.org/10.1101/2020.11.27.20238147; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Figure 4 Decomposition of odds ratios for COVID-19 mortality amongst older adults (aged ≥65 years) across ethnic groups, stratified by sex



Note: The overall height of the bar corresponds to the odds ratio (OR), relative to the White population, based on a logistic regression model adjusted for age. Error bars are 95% confidence intervals. The proportion of the age-adjusted ORs explained by living in a multi-generational household were calculated through a mediation analysis. The unexplained part corresponds to the ORs from a model adjusted for age, geographical factors (region, population density, urban/rural classification), socio-economic characteristics (IMD decile, household deprivation, educational attainment, social grade, household tenancy), health: (self-reported health and disability from the Census, pre-existing conditions based on hospital contacts, number of hospital admissions, total days spent in hospital)

Discussion

Principal findings

This paper makes two contributions to the research on COVID-19. First, we find that, among older adults, household composition is associated with COVID-19 mortality, even after adjusting for a range of socio-demographic factors and measures of health. Our results indicate that compared to those living in a two older adult household, older adults, especially females, living in a multi-generational household are at greater risk of COVID-19 death. Living alone is also associated with elevated COVID-19 mortality. Second, we find that living in a multi-generational household explains between 10% and 15% of the excess COVID-19 mortality risk for females of South Asian background, but very little for males or people from other ethnic groups.

Comparison with related studies

Our results are consistent with emerging evidence that household size is associated with the risk of infection [18, 19], and that older adults tend to be at greater risk of household transmission [20, 21]. Older people living in large household tend to live in multi-generational households, co-habiting with younger adults and children. There is some evidence that, amongst older adults, living with dependent children is not strongly associated with the risk of COVID-19 infection or adverse outcomes [22]. Whilst our results indicate that older adults living in a multi-generational household are at greater risk of COVID-19 death compared to those living with another older adult, we find

medRxiv preprint doi: https://doi.org/10.1101/2020.11.27.20238147; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

little difference in risk between older people living in a multi-generational household with or without young children.

Several studies have analysed ethnic differences in COVID-19 infection and mortality [3, 7, 6, 8, 4]. Although we focus on older adults only, we find that almost all ethnic minority groups were at higher risk of COVID-19 deaths compared to the White population, and that the differences were attenuated once we adjusted for a range of geographical factors, socio-demographic characteristics and comorbidities. We improve the existing evidence on ethnic inequalities in COVID-19 mortality by using a causal mediation approach to quantify the importance of living in a multi-generational household.

Mechanisms

Our results suggest that older people are placed at increased exposure to infection by living with younger adults rather than young children. After adjusting for confounding factors, we find that the risk of COVID-19 death is similar amongst older adults living in a household with young children and those living in a household with younger adults only. The increased risk is likely to be driven by co-residing with younger adults, who have a higher risk of infection than older people [19]. Younger adults are likely to be at increased risk of exposure because of work, as evidence suggests that in England people who are working were at greater risk of infection compared to people not in employment, especially if they were working in patient or client-facing occupations [19, 23, 24].

Older adults living by themselves were also found to be at greater risk of COVID-19 death than those living with another older adult. During the COVID-19 pandemic, older people living alone were more likely to have received help from carers, including informal helpers, than people living with another older adult [25]. These frequent contacts with people from different households could increase the risk of being exposed to the virus.

We find that living in a multi-generational household explains between 10% and 15% of the excess COVID-19 mortality risk for females of South Asian background, but very little for males, despite a similar proportion of them living in a multi-generational household. Women spend more time at home than men and still do the majority of unpaid housework [26], which could increase the risk of household transmission.

Strengths and limitations

The primary strength of our study lies in the use of a unique linked population-level dataset which combine the 2011 Census with death registration data and hospital records. Unlike data based solely on health records, our study dataset contains a broad range of information on demographic, socio-economic, and household characteristics, including occupation. Unlike sample survey data, it contains millions of observations covering the entire population of interest, allowing us to examine both the association between household composition and COVID-19 mortality and also whether living in a multi-generational household explains some of the disparity in COVID-19 mortality between ethnic groups. We were able to examine differences between disaggregated ethnic minority groupings rather than high-level categories of South Asian, Black, and Other.

The main limitation of our study is that household composition is likely to be imprecisely measured. Whilst household composition is based on a detailed and accurate measurement taken in 2011, we could only identify changes since then due to death of household members or a move to a care home. Whist we took several steps to limit the measurement error, such as focusing on older adults,

medRxiv preprint doi: https://doi.org/10.1101/2020.11.27.20238147; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

including only adults aged 25 or over and children aged 0 to 9 at the time of the census in our definition of household composition, our household composition measure may not reflect current living circumstances of everybody in our population of interest. To mitigate concerns about measurement error, we showed that our results are robust to using different definitions of household composition. Nonetheless, measurement error is likely to attenuate the explanatory power of household composition in our models. In addition, while we have used a causal mediation approach, our analysis remains based on observational data and therefore residual confounding is likely. Another limitation is that our statistical approach assumes that the effect of living in different types of household composition is the same across ethnic groups.

Conclusions

Older adults living in multi-generational households are at elevated risk of experiencing harms from COVID-19 compared to older adults living with people of the same age. However, there has been little focus on implementing effective interventions (such as creating plans to effectively isolate and improving ventilation within the home) to reduce transmission risk within the household [26]. Relevant public health interventions should be directed at communities where multi-generational households are highly prevalent. Living in a multi-generational household explains some of the excess COVID-19 mortality risk for females of South Asian background, but very little for males or people from other ethnic groups. Further research is needed to explain the difference in COVID-19 mortality between ethnic groups.

References

[1] ONS, “Updating ethnic contrasts in deaths involving the coronavirus (COVID-19), England and Wales: deaths occurring 2 March to 28 July 2020,” 2020. [Online]. Available: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/articles/updatingethniccontrastsindeathsinvolvingthecoronaviruscovid19englandandwales/deathsoccurring2marchto28july2020.

[2] Public Health England, “Disparities in the risk and outcomes of COVID-19,” 2020. [Online]. Available: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/892085/disparities_review.pdf.

[3] R. W. Aldridge, D. Lewer, S. V. Katikireddi, R. Mathur, N. Pathak, R. Burns, E. B. Fragaszy, A. M. Johnson, D. Devakumar, I. Abubakar and A. Hayward, “Black, Asian and Minority Ethnic groups in England are at increased risk of death from COVID-19: indirect standardisation of NHS mortality data,” *Wellcome Open Research,* vol. 5, p. 88, 5 2020.

[4] C. L. Niedzwiedz, C. A. O'Donnell, B. D. Jani, E. Demou, F. K. Ho, C. Celis-Morales, B. I. Nicholl, F. S. Mair, P. Welsh, N. Sattar, J. P. Pell and S. V. Katikireddi, “Ethnic and socioeconomic differences in SARS-CoV-2 infection: Prospective cohort study using UK Biobank,” *BMC Medicine,* vol. 18, 5 2020.

[5] D. Pan, S. Sze, J. S. Minhas, M. N. Bangash, N. Pareek, P. Divall, C. M. L. Williams, M. R. Oggioni, I. B. Squire, L. B. Nellums, W. Hanif, K. Khunti and M. Pareek, “The impact of

medRxiv preprint doi: https://doi.org/10.1101/2020.11.27.20238147; this version posted December 2, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

ethnicity on clinical outcomes in COVID-19: A systematic review," *EClinicalMedicine,* vol. 23, 2020.

[6] D. Ayoubkhani, V. Nafilyan, C. White, P. Goldblatt, C. Gaughan, L. Blackwell, N. Rogers, A. Banerjee, K. Khunti, M. Glickman, B. Humberstone and I. Diamond, "Ethnic minority groups in England and Wales - factors affecting the size and timing of elevated COVID-19 mortality : a retrospective cohort study linking Census and death records," *Medrxiv,* 2020.

[7] E. J. Williamson, A. J. Walker, K. Bhaskaran, S. Bacon, C. Bates, C. E. Morton, H. J. Curtis, A. Mehrkar, D. Evans, P. Inglesby, J. Cockburn, H. I. McDonald, B. MacKenna, L. Tomlinson, I. J. Douglas, C. T. Rentsch, R. Mathur, A. Y. S. Wong, R. Grieve, D. Harrison, H. Forbes, A. Schultze, R. Croker, J. Parry, F. Hester, S. Harper, R. Perera, S. J. W. Evans, L. Smeeth and B. Goldacre, "OpenSAFELY: factors associated with COVID-19 death in 17 million patients.," *Nature,* 8 7 2020.

[8] R. Mathur, C. T. Rentsch, C. E. Morton, W. J. Hulme, A. Schultze, B. MacKenna, R. Eggo, K. Bhaskaran, A. Y. S. Wong, E. J. Williamson, H. Forbes, K. Wing, H. I. McDonald, C. Bates, S. Bacon, A. J. Walker, D. Evans, P. Inglesby, A. Mehrkar, H. J. Curtis, N. J. DeVito, R. Croker, H. Drysdale, J. Cockburn, J. Parry, F. Hester, S. Harper, I. J. Douglas, L. Tomlinson, S. J. W. Evans, R. Grieve, D. Harrison, K. Rowan, K. Khunti, N. Chaturvedi, L. Smeeth and B. Goldacre, "Ethnic differences in COVID-19 infection, hospitalisation, and mortality: an OpenSAFELY analysis of 17 million adults in England," *medRxiv,* p. 2020.09.22.20198754, 2020.

[9] ONS, "Ethnicity facts and figures: Families and households," 2019. [Online]. Available: https://www.ethnicity-facts-figures.service.gov.uk/uk-population-by-ethnicity/demographics/families-and-households/latest.

[10] P. Muennig, B. Jiao and E. Singer, "Living with parents or grandparents increases social capital and survival: 2014 General Social Survey-National Death Index," *SSM - Population Health,* vol. 4, p. 71–75, 4 2018.

[11] I. Kawachi and L. F. Berkman, "Social Capital, Social Cohesion, and Health," in *Social Epidemiology*, Oxford University Press, 2015.

[12] G. Harlem and M. Lynn, "Descriptive analysis of social determinant factors in urban communities affected by COVID-19," *Journal of Public Health,* 6 2020.

[13] S. Haroon, J. S. Chandan, J. Middleton and K. K. Cheng, *Covid-19: Breaking the chain of household transmission,* vol. 370, BMJ Publishing Group, 2020.

[14] Department for Communities and Local Government, "The English Indices of Deprivation 2015," 2015.

[15] Welsh Government, "Welsh Index of Multiple Deprivation (WIMD) 2014," 2014.

[16] K. Imai, L. Keele and D. Tingley, "A General Approach to Causal Mediation Analysis," *Psychological Methods,* vol. 15, p. 309–334, 12 2010.

[17] K. Imai, L. Keele and T. Yamamoto, "Identification, inference and sensitivity analysis for causal mediation effects," *Statistical Science,* vol. 25, p. 51–71, 2 2010.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.27.20238147; this version posted December 2, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

[18] C. A. Martin, D. R. Jenkins, J. S. Minhas, L. J. Gray, J. Tang, C. Williams, S. Sze, D. Pan, W. Jones, R. Verma, S. Knapp, R. Major, M. Davies, N. Brunskill, M. Wiselka, C. Brightling, K. Khunti, P. Haldar and M. Pareek, "Socio-demographic heterogeneity in the prevalence of COVID-19 during lockdown is associated with ethnicity and household size: Results from an observational cohort study," *EClinicalMedicine,* vol. 25, 8 2020.

[19] H. Ward, C. J. Atchison, M. Whitaker, K. E. C. Ainslie, J. Elliott, L. C. Okell, R. Redd, D. Ashby, C. A. Donnelly, W. Barclay, A. Darzi, G. Cooke, S. Riley and P. Elliott, "Antibody prevalence for SARS-CoV-2 in England following first peak of the pandemic: REACT2 study in 100,000 adults," *medRxiv,* p. 2020.08.12.20173690, 1 2020.

[20] Q. L. Jing, M. J. Liu, Z. B. Zhang, L. Q. Fang, J. Yuan, A. R. Zhang, N. E. Dean, L. Luo, M. M. Ma, I. Longini, E. Kenah, Y. Lu, Y. Ma, N. Jalali, Z. C. Yang and Y. Yang, "Household secondary attack rate of COVID-19 and associated determinants in Guangzhou, China: a retrospective cohort study," *The Lancet Infectious Diseases,* vol. 20, p. 1141–1150, 10 2020.

[21] H. F. Fung, L. Martinez, F. Alarid-Escudero, J. A. Salomon, D. M. Studdert, J. R. Andrews, J. D. Goldhaber-Fiebert and S. C.-C. M. Group, "The household secondary attack rate of SARS-CoV-2: A rapid review," *Clinical Infectious Diseases,* 10 2020.

[22] H. Forbes, C. E. Morton, S. Bacon, H. I. McDonald, C. Minassian, J. P. Brown, C. T. Rentsch, R. Mathur, A. Schultze, N. J. DeVito, B. MacKenna, W. J. Hulme, R. Croker, A. J. Walker, E. J. Williamson, C. Bates, A. Mehrkar, H. J. Curtis, D. Evans, K. Wing, P. Inglesby, H. Drysdale, A. Y. S. Wong, J. Cockburn, R. McManus, J. Parry, F. Hester, S. Harper, I. J. Douglas, L. Smeeth, S. J. W. Evans, K. Bhaskaran, R. M. Eggo, B. Goldacre, L. A. Tomlinson and T. DataLab, "Association between living with children and outcomes from COVID-19: an OpenSAFELY cohort study of 12 million adults in England," *medRxiv,* p. 2020.11.01.20222315, 2020.

[23] M. Mutambudzi, C. L. Niedzwiedz, E. B. Macdonald, A. H. Leyland, F. S. Mair, J. J. Anderson, C. A. Celis-Morales, J. Cleland, J. Forbes, J. M. R. Gill, C. Hastie, F. K. Ho, B. D. Jani, D. F. Mackay, B. I. Nicholl, O&#039, C. A. Donnell, N. I. Sattar, P. I. Welsh, J. P. Pell, S. V. Katikireddi and E. Demou, "Occupation and risk of severe COVID-19: prospective cohort study of 120,075 UK Biobank participants," *medRxiv,* p. 2020.05.22.20109892, 1 2020.

[24] A. S. V. Shah, R. Wood, C. Gribben, D. Caldwell, J. Bishop, A. Weir, S. Kennedy, M. Reid, A. Smith-Palmer, D. Goldberg, J. McMenamin, C. Fischbacher, C. Robertson, S. Hutchinson, P. McKeigue, H. Colhoun and D. A. McAllister, "Risk of hospital admission with coronavirus disease 2019 in healthcare workers and their households: nationwide linkage cohort study," *BMJ,* vol. 371, p. m3582, 10 2020.

[25] M. Evandrou, J. Falkingham, M. Qin and A. Vlachantoni, "Older and 'staying at home'during lockdown: informal care receipt during the COVID-19 pandemic amongst people aged 70 and over in the UK," *SocArXiv,* 2020.

[26] A. McMunn, L. Bird, E. Webb and A. Sacker, "Gender Divisions of Paid and Unpaid Work in Contemporary UK Couples," *Work, Employment and Society,* vol. 34, no. 2, 2020.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.27.20238147; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

[27] P. Little, R. C. Read, R. Amlôt, T. Chadborn, C. Rice, J. Bostock and L. Yardley, "Reducing risks from coronavirus transmission in the home—the role of viral load," *BMJ,* vol. 369, p. m1728, 5 2020.

[28] NIHR, "Multimillion investment in new research projects to investigate higher COVID-19 risk among certain ethnic groups," 2020. [Online]. Available: https://www.nihr.ac.uk/news/multimillion-investment-in-new-research-projects-to-investigate-higher-covid-19-risk-among-certain-ethnic-groups/25333.

Footnotes

Funding

This research was funded by the Office for National Statistics. VN is also funded by Health Data Research UK (HDR-UK). HDR-UK is an initiative funded by the UK Research and Innovation, Department of Health and Social Care (England) and the devolved administrations, and leading medical research charities. SVK acknowledges funding from a NRS Senior Clinical Fellowship (SCAF/15/02), the Medical Research Council (MC_UU_12017/13) and the Scottish Government Chief Scientist Office (SPHSU13)."

Data availability statement

Under the provisions of the Statistics and Registration Service Act 2007, the linked 2011 Census data used in this study are not permitted to be shared.

Ethics approval

Following assessment using the the National Statistician's Data Ethics Advisory Committee (NSDEC)'s tool, we engaged with the UK Statistics Authority Data Ethics team and it was decided that ethical approval was not required. This is standard practice for analysing national Census data.

Author contributors

VN, NI, DA, CL and KK contributed to the study conceptualisation and design. VN lead the preparation of the study data and performed the statistical analyses. All authors contributed to interpretation of the results. VN drafted the manuscript. DA, NI, CG, VK, RM, MA, PG, AJ, AB, SG and KK contributed to the critical revision of the manuscript. All authors approved the final manuscript. VN is the guarantor for the study. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

Conflict of interest statement

Prof. Khunti is a member of the Independent Scientific Advisory Group for Emergencies (SAGE), a Trustee of the SAHF, Director of the NIHR Applied Research Collaboration (ARC) East Midlands, and Director of the Centre for Black and Minority Ethnic (BME) Health. The other authors declare no competing interests.

Appendix

medRxiv preprint doi: https://doi.org/10.1101/2020.11.27.20238147; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Estimating the Average Causal Mediated Effect

The Average Causal Mediated Effect (ACME) of living in a multi-generational household for ethnic group k compared to the white group, $\delta(k)$ is estimated as:

$$\overline{\delta(k)} = \mathrm{E}\{[\Pr(MGH = 1|Ethnic = k\ ,X) - \Pr(MGH = 1|Ethnic = white\ ,X\ )] \cdot$$

$$[\ \Pr(Cdeath = 1|\ MGH = 0, Ethnic = k, X) - \Pr(Cdeath = 1|\ MGH = 0, Ethnic = k, X)]\}$$

Where MGH indicates if individual i lives in a multi-generational household, $Ethnic$ is the ethnic group and X is a vector of factors likely to confound the relationship between the mediator and the outcome. X includes geographical factors (region, population density, urban/rural classification), socio-economic characteristics (IMD decile, household deprivation, educational attainment, social grade, household tenancy), health: (self-reported health and disability from the Census, pre-existing conditions based on hospital contacts, number of hospital admissions, total days spent in hospital).

To estimate each component of $\overline{\delta(k)}$, we use predicted probabilities based on logistic regression models.

The total estimated difference in probability of COVID-19 death between ethnic group k and the White ethnic group is given by:

$$\overline{\tau(k)} = E\{\ \Pr(Cdeath = 1|\ Ethnic = k, age) - \Pr(Cdeath = 1|\ , Ethnic = white, age)\}$$

The proportion of the difference in the probability to die from COVID-19 between ethnic groups that is mediated by living in a multi-generational household is given by the ACME, $\overline{\delta(k)}$, as a proportion of the total effect $\overline{\tau(k)}$. We then used this proportion to decompose the age-adjusted odds ratios.

Supplementary Table 1: Sample selection and number of participants

Sample	Number of people
All	53,483,456
Excluding non-usual residents	52,637,675
Excluding those who died before 2 March 2020	48,468,645
Excluding those who lived in Wales in 2019	45,842,599
Excluding those living in care home in 2019	45,306,953
Excluding those over 110 years old	45,306,378
Excluding those less than 65	10,078,568

Supplementary Table 2: Distributions of study variables, stratified by household composition:

		All		Single	Two-adults hh	MGH without young children	MGH with young children	3+ older adults
		Mean	SD	Mean	Mean	Mean	Mean	Mean
Mortality	COVID-19 death	0.0028	0.0526	0.0040	0.0021	0.0023	0.0030	0.0023

medRxiv preprint doi: https://doi.org/10.1101/2020.11.27.20238147; this version posted December 2, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

	Other death	0.0123	0.1100	0.0178	0.0098	0.0087	0.0078	0.0115
Household composition	Two older adults	0.5506	0.4974					
	Single	0.3250	0.4684					
	MGH without young children	0.0983	0.2978					
	MGH with young children	0.0196	0.1387					
	Other 3+ adults	0.0065	0.0805					
Age	Age in 2020	75.2	7.6	78.0	74.2	72.7	72.6	76.0
Sex	Female	0.5385	0.4985	0.6558	0.4823	0.4763	0.4926	0.5062
	Population density	3531.1	3816.0	3922.9	3102.8	4195.7	5499.0	4216.3
Ethnicity	White	0.9387	0.2399	0.9547	0.9587	0.8408	0.6209	0.8852
	Bangladeshi	0.0023	0.0477	0.0010	0.0011	0.0063	0.0374	0.0023
	Black African	0.0045	0.0669	0.0046	0.0028	0.0091	0.0273	0.0093
	Black Caribbean	0.0088	0.0933	0.0116	0.0060	0.0128	0.0186	0.0125
	Chinese	0.0035	0.0590	0.0025	0.0029	0.0084	0.0097	0.0066
	Indian	0.0181	0.1334	0.0087	0.0122	0.0627	0.1088	0.0382
	Mixed	0.0044	0.0663	0.0053	0.0035	0.0051	0.0077	0.0120
	Other	0.0120	0.1089	0.0085	0.0088	0.0300	0.0631	0.0271
	Pakistani	0.0077	0.0875	0.0031	0.0039	0.0247	0.1064	0.0068
Region	East	0.1194	0.3243	0.1155	0.1242	0.1110	0.0976	0.1001
	East Midlands	0.0914	0.2881	0.0877	0.0964	0.0787	0.0775	0.0816
	London	0.1024	0.3031	0.1106	0.0804	0.1695	0.2267	0.1590
	North East	0.0517	0.2214	0.0545	0.0522	0.0442	0.0337	0.0392
	North West	0.1339	0.3406	0.1397	0.1314	0.1329	0.1180	0.1217
	South East	0.1722	0.3776	0.1672	0.1772	0.1661	0.1457	0.1763
	South West	0.1217	0.3269	0.1186	0.1290	0.0979	0.0847	0.1327
	West Midlands	0.1072	0.3094	0.1047	0.1066	0.1156	0.1266	0.1027
	Yorkshire and the Humbers	0.1001	0.3001	0.1014	0.1027	0.0840	0.0895	0.0867
Household deprivation	Not deprived	0.3524	0.4777	0.2349	0.4197	0.3835	0.3082	0.1835
	Deprived along 1 dimension	0.3836	0.4863	0.4175	0.3700	0.3533	0.3540	0.3873
	Deprived along 2 dimensions	0.2161	0.4116	0.2907	0.1737	0.2004	0.2295	0.2693
	Deprived along 3 dimensions	0.0447	0.2067	0.0530	0.0346	0.0579	0.0935	0.1389
	Deprived along 4 dimensions	0.0032	0.0566	0.0039	0.0019	0.0049	0.0148	0.0211
Urban/rural classification	Rural hamlets and isolated dwellings	0.0433	0.2036	0.0308	0.0509	0.0420	0.0398	0.0592
	Rural hamlets and isolated dwellings in a sparse setting	0.0044	0.0664	0.0035	0.0052	0.0037	0.0036	0.0034
	Rural town and fringe	0.1070	0.3091	0.1026	0.1160	0.0818	0.0610	0.0836
	Rural town and fringeï¿½in a sparse setting	0.0050	0.0709	0.0055	0.0053	0.0027	0.0021	0.0022
	Rural village	0.0734	0.2607	0.0608	0.0843	0.0588	0.0463	0.0754
	Rural villageï¿½in a sparse setting	0.0051	0.0715	0.0044	0.0060	0.0032	0.0032	0.0061

medRxiv preprint doi: https://doi.org/10.1101/2020.11.27.20238147; this version posted December 2, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

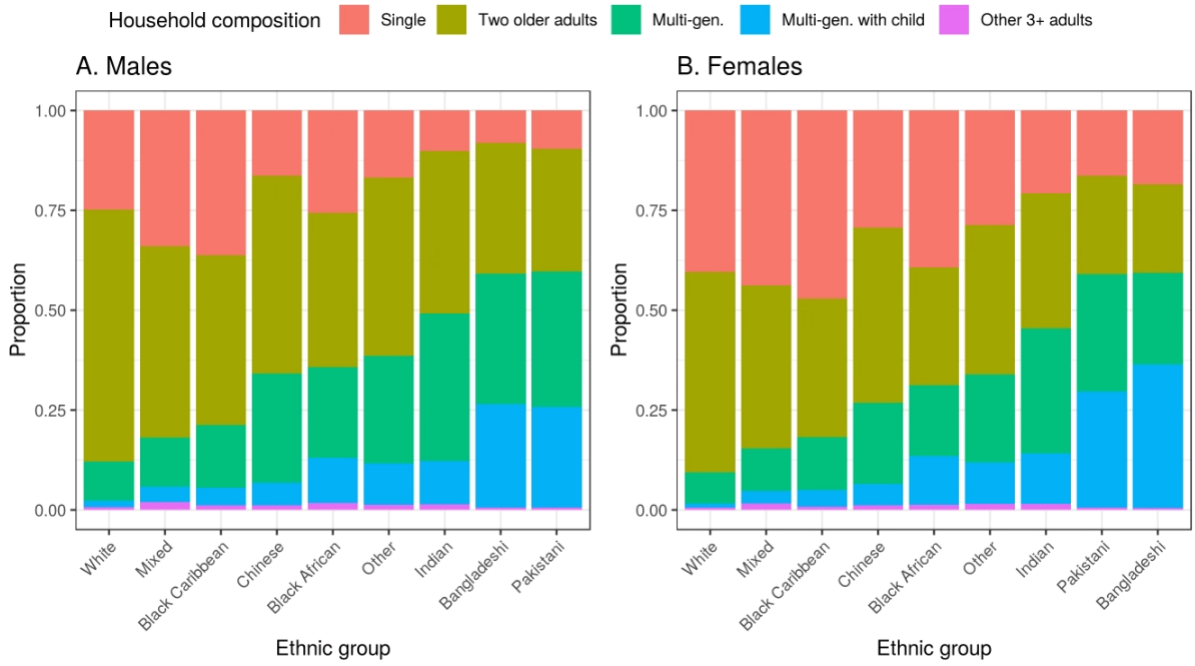
	Urban city and town	0.4359	0.4959	0.4459	0.4408	0.3945	0.3551	0.3930
	Urban city and town in a sparse setting	0.0022	0.0473	0.0025	0.0023	0.0015	0.0006	0.0010
	Urban major conurbation	0.2899	0.4537	0.3087	0.2556	0.3815	0.4626	0.3405
	Urban minor conurbation	0.0337	0.1803	0.0351	0.0337	0.0302	0.0257	0.0356
Approximated social grade	AB	0.1966	0.3974	0.1476	0.2253	0.1953	0.2193	0.1591
	C1	0.2968	0.4569	0.3134	0.2932	0.2667	0.2781	0.2904
	C2	0.2010	0.4007	0.1498	0.2210	0.2532	0.2204	0.2143
	D	0.2665	0.4421	0.3135	0.2399	0.2629	0.2477	0.2739
	E	0.0392	0.1940	0.0758	0.0205	0.0218	0.0345	0.0623
Health	Very good	0.2202	0.4144	0.1798	0.2441	0.2253	0.1949	0.2138
	Good	0.4469	0.4972	0.4159	0.4660	0.4554	0.3949	0.4058
	Fair	0.2523	0.4343	0.3032	0.2229	0.2426	0.2793	0.2618
	Poor	0.0662	0.2486	0.0835	0.0550	0.0623	0.1048	0.0874
	Very poor	0.0144	0.1193	0.0176	0.0120	0.0143	0.0261	0.0312
Long-term health problem or disability	Daily activities no limited	0.1233	0.3287	0.1606	0.1009	0.1107	0.1708	0.1944
	Daily activities limited a little	0.1935	0.3950	0.2359	0.1725	0.1712	0.1943	0.1837
	Daily activities limited a lot	0.6833	0.4652	0.6035	0.7266	0.7181	0.6349	0.6220
Hospital based comorbidities	Cancer	0.1074	0.3096	0.1031	0.1130	0.0978	0.0817	0.0778
	Cardiovascular disease	0.2997	0.4581	0.3310	0.2868	0.2750	0.2939	0.2170
	Digestive disorder	0.0201	0.1404	0.0212	0.0196	0.0197	0.0215	0.0120
	Mental health condition	0.0229	0.1496	0.0346	0.0175	0.0159	0.0162	0.0250
	Metabolic disorder	0.0883	0.2837	0.0922	0.0824	0.0999	0.1373	0.0681
	Musculoskeletal disorder	0.1035	0.3046	0.1300	0.0923	0.0835	0.0880	0.0755
	Neurological disorder	0.0184	0.1343	0.0198	0.0181	0.0156	0.0152	0.0172
	Renal disorder	0.0721	0.2586	0.0943	0.0617	0.0585	0.0653	0.0588
	Respiratory disorder	0.1228	0.3282	0.1459	0.1114	0.1113	0.1268	0.0932
Overcrowded		0.0313	0.1742	0.0373	0.0143	0.0631	0.2060	0.1699
IMD Decile	1	0.0670	0.2500	0.0910	0.0496	0.0706	0.1325	0.0930
	2	0.0717	0.2581	0.0896	0.0576	0.0806	0.1213	0.0925
	3	0.0810	0.2728	0.0946	0.0699	0.0907	0.1114	0.0943
	4	0.0911	0.2877	0.0994	0.0842	0.0979	0.1062	0.1057
	5	0.1020	0.3027	0.1041	0.1008	0.1028	0.0972	0.1015
	6	0.1111	0.3143	0.1079	0.1140	0.1091	0.0916	0.1199
	7	0.1178	0.3223	0.1088	0.1248	0.1127	0.0913	0.1198
	8	0.1191	0.3239	0.1059	0.1298	0.1108	0.0894	0.0967
	9	0.1193	0.3242	0.1022	0.1325	0.1121	0.0802	0.0927
	10	0.1198	0.3248	0.0965	0.1368	0.1128	0.0789	0.0839
Level of highest qualification	No qualification	0.3738	0.4838	0.4581	0.3274	0.3457	0.4092	0.4122
	Level 1: 1-4 GCSE/O-Level	0.0930	0.2905	0.0852	0.0972	0.0987	0.0818	0.0838
	Level 2: 5+ GCSE/O levels	0.1050	0.3065	0.1000	0.1088	0.1039	0.0849	0.1035
	Apprenticeship	0.0592	0.2360	0.0425	0.0695	0.0617	0.0384	0.0484
	Level 3: 2+ A Levels or equivalent	0.0655	0.2474	0.0521	0.0721	0.0737	0.0589	0.0710
	Level 4+: degree or above	0.2377	0.4257	0.2040	0.2611	0.2281	0.1983	0.2105

medRxiv preprint doi: https://doi.org/10.1101/2020.11.27.20238147; this version posted December 2, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Tenure of household	Other	0.0657	0.2478	0.0581	0.0639	0.0882	0.1286	0.0706
	Owned outright	0.6187	0.4857	0.5915	0.6605	0.5405	0.3255	0.5041
	Owned with a mortgage	0.1955	0.3966	0.1157	0.2114	0.3203	0.4322	0.2394
	Shared ownership	0.0042	0.0647	0.0061	0.0034	0.0025	0.0044	0.0058
	Social rented from council	0.0657	0.2477	0.1074	0.0432	0.0507	0.0736	0.0851
	Other social rented	0.0557	0.2294	0.0948	0.0356	0.0380	0.0594	0.0644
	Private rented	0.0493	0.2164	0.0664	0.0385	0.0413	0.0960	0.0790
	Living rent free	0.0109	0.1041	0.0181	0.0074	0.0067	0.0089	0.0222

medRxiv preprint doi: https://doi.org/10.1101/2020.11.27.20238147; this version posted December 2, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Supplementary Figure 1 Household composition by ethnic group for people in England aged ≥65 years, stratified by sex



Linked 2011 Census and mortality registration data for people in England aged ≥65 years, excluding those living in a care home in 2019. The number of adults in the household was calculated as the number of people aged ≥16 years who lived in the household at the time of the Census, minus those who died between 27 March 2011 and 1 March 2020.

Supplementary Table 3: Hazard ratios for COVID-19 related death for older adults (aged ≥65 years) in England, compared to living in a household with one other older adult, using different definitions of household composition, stratified by sex

	Men		Women	
	Def. A1	Def. A2	Def. A1	Def. A2
Single	1.164 [1.123 - 1.206]	1.164 [1.124 - 1.207]	1.12 [1.071 - 1.170]	1.12 [1.071 - 1.171]
Multi-generational household without children	1.032 [0.975 - 1.092]	1.027 [0.974 - 1.083]	1.207 [1.115 - 1.308]	1.226 [1.135 - 1.323]
Multi-generational household with children	1.129 [1.007-1.266]	1.128 [1.006-1.266]	1.165 [1.007 - 1.348]	1.176 [1.016 - 1.361]
3+ older adults	0.873 [0.702 - 1.085]	0.92 [0.749 - 1.131]	0.761 [0.577 - 1.002]	0.715 [0.537 -0.953]
Observations	432,894		489,192	

Note:. Def A1: we derived household composition in 2020 based on the number of adults aged 20 years (instead of 25 years); Def A2: multi-generational household as household with someone aged 65 years or over in 2020 co-resided with at least one other adult aged more than 15 years (instead of 20 years) younger. Hazard ratios compared to living in a household with one other older adult. Fully adjusted Cox regression models include geographical factors (region, population density, urban/rural classification), ethnicity, socio-economic characteristics (IMD decile, household deprivation, educational attainment, social grade, household tenancy), health (self-reported health and disability from the Census, pre-existing conditions based on hospital contacts, number of hospital admissions, total days spent in hospital), a measure for overcrowding, and property type.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.27.20238147; this version posted December 2, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Supplementary Table 4: Proportion of difference in COVID-19 mortality rates between ethnic groups mediated by living in a multi-generational household, stratified by sex

	Multi-gen with young children		Multi-gen without young children		Multi-gen (total)	
	Prop.	95% CI	Prop.	95% CI	Prop.	95% CI
Females						
Bangladeshi	0.0856	[0.0057 - 0.1738]	0.0379	[0.0195 - 0.0647]	0.1234	[0.0444 - 0.2155]
Black African	0.0296	[0.0020 - 0.0615]	0.0261	[0.0138 - 0.0414]	0.0557	[0.0254 - 0.0937]
Black Caribbean	0.0083	[0.0005 - 0.0169]	0.0150	[0.0078 - 0.0238]	0.0233	[0.0123 - 0.0368]
Chinese	0.02750	[-0.1331 - 0.2331]	0.1011	[-0.4773 - 0.6481]	0.1286	[-0.6403 – 0.8909]
Indian	0.0371	[0.0023 - 0.0743]	0.0852	[0.0465 - 0.1282]	0.12223	[0.0726 - 0.1834]
Mixed	0.0107	[-0.0128 - 0.0636]	0.0100	[-0.0137 - 0.0499]	0.0207	[-0.0339 - 0.1001]
Other	0.0305	[0.0018 - 0.0630]	0.0489	[0.0255 - 0.0764]	0.0794	[0.0430 - 0.1260]
Pakistani	0.0745	[0.0045 - 0.1519]	0.0630	[0.0343 - 0.0969]	0.1375	[0.0651 - 0.2230]
Males						
Bangladeshi	0.0361	[-0.0062 - 0.0810]	0.0085	[-0.0150 - 0.0270]	0.0446	[-0.0070 - 0.0953]
Black African	0.0125	[-0.0024 - 0.0312]	0.0038	[-0.0074 - 0.0132]	0.0163	[-0.0035 - 0.0381]
Black Caribbean	0.0033	[-0.0007 - 0.0088]	0.0019	[-0.0035 - 0.0061]	0.0051	[-0.0015 - 0.0128]
Chinese	0.0209	[-0.2011 - 0.4050]	0.0254	[-0.1984 - 0.4917]	0.0462	[-0.3185 - 0.8844]
Indian	0.0192	[-0.0034 - 0.0449]	0.0138	[-0.0246 - 0.0421]	0.0330	[-0.0112 - 0.0704]
Mixed	0.0033	[-0.0006 - 0.0090]	0.0007	[-0.0016 - 0.0033]	0.0040	[-0.0008 - 0.0106]
Other	0.0157	[-0.0030 - 0.0369]	0.0157	[-0.0132 - 0.0239]	0.0232	[-0.0057 - 0.0521]
Pakistani	0.0407	[-0.0077 - 0.0926]	0.0407	[-0.0167 - 0.0300]	0.0505	[-0.0083 - 0.1085]

Note: Proportion of difference in COVID-19 mortality between ethnic group mediated by living in a multi-generational household is estimated as the Average Causal Mediated Effect (ACME) as a proportion of the age-adjusted difference in the probability of COVID-19 death. The ACME is derived based on models that adjust for geographical factors (region, population density, urban/rural classification), socio-economic characteristics (IMD decile, household deprivation, educational attainment, social grade, household tenancy), and health: (self-reported health and disability from the Census, pre-existing conditions based on hospital contacts). 95% Confidence intervals are obtained via bootstrapping, using 500 replications.

Supplementary Table 5: Proportion of difference in COVID-19 mortality rates between ethnic groups mediated by living in a multi-generational household using different definitions of household composition, stratified by sex

	Men			Women		
Household composition	Main def	Def. A1	Def. A2	Main def	Def. A1	Def. A2
Bangladeshi	4.5%	4.5%	4.6%	12.3%	12.3%	15.2%
Black African	1.6%	1.6%	1.7%	5.6%	5.6%	6.8%
Black Caribbean	0.5%	0.5%	0.5%	2.3%	2.3%	2.5%
Chinese	4.6%	4.7%	4.5%	12.9%	12.9%	14.9%
Indian	3.3%	3.3%	3.1%	12.2%	12.2%	13.3%
Mixed	0.4%	0.4%	0.4%	2.1%	2.2%	2.5%
Other	2.3%	2.3%	2.3%	7.9%	7.9%	9.6%
Pakistani	5.1%	5.1%	5.1%	13.8%	13.7%	15.7%

medRxiv preprint doi: https://doi.org/10.1101/2020.11.27.20238147; this version posted December 2, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Note: Main definition of multi-generational household: someone aged 65 years or over on 2 March in 2020 co-resided with at least one other adult aged more than 20 years younger (and at least 25 in 2011), or with at least one child. Def A1: we derived household composition in 2020 based on the number of adults aged 20 years (instead of 25 years); Def A2: multi-generational household as household with someone aged 65 years or over in 2020 co-resided with at least one other adult aged more than 15 years (instead of 20 years) younger. Proportion of difference in COVID-19 mortality between ethnic group mediated by living in a multi-generational household is estimated as the Average Causal Mediated Effect (ACME) as a proportion of the age-adjusted difference in the probability of COVID-19 death. The ACME is derived based on models that adjust for geographical factors (region, population density, urban/rural classification), socio-economic characteristics (IMD decile, household deprivation, educational attainment, social grade, household tenancy), and health: (self-reported health and disability from the Census, pre-existing conditions based on hospital contacts).