

Greater risk of severe COVID-19 in non-White ethnicities is not explained by cardiometabolic, socioeconomic, or behavioural factors, or by 25(OH)-vitamin D status: study of 1,326 cases from the UK Biobank

Dr. Zahra Raisi-Estabragh^{1,2}, Ms. Celeste McCracken¹, Dr. Mae S. Bethell³, Ms. Jackie Cooper¹, Prof. Cyrus Cooper^{4,5,6}, Prof Mark J. Caulfield¹, Prof Patricia B. Munroe¹, Prof. Nicholas C. Harvey^{4,5}, Prof. Steffen E. Petersen^{*1,2}

1. William Harvey Research Institute, NIHR Barts Biomedical Research Centre, Queen Mary University of London, London, UK

2. Barts Heart Centre, St Bartholomew's Hospital, Barts Health NHS Trust, London, UK

3. North West Anglia NHS Foundation Trust, Hinchingsbrooke Hospital, Huntingdon, UK

4. MRC Lifecourse Epidemiology Unit, University of Southampton, Southampton, UK

5. NIHR Southampton Biomedical Research Centre, University of Southampton and University Hospital Southampton NHS Foundation Trust, Southampton, UK

6. NIHR Biomedical Research Centre, University of Oxford, Oxford, UK

***Corresponding author:** Professor Steffen E. Petersen. William Harvey Research Institute, NIHR Barts Biomedical Research Centre, Queen Mary University of London, London, UK; Email: s.e.petersen@qmul.ac.uk; Telephone: +44-2078826902

Abstract

Background

We examined whether the greater severity of coronavirus disease 2019 (COVID-19) amongst men and non-White ethnicities is explained by cardiometabolic, socio-economic, or behavioural factors.

Methods

We studied 4,510 UK Biobank participants tested for COVID-19 (positive, $n=1,326$). Multivariate logistic regression models including age, sex, and ethnicity were used to test whether addition of: 1) cardiometabolic factors (diabetes, hypertension, high cholesterol, prior myocardial infarction, smoking, BMI); 2) 25(OH)-vitamin D; 3) poor diet; 4) Townsend deprivation score; 5) housing (home type, overcrowding); or 6) behavioural factors (sociability, risk taking) attenuated sex/ethnicity associations with COVID-19 status.

Results

There was over-representation of men and non-White ethnicities in the COVID-19 positive group. Non-Whites had, on average, poorer cardiometabolic profile, lower 25(OH)-vitamin D, greater material deprivation, and were more likely to live in larger households and flats/apartments. Male sex, non-White ethnicity, higher BMI, Townsend deprivation score, and household overcrowding were independently associated with significantly greater odds of COVID-19. The pattern of association was consistent for men and women; cardiometabolic, socio-demographic and behavioural factors did not attenuate sex/ethnicity associations.

Conclusions

Sex and ethnicity differential pattern of COVID-19 is not adequately explained by variations in cardiometabolic factors, 25(OH)-vitamin D levels, or socio-economic factors. Investigation of alternative biological pathways and different genetic susceptibilities is warranted.

Introduction

The coronavirus disease 2019 (COVID-19) pandemic has to date resulted in over 5.8 million cases and 362,000 deaths worldwide¹. An increasing number of reports have highlighted men and Black and Minority Ethnic (BAME) cohorts as at higher risk of adverse COVID-19 outcomes^{2,3}. Variations in cardiometabolic disease burden⁴, oestrogen pathway activity⁵, vitamin D levels⁶, and angiotensin-converting enzyme (ACE) 2 receptor expression⁷ have been proposed as potential explanations for the differential pattern of disease severity. Furthermore, disparities in socio-economic standards, housing conditions, socialisation habits, and risk perception are potentially relevant with likely implications for risk of exposure and transmission. Understanding the significance of these factors is urgently needed to inform public health and research efforts.

We therefore investigated, in the UK Biobank (UKB) cohort, whether differential patterns of COVID-19 incidence and severity by sex and ethnicity might be explained by cardiometabolic, socio-economic, lifestyle, and behavioural exposures.

Methods

Setting and study population

UKB is a prospective cohort study of over half a million men and women from across the UK covering a range of urban and rural settings. Recruitment was between 2006-2010 through postal invite of individuals aged 40-69 years-old identified through National Health Service (NHS) registers. Baseline assessment included detailed characterisation of socio-demographics, lifestyle, health, a series of physical measures, and blood biochemistry. The protocol is publicly available⁸. Data linkage with Hospital Episode Statistics (HES) enables prospective tracking of health outcomes for all participants with conditions recorded according to international classification of disease (ICD). Incidence of key events, such as myocardial infarction (MI), are algorithmically defined by cross-checking over multiple data sources⁹. Linkage with Public Health England infection data has enabled rapid release of linked COVID-19 test results for UKB participants to researchers¹⁰. As such, near-real time study of the biological and epidemiological determinants of the disease is possible. The

latest data release (29/05/2020) includes test results from 16/03/2020 to 18/05/2020. As UK testing during this period was almost entirely restricted to hospitalised patients, researchers have been advised that COVID-19 positive status in UKB can be taken as surrogate for severe disease¹¹.

Exposures

We considered relevant demographic (age, sex, ethnicity), biological (cardiometabolic, 25(OH)-vitamin D status, poor diet quality), socio-economic (material deprivation, type of home, household overcrowding), and behavioural (sociability, attitude to risk) disease determinants (Supplementary Table 1).

We used age and sex as recorded at baseline. Ethnicity is reported as Caucasian (any White background) and non-Caucasian. For consistency with wider UK classification, we document ethnicity as White and non-White. For the latter we report breakdown of ethnicities as per existing UKB categories: Black (Caribbean, African, any other Black background), Asian (Indian, Pakistani, Bangladeshi, any other Asian background), Chinese, Mixed (White and Black Caribbean, White and Black African, White and Asian, Any other mixed background), and “other”. Townsend deprivation score is reported by the UKB as a measure of material deprivation calculated at baseline: zero, positive, and negative scores correspond to average, higher and lower levels of deprivation respectively, relative to national averages¹². We used type of housing as a binary variable comprising communal living spaces (flat, apartment, sheltered accommodation) vs stand-alone housing (house, bungalow). We considered household overcrowding based on self-report of household size (number of people in the same home) and intergenerational cohabitation (number of generations in the same home). Socialisation habits were defined per self-reports of frequency of family/friend visits and participation in regular leisure activities outside the home (sports, pub, religious group, adult education class). Attitude to risk was assessed using self-report of tendency “to take risks” (binary yes/no). BMI was calculated from height and weight recorded at baseline. Smoking status was based on self-report. Hypertension, diabetes, and hypercholesterolaemia were defined through cross-checking across self-report and HES data. A list of ICD codes used is available in Supplementary

Table 2. Prior MI was obtained from UKB algorithmically defined health outcomes. We used serum 25(OH)-vitamin D levels measured at baseline (CLIA analysis on a DiaSorin Ltd. LIASON XL), limiting to results between 10-375 nmol/L based on the manufacturer's analytic reportable range¹³. We adjusted for seasonality by regressing vitamin D on month of sampling as a factor, this allowed derivation of vitamin D adjusted to the same month for each participant. Preliminary analysis demonstrated significant difference in vitamin D by ethnicity and variation in the degree of seasonal variation by ethnicity (Figure 1D). We therefore performed seasonality adjustment separately for White and non-White ethnicities and added the intercept to the adjusted variables to maintain the difference between the two ethnic groups (i.e. regression residuals + intercept). We used processed meat intake as a marker of poor diet quality. We converted self-reported weekly intake frequencies into probabilities of daily intake and multiplied by portion size to derive a continuous measure of daily consumption in grams, as per previous publications using this dataset^{14,15}.

Ethics

This study was covered by the ethics approval for UKB studies from the NHS National Research Ethics Service on 17th June 2011 (Ref 11/NW/0382) and extended on 10th May 2016 (Ref 16/NW/0274).

Statistical analysis

Statistical analysis was performed using R Version 3.6.2 [R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>], and RStudio Version 1.2.5019 [RStudio Team (2015). RStudio: Integrated Development for R. RStudio, Inc., Boston, MA URL <http://www.rstudio.com/>].

UKB participants were grouped according to COVID-19 status: test positive, test negative, and untested. In analysis of an earlier data release, we demonstrated similar associations when comparing the untested cohort with both the test negatives and test positives; suggesting that comparison with the whole cohort in this way reveals associations with general hospitalisation rather than specifically with

COVID-19¹⁶. Therefore, to avoid bias relating to hospitalisation, in the present study, we limited logistic regression models to within the tested cohort i.e. test positives vs test negatives. We performed analyses in the whole tested sample, and separately in men and women. Logistic regression models were used to examine univariate and multivariate associations. We undertook individual multivariate models for each hypothesis to maximise sample size. We tested for multicollinearity setting a variance inflation factor (VIF) cut-off of 2.5. We present odds ratio (OR) for each exposure with the corresponding 95% confidence interval (CI) and p-value.

RESULTS

Population characteristics

Sex and ethnicity

Test results for 4,510 participants were available (positive, $n=1,326$; negative, $n=3,184$). Baseline characteristics are summarised in Table 1. Comparisons with the untested cohort ($n=497,996$) and characteristics by sex and ethnicity are summarised in Supplementary Tables 3, 4 and 5. There was over-representation of men and non-White ethnicities in the test positive cohort (Figure 1A, Figure 1B). Individuals of Black and Asian ethnicity were most disproportionately affected with Black ethnicities contributing over 3.5 times the number of positive cases than their representation in the untested cohort (Supplementary Table 3, Figure 1B).

Cardiometabolic factors and vitamin D

Men and non-White ethnicities had overall greater burden of cardiometabolic morbidities compared to women and White cohorts respectively (Figure 1E, Figure 1F). There were greater rates of smoking in White compared to Non-White ethnicities. Serum 25(OH)-vitamin D levels were, on average, higher in men and White ethnicities than women and non-White cohorts (Figure 1D).

Socio-demographic and behavioural factors

In comparison to the test negatives, those with a positive test had greater levels of social deprivation, were more likely to live in crowded households (Figure 1C). Non-Whites had higher levels of

material deprivation by Townsend score compared to those of White ethnicity; levels of deprivation between men and women were similar (Supplementary Table 4). The frequency of family/friend visits and leisure activities outside the home were comparable between the test positive and test negative groups. There was greater tendency to risk taking behaviour in the test positive cohort, which was greater in men vs women and in non-White vs White ethnicities.

Univariate associations of exposures with COVID-19 positive status

We tested the univariate association of all defined exposures with COVID-19 positive status within the tested cohort (Supplementary Table 6). Male sex, non-White ethnicity, higher BMI, greater material deprivation, and greater household overcrowding (household size, generations in household) were associated with increased odds of COVID-19 positive test. More frequent visits from family/friends were associated with lower risk of COVID-19 hospitalisation (given that a positive test implied hospital attendance), perhaps reflecting the role of social support in enabling individuals to remain at home when ill. There was a negative association between age and COVID-19 positivity, which may reflect the narrow range and distribution of ages in the sample. Testing separately in men, non-White ethnicity, greater material deprivation, and higher BMI were the only statistically significant exposures. For women, additionally, lower 25(OH)-vitamin D status, greater household overcrowding (household size, generations in household), and greater risk taking behaviour were associated with COVID-19 positivity.

Independent associations of specific exposures with COVID-19 status

Cardiometabolic factors

We undertook multivariate logistic regression models incorporating sex, age, ethnicity, smoking, BMI, diabetes, hypertension, high cholesterol, and prior MI (Table 2, Model 1). Male sex and non-White ethnicity were associated with greater odds of COVID-19 positive status with OR 1.28 (1.12, 1.46) and 1.78 (1.43, 2.20) respectively. Every 1kg/m² of BMI was associated with 1.03 (1.01, 1.04) greater odds of COVID-19 positivity. There was a borderline negative association with age 0.99 [0.98, 1.00], which remained significant in sex-stratified analysis for women only. Non-White

ethnicity and higher BMI were statistically significant associations for both men and women, with no evidence of attenuation (compared with the crude models).

25(OH)-vitamin D status and poor diet quality

In multivariate logistic regression models incorporating sex, age and ethnicity, there was no significant association between season-adjusted 25(OH)-vitamin D status and COVID-19 positivity (Table 2, Model 2). Similarly, in a separate model, adjustment for sex, age, and ethnicity demonstrated no statistically significant association between processed meat consumption and COVID-19 status (Table 2, Model 3). In both models, male sex was associated with higher odds of COVID-19 positive test and non-White ethnicity was consistently associated with COVID-19 positive test across men and women, with no evidence of attenuation.

Material deprivation

We tested the effect of material deprivation in multivariate models with mutual adjustment for sex, age, and ethnicity (Table 3, Model 4). There was a small, but statistically significant association between greater material deprivation and odds of COVID-19 positivity [OR 1.03 (1.01, 1.05)]. There remained strong and significant associations with male sex [OR 1.27 (1.11, 1.45)] and non-White ethnicity [OR 1.67 (1.34, 2.07)].

Housing conditions

We considered the effect of housing conditions in multivariate logistic regression models including sex, age, ethnicity, home type, and household size. In the whole sample, male sex, non-White ethnicity, and greater household size were associated with greater odds of COVID-19 positivity (Table 3, Model 5). Testing separately in men and women, non-White ethnicity was the only exposure which remained significantly associated with COVID-19 status. Attenuation of associations with household size is likely due to the small effect size and limited heterogeneity of the exposure in each of the sexes individually.

Socialisation habits and attitudes to risk

We undertook separate multivariate logistic regression models testing for associations between COVID-19 status, socialisation habits (Table 3, Model 6) and risk-taking attitude (Table 3, Model 7) whilst adjusting for age, sex, and ethnicity. Statistically significant associations were observed with male sex and non-White ethnicity which were not attenuated from crude models by adjustment for socialisation or risk-taking behaviour.

Final model

We built a final multivariate logistic regression model, with covariates selected based on previous model permutations including sex, ethnicity, BMI, Townsend score, and household size (Supplementary Table 7). All covariates apart from age were statistically significant in this model. Male sex and non-White ethnicity were associated with greater odds of COVID-19 positivity: OR 1.23 (1.08, 1.41) and 1.59 (1.26, 1.99) respectively. Every 1kg/m² increase in BMI was associated with 1.02 (1.01, 1.03) greater odds of COVID-19 positivity and for every additional person living in the same household the odds increased by 1.09 (1.03, 1.16).

Discussion

Main finding of this study

In 4,510 UKB participants tested for COVID-19 in a hospital setting (positive, $n=1,326$), male sex, non-White ethnicity, higher BMI, and greater household size were associated with significantly greater odds of COVID-19 positive test. Despite variation in burden of cardiometabolic morbidities, 25(OH)-vitamin D levels, and material deprivation by sex and ethnicity, these factors were not significantly associated with COVID-19 positive status and did not explain the strong association with ethnicity. The pattern of tested associations did not vary between men and women.

What is already known on this topic

There is mounting evidence for disproportionate adverse effects of COVID-19 in non-White ethnicities². UK national audit data demonstrates that up to one-third of COVID-19 patients requiring intensive care are from BAME backgrounds, a rate which is far greater than their representation in the general population¹⁷. Similarly, an analysis of COVID-19 deaths amongst NHS staff, found that 64% of deaths were in individuals from BAME cohorts, markedly disproportionate to their 20% contribution to the NHS workforce¹⁸. The latest report from the Office of National Statistics (ONS) also demonstrates greater risk of COVID-19 mortality in BAME groups¹⁹; individuals of Black ethnicity appeared at highest risk with over 3.5 times greater risk of COVID-19 death compared to Caucasians, followed by Asian ethnicities¹⁹. Similarly, in the USA, there has been growing concern over the disproportionate impact of COVID-19 deaths amongst African Americans. In Chicago, 54.5% of all COVID-19 deaths are in Black people, whilst they constitute only 30% of the population²⁰. These patterns are echoed across Europe, with Nordic countries reporting as much as ten times greater risk of COVID-19 in Somali populations²¹. We had previously documented this preponderance of cases amongst BAME individuals in our analysis of the initial UKB data release¹⁶; here we have confirmed the observation in the larger dataset, and importantly demonstrated a non-uniform impact across different BAME groups with highest rates amongst Black followed by Asian ethnicities.

The greater cardiometabolic burden in both BAME and male cohorts has been proposed as potentially important in driving adverse COVID-19 outcomes. In our sample, these cohorts had broadly poorer cardiometabolic profile than their female and White counterparts. However, in multivariate logistic regression models including multiple cardiometabolic factors statistically significant associations with COVID-19 status were observed with male sex, non-White ethnicity, and higher BMI only. This suggests that the greater cardiometabolic burden in non-White ethnicities does not account for the adverse COVID-19 outcomes in this group.

Consistent with our findings, UK and USA statistics highlight obesity as a marker of poor COVID-19 outcomes, such as requirement for intensive care²². There are increasing suggestions of a possible pathophysiological link between adiposity and COVID-19 severity. Wide expression of ACE2 receptors within adipose tissue is thought to promote binding and cellular entry of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)²³. It has been suggested that adipose tissue may act as a “viral reservoir” for SARS-CoV-2, perhaps contributing to a more prolonged and severe illness²³. In addition, adipose tissue is a known source of inflammatory cytokines, such as Interleukin 6²⁴. This is hypothesised to be linked to the association of adiposity with greater likelihood of cytokine storms and the consequent risk of severe respiratory complications in the context of COVID-19. Indeed, studies have demonstrated association of higher Interleukin 6 levels with respiratory failure and requirement for mechanical ventilation in COVID-19 patients²⁵. Greater adiposity, as well as non-White ethnicity, is associated with lower 25(OH)-vitamin D status. Although the active 1,25(OH)₂-vitamin D form has immune system functions²⁶, evidence linking low 25(OH)-vitamin D (the circulating storage form, and poorly correlated with 1,25(OH)₂-vitamin D) with COVID-19 disease have been contradictory²⁷. In our study, our finding of no independent associations between 25(OH)-vitamin D status and COVID-19 disease suggest that the association is confounded by ethnicity and BMI. Interestingly, the BMI association was retained in multivariate models, suggesting a possible independent role for adiposity, which clearly deserves further investigation.

Socio-economic deprivation is associated with poorer global health outcomes²⁸ with specific implications for infectious disease due to its association with overcrowded living conditions. It has been suggested that ethnic differences in COVID-19 severity may relate to clustering of material deprivation with BAME status²⁹. In the UKB material deprivation is reported using the Townsend score, which is based on four factors- employment, car ownership, home ownership, and household overcrowding. Consistent with national reports, we found higher material deprivation in non-White participants. In multivariate models including age, sex, ethnicity, and Townsend score, there were significantly greater odds of COVID-19 with greater material deprivation whilst the association with

ethnicity appeared strong and significant. Testing separately for the effect of household overcrowding, this exposure appeared significant independent of sex, ethnicity, age, and home type. This suggests that in modern western societies, it is not global economic deprivation, but specific aspects relating to household overcrowding that have relevance to COVID-19. Consistent with these observations, a survey of COVID-19 cases from New York reports the highest number of cases occurring in areas with the largest average household size³⁰. Furthermore, analysis of UK cases by the ONS also demonstrates that material deprivation does not adequately explain the ethnic disparities in COVID-19 outcomes¹⁹.

Behavioural factors, in particular attitudes that may compromise adherence to lockdown measures, have been proposed as potentially important in determining risk of exposure to SARS-CoV-2. There have been suggestions that the higher rates of COVID-19 in men may relate to their greater tendency to take risks than women³¹ with a recent UK survey identifying young men as the most likely demographic to break lockdown measures³². Cultural differences in socialisation habits and more frequent contact with friends and extended family are also potentially relevant factors. In our analysis, we did not find these factors to be significantly important in conferring COVID-19 positive status with the increased odds associated with male sex and non-White ethnicity occurring independent of these factors.

What this study adds

This study is consistent with growing reports of higher risk of severe COVID-19 in men and non-White ethnicities, in particular Black populations. We identified higher BMI, greater material deprivation, and household overcrowding as independent risk factors for COVID-19. In addition, we demonstrated that the ethnic and sex differential pattern of COVID-19 is not adequately explained by greater cardiometabolic disease burden, lower vitamin D levels, greater social deprivation, or behavioural factors. Thus, alternative biological pathways and different genetic susceptibilities warrant consideration.

Limitations of this study

Aggregating all BAME populations into one cohort may overlook important differences between non-Caucasian ethnicities. Further study in samples with greater ethnic diversity is needed. Given the observational nature of the study, we cannot discern causal relationships, and although we controlled for a wide range of covariates, the possibility of residual confounding should be considered.

Furthermore, the current dataset does not allow assessment of specific COVID-19 health outcomes.

Conclusion

There is greater risk of severe COVID-19 amongst men and non-White ethnicities. The augmented risk in BAME populations is non-uniform and disproportionately affects Black and Asian ethnicities. The sex and ethnicity differential pattern of COVID-19 is not adequately explained by variations in cardiometabolic factors, 25(OH)-vitamin D levels, or socio-economic factors. Thus, investigation of alternative biological pathways and genetic susceptibilities is warranted.

Funding statement

ZRE is supported by a British Heart Foundation Clinical Research Training Fellowship (FS/17/81/33318). SEP, PBM, MJC acknowledge support from the Barts Biomedical Research Centre funded by the National Institute for Health Research (NIHR). NCH and CC acknowledge support from the UK Medical Research Council (MRC #405050259; #U105960371), NIHR Southampton Biomedical Research Centre, University of Southampton and University Hospital Southampton, and NIHR Oxford Biomedical Research Centre, University of Oxford.

Acknowledgements

This study was undertaken using the UK Biobank resource, Access Application 2964.

References

1. World Health Organization (WHO). Coronavirus disease (COVID-19) Situation Report-131.
2. Kirby T. Evidence mounts on the disproportionate effect of COVID-19 on ethnic minorities. *Lancet Respir Med* 2020;**S2213-2600**:30228–30229.
3. Wenham C, Smith J, Morgan R. COVID-19: the gendered impacts of the outbreak. *Lancet* 2020;**395**:846–848.
4. Chaturvedi N. Ethnic differences in cardiovascular disease. *Heart* 2003;**89**:681–686.
5. Channappanavar R, Fett C, Mack M, Eyck PP Ten, Meyerholz DK, Perlman S. Sex-Based Differences in Susceptibility to Severe Acute Respiratory Syndrome Coronavirus Infection. *J Immunol* 2017;**198**:4046–4053.
6. Herrick KA, Storandt RJ, Afful J, et al. Vitamin D status in the United States, 2011–2014. *Am J Clin Nutr* 2019;**110**:150–157.
7. Cao Y, Li L, Feng Z, et al. Comparative genetic analysis of the novel coronavirus (2019-nCoV/SARS-CoV-2) receptor ACE2 in different populations. *Cell Discov* 2020;**6**:1–4.
8. UK Biobank: Protocol for a large-scale prospective epidemiological resource. 2007. <https://www.ukbiobank.ac.uk/wp-content/uploads/2011/11/UK-Biobank-Protocol.pdf> (13 May 2020)
9. Schnier C, Bush K, Nolan J, Sudlow C. Definitions of Acute Myocardial Infarction and Main Myocardial Infarction Pathological Types UK Biobank Phase 1 Outcomes Adjudication Documentation on behalf of UK Biobank Outcome Adjudication Group Definitions of Acute Myocardial Infarction. 2017. http://biobank.ndph.ox.ac.uk/showcase/showcase/docs/alg_outcome_mi.pdf (26 May 2020)
10. Bugbank Homepage. <http://www.bugbank.uk/index.html> (14 May 2020)
11. UK Biobank data showcase: Records of COVID-19 test results (Data-Field 40100). 2020. <http://biobank.ndph.ox.ac.uk/showcase/field.cgi?id=40100> (14 May 2020)
12. Townsend P, Phillimore P, Beattie A. Health and Deprivation: Inequality and the North. *Nurs Stand* 1988;**2**:34–34.
13. Fry D, Almond R, Moffat S, Gordon M, Singh P. UK Biobank Biomarker Project Companion Document to Accompany Serum Biomarker Data. 2019. <http://www.ukbiobank.ac.uk/uk-biobank-biomarker-panel/>. (14 May 2020)
14. Schenker S. Portion sizes Food Fact Sheet. 2016. www.bda.uk.com/foodfacts (25 April 2020)
15. Anderson JJ, Darwis NDM, Mackay DF, Celis-Morales CA, Lyall DM, Sattar N, Gill JMR, Pell JP. Red and processed meat consumption and breast cancer: UK Biobank cohort study and meta-analysis. *Eur J Cancer* 2018;**90**:73–82.
16. Raisi-Estabragh Z, McCracken C, Ardissino M, Bethell MS, Cooper J, Cooper C, Harvey NC, Petersen SE. Non-white ethnicity, male sex, and higher body mass index, but not medications

- acting on the renin-angiotensin system are associated with coronavirus disease 2019 (COVID-19) hospitalisation: review of the first 669 cases from the UK biobank. *medRxiv* 2020; doi: 2020.05.10.20096925.
17. ICNARC report on COVID-19 in critical care. ICNARC COVID-19 Study Case Mix Program. Database. 2020. p. 1–16. <https://www.icnarc.org/Our-Audit/Audits/Cmp/Reports> (26 May 2020)
 18. Cook T, Kursumovic E, Lennane S. Exclusive: deaths of NHS staff from covid-19 analysed. *Health Serv J* 2020; <https://www.hsj.co.uk/exclusive-deaths-of-nhs-staff-from-covid-19-analysed/7027471.article> (26 May 2020)
 19. White C. Coronavirus (COVID-19) related deaths by ethnic group , England and Wales □: 2 March 2020 to 10 April 2020- Office for National Statistics. 2020.
 20. City of Chicago: Latest Data. <https://www.chicago.gov/city/en/sites/covid-19/home/latest-data/2020-04-21.html> (24 May 2020)
 21. COVID-19 takes unequal toll on immigrants in Nordic region. <https://uk.reuters.com/article/uk-health-coronavirus-norway-immigrants/covid-19-takes-unequal-toll-on-immigrants-in-nordic-region-idUKKCN2260Y2> (24 May 2020)
 22. Kass DA, Duggal P, Cingolani O. Obesity could shift severe COVID-19 disease to younger ages. *Lancet* 2020;**395**:1544–1545.
 23. Kruglikov IL, Scherer PE. The role of adipocytes and adipocyte □like cells in the severity of COVID □19 infections. *Obesity* 2020; doi:10.1002/oby.22856
 24. Michalakis K, Ilias I. SARS-CoV-2 infection and obesity: Common inflammatory and metabolic aspects. *Diabetes Matab Syndr* 2020;**14**:469–471.
 25. Herold T, Jurinovic V, Arnreich C, et al. Level of IL-6 predicts respiratory failure in hospitalized symptomatic COVID-19 patients. *medRxiv* 2020; doi: 2020.04.01.20047381.
 26. Aranow C. Vitamin D and the Immune System. *J Investig Med* 2011;**59**:881–886.
 27. Lanham-New SA, Webb AR, Cashman KD, et al. Vitamin D and SARS-CoV-2 virus/COVID-19 disease. *BMJ Nutr Prev Heal* *BMJ*; 2020; doi: bmjnph-2020-000089.
 28. Health state life expectancies by national deprivation deciles, England and Wales, 2016 to 2018-Office for National Statistics.
 29. People living in deprived neighbourhoods - GOV.UK Ethnicity facts and figures. <https://www.ethnicity-facts-figures.service.gov.uk/uk-population-by-ethnicity> (24 May 2020)
 30. Wadhwa RK, Wadhwa P, Gaba P, et al. Variation in COVID-19 Hospitalizations and Deaths Across New York City Boroughs. *JAMA* 2020; doi: 10.1001/jama.2020.7197.
 31. Pawlowski B, Atwal R, Dunbar RIM. Evolutionary Psychology Sex Differences in Everyday Risk-Taking Behavior in Humans. *Evol Psychol* 2008;**6**:29–42.
 32. Kivimäki M, Batty GD, Pentti J, et al. Association between socioeconomic status and the development of mental and physical health conditions in adulthood: a multi-cohort study.

Lancet Public Heal 2020;**5**:e140–e149.

Table 1. Baseline demographics by COVID-19 status

	Test positive (n=1,326)	Test negative (n=3,184)
Men	696 (52.5%)	1,505 (47.3%)
Age	68.11 (\pm 9.23)	68.91 (\pm 8.72)
White ethnicity	1,141 (86.0%)	2,927 (91.9%)
Non-White ethnicity	174 (13.1%)	241 (7.6%)
Black ethnicity	76 (5.7%)	91 (2.9%)
Asian ethnicity	60 (4.5%)	78 (2.4%)
Chinese ethnicity	6 (0.5%)	3 (0.1%)
Mixed ethnicity	9 (0.7%)	24 (0.8%)
Other ethnicity*	34 (2.6%)	61 (1.9%)
Smoking (current or previous)	683 (51.5%)	1,653 (51.9%)
Processed meat intake (g/day)	17.08 (\pm 15.67)	16.33 (\pm 15.00)
BMI (kg/m ²)	28.04 [\pm 6.47]	27.41 [\pm 6.37]
Diabetes	217 (16.4%)	449 (14.1%)
Hypertension	624 (47.1%)	1,457 (45.8%)
High cholesterol	437 (33.0%)	1,034 (32.5%)
Prior myocardial infarction	96 (7.2%)	242 (7.6%)
Vitamin D (nmol/L)**	33.88 [\pm 27.01]	35.45 [\pm 26.78]
Townsend deprivation score	-0.91 [\pm 5.34]	-1.55 [\pm 5.00]
Home type (flat/apartment)	191 (14.4%)	455 (14.3%)
Household size	2.50 (\pm 1.31)	2.32 (\pm 1.22)
Number of generations in household	1.41 (\pm 0.52)	1.35 (\pm 0.50)
Family/friend visits	975 (73.5%)	2,438 (76.6%)
Regular leisure activity	897 (67.6%)	2,124 (66.7%)
Tendency to take risks	404 (30.5%)	916 (28.8%)

Table 1 footnote: Results are number (percentage) for categorical and mean (standard deviation) or median [interquartile range] for continuous variables. *Ethnicity was missing for n=11 test positive and n=16 test negative participants, these participants are included as part of “other ethnicity” in this table but have been excluded from subsequent modelling. **Vitamin D has been adjusted for seasonality.

Table 2. Multivariate logistic regression models testing the role of cardiometabolic factors (Model 1), vitamin D (Model 2), and poor diet (Model 3) in determining risk of COVID-19

	Exposures	Whole tested sample <i>n</i> =4,510	Men <i>n</i> =2,201	Women <i>n</i> =2,309
<i>Model 1: sex, age, ethnicity, smoking, BMI, diabetes, hypertension, high cholesterol, prior MI</i>	Male sex	1.28* [1.12, 1.46]	–	–
		4.05×10^{-4}	–	–
	Age	0.99* [0.98, 1.00]	1.00 [0.98, 1.01]	0.99* [0.97, 1.00]
	Non-White ethnicity	0.0157	0.5128	0.0097
		1.78* [1.43, 2.20]	2.07* [1.50, 2.84]	1.55* [1.15, 2.09]
	Smoking (previous/current)	1.88×10^{-7}	7.90×10^{-6}	0.0040
		1.02 [0.89, 1.16]	1.12 [0.92, 1.36]	0.91 [0.75, 1.10]
	BMI (kg/m ²)	0.7961	0.2533	0.3352
		1.02* [1.01, 1.03]	1.03* [1.01, 1.05]	1.02 [1.00, 1.03]
	Diabetes	0.0015	0.0051	0.0537
		1.08 [0.88, 1.32]	1.06 [0.82, 1.38]	1.08 [0.77, 1.49]
	Hypertension	0.4781	0.6529	0.6665
		1.01 [0.86, 1.18]	0.93 [0.74, 1.16]	1.11 [0.89, 1.40]
	High cholesterol	0.8875	0.5004	0.3563
		0.97 [0.82, 1.15]	1.04 [0.83, 1.31]	0.89 [0.68, 1.15]
Prior MI	0.7479	0.7108	0.3690	
	0.89 [0.68, 1.16]	0.85 [0.62, 1.15]	0.97 [0.55, 1.65]	
	0.4041	0.2961	0.8990	
<i>Model 2: age, sex, ethnicity, vitamin D</i>	Male sex	1.31* [1.14, 1.50]	–	–
		1.85×10^{-4}	–	–
	Age	0.99* [0.98, 1.00]	1.00 [0.99, 1.01]	0.99* [0.97, 1.00]
	Non-White ethnicity	0.0166	0.5500	0.0073
		1.77* [1.41, 2.22]	2.02* [1.45, 2.82]	1.60* [1.16, 2.18]
	Vitamin D	9.27×10^{-7}	3.51×10^{-5}	0.0038
1.00 [1.00, 1.00]		1.00 [1.00, 1.01]	1.00 [1.00, 1.01]	
	0.7185	0.7464	0.9288	
<i>Model 3: age, sex, ethnicity, processed meat</i>	Male sex	1.26* [1.10, 1.44]		
		8.55×10^{-4}		
	Age	0.99* [0.98, 1.00]	1.00 [0.99, 1.01]	0.99* [0.98, 1.00]
	Non-White ethnicity	0.0144	0.4993	0.0082
		1.81* [1.46, 2.24]	2.08* [1.52, 2.85]	1.62* [1.21, 2.17]
	Processed meat intake (100 grams/day)	4.18×10^{-8}	4.95×10^{-6}	0.0011
		1.26 [0.81, 1.94]	1.01 [0.57, 1.77]	1.83 [0.91, 3.66]
	0.3032	0.9742	0.0871	

Table 2 footnote: Results are odds ratios, 95% confidence interval, and p-values for each exposure from three separate models (1, 2, and 3). Exposures are mutually adjusted.

Table 3. Multivariate logistic regression models testing the role of material deprivation (Model 4), housing conditions (Model 5), socialisation habits (Model 6), and attitude to risk (Model 7) in determining risk of COVID-19

	Exposures	Whole tested sample n=4,510	Men n=2,201	Women n=2,309
<i>Model 4: age, sex, ethnicity, Townsend deprivation score</i>	Male sex	1.27* [1.11, 1.45]	–	–
		3.87×10^{-4}	–	–
	Age	0.99* [0.98, 1.00]	1.00 [0.99, 1.01]	0.99* [0.98, 1.00]
	Non- White ethnicity	0.0222	0.6323	0.0089
		1.67* [1.34, 2.07]	1.92* [1.39, 2.64]	1.49* [1.11, 2.01]
	Townsend deprivation score	3.94×10^{-6}	6.15×10^{-5}	0.0084
		1.03* [1.01, 1.05]	1.03* [1.00, 1.06]	1.03* [1.00, 1.06]
	0.0024	0.0402	0.0232	
<i>Model 5: age, sex, ethnicity, home type, household size*</i>	Male sex	1.24* [1.09, 1.42]	–	–
		0.0016	–	–
	Age	1.00 [0.99, 1.01]	1.00 [0.99, 1.01]	0.99 [0.98, 1.00]
	Non- White ethnicity	0.3827	0.8207	0.1655
		1.73* [1.39, 2.17]	1.86* [1.33, 2.59]	1.66* [1.22, 2.24]
	Home type	1.36×10^{-6}	2.60×10^{-4}	0.0011
		0.98 [0.80, 1.20]	1.05 [0.80, 1.38]	0.90 [0.66, 1.22]
Household size	0.8650	0.7044	0.4918	
	1.08* [1.02, 1.14]	1.08 [0.99, 1.18]	1.07 [0.99, 1.16]	
	0.0140	0.0764	0.0941	
<i>Model 6: age, sex, ethnicity, socialisation habits</i>	Male sex	1.27* [1.11, 1.46]	–	–
		4.78×10^{-4}	–	–
	Age	0.99* [0.98, 1.00]	1.00 [0.99, 1.01]	0.99* [0.98, 1.00]
	Non-White ethnicity	0.0129	0.4313	0.0099
		1.77* [1.43, 2.20]	2.12* [1.54, 2.91]	1.53* [1.13, 2.06]
	Family/friend visits	2.03×10^{-7}	3.31×10^{-6}	0.0050
		0.91 [0.78, 1.06]	0.88 [0.71, 1.08]	0.95 [0.75, 1.21]
	Regular leisure activity	0.2340	0.2073	0.6886
	1.05 [0.91, 1.21]	1.15 [0.94, 1.41]	0.95 [0.78, 1.16]	
	0.5104	0.1633	0.6149	
<i>Model 7: age, sex, ethnicity, risk taking</i>	Male sex	1.27* [1.11, 1.45]	–	–
		4.07×10^{-4}	–	–
	Age	0.99* [0.98, 1.00]	1.00 [0.99, 1.01]	0.99* [0.98, 1.00]
	Non-White ethnicity	0.0191	0.5423	0.0114
		1.79* [1.45, 2.21]	2.06* [1.51, 2.82]	1.60* [1.19, 2.13]
	Risk taking	5.99×10^{-8}	4.71×10^{-6}	0.0016
		1.02 [0.88, 1.17]	0.93 [0.77, 1.12]	1.14 [0.92, 1.42]
	0.8377	0.4340	0.2221	

Table 3 footnote: Results are odds ratios, 95% confidence interval, and p-values for each exposure from three separate models (4, 5, and 6). * Initial analyses additionally included number of generations in household, however, we observed significant multicollinearity between this variable and household size with higher VIF in the latter, hence it was removed from the final model. Exposures are mutually adjusted.

Figure legends:

Figure 1: Title: Baseline participant characteristics. **Figure 1 footnote:** **Panel A:** Male: Female split by COVID-19 status; **Panel B:** Percentage of participants from different non-Caucasian ethnicities by COVID-19 status; **Panel C:** Townsend deprivation score by ethnicity and COVID-19 status; **Panel D:** Vitamin D levels by month of measurement stratified by sex and ethnicity; **Panel E:** Cardiometabolic profile stratified by ethnicity; **Panel F:** Cardiometabolic profile stratified by sex. COVID-19: coronavirus disease 2019; MI: myocardial infarction.

