Social Determinants of Ethnic Disparities in SARS-CoV-2 Infection: UK Biobank Seroprevalence Study

Supplementary Appendix, Table of contents

Table S1: Variables used to select participants

Selection variable	Subcategories	Definition or source
Geographic region (postcode area)	Aberdeen / Bath / Birmingham / Blackburn / Blackpool / Bolton / Bournemouth / Bradford / Brighton / Bristol / Bromley / Cambridge / Canterbury / Cardiff / Carlisle / Central London / Central London / Chelmsford / Chester / Cleveland / Colchester / Coventry / Crewe / Croydon / Darlington / Dartford / Derby / Doncaster / Dorchester / Dudley / Dumfries and Galloway / Dundee / Durham / East London / Edinburgh / Enfield / Exeter / Falkirk and Stirling / Galashiels / Glasgow / Gloucester / Guildford / Halifax / Harrogate / Harrow / Hemel Hempstead / Hereford / Huddersfield / Hull / Ilford / Inverness / Ipswich / Kilmarnock / Kingston upon Thames / Kirkcaldy / Kirkwall / Lancaster / Leeds / Leicester / Lincoln / Liverpool / Llandrindod Wells / Llandudno / Luton / Manchester / Milton Keynes / Motherwell / Newcastle upon Tyne / Newport / North London / North West London / Northampton / Northern Ireland / Norwich / Nottingham / Oldham / Outer Hebrides / Oxford / Paisley / Perth / Peterborough / Plymouth / Portsmouth / Preston / Reading / Redhill / Rochester / Romford / Salisbury / Sheffield / Shrewsbury / Slough / South East London / South West London / Southall / Southampton / Southend-on-Sea / St Albans / Stevenage / Stockport / Stoke-on-Trent / Sunderland / Sutton / Swansea / Swindon / Taunton / Telford / Tonbridge / Torquay / Truro / Twickenham / Wakefield / Walsall / Warrington / Watford / West London / Wigan / Wolverhampton / Worcester / York	Population estimates sourced from the 2011 Census were used to identify the number of participants required from each postcode area to achieve a geographically representative sample of the UK.(Office for National Statistics)
Socioeconomic status	More deprived / average / less deprived	Townsend Deprivation Score tertiles were calculated across UK Output Areas and assigned to participants based on their postcode. These were used to determine socioeconomic status. Data were sourced from the UK Data Service based on the 2011 Census.
Age (years)	<30 / 30 – <40 / 40 – <50 / ≥50 (offspring) <60 / 60 – <70 / ≥70 (UK Biobank participants)	Defined as age on 1st January 2020

2

Sex	female / male	Specified by participant
Ethnicity	white / non-white	Participants were defined as white if they specified their ethnicity as 'white', 'white British', 'white Irish' or 'any other white background'
Urban-rural classification	urban / rural	The 2011 Rural-Urban Classification (used for English and Welsh addresses) defines urban Output Areas as built-up areas with a population of 10,000 or more. The Scottish Government Urban Rural Classification version 2016 (used for Scottish addresses) defines urban Output Areas as settlements with a population of 10,000 or more. (Department for Environment, 2013; Geographic Information Science & Analysis Team, 2018)
London-based	London-based / based outside London	Participants were defined as London-based if their address was located within any of the following postcode areas: Bromley, Central London, Croydon, Dartford, East London, Enfield, Harrow, Ilford, Kingston upon Thames, North London, North West London, South East London, South West London, Southall, Sutton, Twickenham, Watford, West London

Table S2: Number of participants that returned a blood sample each month, and whether they were positive for SARS-CoV-2 antibodies

Month of sample collection	Number of participants returning samples	Number of participants who tested positive	Number of new positives		
			White ethnicity	Non-white ethnicity	All*
May	1479	79	73	5	79
June	14254	956	758	197	956
July	17384	1229	303	44	347
August	16740	1134	20	4	24
September	16481	1131	45	8	53
October	16202	1247	86	15	101
November	12172	1067	115	21	136
December	56	13	2	1	3
Total	18887	1669	1402	295	1669

Analyses among all participants recruited into the UK Biobank SARS-CoV-2 Seroprevalence Study. *Includes participants of unknown ethnicity.

4

Table S3: Operational definitions of variables, and categories used in analyses

Variables	Operational definition	Categories using in main analyses
Ethnicity	Self-reported at recruitment	White (White British, White Irish, White, Any other white background); Black (African, Any other black background, Caribbean); South Asian (Indian, Pakistani, Asian or Asian British, Bangladeshi); Mixed (White and Asian, White and Black African, White and Black Caribbean, Mixed, Any other mixed background); Chinese (Chinese); and Other (Any other Asian background, Other ethnic group).
Age	Date of sample collection minus date of birth	< 55, 55-59, 60-64, 65-69, 70-74, ≥75 years
Sex	Self-reported at recruitment	Female / male
Urban-rural classification	The 2011 Rural-Urban Classification (used for English and Welsh addresses) defines urban Output Areas as built-up areas with a population of 10,000 or more. The Scottish Government Urban Rural Classification version 2016 (used for Scottish addresses) defines urban Output Areas as settlements with a population of 10,000 or more. (Department for Environment, 2013; Geographic Information Science & Analysis Team, 2018)	Urban / rural
UK region	Region of residence at recruitment	East Midlands, East of England, London, North East, North West, Scotland, South East, South West, Wales, West Midlands, Yorkshire and the Humber
Work location	Derived from questionnaire on social factors administered at the end of the study. Response to the question "What was your employment situation in time period 1 (i.e. mid-March to end of June 2020)?" Participants were classified as: (1) working from home if they selected the option "employed/self-employed and working from home"; (2) working at a work place if they selected either "employed /self-employed and working with others at a workplace" or "employed/self-employed and working alone at a workplace; or (3) Retired	Work at a workplace; work from home; retired, unemployed & others; missing

5

	and others if they selected any of the other options.	
Work in NHS or social care	Derived from questionnaire on social factors administered at the end of the study. Derived from responses to two questions: (1) "At any time between 01 March and 30 November 2020 were you designated a key worker?" and (2) "Which keyworker occupation group was this role in?" Participants were classified as working in NHS or social care if they indicated they were key workers and selected the option NHS and social care when asked about the keyworker occupation group. Those who indicated that they were not keyworkers and keyworkers in other occupations groups were considered as non-NHS/social care workers.	Yes, no, missing
Work in close proximity with others	Derived from questionnaire on social factors administered at the end of the study. Participants who indicated "Yes" in response to the question "At any time between 01 March and 30 November 2020 were you typically in close contact (defined as face-to-face within 1 metre for 10 minutes or more) with the public or co-workers as part of your job? , with options "Yes", "No" and "Prefer not to answer" (classified as 'missing in analyses).	Yes, no, missing
Number of co-residents aged 18 years or under	Derived from questionnaire on social factors administered at the end of the study. Created by adding together the number of household members (since 01 January 2020; excluding respondent) in the following age groups: 0-4 years, 5-11 years and 12-18 years. Participants who indicated they live alone were assigned the value "0".	0,1,2,≥3, missing
Number of co-residents aged > 18 years	Derived from questionnaire on social factors administered at the end of the study. Created by adding together the number of household members (since 01 January; excluding respondent) in the following age groups: 19-29 years, 30-39 years, 40-49 years, 50-59 years, 60-69 years, 70-79 years and 80+ years. Participants who indicated they live alone were assigned the value "0".	0,1,2,≥3, missing
Has a co-resident who works outside the home	Derived from questionnaire on social factors administered at the end of the study. Participants who indicated "Yes" in response to the question "Do you live with someone who worked outside the home?"	No (including 'no co-resident'), yes, missing
Public transport is usual mode of transport	Derived from questionnaire on social factors administered at the end of the study. Participants who selected any of the following options when asked about their main mode of transportation when they were outside the	Yes, no, missing

6

Supplemental material

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

J Epidemiol Community Health

	home in time period 1 (i.e. mid-March to end of June 2020): bus, taxi/uber, train, tube or personal vehicle – with non-household member.	
Frequency of shopping	Derived from questionnaire on social factors administered at the end of the study. Derived from response to the question: "In time period 1 ((i.e. mid-March to end of June 2020), how many days per week did you typically leave your home to go shopping?"	Never/ Less than once a week; 1 to 2 days per week; 3 to 4 days per week; missing
Deprivation	Derived from participant's postcode at recruitment. Deprivation defined by Townsend deprivation score for postcode; more deprived had score ≥ 2.	More deprived; average deprivation; less deprived
Highest level of Education	Derived from questionnaire on social factors administered at the end of the study. Derived from response to the question: "Which of the following qualifications do you have?"	University or college; A-levels; O-levels or CSE; others; none; missing

7

Table S4: Demographic characteristics of the participants of the UK Biobank Seroprevalence Study and, for comparison, the 2021 UK Population

Characteristic	UK Biobank Serology Study – all recruited participants (n=20,203)	UK Biobank Serology Study – included in the main analyses (n=14,571)	2021 UK Population (excluding N. Ireland)*
Age (years)			
<30	10.3	7.4	18.8
30-39	15.9	12.7	17.2
40-49	12.8	11.9	16.0
50-59	18.8	20.8	17.3
60-69	20.1	23.3	13.6
≥70	22.0	23.8	17.2
Sex			
Female	56.0	57.8	51.6
Male	44.0	42.1	48.4
Ethnicity			
Black	2.5	2.5	3.7
Chinese	0.7	0.9	0.8
South Asian	2.8	2.7	7.8
Mixed	3.7	3.6	1.9
White	87.3	87.3	83.9
Other	2.7	2.9	2.0
Region			
East of England	4.7	4.8	9.7
South West	6.9	6.9	8.8
London	29.8	29.5	13.5
East Midlands	6.0	6.1	7.5
South East	12.9	13.3	14.3
Scotland	6.2	6.4	8.4
West Midlands	6.9	6.8	9.2
North West	10.8	10.5	11.4
Yorkshire and The Humber	7.8	7.5	8.4
North East	4.1	4.1	4.0
Wales	3.9	4.0	4.8

*Participants were not recruited from Northern Ireland into UK Biobank. Data: Office for National Statistics. 2021 Census. https://census.gov.uk/census-2021-results Accessed 10-06-2023

8

Table S5: Participants characteristics and their social circumstances during the first wave of the pandemic, by ethnicity

	All (n = 14571)	White (n = 12725)	Black (n = 369)	South Asian (n = 396)	Chinese (n = 126)	Mixed (n = 526)	Other (n = 429)
Age (years), mean (SD)	55.6 (15.6)	55 (16.1)	60.9 (7.3)	61.8 (9.2)	62.5 (8.0)	54.0 (14.0)	63.5 (8.8)
Age (years), n (%)							
< 55	6038 (41.4)	5603 (44)	65 (17.6)	74 (18.7)	16 (12.7)	213 (40.5)	67 (15.6)
55-59	1671 (11.5)	1273 (10)	105 (28.5)	87 (22)	28 (22.2)	111 (21.1)	67 (15.6)
60-64	1599 (11)	1232 (9.7)	93 (25.2)	73 (18.4)	31 (24.6)	83 (15.8)	87 (20.3)
65-69	1790 (12.3)	1486 (11.7)	54 (14.6)	83 (21)	28 (22.2)	49 (9.3)	90 (21)
70-74	2211 (15.2)	1980 (15.6)	36 (9.8)	51 (12.9)	14 (11.1)	50 (9.5)	80 (18.6)
75+	1262 (8.7)	1151 (9)	16 (4.3)	28 (7.1)	9 (7.1)	20 (3.8)	38 (8.9)
Sex, n (%)							
Female	8431 (57.9)	7298 (57.4)	224 (60.7)	189 (47.7)	95 (75.4)	366 (69.6)	259 (60.4)
Male	6140 (42.1)	5427 (42.6)	145 (39.3)	207 (52.3)	31 (24.6)	160 (30.4)	170 (39.6)
Urban-rural classification, n (%)							
Rural	2052 (14.1)	1923 (15.1)	14 (3.8)	25 (6.3)	6 (4.8)	60 (11.4)	24 (5.6)
Urban	12519 (85.9)	10802 (84.9)	355 (96.2)	371 (93.7)	120 (95.2)	466 (88.6)	405 (94.4)
UK Regions							
East Midlands	894 (6.1)	776 (6.1)	22 (6)	29 (7.3)	7 (5.6)	38 (7.2)	22 (5.1)
East of England	703 (4.8)	671 (5.3)	7 (1.9)	4 (1)	4 (3.2)	9 (1.7)	8 (1.9)
London	4292 (29.5)	3556 (27.9)	161 (43.6)	163 (41.2)	48 (38.1)	165 (31.4)	199 (46.4)
North East	596 (4.1)	528 (4.1)	10 (2.7)	16 (4)	5 (4)	24 (4.6)	13 (3)
North West	1530 (10.5)	1358 (10.7)	37 (10)	32 (8.1)	14 (11.1)	53 (10.1)	36 (8.4)
Scotland	937 (6.4)	888 (7)	3 (0.8)	7 (1.8)	5 (4)	18 (3.4)	16 (3.7)
South East	1937 (13.3)	1725 (13.6)	34 (9.2)	38 (9.6)	20 (15.9)	71 (13.5)	49 (11.4)
South West	1000 (6.9)	896 (7)	16 (4.3)	21 (5.3)	8 (6.3)	40 (7.6)	19 (4.4)
Wales	587 (4)	548 (4.3)	3 (0.8)	6 (1.5)	2 (1.6)	21 (4)	7 (1.6)
West Midlands	995 (6.8)	831 (6.5)	49 (13.3)	49 (12.4)	7 (5.6)	37 (7)	22 (5.1)
Yorkshire and the Humber	1100 (7.5)	948 (7.4)	27 (7.3)	31 (7.8)	6 (4.8)	50 (9.5)	38 (8.9)

Analyses among 14,571 participants; exclusions as in figure 1. *UK Biobank urban and rural area classification, based on postcode information that is matched to 2011 census information on population density.

9

Table S5 continued:

	All (n = 14571)	White (n = 12725)	Black (n = 369)	South Asian (n = 396)	Chinese (n = 126)	Mixed (n = 526)	Other (n = 429)
Work location							
Work at a workplace	3108 (21.3)	2635 (20.7)	134 (36.3)	109 (27.5)	21 (16.7)	122 (23.2)	87 (20.3)
Work from home	4675 (32.1)	4122 (32.4)	85 (23)	106 (26.8)	36 (28.6)	219 (41.6)	107 (24.9)
Retired, Unemployed & Others	6711 (46.1)	5906 (46.4)	146 (39.6)	176 (44.4)	69 (54.8)	180 (34.2)	234 (54.5)
Missing	77 (0.5)	62 (0.5)	4 (1.1)	5 (1.3)	-	5 (1)	1 (0.2)
Work in NHS or social care, n (%)							
No	12350 (84.8)	10863 (85.4)	270 (73.2)	318 (80.3)	108 (85.7)	436 (82.9)	355 (82.8)
Yes	1235 (8.5)	1038 (8.2)	61 (16.5)	39 (9.8)	7 (5.6)	57 (10.8)	33 (7.7)
Missing	986 (6.8)	824 (6.5)	38 (10.3)	39 (9.8)	11 (8.7)	33 (6.3)	41 (9.6)
Work in close proximity with others, n (%)							
No	9760 (67.0)	8645 (67.9)	189 (51.2)	232 (58.6)	88 (69.8)	343 (65.2)	263 (61.3)
Yes	3372 (23.1)	2841 (22.3)	142 (38.5)	117 (29.5)	25 (19.8)	140 (26.6)	107 (24.9)
Missing	1439 (9.9)	1239 (9.7)	38 (10.3)	47 (11.9)	13 (10.3)	43 (8.2)	59 (13.8)
Number of co-residents aged ≤18 years							
0	11051 (75.8)	9590 (75.4)	281 (76.2)	312 (78.8)	112 (88.9)	391 (74.3)	365 (85.1)
1	1440 (9.9)	1242 (9.8)	46 (12.5)	45 (11.4)	10 (7.9)	62 (11.8)	35 (8.2)
2	1594 (10.9)	1470 (11.6)	22 (6)	25 (6.3)	2 (1.6)	60 (11.4)	15 (3.5)
3+	418 (2.9)	370 (2.9)	15 (4.1)	8 (2)	2 (1.6)	12 (2.3)	11 (2.6)
Missing	68 (0.5)	53 (0.4)	5 (1.4)	6 (1.5)	-	1 (0.2)	3 (0.7)
Number of co-residents aged >18 years							
0	3116 (21.4)	2679 (21.1)	116 (31.4)	65 (16.4)	38 (30.2)	120 (22.8)	98 (22.8)
1	8219 (56.4)	7401 (58.2)	146 (39.6)	146 (36.9)	53 (42.1)	276 (52.5)	197 (45.9)
2	1961 (13.5)	1632 (12.8)	58 (15.7)	97 (24.5)	16 (12.7)	83 (15.8)	75 (17.5)
3+	1215 (8.3)	967 (7.6)	44 (11.9)	82 (20.7)	19 (15.1)	46 (8.7)	57 (13.3)
Missing	60 (0.4)	46 (0.4)	5 (1.4)	6 (1.5)	-	1 (0.2)	2 (0.5)

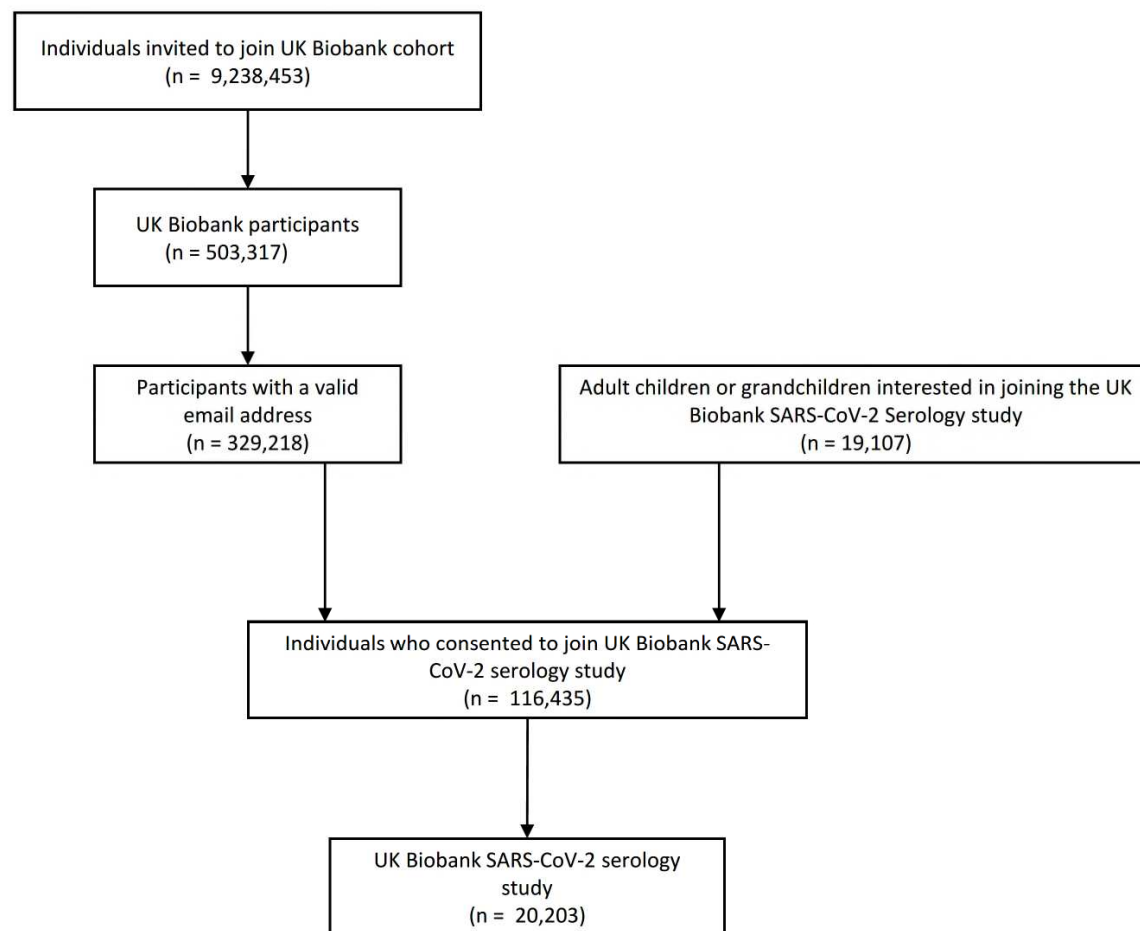Analyses among 14,571 participants; exclusions as in figure 1.

10

Table S5 continued:

	All (n = 14571)	White (n = 12725)	Black (n = 369)	South Asian (n = 396)	Chinese (n = 126)	Mixed (n = 526)	Other (n = 429)
Has a co-resident who works outside the home							
No / No co-resident	10069 (69.1)	8860 (69.6)	230 (62.3)	246 (62.1)	95 (75.4)	354 (67.3)	284 (66.2)
Yes	4430 (30.4)	3807 (29.9)	134 (36.3)	144 (36.4)	31 (24.6)	171 (32.5)	143 (33.3)
Missing	72 (0.5)	58 (0.5)	5 (1.4)	6 (1.5)	-	1 (0.2)	2 (0.5)
Public transport is usual mode of transport							
No	11242 (77.2)	10062 (79.1)	211 (57.2)	241 (60.9)	68 (54)	392 (74.5)	268 (62.5)
Yes	474 (3.3)	377 (3)	38 (10.3)	9 (2.3)	5 (4)	25 (4.8)	20 (4.7)
Missing	2855 (19.6)	2286 (18)	120 (32.5)	146 (36.9)	53 (42.1)	109 (20.7)	141 (32.9)
Frequency of shopping							
Never/ Less than once a week	3705 (25.4)	3309 (26)	75 (20.3)	91 (23)	19 (15.1)	119 (22.6)	92 (21.4)
1 to 2 days per week	6122 (42)	5455 (42.9)	134 (36.3)	115 (29)	40 (31.7)	232 (44.1)	146 (34)
3 to 4 days per week	1159 (8)	1036 (8.1)	26 (7)	23 (5.8)	9 (7.1)	40 (7.6)	25 (5.8)
Missing	3585 (24.6)	2925 (23)	134 (36.3)	167 (42.2)	58 (46)	135 (25.7)	166 (38.7)
Deprivation level, n (%)†							
More deprived	3415 (23.4)	2890 (22.7)	149 (40.4)	83 (21)	24 (19)	133 (25.3)	136 (31.7)
Average	5969 (41)	5267 (41.4)	139 (37.7)	161 (40.7)	49 (38.9)	194 (36.9)	159 (37.1)
Less deprived	5187 (35.6)	4568 (35.9)	81 (22)	152 (38.4)	53 (42.1)	199 (37.8)	134 (31.2)
Highest level of education							
University or College	9131 (62.7)	8020 (63)	193 (52.3)	246 (62.1)	74 (58.7)	330 (62.7)	268 (62.5)
A-level	1602 (11)	1439 (11.3)	23 (6.2)	29 (7.3)	13 (10.3)	65 (12.4)	33 (7.7)
O-level or CSE	1850 (12.7)	1630 (12.8)	62 (16.8)	44 (11.1)	14 (11.1)	63 (12)	37 (8.6)
Others	1145 (7.9)	924 (7.3)	67 (18.2)	48 (12.1)	11 (8.7)	42 (8)	53 (12.4)
None	697 (4.8)	593 (4.7)	17 (4.6)	24 (6.1)	12 (9.5)	19 (3.6)	32 (7.5)
Missing	146 (1)	119 (0.9)	7 (1.9)	5 (1.3)	2 (1.6)	7 (1.3)	6 (1.4)

Analyses among 14,571 participants; exclusions as in figure 1. †Deprivation defined by Townsend deprivation score for postcode; more deprived had score ≥ 2.

11

Supplemental material

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

J Epidemiol Community Health

Table S6: Odds ratios for SARS-CoV-2 seropositivity by ethnicity, progressively adjusted for potential explanatory factors (among participants who returned a sample in May, June or July 2020*)

	Odds ratio (95%CI)						Percent reduction in LR statistics for ethnicity
	White	Black	South Asian	Chinese	Mixed	Other	
Baseline model (i.e. ethnicity + age + sex)	1.00 (0.92-1.08)	2.87 (2.16-3.80)	1.67 (1.20-2.32)	0.94 (0.44-2.02)	1.42 (1.06-1.9)	1.84 (1.35-2.51)	
+ geographic factors	1.00 (0.92-1.09)	2.37 (1.78-3.15)	1.42 (1.02-1.98)	0.84 (0.39-1.81)	1.35 (1.01-1.81)	1.56 (1.14-2.13)	38
+ occupational factors	1.00 (0.92-1.09)	2.09 (1.57-2.79)	1.31 (0.94-1.83)	0.88 (0.41-1.90)	1.34 (1.00-1.79)	1.48 (1.08-2.03)	54
+ household factors	1.00 (0.92-1.09)	2.08 (1.55-2.78)	1.22 (0.87-1.72)	0.87 (0.41-1.89)	1.35 (1.00-1.80)	1.42 (1.04-1.95)	57
+ shopping frequency & use of public transport	1.00 (0.91-1.09)	2.02 (1.51-2.71)	1.20 (0.85-1.68)	0.84 (0.39-1.81)	1.33 (0.99-1.78)	1.39 (1.01-1.90)	61
+ deprivation & education	1.00 (0.91-1.10)	1.98 (1.48-2.66)	1.20 (0.85-1.68)	0.84 (0.39-1.81)	1.32 (0.99-1.77)	1.37 (1.00-1.89)	64

Odds ratios relative to participants of White ethnicity, adjusted for age and sex (baseline model), with progressive adjustment for social factors where indicated; group-specific variances used to calculate 95% CIs. *Seropositivity for SARS-CoV-2 defined as positive antibody test in any month. See table S3 for further details on characteristics. LR= likelihood-ratio. LR statistics for ethnicity in the baseline model = 59.8.

Figure S1: Flowchart of participant recruitment

REFERENCES

Department for Environment, F. R. A. (2013). *2011 Rural Urban Classification*.

Geographic Information Science & Analysis Team. (2018). *Scottish Government Urban Rural Classification 2016*. The Scottish Government

Office for National Statistics. *2011 Census*. Retrieved 14-02-2022 from