UK Biobank SARS-CoV-2 Serology Study

15 Jan 2021

Table of Contents

1. Executive Summary

UK Biobank collected blood samples from approximately 20,200 individuals on a monthly basis for six months to determine the extent of past infection with SARS-CoV-2 in different population subgroups across the United Kingdom.

Between 27th May and 4th December 2020, 18,893 individuals (93.5%) provided at least one sample that was successfully assayed at the Target Discovery Institute (University of Oxford). A total of 1,699 (9%) individuals had at least one sample that was seropositive for SARS-CoV-2 anti-spike antibodies during this six-month study period. There were significant differences in seroprevalence by age, ethnicity, geographic region and socioeconomic status (see Section 3).

A total of 1,264 individuals were seropositive in month one. During the study period (end-May to early-December 2020), 453 participants were newly identified as having a positive test (i.e., had a positive test result following prior negative test results); most of these positive tests occurred from September onwards, coinciding with the start of the second wave (see Section 6).

The duration of SARS-CoV-2 sero-positivity was assessed among participants who were seropositive in month one and had reported symptoms or had a positive PCR test (ascertained through linkage to testing data). Of these 705 participants, 86 (12.2%) sero-reverted (i.e. had a subsequent negative antibody test) within six months of symptom onset. Only five participants (<1%) became negative within three months. Hence, 87.8% of participants remained seropositive for at least six months after infection (see Section 6).

The most common symptom associated with sero-positivity was a loss of sense of taste/smell (OR = 29.8, 95% CI = 22.8 – 38.9), followed by fever (OR = 8.2, 95% CI, 6.7 – 10.0) and chills (OR = 6.4, 95% CI, 5.4 – 7.7). However, all symptoms were associated with sero-positivity to a greater or lesser extent. Having at least one of the three classic COVID-19 symptoms (fever, persistent dry cough or loss of sense of taste or smell) was associated was a 12-fold higher risk of being seropositive (OR = 12.1, 95% CI, 10.2 – 14.3). Overall, 24% of seropositive participants were completely asymptomatic (compared with 54% of those who were seronegative) and 40% did not have one of the three 'classic' COVID-19 symptoms (see section 7).

2. Study population

An overview of the participant characteristics is shown in Table 1.

Table 1 Participant characteristics for 18,893 participants who had provided at least one sample

Characteristics		N	%
Gender	Men	8,220	43.5
	Women	10,673	56.5
Age Group (years)	<30	2,050	10.9
	30-39	2,898	15.3
	40-49	2,343	12.4
	50-59	3,722	19.7
	60-69	3,882	20.5
	70+	3,998	21.2
Ethnicity*	White	16,479	87.4
	Black	466	2.5
	Chinese	146	0.8
	Mixed	708	3.8
	Other	523	2.8
	South Asian	532	2.8
Region	East Midlands	1,133	6.0
	East of England	889	4.7
	London	5,633	29.8
	North East	780	4.1
	North West	2,037	10.8
	Scotland	1,168	6.2
	South East	2,441	12.9
	South West	1,302	6.9
	Wales	742	3.9
	West Midlands	1,301	6.9
	Yorkshire	1,467	7.8
Location of Residence	Rural	2,596	13.7
	Urban	16,297	86.3
Townsend Deprivation Index	Less Deprived	6,577	34.8
	Average	7,730	40.9
	More Deprived	4,586	24.3

*Ethnicity counts exclude 39 individuals with missing information.

3. Seroprevalence of SARS-CoV-2 infection

Of the 18,893 participants who provided at least one sample between May and December 2020, 9% (*N*=1,699) had at least one sample that was seropositive for SARS-CoV-2 (i.e. they had at least one positive test result during the study period); 91% (*N*=17,194) were seronegative (i.e. all samples submitted were seronegative).

The seroprevalence of SARS-CoV-2 infection was 6.6% at the start of the study period (end-May to June), increasing to 8.8% by the end of November/early December (Figure 1).

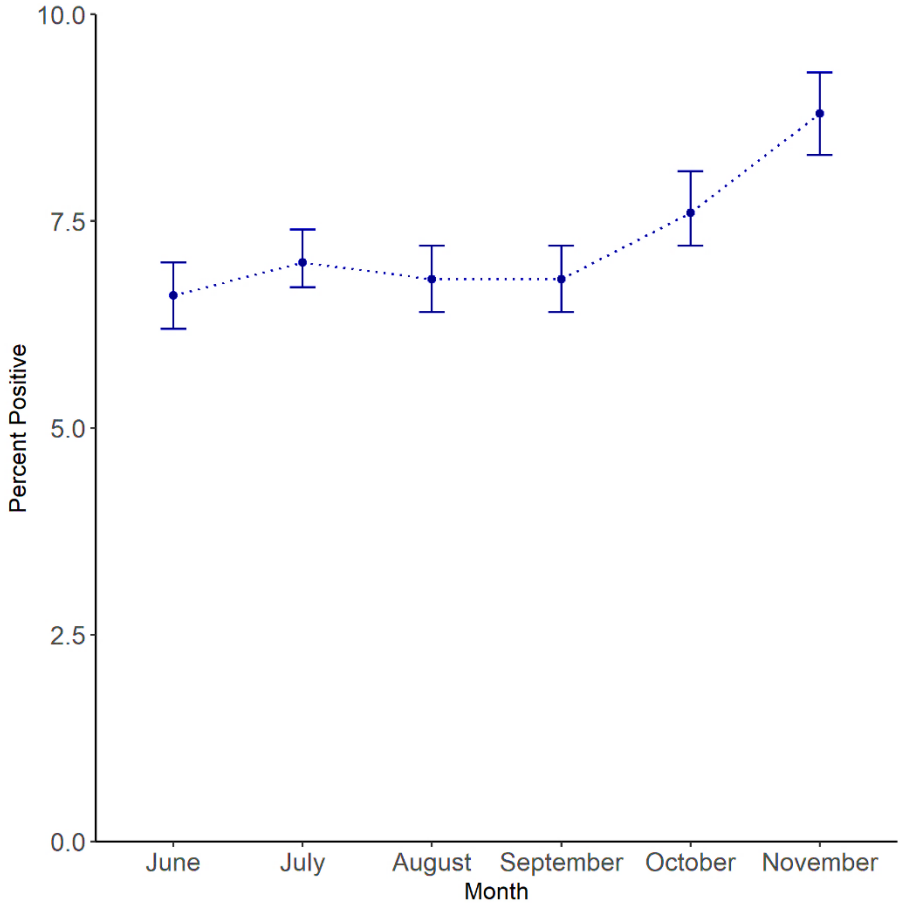

Figure 1 Seroprevalence of SARS-CoV-2 across the study months.
* Prevalence estimate with 95% confidence intervals.
**The last week of May is combined with June. The first week of December is combined with November.
***Where individuals have multiple samples provided per calendar month only one sample is included in the monthly prevalence estimates.

3.1 Gender

There was no evidence of a difference in seroprevalence by gender ($P_{\text{test for heterogeneity}}$= 0.32) with 9.2% of women and 8.8% of men being seropositive for SARS-CoV-2 infection (Figure 2).

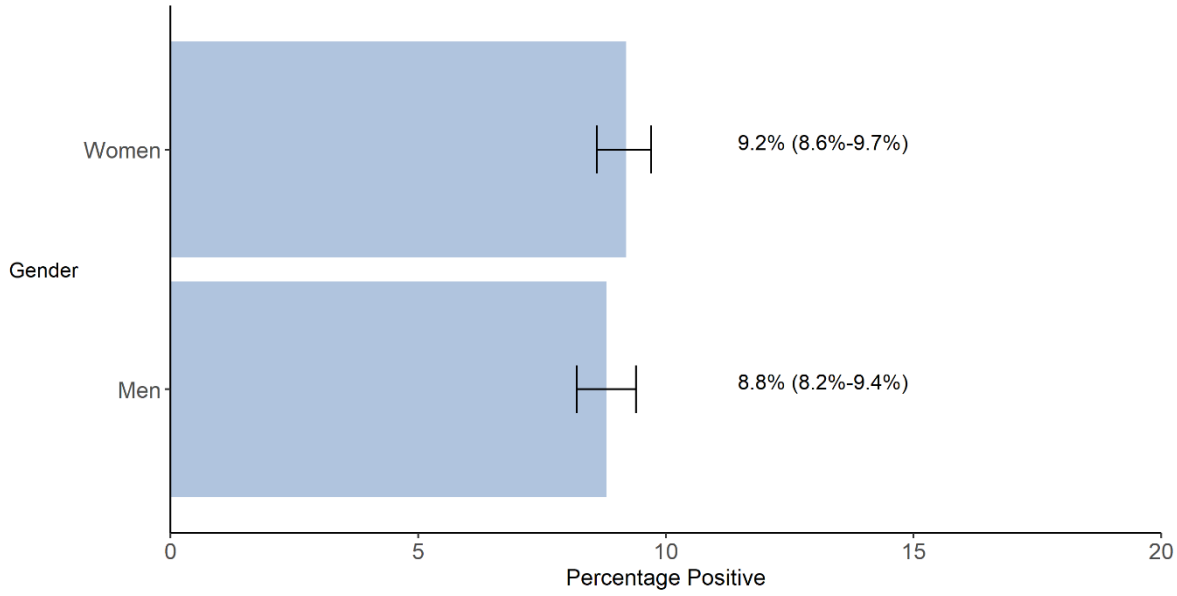

Figure 2 Percentage* of individuals ever seropositive for SARS-CoV-2 infection by gender.
* The black lines indicate the 95% confidence intervals.

3.2 Age

Seroprevalence of SARS-CoV-2 infection differed by age ($P_{\text{test for heterogeneity}}$<0.001), ranging from 13.5% among those aged <30 years to 6.7% among those aged 70+ years (Figure 3a). The seroprevalence of SARS-CoV-2 infection was consistently lower among those over 50 years of age across all the study months (Figure 3b).

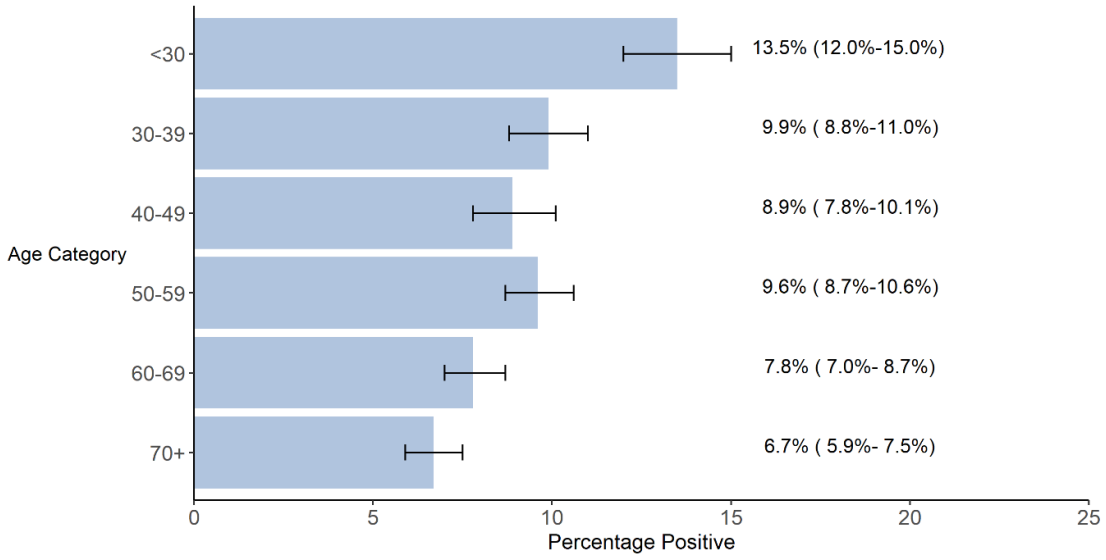

Figure 3a Percentage* of individuals ever seropositive for SARS-CoV-2 infection by age.

* The black lines indicate the 95% confidence intervals.

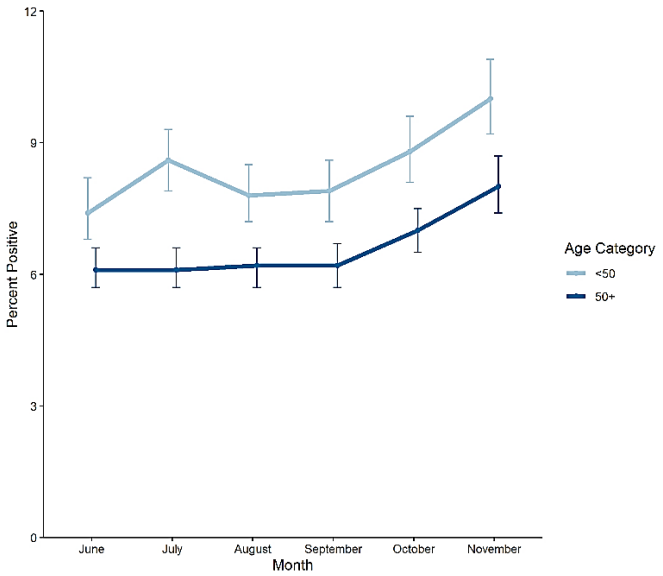

Figure 3b Percentage* of individuals seropositive for SARS-CoV-2 infection by age per study month.

* Prevalence estimate with 95% confidence intervals.

**The last week of May is combined with June. The first week of December is combined with November.

***Where individuals have multiple samples provided per calendar month only one sample is included in the monthly prevalence estimates.

3.3 Ethnicity

Seroprevalence of SARS-CoV-2 differed by ethnicity ($P_{\text{test for heterogeneity}}$<0.001), being highest among those of Black ethnicity (16.3%) and lowest among those of White (8.5%) and Chinese ethnicities (7.5%; Figure 4a). The seroprevalence of SARS-CoV-2 infection was consistently lower among those of White ethnicity across all the study months (Figure 4b).

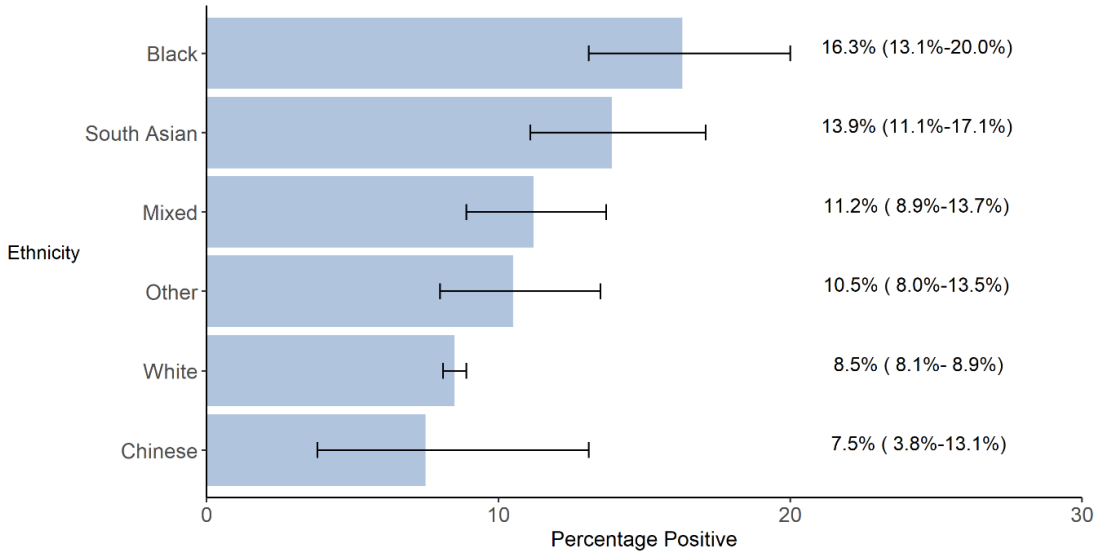

Figure 4a Percentage* of individuals ever seropositive for SARS-CoV-2 infection by ethnic group.

* The black lines indicate the 95% confidence intervals.

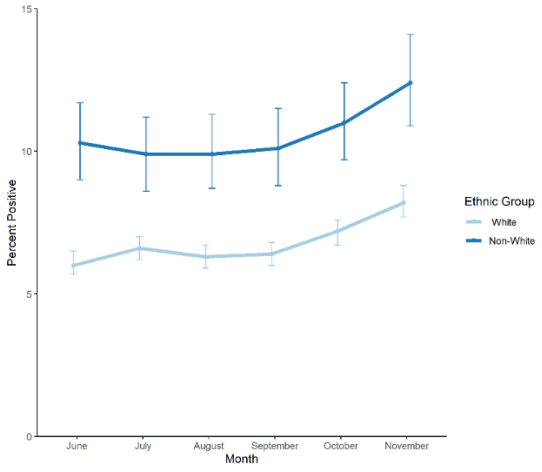

Figure 4b Percentage* of individuals seropositive for SARS-CoV-2 infection by ethnic group per study month.

* Prevalence estimate with 95% confidence intervals.

**The last week of May is combined with June. The first week of December is combined with November.

***Where individuals have multiple samples provided per calendar month only one sample is included in the monthly prevalence estimates.

3.4 Region

Seroprevalence of SARS-CoV-2 infection varied by region ($P_{\text{test for heterogeneity}}<0.001$), being highest in London (12.4%) and lowest in Scotland (5.5%; Figure 5a and 5b).

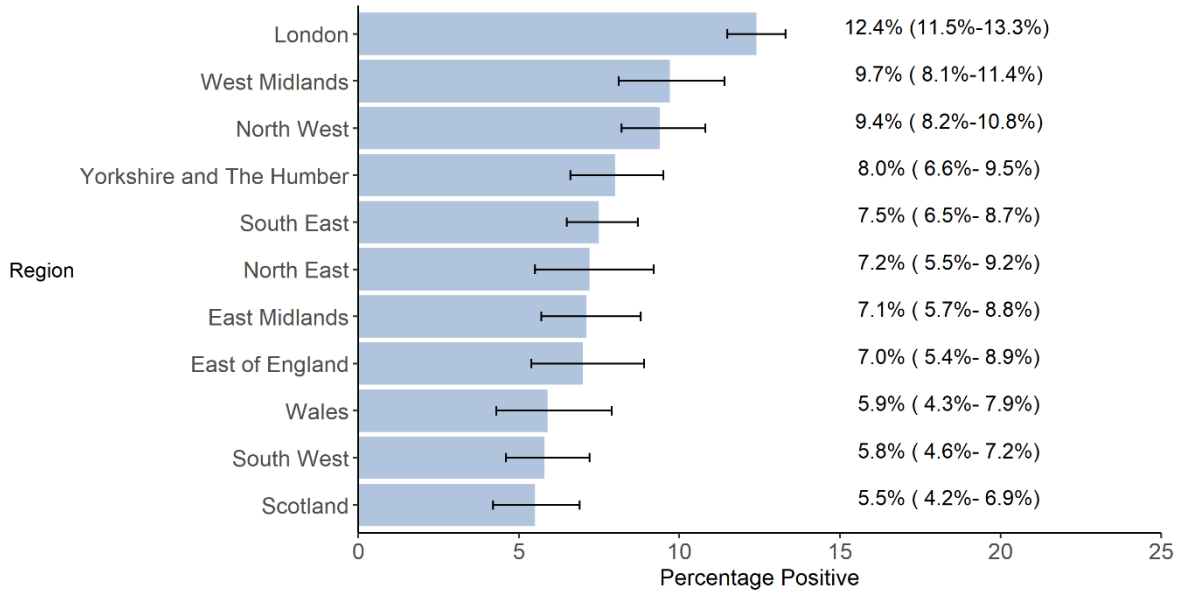

Figure 5a Percentage* of individuals ever seropositive for SARS-CoV-2 infection by region.

* The black lines indicate the 95% confidence intervals.

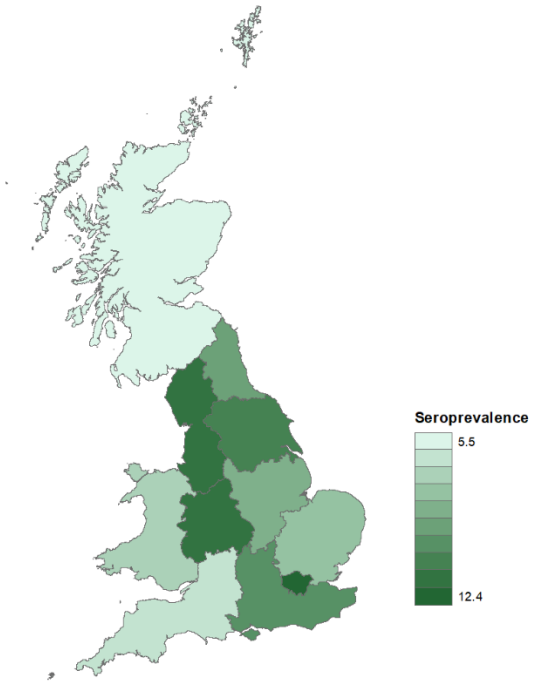

Figure 5b Map of the percentage of individuals ever seropositive for SARS-CoV-2 infection by region.

3.5 Rural/Urban

Seroprevalence of SARS-CoV-2 infection varied by urban-rural location of residence ($P_{\text{test for heterogeneity}}$<0.001), being higher in those living in urban (9.4%) compared to rural areas (6.2%; Figure 6).

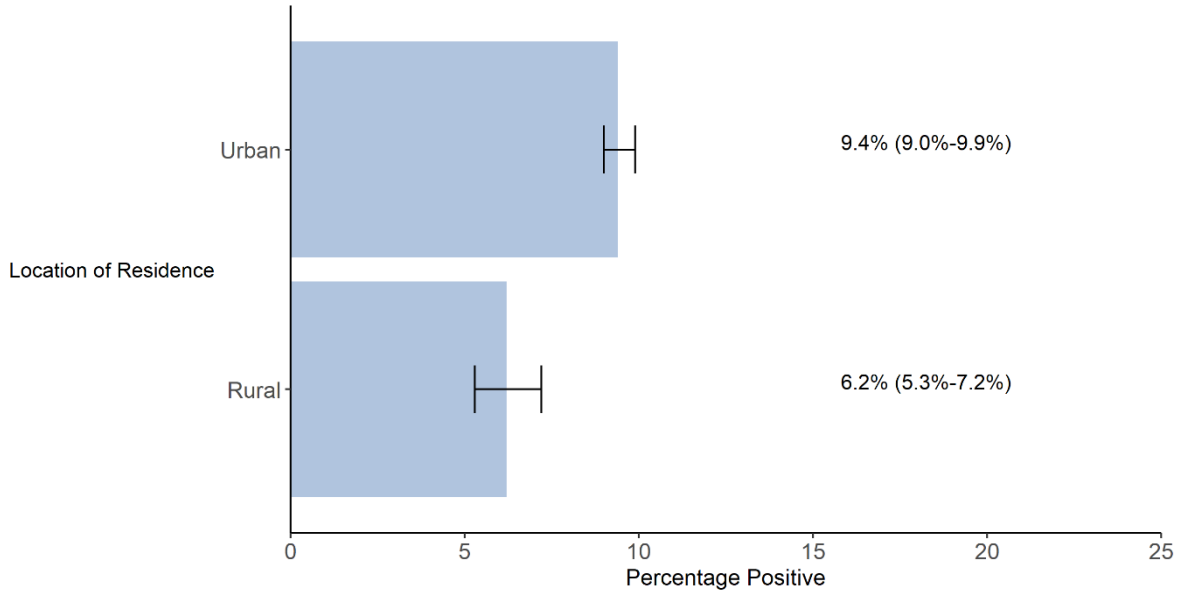

Figure 6 Percentage* of individuals ever seropositive for SARS-CoV-2 infection by location of residence.

* The black lines indicate the 95% confidence intervals.

3.6 Townsend Deprivation Index

Seroprevalence of SARS-CoV-2 infection varied significantly by Townsend Deprivation Index ($P_{\text{test for heterogeneity}}<0.001$), being 11.4% in areas of higher socio-economic deprivation, compared with 7.8% in areas of lower socio-economic deprivation (Figure 7).

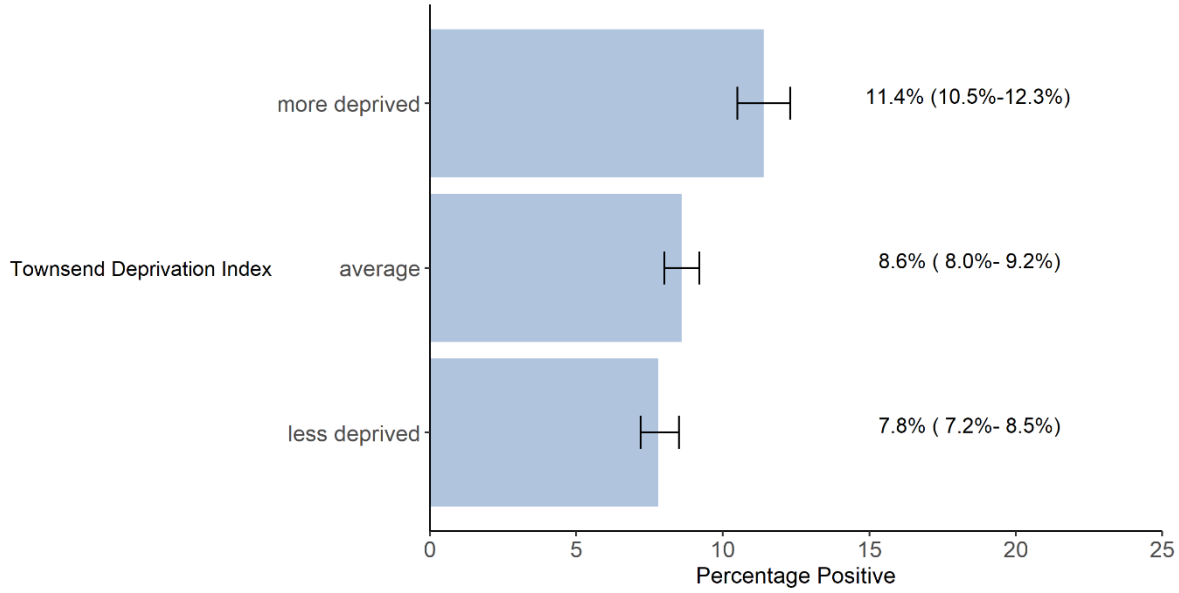

Figure 7 Percentage* of individuals ever seropositive for SARS-CoV-2 infection by Townsend Deprivation Index**.

* The black lines indicate the 95% confidence intervals.

**Townsend Deprivation Index categories are defined as: <-2 (less deprived), -2 to <2 (average), 2+ (more deprived)

4. Seroprevalence of SARS-CoV-2 infection in London

Samples from participants who live in London accounted for 29.8% of all samples analysed (*N*=5,633).

4.1 Seroprevalence across boroughs (London only)

Seroprevalence estimates differed across London, being highest in East London (14.7%) and lowest in West London (10.5%; $P_{\text{test for heterogeneity}}$=0.001; Figures 8a and 8b), although even in West London, the seroprevalence estimate was higher than for the UK as a whole.

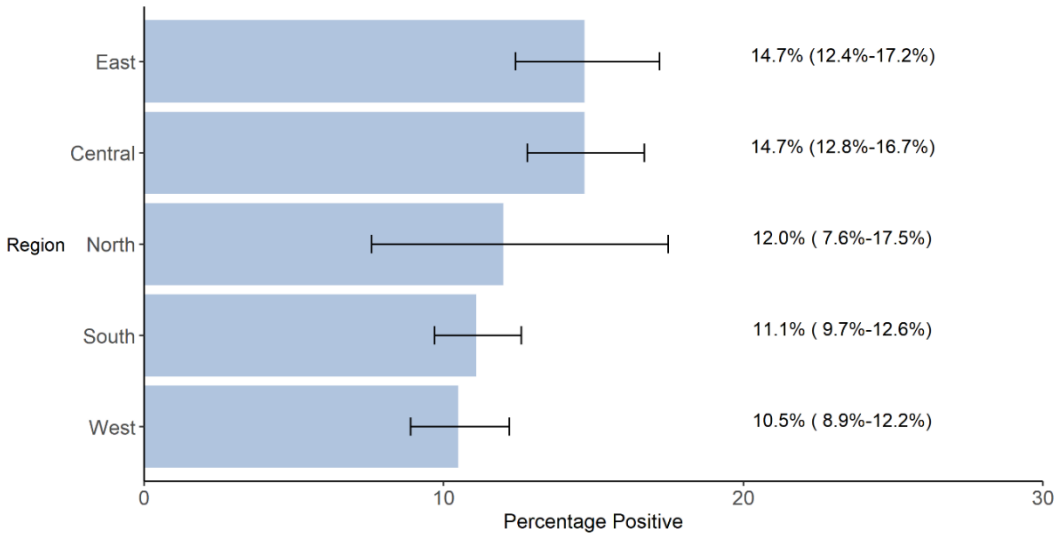

Figure 8a Percentage* of individuals ever seropositive for SARS-CoV-2 infection by region of London. * The black lines indicate the 95% confidence intervals.

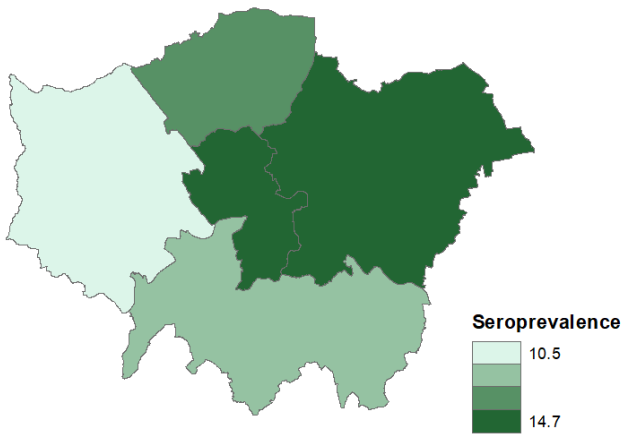

Figure 8b Map of percentage of individuals ever seropositive for SARS-CoV-2 infection by region of London

5. Ethnicity and SARS-CoV-2 seropositivity

Individuals belonging to Black, Asian and minority ethnic (BAME) groups had a greater likelihood of being seropositive for SARS-CoV-2. Seropositivity was highest among Black individuals (16.3%) and lowest among White (8.5%) and Chinese (7.5%) individuals.

We have investigated whether these differences are explained by known socio-demographic factors using logistic regression. Table 2 shows the odds of being seropositive in different ethnic groups compared to participants of White ethnicity. Those with a Black ethnic background were approximately twice as likely to be seropositive compared to their White counterparts (OR of 2.09, 95% CI: 1.62-2.68). After adjusting for age, sex, socio-economic deprivation, region and urban/rural status, the risk associated with being seropositive among Black ethnic minority participants was reduced but remained significantly higher than that of White participants (OR: 1.96; 95% CI: 1.49-2.53). Individuals of South Asian ethnicity were all also more likely to be seropositive compared to their White counterparts following adjustment for sociodemographic factors (Table 2).

Table 2 Association between ethnic group and SARS-CoV-2 infection status before and after adjustment for sociodemographic factors.

Ethnicity	**Unadjusted**		**Adjusted for age, sex, Townsend Deprivation Index, region and urban/rural status**	
	OR (95% CI)	**X^2**	**OR (95% CI)**	**X^2**
White	1.00 (ref)	49.21	1.00 (ref)	289.16
Black	2.10 (1.63-2.70)		1.95 (1.50-2.54)	
South Asian	1.74 (1.35-2.23)		1.73 (1.33-2.25)	
Chinese	0.88 (0.47-1.62)		0.92 (0.49-1.71)	
Mixed	1.35 (1.06-1.72)		1.26 (0.96-1.72)	
Other	1.26 (0.95-1.68)		1.29 (0.96-1.72)	

Model comparison: X^2=239.95 p<0.001
Abbreviations: OR (odds ratio), 95% CI (95% Confidence Interval)

6. New SARS-CoV-2 infections and duration of seroprevalence

6.1 New SARS-CoV-2 infections during the study period

From end-May to early-December 2020, there were 453 new positive tests for SARS-CoV-2 antibodies, with most of these occurring from the start of September, corresponding to the start of the second wave.

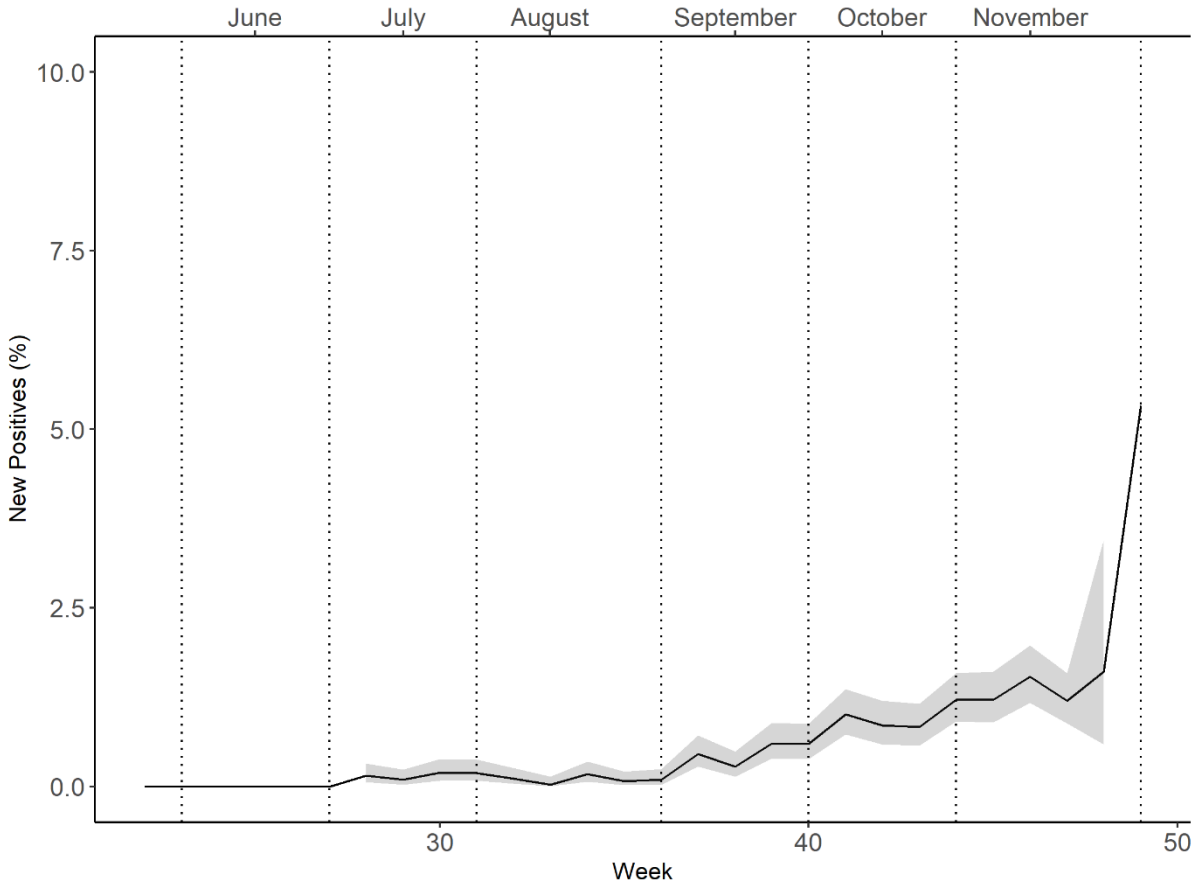

Figure 11 Percentage* of individuals newly seropositive for SARS-CoV-2 infection by week of study.

* The grey shading indicates the 95% confidence interval.

** The last week of the study has a smaller number of individuals than earlier weeks, resulting in a higher percentage of new positives with wider confidence intervals.

6.2 Duration of SARS-CoV-2 sero-positivity

The duration of the presence of detectable antibodies was investigated among participants who were seropositive in the first month of the study and who had provided at least one subsequent sample and who had reported a date of symptom onset and/or had a PCR test result, in order to estimate an approximate date of infection (N=705). Duration of sero-positivity was calculated as the time between the date of the onset of self-reported symptoms or a positive PCR test (whichever was first) until the date of the first negative test or the last available positive test. For participants that reported symptoms prior to 2020 (N=26), the duration of time was calculated from the 1st January 2020.

Of the 705 participants included in this analysis, 86 (12.2%) sero-reverted (i.e. had a subsequent negative antibody test) within six months of symptom onset. Only five participants (<1%) became negative within three months. Hence, 87.8% of participants remained seropositive for at least six months after infection (Figure 12).

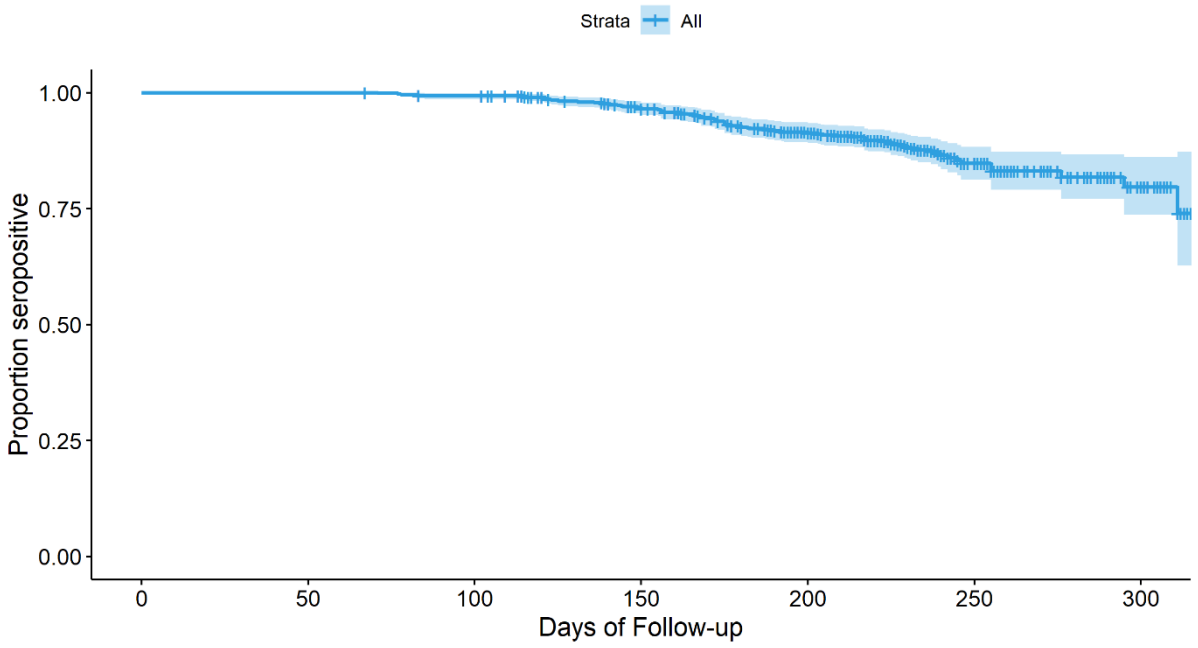

Figure 12 Proportion of seropositive participants remaining seropositive over time

7. Symptom survey analysis

All study participants were invited to complete a symptom survey at the time of blood collection. In month 1 (end May-June), participants were asked about symptoms that had occurred since the start of 2020, whereas subsequent surveys asked about symptoms occurring in the previous month.

Over the six months of data collection (end-May to early-December), 99,464 surveys had a corresponding assay result (13,968 in month 1, 20,461 in month 2, 17,043 in month 3, 18,458 in month 4, 17,719 in month 5 and 11,657 in month 6). A matched case-control dataset was generated whereby the survey for seropositive samples (cases) was matched to a random subset of surveys for seronegative samples (controls) by month, to allow for seasonal variation in symptoms. Three negative surveys were selected for each positive survey. Seropositive participants that reported no symptoms at time of seroconversion had their previous two surveys examined in case of a delay between infection and seroconversion. Surveys that included reported symptoms prior to 01 Jan 2020 were excluded from further analysis (n=1,094).

The matched case-control dataset included 6,648 surveys across the 6 months: 1,657 from seropositive cases (860 from month 1, 478 from month 2, 32 from month 3, 45 from month 4, 105 from month 5 and 137 from month 6) and 4,971 from seronegative controls (2,580 in month 2, 1,434 in month 2, 96 in month 3, 135 from month 4, 315 in month 5 and 411 in month 6).

7.1. Distribution of symptoms

The most common symptom associated with sero-positivity was a loss of sense of taste/smell (OR = 29.8 95% CI = 22.8 – 38.9), followed by fever (OR = 8.2, 95% CI, 6.7 – 10.0) and chills (OR = 6.4, 95% CI, 5.4 – 7.7). However, all symptoms were associated with sero-positivity to a greater or lesser extent. Having at least one of the three classic COVID-19 symptoms (fever, persistent dry cough or loss of sense of taste or smell) was associated was a 12-fold higher risk of being seropositive (OR = 12.1, 95% CI, 10.2 – 14.3; Table 1).

24% of seropositive participants were asymptomatic (compared with 54% of those who were seronegative); 40% did not have one of the three ‘classic’ COVID-19 symptoms.

Just under half of seropositive participants (47.3%) reported symptoms in March-April, with 7.8% reporting symptoms in January-February. In contrast, there was no obvious spike in symptoms among those who were seronegative (Fig 1).

No differences were observed in the association of loss of sense of taste/smell, fever or cough by sero-positivity status by age, sex or socioeconomic status (Fig 2). Similarly, there was no difference in the association of being asymptomatic by age, sex or socio-economic status (Fig. 3).

Table 3: Distribution and association of reported symptoms by serostatus

Symptom	Serostatus				Odds Ratio[1]	95% Confidence Interval
	Negative		Positive			
Loss of taste or smell	133	2.7%	708	42.7%	29.75	22.75 – 38.90
at least 1 of the three 'Classic' COVID-19 symptoms[2]	576	11.6%	973	58.7%	12.07	10.19 – 14.30
Fever	238	4.8%	469	28.3%	8.19	6.73 – 9.97
Chills	292	5.9%	480	29.0%	6.42	5.37 – 7.68
Lethargy	1,055	21.2%	927	55.9%	4.96	4.32 – 5.69
Muscle ache	614	12.4%	665	40.1%	4.76	4.12 – 5.51
Dry Cough	358	7.2%	438	26.4%	4.39	3.72 – 5.17
Shortness of Breath	382	7.7%	407	24.6%	3.84	3.25 – 4.52
Chest Pain	192	3.9%	207	12.5%	3.36	2.70 – 4.18
Headache	1,070	21.5%	737	44.5%	2.89	2.52 – 3.30
Nausea or Vomiting	248	5.0%	185	11.2%	2.27	1.83 – 2.81
Diarrhoea	486	9.8%	331	20.0%	2.23	1.89 – 2.63
Wheezing	243	4.9%	162	9.8%	1.98	1.59 – 2.46
Sore Throat	895	18.0%	497	30.0%	1.86	1.62 – 2.15
Abdominal Pain	278	5.6%	168	10.1%	1.85	1.49 – 2.29
Productive Cough	271	5.5%	154	9.3%	1.85	1.50 – 2.30
Runny Nose	874	17.6%	437	26.4%	1.61	1.40 – 1.86
Asymptomatic	2,681	53.9%	396	23.9%	0.25	0.22 – 0.29

[1]Adjusted for age, sex, ethnicity and townsend deprivation score

[2] Fever, persistent dry cough, loss of sense of taste or smell

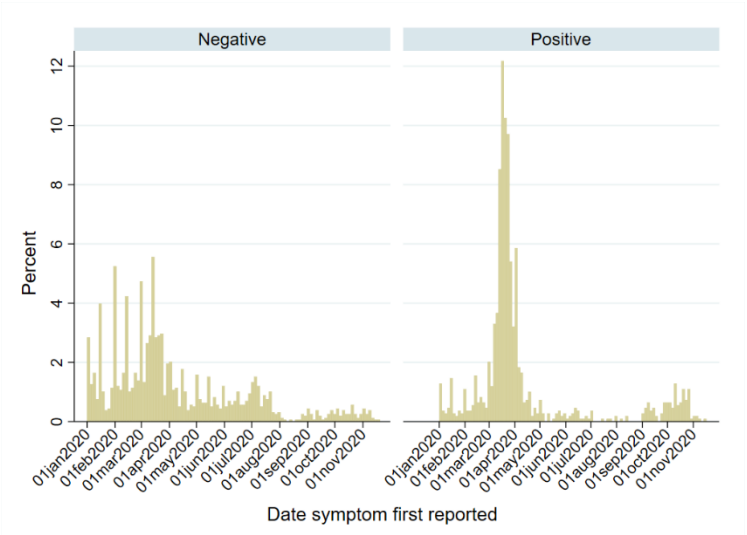

Figure 13: Distribution of date when symptoms first started across months 1-6 by serostatus

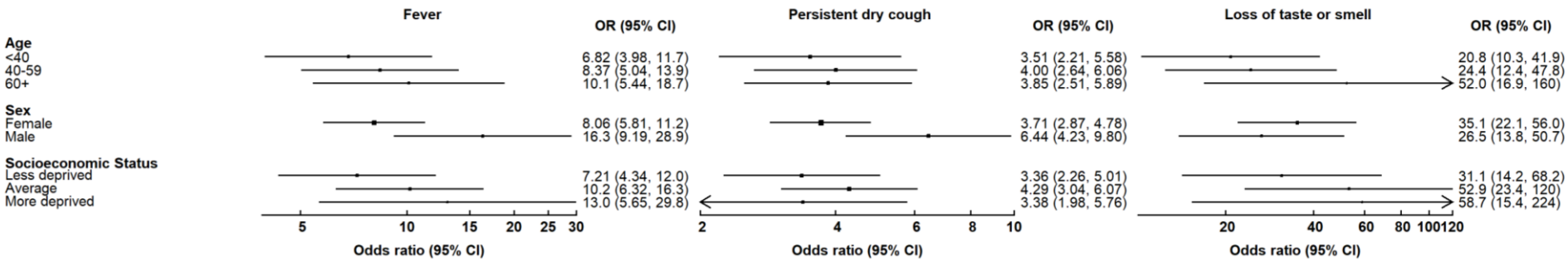

Figure 14: Odds[1] of the three 'classic' Covid-19 symptoms given serostatus, by age, sex and socioeconomic status.

1 – adjusted for age, sex, ethnicity and SES

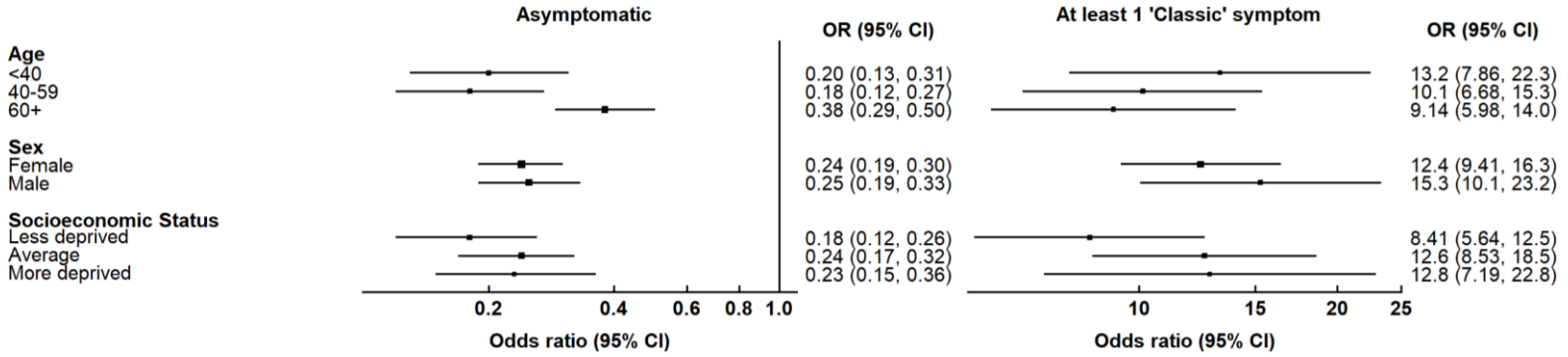

Figure 15: Odds[1] of being asymptomatic or having at least 1 of the 'classic' Covid-19 symptoms given serostatus, by age, sex and socioeconomic status

1 – adjusted for age, sex, ethnicity and SES

7.2 Sensitivity and Specificity of symptoms

Linkage to Public Health England's SARS-CoV-2 PCR testing result database has been undertaken for the full UK Biobank cohort, so these results are available for the 10,000 participants that are also UK Biobank participants. PCR test results were linked to surveys in the following manner:

- If a survey reported a date that symptoms started, any PCR test taken from 4 days before to 7 days after (to allow for recall bias on date of symptom starting) was linked to those symptoms.
- Any PCR test taken prior to the first survey was linked to the symptoms from the first survey. If multiple PCR test had been taken, the first PCR test was linked, unless there was a combination of positive and negative tests, in which case the first positive test was linked.
- Any PCR taken between two surveys was linked to the subsequent survey, as any symptoms would be reported in the subsequent survey.

This resulted in 795 surveys being associated with PCR test results, taken by 620 participants. Sensitivity ranged from 7.6% - 73.4% depending on how many symptoms were considered the cut-off point. Specificity ranged from 91.9% - 99.6% (see Table 4). The area under the Receiver Operator Curve was 0.83 (95% Confidence Interval 0.80 – 0.86; see Fig. 16).

Table 4: Sensitivity and specificity of number of symptoms[1] reported when compared to SARs-CoV-2 PCR test results.

		PCR Result		Sensitivity	Specificity
		Negative	Positive		
Number of symptoms[1]	1	44	31	73.42%	91.9%
	2	11	21	34.18%	98.04%
	3	3	6	7.59%	99.58%

1 – Symptoms included: Fever, persistent dry cough and loss of sense of taste or smell

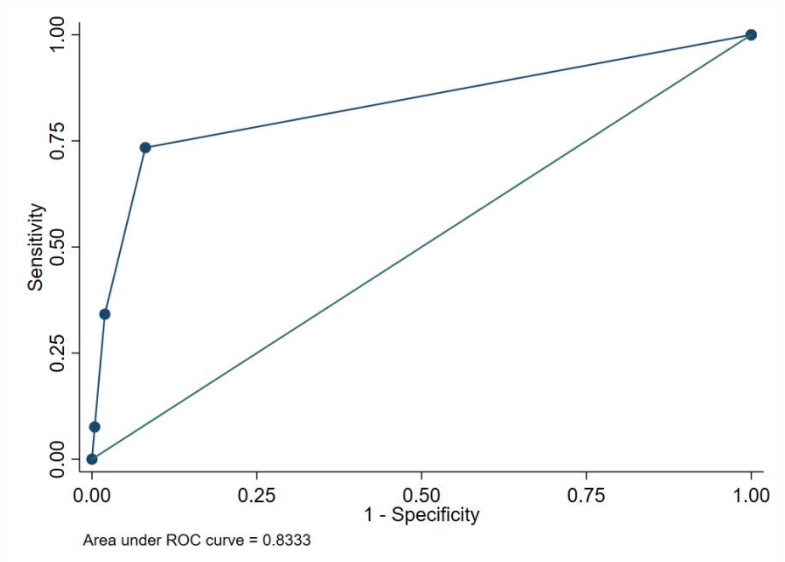

Figure 16: Receiver Operator Curve for sensitivity and specificity of number of symptoms reported[1] compared to SARs-CoV-2 PCR test results.

1 Symptoms included – Fever, persistent dry cough and loss of sense of taste or smell.