

Coronavirus (COVID-19) Infection Survey quality report: December 2022

Information on the Coronavirus (COVID-19) Infection Survey data collection method change from study worker home visit to remote data collection. This survey is being delivered in partnership with University of Oxford, University of Manchester, UK Health Security Agency (UKHSA) and Wellcome Trust, working with the University of Oxford and partner laboratories to collect and test samples.

Contact:
Eleanor Fordham and Elizabeth
Fuller
health.data@ons.gov.uk
+44 1633 560499

Release date:
30 March 2023

Next release:
To be announced

Table of contents

1. [Main points](#)
2. [Change in Coronavirus \(COVID-19\) Infection Survey data collection method](#)
3. [Likelihood of testing positive for COVID-19 by data collection method](#)
4. [Likelihood of strong positive COVID-19 cases reporting symptoms by data collection method](#)
5. [Representativeness of the Coronavirus \(COVID-19\) Infection Survey population sample by data collection method](#)
6. [Summary of findings](#)
7. [How the data are measured](#)
8. [Coronavirus \(COVID-19\) Infection Survey, Quality Report data](#)
9. [Collaboration](#)
10. [Glossary](#)
11. [Related links](#)
12. [Cite this article](#)

1 . Main points

Following on from our [September quality report](#), we have continued to assess the impact of the change to our data collection method from study worker home visits to remote data collection on our coronavirus (COVID-19) estimates.

We compared the estimated likelihood of testing positive for COVID-19 on a nose and throat swab, as well as the likelihood of people with a strong positive test reporting symptoms, by data collection method while adjusting for several variables. We also assessed the representativeness of our remote data collection and study worker population samples by comparing them with demographic data from the 2021 Census.

- In the first week of the period studied (11 to 31 July 2022), participants who provided a swab sample by remote data collection were more likely to test positive compared with those who provided a swab sample with a study worker home visit; however, after this there was no difference between the groups in their likelihood of testing positive.
- Analysis using data from 19 July to 1 August 2022 showed that remote data collection led to participants with a strong positive test being 2.7 (95% confidence interval: 2.2 to 3.3) times more likely to report symptoms compared with those who had study worker home visit data collection.
- Data from the COVID-19 Infection Survey on the percentage of people with a strong positive test reporting symptoms should not be considered equivalent across the two data collections methods; however, data from both collection methods still provide valuable insights on the most commonly reported symptoms and trends in reported symptoms when analysed separately.
- Remote data collection and study worker population samples are representative of the Census 2021 population by sex, age and region.
- The demographic profiles of remote data collection and study worker home visit population samples are very similar to each other.

2 . Change in Coronavirus (COVID-19) Infection Survey data collection method

The Coronavirus (COVID-19) Infection Survey estimates:

- how many people across England, Wales, Northern Ireland and Scotland would have tested positive for a COVID-19 infection on a nose and throat swab, regardless of whether they report experiencing symptoms
- the number of people who would have tested positive for antibodies against SARS-CoV-2 (the virus that causes COVID-19) at different levels on a blood sample

We also analyse the characteristics of people testing positive for COVID-19, including the percentage of people with a strong positive COVID-19 test result reporting symptoms.

Since the start of the survey in April 2020, until July 2022, the Coronavirus (COVID-19) Infection Survey questionnaire data, and swab and blood samples had been collected by study worker home visits to participants. From July 2022, we changed the way that we collect our data, moving from study worker home visits to a more flexible remote data collection approach. We introduced a digitalised questionnaire, which participants can complete online or by telephone, as well as participants returning swab and blood sample kits through the post (or by courier for some participants).

Our [August 2022 Coronavirus \(COVID-19\) Infection Survey quality report](#) provides more information on this data collection method change, as well as initial analysis comparing the two data collection methods. Findings from this first analysis show there were minimal differences between estimates of COVID-19 positivity produced using the two data collection methods.

Our [September 2022 Coronavirus \(COVID-19\) Infection Survey quality report](#) provides information on analysis of the effects of the two data collection methods on the percentages who would have levels of antibodies against SARS-CoV-2 (the virus that causes COVID-19) above different thresholds.

This quality report includes analysis comparing the likelihood of testing positive for COVID-19 on a nose and throat swab as well as the likelihood of those with a strong positive test result reporting symptoms according to data collection method, while adjusting for several variables. It also includes comparisons between the achieved population samples from our remote and study worker home visit data collection methods and the Census 2021 population to assess how representative these population samples are of the England and Wales population.

The analysis in [Section 3: Likelihood of testing positive for COVID-19 by data collection method](#) and in [Section 4: Likelihood of strong positive COVID-19 cases reporting symptoms by data collection method](#) includes Coronavirus (COVID-19) Infection Survey participants aged 2 years and over in the UK. Analysis in [Section 5: Representativeness of the Coronavirus \(COVID-19\) Infection Survey population sample by data collection method](#) includes Coronavirus (COVID-19) Infection Survey participants aged 2 years and over in England and Wales. Data from the Coronavirus (COVID-19) Infection Survey are based on those living in private households, excluding those living in care homes or other communal establishments.

3 . Likelihood of testing positive for COVID-19 by data collection method

In our [August quality report](#) we compared modelled estimates of the percentage of people testing positive for coronavirus (COVID-19) on a nose and throat swab by data collection method. The estimates were produced using the same methods as those in our weekly [Coronavirus \(COVID-19\) Infection Survey bulletin](#), which use a Bayesian multi-level regression (MRP) model that adjusts for age, sex and region. More information on the methods used to produce COVID-19 positivity rates in our weekly bulletin can be found in our [methods article](#).

To further assess the impact of the change in how the data are collected, we compared the estimated likelihood of testing positive for COVID-19 on a nose and throat swab by data collection method and calendar date while adjusting for several variables. The analysis is based on regression models similar to those presented in our [Analysis of populations in the UK by risk of testing positive for coronavirus \(COVID-19\)](#) September 2021 publication, which provides a more detailed explanation of the methods used.

The models presented here additionally include an interaction between data collection method and calendar date to test for variation in any effect of data collection method on the likelihood of testing positive for COVID-19 over time.

The analysis in this section uses data from 11 to 31 July 2022 and includes 4,104 positive results from 89,030 people who provided a nose and throat swab to study workers, and 2,430 positive results from 53,718 people who provided a nose and throat swab via post or courier. Our first regression model allowed us to test the effect of data collection method by calendar date on the likelihood of testing positive for COVID-19 on a nose and throat swab, while controlling for the following demographic variables:

- age
- sex
- geographical region the participant lives in
- ethnicity
- [deprivation score](#)
- household size
- whether the household was multigenerational
- urban or rural classification of the participant's address
- [effect of a disability](#) (from not having a disability to affected "a lot" by a disability)

The likelihood of testing positive for COVID-19 for those who provided a nose and throat swab sample remotely, compared with those who provided a swab sample at study worker home visits, by calendar date from 11 to 31 July 2022, is shown in Figure 1.

Results show that between 11 and 17 July 2022, participants who provided a swab sample by remote data collection were more likely to test positive than those who provided a swab sample at a study worker home visit. It is possible that those who provided a swab sample remotely in the initial few days of its launch were different in a way that meant that their risk of testing positive was higher than those who provided a swab sample through a study worker home visit in the same time period. For example, symptomatic participants may have provided a swab sample remotely sooner in these earlier days of the online survey launch than participants with no symptoms so that they could know their infection status.

All participants who provided a swab sample by remote data collection at the beginning of the online survey launch were at the start of their 14-day data collection window. Subsequent samples could have been taken at any time during a participant's 14-day data collection window, and invitations to participants to move to the remote data collection approach were also staggered. This means that participants could provide a sample at any point during their testing window, leading to overlap in times from the start of their testing window. This is why the behaviour of symptomatic participants may have affected the results in the first week, when there was no overlap with other data collection windows.

Between 18 and 31 July 2022, there was no statistical evidence of a difference between those who provided a swab sample by remote data collection and those who provided a swab sample at a study worker home visit in the likelihood of testing positive for COVID-19 on a nose and throat swab. This finding supports overall comparability of the results obtained from remote data collection and study worker home visits.

The [odds ratios](#) for this analysis are shown in Figure 1. An odds ratio of greater than 1 indicates a greater likelihood of an outcome in the specified group compared with the reference group, and an odds ratio of less than 1 indicates a lower likelihood. In this case, an odds ratio of greater than 1 indicates an increased likelihood of testing positive for COVID-19 for those who provided a swab sample remotely compared with those who provided a swab sample with a study worker home visit. An odds ratio of less than 1 indicates a decreased likelihood of testing positive for COVID-19.

Figure 1: There was no statistical evidence of a difference in the likelihood of testing positive for coronavirus (COVID-19) between remote and study worker home visit data collection methods, from 18 to 31 July 2022

Estimated likelihood of testing positive for COVID-19 on nose and throat swabs by day for those that provided a swab sample remotely compared with those who provided a swab sample at a study worker home visit, UK, 11 to 31 July 2022

Notes:

1. An odds ratio of greater than 1 indicates a greater likelihood of an outcome in the specified group compared with the reference group, and an odds ratio of less than 1 indicates a lower likelihood.
2. This model controls for age, sex, geographical region the participant lives in, ethnicity, deprivation score, household size, whether the household was multigenerational, urban or rural classification of participant's address, and the effect of a disability.

Download the data

[.xlsx](#)

Sensitivity analysis was produced using a second regression model, which controlled for the variables mentioned previously, as well as other variables that are associated with COVID-19 positivity, such as COVID-19 vaccinations, previous COVID-19 infection and recent contact with hospitals. When controlling for these additional variables the results comparing the two data collection methods were very similar. Odds ratios from this model and the previous model can be found in Tables 1a and 1b of the [Coronavirus \(COVID-19\) Infection Survey quality report: December 2022 dataset](#).

All variables used and variables considered for these models can be found in [Section 7: How the data are measured](#).

4 . Likelihood of strong positive COVID-19 cases reporting symptoms by data collection method

This section considers the effect of the data collection method on the reporting of symptoms in people with a strong positive coronavirus (COVID-19) test ([cycle threshold \(Ct\)](#) value less than 30). This analysis uses data from 19 July to 1 August 2022 where both remote and study worker home visit data collection methods were used. Participants across the UK were asked to report whether they experienced the following symptoms in the seven days before they were tested, and separately whether they felt that they had symptoms compatible with a COVID-19 infection in the last seven days:

- fever
- muscle ache (myalgia)
- fatigue (weakness or tiredness)
- sore throat
- cough
- shortness of breath
- headache
- nausea or vomiting
- abdominal pain
- diarrhoea
- loss of taste or loss of smell

Symptoms were self-reported and were not professionally diagnosed.

Among those who tested positive for COVID-19 with a strong positive test, the percentage who reported symptoms was 60% (95% confidence interval: 57% to 62%) for data collected by study worker home visits, and 78% (95% confidence interval: 76% to 80%) for data collected remotely.

To further assess the effect of how the data were collected, we used a logistic regression model to compare the estimated likelihood of reporting symptoms by data collection method, among those with a strong positive test. The model controlled for age, sex, region, ethnicity, long-term health condition, work sector and [deprivation](#) score.

The results showed that remote data collection led to participants with a strong positive test being 2.7 (95% confidence interval: 2.2 to 3.3) times more likely to report symptoms compared with those who had study worker home visit data collection.

There are several potential reasons why the reporting of symptoms may differ by data collection method. This analysis only included participants who tested positive for COVID-19 with a strong positive test, and participants may be more likely to choose to complete the survey and test using remote data collection while experiencing symptoms. In contrast, study worker visits were scheduled independently, so participants did not have the same choice about when they occurred. Other potential reasons may relate to differences in how the questionnaire was interpreted when completed remotely compared with completion with a study worker.

These results show that data from the COVID-19 Infection Survey on the percentage of people with a strong positive test reporting symptoms should not be considered equivalent across the two data collections methods. Data from both collection methods provide valuable insights on the most commonly reported symptoms and trends in reported symptoms among the population when analysed separately.

Analysis on symptoms presented in our [Coronavirus \(COVID-19\) Infection Survey: characteristics of people testing positive for COVID-19, UK](#) publications used a different method. This method considered the percentage of people that reported symptoms at survey visits within 35 days of the first positive test in a positive episode where any test was a strong positive. For this reason, symptoms analysis presented in this publication is not directly comparable with our previous publications.

5 . Representativeness of the Coronavirus (COVID-19) Infection Survey population sample by data collection method

This section analyses the representativeness of the achieved population samples from remote data collection and from study worker home visits by assessing whether there is any evidence that participants providing swab samples through remote data collection are different to participants providing swab samples at study worker home visits. It also compares how well the Coronavirus (COVID-19) Infection Survey population samples from remote data collection represent the England and Wales population using unadjusted and adjusted estimates.

The unadjusted population sample is the actual number of people taking part in the survey during the time period specified, whereas the adjusted population sample has been weighted to be representative of the target population. The weights applied to these population samples are calibrated to population projections that are based on Census 2011 as well as births, deaths and migration that have occurred since 2011.

Census 2011, rather than Census 2021, was used to adjust the population samples because these data are available from all four countries of the UK. In the case of ethnicity, proportions have been taken from the Annual Population Survey (APS) and applied to the overall total for the projections. This weighting method is similar to the post-stratification used to produce our headline estimates in our weekly Coronavirus (COVID-19) Infection Survey bulletin, which adjusts the results from the population sample to be representative of the overall population in terms of age, sex, and region (region is only adjusted for in the England model). For more information, see our [Coronavirus \(COVID-19\) Infection Survey: methods and further information methodology](#) publication.

The [census](#) is undertaken every 10 years and gives us a picture of all the people and households in England and Wales; it is the most complete source of information about the population that we have. We assessed how well our Coronavirus (COVID-19) Infection Survey data for England and Wales represented the population by comparing it against Census 2021 data, focusing on England and Wales only to align with the coverage of Census 2021.

While data from our Coronavirus (COVID-19) Infection Survey sample exclude those living in care homes and other communal establishments, and those aged under 2 years old, the Census 2021 population data do not. In addition, the 2021 Census was carried out in the first half of 2021, and our Coronavirus (COVID-19) Infection Survey samples are for 14-day periods in 2022. This means that we would expect there to be some differences between the two data sources, as the demographic profile of those living in care homes and other communal establishments may differ to the private residential population, and demographic events such as migration will have affected the demographic profile of the England and Wales population since the 2021 Census.

Data from the Coronavirus (COVID-19) Infection Survey participants in England and Wales were aggregated over 14-day periods from 6 to 19 May 2022 for those providing samples through study worker home visits and from 10 to 23 August 2022 for those providing samples through remote data collection. Participants of the Coronavirus (COVID-19) Infection Survey take part only once per month, and so around half the total participants were assessed in the 14-day period.

We compared the unadjusted and adjusted profiles (percentages of participants appearing across different categories of a variable) of these population samples with each other, and to the Census 2021 population, by sex, age, ethnic group, household size and region. Tables 1 to 5 provide the absolute differences between the profiles of these population samples. The data show that:

- participants providing samples through remote data collection and study worker home visits share generally very similar profiles, both when unadjusted and when adjusted, with the majority of differences between these population samples being below 1 percentage point
- when compared with the Census 2021 population, participants providing samples through both remote data collection and study worker home visits are broadly similar to the England and Wales population when unadjusted, and more so when adjusted
- when compared with the Census 2021 population, participants providing samples through both remote data collection and study worker home visits overrepresent people from a White ethnic group and two-person households before adjustment, with the same overrepresentation being evident but smaller after adjustment

Table 1: Difference in sex profiles between Coronavirus (COVID-19) Infection Survey (CIS) participants providing swab samples through remote data collection and study worker home visits, and the Census 2021 population. Absolute differences between unadjusted and adjusted Coronavirus (COVID-19) Infection Survey population samples and Census 2021 population by sex, England and Wales, 6 to 19 May 2022 for remote data collection and 10 to 23 August 2022 for study worker home visits

Demographic Variable	Absolute difference in percentage points between CIS remote data collection and CIS study worker home visit population samples		Absolute difference in percentage points between CIS remote data collection population sample and Census 2021		Absolute difference in percentage points between CIS study worker home visit population sample and Census 2021	
	Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted
Sex						
Male	0.2	0.0	-2.8	0.4	-3.0	0.4
Female	-0.2	0.0	2.8	-0.4	3.0	-0.4

Source: Office for National Statistics – Coronavirus (COVID-19) Infection Survey and Census 2021

Table 2: Difference in age profiles between Coronavirus (COVID-19) Infection Survey (CIS) participants providing swab samples through remote data collection and study worker home visits, and the Census 2021 population. Absolute differences between unadjusted and adjusted Coronavirus (COVID-19) Infection Survey population samples and Census 2021 population by age, England and Wales, 6 to 19 May 2022 for remote data collection and 10 to 23 August 2022 for study worker home visits

Demographic Variable	Absolute difference in percentage points between CIS remote data collection and CIS study worker home visit population samples		Absolute difference in percentage points between CIS remote data collection population sample and Census 2021		Absolute difference in percentage points between CIS study worker home visit population sample and Census 2021	
	Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted
Age						
2 to 7 years	0.0	-0.3	-4.6	-2.2	-4.6	-1.9
8 to 11 years	0.3	0.1	-1.4	2.2	-1.7	2.1
12 to 15 years	0.4	0.3	-0.2	0.6	-0.6	0.3
16 to 24 years	0.1	0.0	-6.0	-0.5	-6.1	-0.5
25 to 34 years	-0.8	-0.1	-7.5	-0.3	-6.7	-0.2
35 to 49 years	-0.4	0.0	-1.3	0.0	-0.9	0.0
50 to 59 years	0.0	-0.3	3.2	-2.5	3.2	-2.2
60 to 64 years	0.5	0.2	4.2	0.8	3.7	0.6
65 to 69 years	0.4	0.0	5.4	1.9	5.0	1.9
70 to 74 years	0.2	0.4	5.0	1.3	4.8	0.9
75 to 79 years	-0.1	0.1	3.8	1.0	3.9	0.9
80 years and above	-0.8	-0.4	-0.7	-2.3	0.1	-1.9

Source: Office for National Statistics – Coronavirus (COVID-19) Infection Survey and Census 2021

Table 3: Difference in ethnicity profiles between Coronavirus (COVID-19) Infection Survey (CIS) participants providing swab samples through remote data collection and study worker home visits, and the Census 2021 population

Absolute differences between unadjusted and adjusted Coronavirus (COVID-19) Infection Survey population samples and Census 2021 population by ethnicity, England and Wales, 6 to 19 May 2022 for remote data collection and 10 to 23 August 2022 for study worker home visits

Demographic Variable	Absolute difference in percentage points between CIS remote data collection and CIS study worker home visit population samples		Absolute difference in percentage points between CIS remote data collection population sample and Census 2021		Absolute difference in percentage points between CIS study worker home visit population sample and Census 2021	
	Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted
Ethnicity						
White	0.3	0.9	10.9	9.7	10.6	8.8
Asian	0.0	-0.4	-5.0	-4.6	-5.0	-4.2
Black	-0.1	-0.3	-3.1	-3.1	-3.0	-2.8
Other	0.0	-0.1	-1.4	-1.3	-1.4	-1.2
Mixed	-0.1	-0.2	-1.3	-0.7	-1.2	-0.5

Source: Office for National Statistics – Coronavirus (COVID-19) Infection Survey and Census 2021

Table 4: Difference in household size profiles between Coronavirus (COVID-19) Infection Survey (CIS) participants providing swab samples through remote data collection and study worker home visits, and the Census 2021 population

Absolute differences between unadjusted and adjusted Coronavirus (COVID-19) Infection Survey population samples and Census 2021 population by household size, England and Wales, 6 to 19 May 2022 for remote data collection and 10 to 23 August 2022 for study worker home visits

Demographic Variable	Absolute difference in percentage points between CIS remote data collection and CIS study worker home visit population samples		Absolute difference in percentage points between CIS remote data collection population sample and Census 2021		Absolute difference in percentage points between CIS study worker home visit population sample and Census 2021	
	Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted
Household size						
1	-2.3	-2.5	1.6	-1.7	3.9	0.8
2	0.3	-0.1	14.4	6.5	14.1	6.6
3	0.4	0.2	-3.7	-1.6	-4.1	-1.8
4	0.5	1.2	-4.2	1.6	-4.7	0.4
5 and above	1.1	1.2	-8.2	-4.9	-9.3	-6.1

Source: Office for National Statistics – Coronavirus (COVID-19) Infection Survey and Census 2021

Table 5: Difference in region profiles between Coronavirus (COVID-19) Infection Survey (CIS) participants providing swab samples through remote data collection and study worker home visits, and the Census 2021 population

Absolute differences between unadjusted and adjusted Coronavirus (COVID-19) Infection Survey population samples and Census 2021 population by region, England and Wales, 6 to 19 May 2022 for remote data collection and 10 to 23 August 2022 for study worker home visits

Demographic Variable	Absolute difference in percentage points between CIS remote data collection and CIS study worker home visit population samples		Absolute difference in percentage points between CIS remote data collection population sample and Census 2021		Absolute difference in percentage points between CIS study worker home visit population sample and Census 2021	
	Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted
Region						
North East	0.4	0.0	0.1	0.1	-0.3	0.1
North West	-0.2	0.0	0.8	-0.1	1.0	-0.1
Yorkshire and The Humber	0.3	0.0	0.8	0.0	0.5	0.0
East Midlands	0.2	0.0	-0.8	0.0	-1.0	0.0
West Midlands	0.1	0.0	-1.6	0.0	-1.7	0.0
East of England	-0.1	0.0	0.3	-0.1	0.4	-0.1
London	-1.4	0.0	0.8	0.3	2.2	0.3
South East	-0.8	0.0	-1.4	-0.2	-0.6	-0.2
South West	0.2	0.0	0.0	-0.1	-0.2	-0.1
Wales	1.2	0.0	1.0	0.1	-0.2	0.1

Source: Office for National Statistics – Coronavirus (COVID-19) Infection Survey and Census 2021

Demographic profiles for those providing samples through study worker home visits and remote data collection, and the Census 2021 population are available alongside the absolute differences between these population samples and Census 2021 population, in our [Coronavirus \(COVID-19\) Infection Survey quality report: December 2022 dataset](#).

To quantify the dissimilarity between the demographic profiles of those providing samples through study worker home visits and remote data collection, and the Census 2021 population, we calculated indices of dissimilarity (Duncan O, Duncan B (1955), [A Methodological Analysis of Segregation Indexes](#), American Sociological Review, Volume 20, Issue 2, pages 210 to 217). The index of dissimilarity can range between 0 and 100, with a value of 0 indicating that the two profiles are completely similar, and a value of 100 indicating that the two profiles are totally dissimilar. The index of dissimilarity represents the percentage of participants that would need to change between categories of a variable for the profiles of both samples or populations to appear similar. For example, an index of 22.0 would mean that 22% of participants would need to change between categories.

The indices of dissimilarity between the demographic profiles of those providing samples through study worker home visits and remote data collection, and the Census 2021 population, are shown in Table 6. Both the remote data collection and study worker home visit population samples are broadly similar to Census 2021 when unadjusted and more so when adjusted. These population samples are also very similar to each other.

Table 6: Indices of dissimilarity for Coronavirus (COVID-19) Infection Survey (CIS) participants providing swab samples through remote data collection and study worker home visits, and Census 2021 profiles by demographic variable

Indices of dissimilarity for unadjusted and adjusted Coronavirus (COVID-19) Infection Survey population samples and Census 2021 population by variable, England and Wales, 6 to 19 May 2022 for remote data collection and 10 to 23 August 2022 for study worker home visits

Demographic Variable	CIS remote data collection population sample Index of Dissimilarity to Census 2021		CIS study worker home visit population sample Index of Dissimilarity to Census 2021		Absolute difference of CIS remote data collection to CIS study worker home visit population samples indices of Dissimilarity	
	Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted
	Sex	2.8	0.4	3.0	0.4	0.2
Age	21.7	7.8	20.7	6.7	2.0	1.1
Ethnicity	10.9	9.7	10.6	8.8	0.2	0.95
Household size	16.1	8.2	18.1	7.9	2.3	2.6
Region	3.8	0.5	4.1	0.5	2.5	0.0
Mean	11.1	5.3	11.3	4.9	1.4	0.7

Source: Office for National Statistics – Coronavirus (COVID-19) Infection Survey and Census 2021

The data in this section show that the achieved remote and study worker home visit population samples are very similar to each other and broadly similar to the Census 2021 population. This provides evidence that the data collection method has not impacted on uptake of the survey and ensures that estimates of positivity are representative of the England and Wales populations.

All data on the representativeness of the Coronavirus (COVID-19) Infection Survey population samples are available in Tables 2a to 2c of our [Coronavirus \(COVID-19\) Infection Survey quality report: December 2022 dataset](#).

6 . Summary of findings

The findings presented in this article, as well as findings from our [August 2022 Coronavirus \(COVID-19\) Infection Survey quality report](#) and [September 2022 Coronavirus \(COVID-19\) Infection Survey quality report](#) indicate that the change to a remote data collection method has had minimal impact on most survey results, including the likelihood of testing positive for COVID-19. However, results show that data from the COVID-19 Infection Survey on the percentage of people with a strong positive test reporting symptoms should not be considered equivalent across the two data collection methods.

7 . How the data are measured

Likelihood of testing positive for coronavirus (COVID-19) by data collection method

The models described in [Section 3: Likelihood of testing positive for COVID-19 by data collection method](#) test the effect of data collection method by day on the likelihood of testing positive for COVID-19, while controlling for several other variables. Variables controlled for in our first model were:

- age
- sex
- geographical region the participant lives in
- ethnicity
- [deprivation score](#)
- household size
- whether the household was multigenerational
- urban or rural classification of the participant's address
- [effect of a disability](#) (from not having a disability to affected "a lot" by a disability)

Variables controlled for in our second model were:

- all of the variables controlled for in our first model
- work status (responses were grouped into "Employed, working", "Employed, not working", "Not working", "Retired" and "Child/student")
- whether the participant was previously infected with COVID-19 based on a positive swab test (in the survey, the English national testing programme or self-reported)
- whether the participant had travelled abroad in the previous 28 days
- COVID-19 vaccinations
- contact with hospitals in the previous 28 days
- contact with care homes in the previous 28 days
- whether the participant currently smoked

Additional variables considered for the model that were not included were:

- whether a child aged 16 years or under lived in the household
- whether an adult aged 70 years or over lived in the household
- days worked outside the home
- whether the participant worked in a patient-facing healthcare role, a health and social care role or a care home
- whether the participant worked in a role that involves direct contact with others
- work sector
- work or school location (at home or elsewhere)
- social distancing at work or school
- how the participant travels to work or school

These variables were not included in the model because our screening process revealed no statistical evidence of association between them and the likelihood of testing positive for COVID-19.

8 . Coronavirus (COVID-19) Infection Survey, Quality Report data

[Coronavirus \(COVID-19\) Infection Survey quality report: December 2022](#)

Dataset | Released 21 December 2022

Quality report data on the Coronavirus (COVID-19) Infection Survey data collection method change from study worker home visit to remote data collection.

9 . Collaboration



The Coronavirus (COVID-19) Infection Survey analysis was produced by the Office for National Statistics (ONS) in collaboration with our research partners at the University of Oxford, the University of Manchester, UK Health Security Agency (UK HSA) and Wellcome Trust. Of particular note are:

- Sarah Walker - University of Oxford, Nuffield Department for Medicine: Professor of Medical Statistics and Epidemiology and Study Chief Investigator
- Koen Pouwels - University of Oxford, Health Economics Research Centre, Nuffield Department of Population Health: Senior Researcher in Biostatistics and Health Economics
- Thomas House - University of Manchester, Department of Mathematics: Reader in Mathematical Statistics

10 . Glossary

Deprivation

Deprivation is based on an [index of multiple deprivation \(IMD\) \(PDF, 2.18MB\)](#) score or equivalent scoring method for the devolved administrations, from 1, which represents most deprived, up to 100, which represents least deprived. The hazard or odds ratio shows how a 10-unit increase in deprivation score, which is equivalent to 10 percentiles or 1 decile, affects the likelihood of testing positive for COVID-19.

SARS-CoV-2

This is the scientific name given to the specific virus that causes COVID-19.

Effect of a disability

To measure how severely a disability affected participants, we asked them if any long-lasting health conditions reduced their ability to carry out day-to-day activities, as part of our Coronavirus (COVID-19) Infection Survey questionnaire. The response options for this question were: "Yes, a lot", "Yes, a little" or "Not at all".

Odds ratio

An odds ratio indicates the likelihood of an individual testing positive for COVID-19 given a particular characteristic or variable. When a characteristic or variable has an odds ratio of 1, this means there is neither an increase nor a decrease in the likelihood of testing positive for COVID-19 compared with the reference category. An odds ratio greater than 1 indicates an increased likelihood of testing positive for COVID-19 compared with the reference category. An odds ratio less than 1 indicates a decreased likelihood of testing positive for COVID-19 compared with the reference category.

Confidence interval

A confidence interval gives an indication of the degree of uncertainty of an estimate, showing the precision of a sample estimate. The 95% confidence intervals are calculated so that if we repeated the study many times, 95% of the time the true unknown value would lie between the lower and upper confidence limits. A wider interval indicates more uncertainty in the estimate. Overlapping confidence intervals indicate that there may not be a true difference between two estimates. For more information, see our methodology page on [statistical uncertainty](#).

Cycle threshold (Ct) values

The strength of a positive coronavirus (COVID-19) test is determined by how quickly the virus is detected, measured by a cycle threshold (Ct) value. The lower the Ct value, the higher the viral load and the stronger the positive test. Positive results with a high Ct value can be seen in the early stages of infection when virus levels are rising, or late in the infection, when the risk of transmission is low.

11 . Related links

[Coronavirus \(COVID-19\) Infection Survey: methods and further information](#)

Methodology | Revised 5 August 2022

This methodology guide is intended to provide information on the methods used to collect the data, process it, and calculate the statistics produced from the Coronavirus (COVID-19) Infection Survey.

[Coronavirus \(COVID-19\) Infection Survey QMI](#)

Methodology | Revised 8 August 2022

Quality and Methodology Information (QMI) for the Coronavirus (COVID-19) Infection Survey (CIS), detailing the strengths and limitations of the data, methods used, and data uses and users.

[Coronavirus \(COVID-19\) Infection Survey. UK](#)

Bulletin | Updated weekly

Estimates for England, Wales, Northern Ireland and Scotland, including regional and age breakdowns. This survey is being delivered in partnership with the University of Oxford, University of Manchester, UK Health Security Agency and Wellcome Trust.

[Coronavirus \(COVID-19\) Infection Survey: characteristics of people testing positive for COVID-19 in countries of the UK](#)

Bulletin | Updated fortnightly

The characteristics of people testing positive for coronavirus (COVID-19) from the COVID-19 Infection Survey. This survey is being delivered in partnership with the University of Oxford, the University of Manchester, UK Health Security Agency and Wellcome Trust.

[Coronavirus \(COVID-19\) Infection Survey: antibody and vaccination data for the UK](#)

Bulletin | Updated fortnightly

Antibody and vaccination data by UK country and English regions from the Coronavirus (COVID-19) Infection Survey. This survey is being delivered in partnership with the University of Oxford, University of Manchester, UK Health Security Agency and Wellcome Trust.

The [Nuffield Department of Medicine study protocol](#) specifies the research for the study.

The Welsh Government publishes [results from the COVID-19 Infection Survey in English](#) and [results from the COVID-19 Infection Survey in Welsh](#) that describe COVID-19 infections in Wales.

The [Department of Health \(Northern Ireland\) publishes results from the COVID-19 Infection Survey](#) that describe COVID-19 infections in Northern Ireland.

The [Scottish Government publishes results from the COVID-19 Infection Survey](#) that describe COVID-19 infections in Scotland.

12 . Cite this article

Office for National Statistics (ONS), published 21 December 2022, ONS website, methodology article, [Coronavirus \(COVID-19\) Infection Survey. Quality Report: December 2022](#)