

# Winter Coronavirus (COVID-19) Infection Study QMI

Quality and Methodology Information (QMI) for the Winter CIS, detailing the strengths and limitations of the data, methods used, and data uses.

Contact:  
Health Statistics and Analysis  
Health.Studies@ons.gov.uk  
+44 8081 961270

Release date:  
26 June 2024

Next release:  
To be announced

## Table of contents

1. [Output information](#)
2. [About this Quality and Methodology Information report](#)
3. [Quality summary](#)
4. [Quality characteristics of the Winter Coronavirus \(COVID-19\) Infection Study \(Winter CIS\)](#)
5. [Test sensitivity and specificity](#)
6. [Methods used to produce the data](#)
7. [Useful links](#)
8. [Cite this methodology](#)

# 1 . Output information

- Survey name: Winter Coronavirus (COVID-19) Infection Study (Winter CIS).
- Time period: the main Winter CIS study runs from November 2023 to March 2024.
- How compiled: estimates are derived from questionnaire responses of Winter CIS participants who were initially enrolled on the ONS Coronavirus (COVID-19) Infection Survey (CIS).
- Geographic coverage: England and Scotland, with selected analysis broken down to the regions of England.
- Participants: identified as eligible for Winter CIS if they had participated in the CIS, listed email as preferred method of contact and had consented to further research; further information on CIS participants is detailed in the [Coronavirus \(COVID-19\) Infection Survey quality report: December 2022](#).
- Number of participants: approximately 290,000 participants were identified as being eligible for Winter CIS and invited to participate in the study.
- Achieved sample size: 139,453 participants.

## 2 . About this Quality and Methodology Information report

The Winter Coronavirus (COVID-19) Infection Study (Winter CIS) is commissioned by the UK Health Security Agency (UKHSA) and jointly delivered by the UKHSA and Office for National Statistics (ONS).

The ONS will be publishing weekly weighted positivity rates to estimate the percentage of the population who would have tested positive via a lateral flow device (LFD) test for the SARS-CoV-2 coronavirus (COVID-19), with or without symptoms. This is different from the methodology used by UKHSA to estimate the prevalence of COVID-19. Prevalence calculations build on positivity rates by taking into account the test sensitivity and specificity to provide a closer reflection of the percentage of the population who actually have COVID-19. The UKHSA provides details of the methods used in their [Quality and Methodology Information](#).

This quality and methodology report contains information on the quality (including the [European Statistical System \(PDF, 3.06MB\)](#) five dimensions of quality) of the statistics produced as outputs from the Winter CIS. The methods used to create Winter CIS statistical outputs are also detailed. The information in this report will help you to:

- understand the strengths and limitations of Winter CIS statistics
- reduce the risk of misusing data
- help you to decide suitable uses for the data
- understand the methods used to create the data

## 3 . Quality summary

### Important points about the Winter Coronavirus (COVID-19) Infection Study (Winter CIS)

The Winter CIS was launched in November 2023, in England and Scotland to:

- estimate how many people test positive for COVID-19 infection using a lateral flow device (LFD), regardless of whether they report experiencing symptoms
- continue surveillance of COVID-19 as part of active community health monitoring
- monitor the impact of COVID-19, long COVID and symptoms of other respiratory infections on the lives of individuals, the community and on health services, and how these are changing
- assess potential pressures to help support the NHS and other services to prepare for future stressors, acting as an early warning system

Only private residential households and their residents are included in the survey. People in hospitals, care homes and/or other communal settings are not included.

In contrast to the previous Coronavirus (COVID-19) Infection Survey (CIS), which used polymerase chain reaction (PCR) tests, participants in Winter CIS were sent LFDs for the detection of the SARS-CoV-2 coronavirus to self-swab, self-complete and report results. No blood samples are requested from participants as part of the study. In addition, no financial incentives are offered for participation.

This study is undertaken jointly by the Office for National Statistics (ONS) and the UK Health Security Agency (UKHSA). The ONS collects information about the UK's society and economy, which provides evidence for policy- and decision-making, and for directing resources to where they are needed most. The UKHSA is an Executive Agency within the Department of Health and Social Care (DHSC). It is responsible for protecting every member of every community from the effects of infectious diseases, chemical, biological, radiological and nuclear incidents, and other health threats.

## Overview of the Winter CIS

The Winter CIS launched in November 2023, in England and Scotland.

The Winter CIS sample is drawn from participants of the previous Coronavirus (COVID-19) Infection Survey (CIS). Eligibility to participate in the Winter CIS was established if a CIS participant had:

- agreed to be approached about other ethically approved research studies
- took part in the digital version of CIS

## Participant recruitment

All eligible CIS participants aged 16 years and over were sent a Winter CIS survey invite by email or letter, depending on their communication preferences in the CIS. All eligible participants aged under 16 years were sent a letter to their parent or guardian's address. The ONS processed and sent all letter communications. GOV.UK Notify, a service offered by the Cabinet Office, sent email communications on our behalf.

As new participants were not recruited into the study, the youngest age of participants is older than the CIS (which covered those aged 2 years and over). No participants under the age of 3 years are in this study.

In many cases, more than one participant from a single household participated in the Winter CIS. This is because the CIS was based initially on a random sample of households to provide a nationally representative survey.

Unlike in the CIS no financial incentives were offered upon return of a completed questionnaire and test result. This could be a factor in the lower sign-up rates to Winter CIS among younger age groups.

## Participant engagement

Once consented to be part of the study (139,453 individuals, out of 290,000 initially approached, consented to take part), participants were sent a set of 14 LFDs and invited every four weeks to complete the Winter CIS questionnaire online.

Completion of the questionnaire was online only (paper copies were not posted out), however, if a participant required help with online completion, a telephone number was provided so the dedicated contact centre could assist the participant and, if needed, complete the questionnaire on their behalf over the phone.

## Uses of Winter CIS

The Winter CIS provides important information about the characteristics of people and households testing positive for COVID-19, as well as for self-reporting long COVID or symptoms of other respiratory infections within the community population. Community in this instance refers to private residential households and it excludes those in hospitals, care homes and/or other communal establishment settings, with the overall sample intended to be applicable as a representative sample of the English and Scottish population.

This will help UK governments understand the impact of COVID-19 infections, self-reported long COVID and symptoms of other respiratory infections on life, the community and health services. It could further assist with service planning and vaccination roll out.

The data can be used for:

- estimating the percentage of the population that would test positive via a lateral flow device (LFD) test for COVID-19
- estimating the prevalence of COVID-19 in the community
- identifying characteristics associated with self-reporting symptoms consistent with influenza-like illnesses (ILI)
- identifying the most common symptoms associated with long COVID experienced by Winter CIS participants

The data cannot be used for:

- comparing this study with the previous [COVID-19 Infection Survey \(CIS\)](#), which ran from 2020 to March 2023, because the two studies used different diagnostic tests and statistical methods, therefore the results are not comparable
- estimating the number of respiratory illness cases or their prevalence in care homes, hospitals and/or other communal settings
- providing information about recovery times for those infected

## Strengths and limitations

Some of the Winter CIS main strengths include:

- a large sample of participants
- high levels of participant engagement
- a questionnaire which examines various symptoms and characteristics across all age groups

An additional strength of our survey is that all participants are a subset of the CIS sample. This enables us to collect and maintain longitudinal data to which we can link back for COVID-19 monitoring, alongside examining other respiratory infections and vaccination histories.

The Winter CIS sample, as with all surveys, is subject to possible bias. More so, as the Winter CIS study population is a subset of the CIS population and selection for the Winter CIS is driven by pre-existing underlying biases within the CIS cohort, for example, participants needed access to an email address to participate. For more details see "Representativeness" in [Section 4: Quality characteristics of the Winter Coronavirus \(COVID-19\) Infection Study \(Winter CIS\)](#).

All estimates presented in our publications contain [uncertainty](#). Although the statistics produced as outputs from the survey data are our best estimates, they should not be regarded as completely accurately reflecting the unknown true numbers we are trying to measure.

## Quality assurance process and data flow

The Winter CIS is a collaboration between the ONS and UKHSA. To ensure robust quality in the analytical outputs of both organisations, additional quality assurance checks are put in place. The consent survey and Winter CIS survey responses are processed, and quality assured by the ONS, before being sent to UKHSA via a secure data transfer. The UKHSA then performs its own quality checks before producing its own analysis. Weekly meetings between the ONS and UKHSA data analysts are also used to carry out additional quality assurance on the data and analysis produced each week.

The ONS collected consent from previous CIS participants to take part in the study via SmartSurvey. These participants formed the base sample on which to run the Winter CIS. Their data were cleaned and processed, removing duplicate entries, and validating or removing entries where participants had entered incorrect information, for instance incorrectly entered postcodes. In cases where the information entered failed the validation checks, data previously provided by the participant as part of the original CIS were checked to see if they could be used to correct the information provided via SmartSurvey. Participant name and address information was then shared with UKHSA through a secure transfer system to enable UKHSA to send LFDs to participants. UKHSA performed their own checks on the address information before sending the tests.

During the running of the study, participant information and responses are collected and processed via internal survey management and secure data systems. Data are regularly reviewed, and participant information is updated daily via the Winter CIS call centre to reflect withdrawals or changes to personal details. A weekly joint delivery meeting is held to discuss operational changes and risks to mitigate any potential issues. Management information is shared on response rates and sample representativeness.

Anonymised survey response files (with participant details, such as name, removed) are generated on a weekly basis with in-built system checks and validations that occur during the main processing of the data on all entry fields. These survey response files are then delivered to the ONS analysis team and to UKHSA via a secure file transfer system. The ONS analysis team then performs additional cleaning and quality assurance of the data before weighting is performed. Checks include identifying duplicate entries, ensuring that the sample information and dates of responses are as expected, and checking the distribution of the survey response fields. Any important issues are flagged to UKHSA. The ONS then analyses the survey responses to create weighted and unweighted responses. The analysis is run via two sets of independently produced code with the results checked against each other to ensure results are robust and consistent. Each of these results is also checked by a secondary quality assurer of the code and outputs. Statistical tests and charts are created to sense check non-response models for weighting, before being presented to senior management for sign off.

Once these checks are complete, a set of data tables with the weighted and unweighted results are produced. These tables go through several further quality assurance and formatting checks before being shared with colleagues in UKHSA via a secure managed file transfer software product. Alongside the file sent to UKHSA, the initial ONS positivity results are compared with UKHSA modelled prevalence estimates by the ONS and UKHSA analysis teams to check the two methods are producing consistent results. Results are then published simultaneously by the ONS and UKHSA.

## 4 . Quality characteristics of the Winter Coronavirus (COVID-19) Infection Study (Winter CIS)

### Relevance

The Winter Coronavirus (COVID-19) Infection Study (Winter CIS) seeks to understand the impact of COVID-19, long COVID and symptoms of other respiratory infections.

Data are collected through self-completion of online questionnaires and tests carried out using lateral flow devices. Data are analysed to understand the impact of COVID-19 and self-reported symptoms on the lives of individuals and the community, including absence from work, and the impact on health services.

These [official statistics in development](#) can be used to highlight potential pressures for the NHS and be used to help support wider services, for example, assisting governments with informed decisions on important policies, such as service planning and vaccination rollouts.

## Uncertainty

Estimates in our publications contain some [uncertainty](#). There are many sources of uncertainty, but the main possible sources in our Winter CIS publications would include:

- quality of data collected in the questionnaire
- participants carrying out the COVID-19 swab tests correctly
- the data are based on a sample of people rather than the whole population, so there is some statistical uncertainty in the estimates
- potential bias, which may not be fully mitigated by the methods used to adjust for this including weighting; this includes factors that move the responding sample away from being fully random such as non-response to CIS, not providing consent to participate in Winter CIS and non-response to Winter CIS

As in any survey, some data can be incorrect or missing. We minimise this during data collection whereby participants are only able to submit one answer for each question.

## Communicating uncertainty

To quantify uncertainty in our analyses, we present 95% [confidence intervals](#) in our data. In addition, the UK Health Security Agency (UKHSA) include credible intervals within their analysis to show the degree of uncertainty.

Confidence intervals give an indication of the degree of uncertainty of an estimate, with a wider interval indicating more uncertainty in the estimate. Overlapping intervals indicate that there may not be a true difference between two estimates.

## Representativeness

Ensuring a representative sample of the general population is important for producing survey-based estimates broken down by characteristics such as age, sex and region. In the survey, this is important to help us understand trends in different population sub-groups across England and Scotland.

Aspects of the study design may affect how representative the study population is compared with the wider population. For example, participants needed access to an email address to participate (this did not necessarily need to be their own and the survey could be completed by others on their behalf). This will exclude a small proportion of the population ([around 6% of households had no internet access as at March 2021](#) (PDF, 7.38MB)) but is unlikely to have a large impact on the estimates produced.

Age is one of the main determinants of internet access but we see an over-representation of the older age groups in the study (Table 1b). Deprivation is also associated with internet access but results from the Winter CIS suggest there is now little association between deprivation and COVID-19 test positivity.

Unlike CIS no incentives were offered to take part in Winter CIS; this may impact response rates in some groups more than others.

The following tables show the representativeness analysis of the Winter CIS sample who initially consented to take part in Winter CIS. The unweighted response population is the actual number of people taking part in the survey.

Within the Winter CIS sample:



- the overall sample is representative of Scotland in terms of population share along with all the English regions
- females are slightly over-represented, while males are slightly under-represented at the UK level (England and Scotland 51% female and 49% male; Winter CIS 57% female and 43% male)
- younger age groups (aged 3 to 17 years, 18 to 34 years and 35 to 44 years) are under-represented when compared with older age groups (aged 55 to 64 years, 65 to 74 years, and 75 years and over), which are over-represented
- those reporting White ethnicity are largely over-represented in England (81%; Winter CIS 95%)

Table 1a: Actual relevant population and Winter CIS response population by sex, England and Scotland

Category	Actual relevant population Population profile		Winter CIS Unadjusted population Sample profile		Difference	
	Number of people	Proportion	Number of people	Proportion	Absolute difference to actual population	Relative difference to actual population
<b>Female</b>	31,628,612	0.51	79,590	0.57	0.06	12%
<b>Male</b>	30,298,136	0.49	59,863	0.43	-0.06	-12%

Source: Winter Coronavirus (COVID-19) Infection Study from the Office for National Statistics

#### Notes

1. Population figures used for the "actual relevant population" are drawn from Census 2021 for England and Scotland Census 2022 for Scotland.
2. Response population figures are unadjusted and represent the actual response numbers.
3. Absolute difference represents the difference in the unweighted proportion in our sample to the actual population.
4. Relative difference represents the proportion of difference attributed to each group. This is calculated by taking the absolute difference and dividing by the proportion of the actual population.



Table 1b: Actual relevant population and Winter CIS response population by age, England and Scotland

Category	Actual relevant population Population profile		Winter CIS Unadjusted population Sample profile		Difference	
	Number of people	Proportion	Number of people	Proportion	Absolute difference to actual population	Relative difference to actual population
<b>3 to 17 years</b>	10,848,346	0.18	5,359	0.04	-0.14	-79%
<b>18 to 34 years</b>	13,564,000	0.23	6,575	0.05	-0.18	-79%
<b>35 to 44 years</b>	8,068,321	0.13	12,993	0.09	-0.04	-31%
<b>45 to 54 years</b>	8,238,488	0.14	21,443	0.15	0.01	12%
<b>55 to 64 years</b>	7,824,403	0.13	32,881	0.24	0.11	81%
<b>65 to 74 years</b>	6,159,724	0.10	38,969	0.28	0.18	172%
<b>75+ years</b>	5,315,444	0.09	21,233	0.15	0.06	72%

Source: Winter Coronavirus (COVID-19) Infection Study from the Office for National Statistics

#### Notes

1. Population figures used for the "actual relevant population" are drawn from Census 2021 for England and ONS mid-year population estimates for Scotland.
2. Response population figures are unadjusted and represent the actual response numbers. ,Absolute difference represents the difference in the unweighted proportion in our sample to the actual population., Relative difference represents the proportion of difference attributed to each group. This is calculated by taking the absolute difference and dividing by the proportion of the actual population.

Table 1c: Actual population and Winter CIS response population by ethnicity, England

Category	Actual relevant population Population profile		Winter CIS Unadjusted population Sample profile		Difference	
	Number of people	Proportion	Number of people	Proportion	Absolute difference to actual population	Relative difference to actual population
<b>White</b>	45,783,401	0.81	119,838	0.95	0.14	17%
<b>Asian/Asian British</b>	5,426,392	0.10	3,671	0.03	-0.07	-70%
<b>Black/African /Caribbean/ Black British</b>	2,381,724	0.04	857	0.01	-0.03	-84%
<b>Mixed/Multiple ethnic groups</b>	1,669,378	0.03	1,388	0.01	-0.02	-63%
<b>Other ethnic groups</b>	1,229,153	0.02	749	0.01	-0.01	-73%

Source: Winter Coronavirus (COVID-19) Infection Study from the Office for National Statistics

#### Notes

1. Population figures used for the “actual relevant population” are drawn from Census 2021 for England. Scotland ethnicity data unavailable from the latest Census at the start of the Winter CIS study period.
2. Response population figures are unadjusted and represent the actual response numbers.
3. Absolute difference represents the difference in the unweighted proportion in our sample to the actual population.
4. Relative difference represents the proportion of difference attributed to each group. This is calculated by taking the absolute difference and dividing by the proportion of the actual population.

## Characteristics

Participants are asked to provide their ethnicity and occupation (among other things) in the participant questionnaire to allow analysis of the characteristics of those completing the survey.

The options provided on the questionnaire for ethnicity are [harmonised](#) to allow for consistency and comparability of statistical outputs from different sources across the UK.

Participants are asked to provide employment data. Occupation is provided in a free-text box, while employment sector is selected from 15 categories, which are coded using the [Standard Occupational Classification](#). This again allows for consistency and comparability of outputs across the UK.

## Accessibility and clarity

The ONS recommended format for accessible content is a combination of HTML web pages for narrative, charts, and graphs, with data being provided in usable formats, such as Excel spreadsheets. Our outputs conform to the ONS [Web accessibility policy](#) in terms of formats and font sizes and the presentation of tables and charts.

More details on related releases can be found on the [release calendar on GOV.UK](#). If there are any changes to the pre-announced release schedule, public attention will be drawn to the change and the reasons for the change will be explained fully.

Winter CIS data will be available in our ONS Trusted Research Environment (TRE); this provides access to microdata and disclosive data, which have the potential to identify individuals. Access to such data requires [Approved Researcher accreditation](#).

## Timeliness and punctuality

This publication provides timely and punctual information from the Winter CIS, detailing our analysis on the impact of COVID-19, self-reported long COVID and other respiratory infections. These data are collected, processed and published within a short time frame. Unweighted estimates are provided up until February 2024 to ensure timely release of data.

For more details on related releases, the [GOV.UK release calendar](#) is available online and provides advance notice of release dates.

## Why you can trust our data

The ONS is the UK's largest independent producer of statistics and its national statistical institute. The [Data Policies and Information Charter](#) details how data are collected, secured and used in the publication of statistics. We treat the data that we hold with respect, keeping the data secure and confidential. We use statistical methods that are professional, ethical and transparent. More information about our [data policies](#) is available.

## Provisional estimates and revisions

The general principle applied to the Winter CIS will be that when data are found to be in error, both the data and any associated analysis that has been published by the ONS will be revised in line with our [revisions and corrections policy](#).

There are several reasons why we may wish to revise the survey estimates once they have been published and /or the datasets disseminated, including errors potentially being discovered in raw or derived variables.

While every effort is made to thoroughly check the data before the data are published or released for dissemination, errors do occasionally occur. When errors occur, corrections are made in a timely manner, announced and clearly explained to users in line with the ONS guide to statistical revisions. Work is also undertaken to mitigate the same error happening again, for example, by reviewing and improving code.

# 5 . Test sensitivity and specificity

## Understanding false-positive and false-negative results

The estimates provided in the Winter Coronavirus (COVID-19) Infection Study (Winter CIS) are for the percentage of the private residential population that would have tested positive for COVID-19 on a lateral flow device (LFD), otherwise known as the positivity rate. We do not report the prevalence rate. To calculate the prevalence rate, we would need an accurate understanding of the swab test's sensitivity (true-positive rate) and specificity (true-negative rate). The UK Health Security Agency (UKHSA) do however provide the prevalence rate and include details in their [methodology](#).

## Test sensitivity

Test sensitivity measures how often the test correctly identifies those who have the virus, so a test with high sensitivity will not have many false-negative results. Studies suggest that sensitivity may be somewhere between 55% and 73%.

In the Winter CIS, once a participant tests positive for SARS-CoV-2, they are asked to take repeat tests every other day until they return two negative tests. This repeat testing data are used to estimate the false positive rates over time of LFDs for the Winter CIS cohort.

The sensitivity of LFDs has been found to be lower than that of PCR tests, which were used in the previous ONS Coronavirus (COVID-19) Infection Survey (CIS). It is therefore not advised to compare positivity rates in Winter CIS with those published in CIS.

## Test specificity

Test specificity measures how often the test correctly identifies those who do not have the virus, so a test with high specificity will not have many false-positive results.

The specificity for the LFDs that are used in the Winter CIS has been shown to be very high, with estimates over 99.9%. In the analysis, it is assumed specificity is between 99.89% and 99.98%, based on the existing evidence.

# 6 . Methods used to produce the data

The data collected by the Winter Coronavirus (COVID-19) Infection Study (Winter CIS) enable us to estimate symptoms of COVID-19 and other respiratory infections, and long COVID by main characteristics, and analysing the impact on work, education and healthcare.

## How we collect the data

### Sampling method

The Overview of the Winter CIS section in [Section 3: Quality summary](#) outlines the criteria for which Winter CIS participation is based upon. Here, Coronavirus (COVID-19) Infection Survey (CIS) participants were considered eligible to be invited to Winter CIS if they had been active CIS participants, and consented to be contacted about participating in future research studies.

More information on the initial CIS sampling method from which our participant pool was selected is available in the [Coronavirus \(COVID-19\) Infection Survey quality report: December 2022](#), which was last updated 30 March 2023.

## Study design

Participants who signed up to take part in the Winter CIS were divided into 12 equal groups. The sample was stratified by region, deprivation and ethnicity at the household level before randomly assigning each household to one of the 12 groups. This ensured each group had the same distribution of main characteristics.

Each week, participants in three groups were sent an email invitation (on Tuesday, Wednesday and Thursday) to take the lateral flow device (LFD) test and complete the online questionnaire. Participants had a seven-day window from the day their email was sent to complete the questionnaire. This study design meant that each participant was emailed every four weeks with an invitation to participate.

Each four-week period was referred to as a "wave" and the study consisted of four waves running from November 2023 to March 2024. The seven-day questionnaire windows overlapped as this helps smooth out the number of responses received each day (participants were more likely to return the questionnaire at the beginning of their window).

## Data we collect

We collect data from each participant by using an online questionnaire. Participants are asked to take a lateral flow device, supplied for the study by the UK Health Security Agency (UKHSA), and to record their result in the questionnaire. They are also asked questions on vaccinations (COVID-19), respiratory infections, symptoms experienced in the past seven days, including long COVID, use of healthcare services and time off work (because of a respiratory infection or for general health reasons), working from home, social contact within a work or healthcare setting, and travel. In the last three weeks of the study (from 13 February 2024), four extra questions were added to collect data on working while ill with a respiratory infection.

A follow-up survey, to understand the length of infection, is sent to participants reporting a positive lateral flow test result in the main questionnaire. The follow-up survey requires the participant to record a lateral flow test result every other day until they returned two negative tests. At the beginning of the study this questionnaire was not well understood by the participants; in the first four weeks a high proportion (55%) were returned without two negative tests being reported. Following changes to the questionnaire wording and instructions, this improved to 20% in the final four weeks.

## Response rates

Of the 294,262 CIS participants invited to sign up to Winter CIS, 139,453 consented to take part (47%). When the first wave started, 132,003 participants had been sent their LFD test kits and their consent data had been checked and cleaned. Of these 132,003, 92% returned at least one questionnaire during the study period, 67% returned a questionnaire in all four waves, 14% returned in three, 6% returned in two and 5% returned in one. Response rates for each wave broken down by demographic characteristics are provided in Table 6 of the [ONS dataset Winter Coronavirus \(COVID-19\) Infection Study, England and Scotland](#).

There were an additional 7,450 participants who were added in wave 4 of the study. These participants had either signed up after the initial consent window was closed (in order to maximise the number of participants who signed up from under-represented groups, the consent window was kept open for six weeks past the initial deadline) or had initially returned incorrect data when consenting to take part in Winter CIS.

## How we analyse the data

The primary objective of the study is to understand the impact of COVID-19 infection rates, self-reported symptoms, including long COVID.

The analysis of the questionnaire data focuses on four primary areas: COVID-19 test positivity, reported symptoms, impact of respiratory infections and long COVID.

## COVID-19 test positivity

The primary objective of the positivity estimates is to estimate the number of people in the population who would test positive for COVID-19, with and without symptoms. This is not the same as being infected, see [Section 5: Test sensitivity and specificity](#).

The analysis of the data is a collaboration between the Office for National Statistics (ONS) and UK Health Security Agency (UKHSA).

## Reported symptoms and impact of respiratory infections

The main aim of the reported symptoms analysis is to identify factors associated with reporting influenza-like illnesses (ILI) and provides information on all respiratory illnesses, not just COVID-19. This is achieved through monitoring changes in self-reported symptoms over time and examining how the likelihood of reporting symptoms varies by main characteristics (for example, age).

Data on self-reported respiratory symptoms are used to determine the characteristics of people who reported:

- symptoms consistent with ILI, as classified by the [United States Centers for Disease Control and Prevention \(CDC\)](#), in the last seven days
- symptoms consistent with ILI, as classified by the [European Centre for Disease Control \(ECDC\)](#), in the last seven days
- one or more days off work because of a respiratory infection, in the last 28 days
- one or more GP appointments for a respiratory infection, in the last 28 days
- knowing or thinking they have had a respiratory infection, in the last 28 days

Participants are asked to report whether they had experienced a number of symptoms over the past seven days, including those such as abdominal pain, cough, diarrhoea, fever, loss of taste, loss of smell and shortness of breath. The symptoms reported may not necessarily be because of an acute infection, some symptoms could be chronic conditions (for example, tiredness or wheezing) although this is less likely with symptoms such as fever and sore throat.

Survey weights are applied to make the sample representative of the population in terms of sex, age, region and ethnicity.

## Long COVID

Long COVID status was self-reported by study participants and so misclassification is possible.

Long COVID duration is also calculated differently in Winter CIS compared with the CIS. The duration of long COVID symptoms within the Winter CIS has been calculated from a self-reported date, while within the CIS the duration was calculated from the first suspected coronavirus infection. We are unable to continue using COVID-19 test results to determine the onset of long COVID because of the end of repeat testing and the increase in the proportion of the population having experienced repeat COVID-19 infections.

Collection of the start date for long COVID is subject to recall error. To help participants, those who could not remember the exact date in the month were instructed to enter the 15th. However, 34% of those with long COVID either did not provide a valid date (5%), selected do not know (28%) or prefer not to say (2%). For those who did provide a valid date, 15% selected the 15th of the month.

The positivity estimates published as part of the Winter CIS showed that COVID-19 infections peaked in mid-December of the 2023 to 2024 winter period. Those testing positive during this period will not have had time to reach 84 days post-infection (the cut-off used to define those with long COVID for 12 weeks or more). In Winter CIS, 0.3% of the sample tested positive for COVID-19 less than 12 weeks before the date of their last survey response and reported long COVID symptoms in their final survey response.

## Weighting

Up until February 2024, publications use unweighted estimates. Weighting is applied to ensure survey results are representative of the target population. This adjusts for differential consent or response rates of demographic groups, which could otherwise cause under-representation of younger age groups, for example.

The weighting strategy accounts for:

- the probability of selection in the Winter CIS
- the probability of responding to the Winter CIS
- known population totals

The Winter CIS sample is a subset of the CIS sample. Therefore, design weights are calculated using CIS design weights adjusted for factors correlated to the likelihood of them consenting to further research using logistic regression. Logistic regression is also used to adjust for the probability of sampled people giving consent to take part and for non-response to the survey.

Initial weights are then calibrated to population totals. For England the calibration groups are age group by sex and region separately. For Scotland the calibration groups are age group and sex separately.

## Probability of being selected to take part in Winter CIS

Design weights are used to account for the probability that a person has been selected to take part in the survey. Differences in the probability that a person is selected to take part in the survey can introduce bias to estimates if not accounted for.

The Winter CIS sample is a subset of the CIS population. Therefore, the probability of being selected for Winter CIS is linked to the initial probability of being selected for CIS. For this reason, CIS design weights are used as a basis for the Winter CIS design weights.

Being selected from the CIS sample for inclusion in Winter CIS is based on fitting certain requirements (for example, having an online preference for questionnaire completion), making their probability of selection 1. However, there are factors that influence their likelihood of fitting those selection criteria, for example, some participants being more likely to respond online or consent to further research. To account for this, logistic regression is used to identify the probability of a participant being selected for Winter CIS given they were in CIS. This includes:

- age
- sex (male and female)
- region
- ethnicity (White and separately all other ethnic groups combined)

For each participant, the CIS design weight is multiplied by the inverse of the probability of selection for Winter CIS consent given they were in CIS, to produce the Winter CIS consent design weights.

The selected CIS sample was then asked to provide consent to be part of the Winter CIS sample. Only participants that provided consent were maintained in the main Winter CIS sample, therefore to calculate design weights for Winter CIS participants, an additional model is used to account for the probability of participants consenting to take part in Winter CIS. Consent is defined as a participant indicating in the consent survey that they voluntarily agree to take part in the study. Non-consent is defined as a participant either not responding to the consent survey or indicating in the consent survey that they do not agree to take part in the study. The consent versus non-consent logistic regression model includes:

- age
- sex (male and female)
- region
- ethnicity (White and separately all other ethnic groups combined)
- previous long COVID status (had long COVID in at least one of the last three CIS visits)

For each participant, the Winter CIS consent design weight is multiplied by the inverse of the probability of consent to Winter CIS to produce Winter CIS design weights.

## Probability of responding to Winter CIS

There is further non-response within the sample that consented to take part in Winter CIS. This could result in biased estimates if not accounted for. There is information we can use from previous survey responses that is not available as population totals, for example, previous long COVID status. Therefore, a logistic regression model is used to calculate the probability of response. The model includes:

- age
- sex (male and female)
- region
- ethnicity (White and separately all other ethnic groups combined)
- previous long COVID status (had long COVID in at least one of the last three CIS visits)

Winter CIS design weights are multiplied by the inverse of the probability for responding to create an initial weight. These weights are the basis used for calibration.



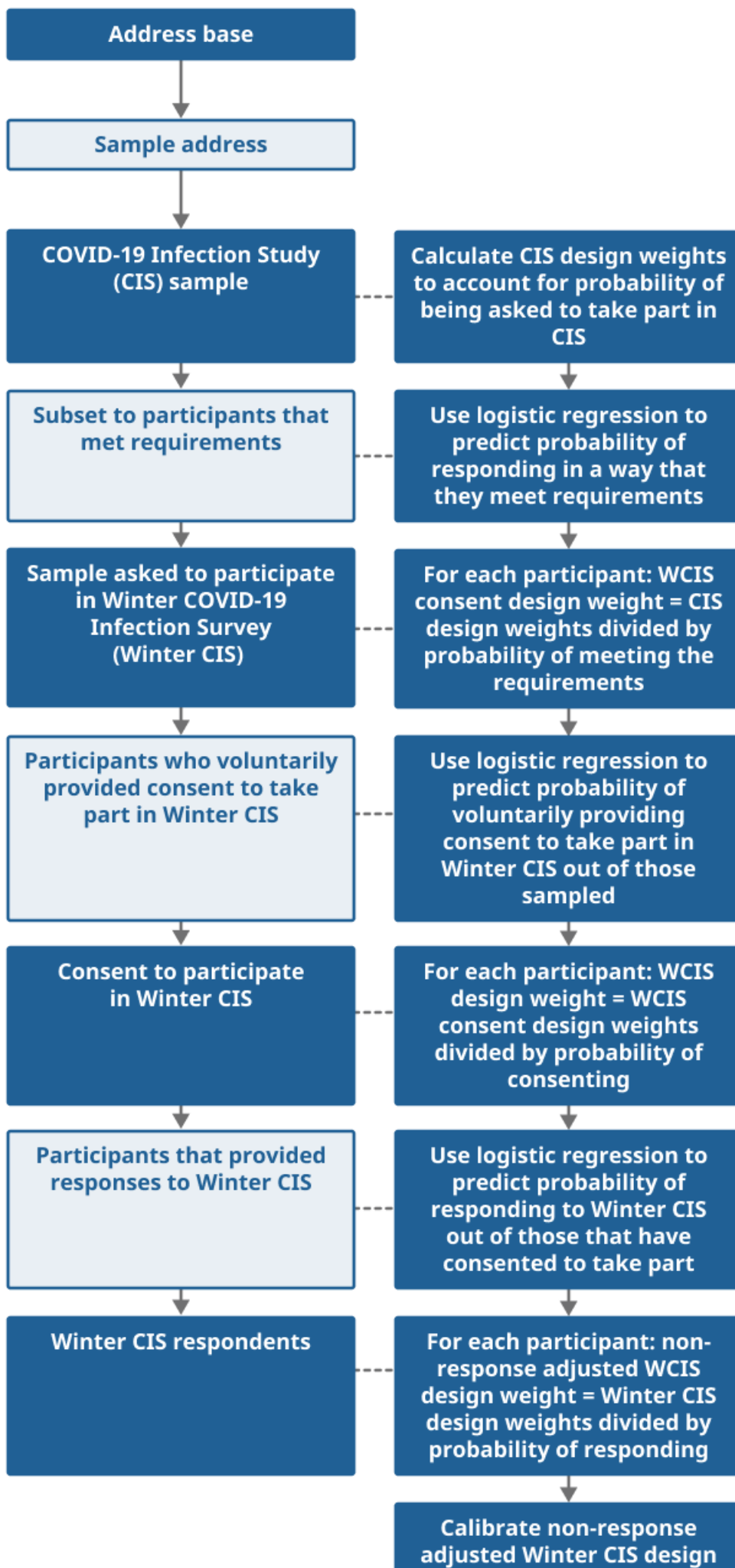
## Calibration and population totals

Calibration is carried out to ensure that the different population groups sum to the known population totals. Calibration can account for more than one set of population totals, summing to the same number, at the same time, for example, population totals based on age by sex, and population totals based on region for the same population. This avoids inferences being made based on small sample sizes, for example, when looking at age by sex by region.

For England the calibration groups are age group by sex and region separately. For Scotland the calibration groups are age group and sex separately.

To obtain population totals for the time period closest to the study dates, population projections were used (with an adjustment to exclude those living in communal establishments). The calibration population totals are provided in Table 5 of the publication. To ensure all data produced for the calibration process are publicly available, the totals for each local authority have been included in the publication; these figures were aggregated to regional level for use in the calibration process.

**Figure 1: Winter CIS sample selection and weighting flow diagram**



Source: Winter Coronavirus (COVID-19) Infection Study from the Office for National Statistics

## 7 . Useful links

The [Winter Coronavirus \(COVID-19\) Infection Study, England and Scotland](#) dataset provides self-reported COVID-19 infections and other respiratory illnesses, including associated symptoms and health outcomes. Joint study with the UK Health Security Agency.

Information about the [Winter Coronavirus \(COVID-19\) Infection Study \(Winter CIS\)](#) and why it was created.

The Coronavirus (COVID-19) Infection Survey (CIS) [methods article](#) provides further information around the CIS survey design, how we process data and how data are analysed. The [study protocol](#) specifies the research for the study.

## 8 . Cite this methodology

Office for National Statistics (ONS), published 24 June 2024, ONS website, methodology, [Winter Coronavirus \(COVID-19\) Infection Study Quality and Methodology Information](#)