Office for National Statistics

Statistical bulletin

# Coronavirus (COVID-19) Infection Survey pilot: England, 24 July 2020

Initial data from the COVID-19 Infection Survey. This survey is being delivered in partnership with IQVIA, Oxford University and UK Biocentre.

Contact:
Emily Connors and Esther Sutherland
infection.survey.analysis@ons.gov.uk
+44 (0)203 9734761

Release date:
24 July 2020

Next release:
To be announced

## Notice

### 14 August 2020

Estimates of positive COVID-19 cases for the South East are incorrect in Coronavirus (COVID-19) Infection Survey pilot bulletins published between 25 June and 7 August 2020. The correct estimates including a back series can be found in publications and accompanying datasets from 14 August 2020 onwards. We apologise for any inconvenience caused.

# Table of contents

1. Main points

2. Number of people in England who had COVID-19

3. Regional analysis

4. Incidence rate

5. Test sensitivity and specificity

6. COVID-19 Infection Survey data

7. Collaboration

8. Glossary

9. Measuring the data

10. Strengths and limitations

11. Related links

# 1 . Main points

- In this bulletin, we refer to the number of current coronavirus (COVID-19) infections within the community population; community in this instance refers to private residential households, and it excludes those in hospitals, care homes or other institutional settings.
- In this bulletin, we use current COVID-19 infections to mean testing positive for SARS-CoV-2, with or without having symptoms, on a swab taken from the nose and throat.
- We estimate around 1 in 2,000 individuals within the community population in England had COVID-19 within the most recent week, from 13 to 19 July 2020.
- This equates to an estimated 27,700 people (95% credible interval: 18,500 to 39,900).
- Modelling of the trend over time suggests that the decline in the number of people in England testing positive on a nose and throat swab has levelled off in recent weeks.
- During the most recent week (13 to 19 July 2020), we estimate there were around 0.52 new COVID-19 infections for every 10,000 people in the community population in England, equating to around 2,800 new cases per day (95% confidence interval: 1,500 to 5,500).
- Modelling of the incidence rate trend suggests that incidence of new infections decreased since mid-May and has now levelled off.

## How the data in this bulletin can be used

The data can be used for:

- estimating the number of current positive cases in the community in England, including cases where people do not report having any symptoms
- identifying differences in numbers of positive cases between different regions
- estimating the number of new cases and change over time in positive cases in England

The data cannot be used for:

- measuring the number of cases and infections in care homes, hospitals and other institutional settings
- estimating the number of positive cases and new infections in smaller geographies, such as towns and cities
- providing information about recovery time of those infected

# 2 . Number of people in England who had COVID-19

## Based on nose and throat swabs, exploratory modelling shows that the number of people in England testing positive has decreased since the start of the study and has now levelled off

Our estimates for rates of people testing positive are based on the result of our trend modelling. This means that these figures cannot be directly compared with those provided in previous bulletins.

During the most recent week of the study[1], we estimate that 27,700 people in England had the coronavirus (COVID-19) (95% credible interval: 18,500 to 39,900)[2]. This equates to 0.05% (95% credible interval: 0.03% to 0.07%) of the population in England or around 1 in 2,000 individuals (95% credible interval: 1 in 3,000 to 1 in 1,400).

These estimates are based on 114,674 swab tests collected over the past six weeks – rather than all swab tests collected over the whole study period as provided in previous publications on 9 July and before – of which 45 individuals from 44 households tested positive. We have updated our model to include only the latest six weeks for which we have data, rather than using data from the start of the study period. As more data are included in the study each week, reducing the time period to six weeks enables us to increase the speed at which we can produce estimates and will allow us to continue to provide timely results in the future.

As this is a household survey, our figures do not include people staying in hospitals, care homes or other institutional settings. In these settings, rates of COVID-19 infection are likely to be different. More information about rates of COVID-19 in care homes can be found in our analysis of the Vivaldi study.

### Figure 1: Modelling shows evidence that the decreases in COVID-19 rates have levelled off

Estimated percentage of the population in England testing positive on nose and throat swabs for the coronavirus (COVID-19) daily since 8 June 2020 (weighted)

### Notes:

1. These statistics refer to infections reported in the community, by which we mean private households. These figures exclude infections reported in hospitals, care homes or other institutional settings.
2. It is important to note that the results are provisional and subject to revision.
3. This analysis was produced by our research partners at the University of Oxford.
4. A break is provided in Figure 1 to show the historical estimates from previous publications alongside the most recent six-week estimates.
5. All estimates are subject to uncertainty; given that a sample is only part of the wider population. The model used to provide these estimates is a Bayesian model: these provide 95% credible intervals. A credible interval gives an indication of the uncertainty of an estimate from data analysis. 95% credible intervals are calculated so that there is a 95% probability of the true value lying in the interval.

Data download

The regression modelling was conducted by our research partners at the University of Oxford. More information about the methods used in the regression model is available in our methodology article.

The model uses all swab test results from the past six weeks. We also present the estimates in non-overlapping 14-day periods. These estimates are available in the dataset which accompanies this bulletin. The 14-day estimates underpin the modelling presented in Figure 1 and are provided for context, but cannot be directly compared. These estimates show a similar trend: that the decrease in the proportion of people testing positive for COVID-19 in previous weeks has levelled off.

To align with the modelling approach used for our headline figures, we have updated the calculation of the 14-day estimates this week. Previously we based these on each participant's most recent test result in the 14-day period, now we have updated this to look at whether they have had any positive result during that period. This has slightly increased the estimates, as some people have a positive test followed by a negative test within the 14-day period. These people would be counted as a positive in our estimates this week (because they have a positive test result during the 2 week period), but a negative (as their last test was negative) last week.

You can find the data for the 14-day period in the data tables.

### Notes: Number of people in England who had COVID-19

1. This is based on model estimates from the week's midpoint, Thursday 16 July.
2. It is important to note that the results are provisional and subject to revision.

# 3 . Regional analysis

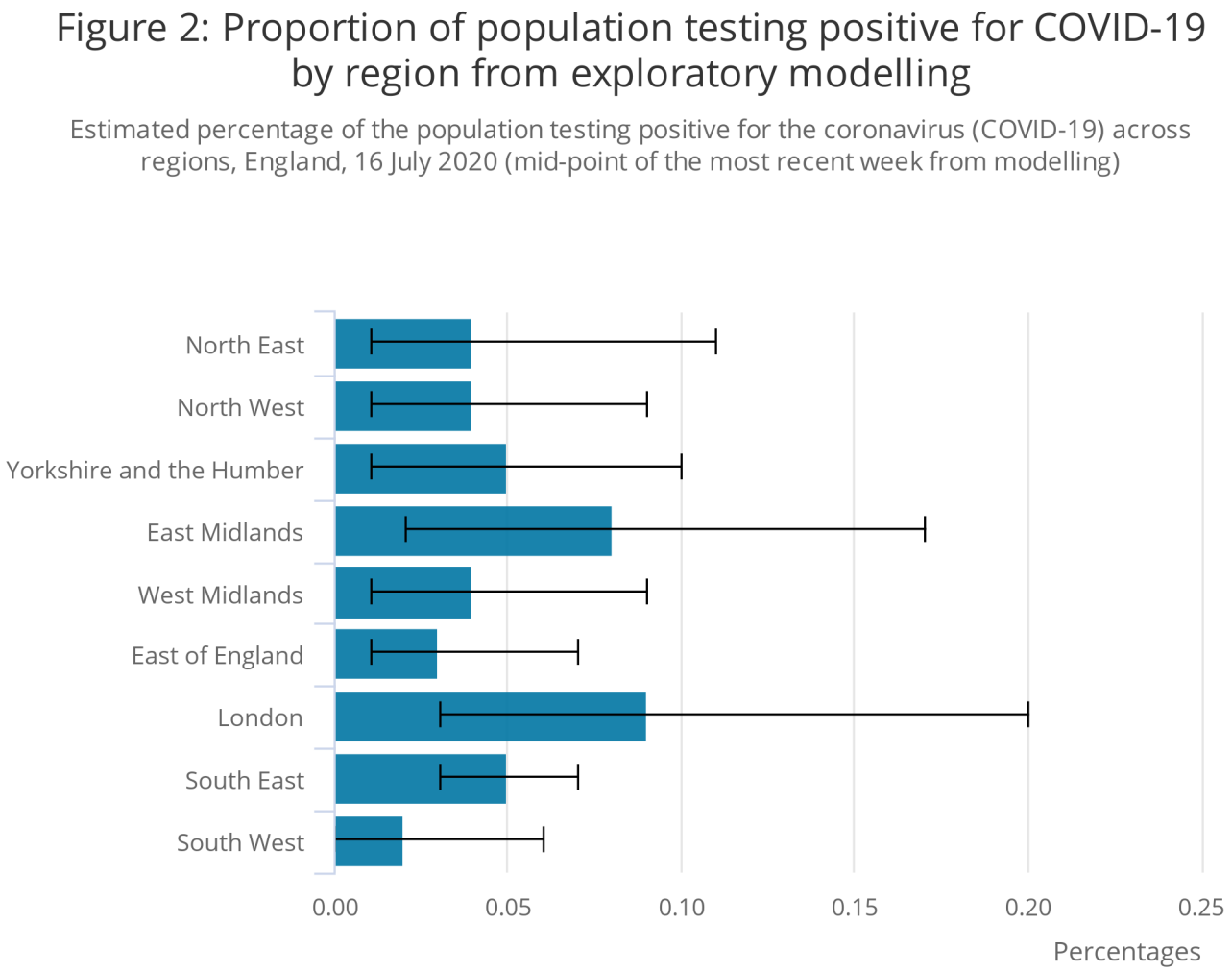
## Based on nose and throat swabs, regional modelling indicates there is no evidence of a difference in the proportion of people testing positive for COVID-19 between regions

The analysis in this section is based on exploratory modelling conducted by our research partners at the University of Oxford.

There is not enough evidence to say with confidence that there is a difference in infection rates between regions. The lower number of people testing positive sampled in the survey within each region means there is high uncertainty in the regional estimates for this period, as indicated by the large credible intervals across most regions.

**Figure 2: Proportion of population testing positive for COVID-19 by region from exploratory modelling**

**Estimated percentage of the population testing positive for the coronavirus (COVID-19) across regions, England, 16 July 2020 (mid-point of the most recent week from modelling)**




**Source: Office for National Statistics – COVID-19 Infection Survey**

**Notes:**

1. These statistics refer to infections reported in the community, by which we mean private households. These figures exclude infections reported in hospitals, care homes or other institutional settings.

2. It is important to note that the results for this period are provisional, as we are still receiving swab test results. This may result in further revisions to the figure.

When comparing between regions over the past six weeks, the rate of people testing positive for COVID-19 in all regions has levelled off. The proportion testing positive by region has been calculated using a similar modelling approach as the national daily trend. This regional modelling is based on estimates from the most recent six-week period.

The analysis is conducted over a six-week period, which means specific positive cases move into and then out of the sample. This causes variability between estimates over time, which is expected given the low number of positive tests by region.

### Figure 3: Exploratory modelling shows that most regions appear level in recent weeks

Estimated percentage of the population in England testing positive on nose and throat swabs for the coronavirus (COVID-19) daily between regions since 8 June 2020

**Notes:**

1. These statistics refer to infections reported in the community, by which we mean private households. These figures exclude infections reported in hospitals, care homes or other institutional settings.

2. It is important to note that the results for this period are provisional, as we are still receiving swab test results. This may result in further revisions to the figure.

Data download

# 4 . Incidence rate

## During the most recent week (12 July to 19 July), we estimate that around 2,800 people became newly infected with COVID-19 per day (95% confidence interval: 1,500 to 5,500)

We have updated our modelling approach for estimating incidence. The model now uses the same Bayesian approach as is used for the positive rate estimates in this bulletin. The model looks at the incidence and positivity rate estimates for each different type of respondent (by age, sex and region) weighted by the percentage of each type in the actual overall population. The data presented in previous bulletins were unweighted. We also now present the incidence rate per day rather than per week.

More information on the methodology of this approach is available.

Using an exploratory modelling approach, we estimate that there were 0.52 new infections per 10,000 people followed for one day (95% confidence interval: 0.28 to 1.00). This also equates to 2,800 new infections per day (95% confidence interval: 1,500 to 5,500).

Our findings suggest that incidence has decreased since the first estimates in mid-May but has since levelled off. Based on the credible intervals, there is no current evidence to suggest changes in incidence rates (Figure 4).

### Figure 4: The latest exploratory modelling shows a decreasing rate of new infections per day since mid-May, which has now levelled off

Estimated numbers of new infections with the coronavirus (COVID-19) per day, England, based on tests conducted daily since 12 May 2020

**Notes:**

1. Credible intervals are large at both ends of the plot because there is less information available. Although we know that individuals have been visited, there is a short delay in getting the associated swab results. The model does not include people when their next swab result is not known, so the sample size for the most recent days is smaller, resulting in wider credible intervals.

2. This model does not control for household clustering.

Data download

The model uses all swab test results. We also present the estimates in non-overlapping 14-day periods. These estimates are available in the dataset that accompanies this bulletin. The 14-day estimates underpin the modelling presented in Figure 4 and are provided for context, but they cannot be directly compared. Our non-overlapping 14-day estimates of incidence are based on only time at risk and new infections in each period, and they are unweighted. These estimates show a similar trend: that the decrease in the proportion of people becoming newly infected with the coronavirus (COVID-19) in previous weeks has levelled off.

We found that the infection rates within households follow a similar trend as for individuals. You can find the household incidence rate in the dataset.

The incidence rate measures the occurrence of new cases of COVID-19. This is calculated by dividing the number of times an individual has a positive test for the first time in the study, having previously tested negative, by the total time everyone is in the study between successive negative tests and up to their first positive test, if any.

To calculate the estimated average number of people becoming newly infected per day, we multiply the incidence rate per week by the community population (54,628,600, see Coverage in Section 9: Measuring the data) and then divide by seven to get the daily rate. We use unrounded numbers to do this, so results will differ if calculated using published rounded numbers.

The incident rate is not the same as the reproduction rate (R), which is the average number of secondary infections produced by one infected person.

# 5 . Test sensitivity and specificity

The estimates provided in Section 2: Number of people in England who had COVID-19 are for the percentage of the private-residential population testing positive for the coronavirus (COVID-19), 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).

While we do not know the true sensitivity and specificity of the test because COVID-19 is a new virus, our data and related studies provide an indication of what these are likely to be. To understand the potential impact of this, we have estimated what prevalence would be in two scenarios using different possible test sensitivity and specificity rates. The results of this show that when these estimated sensitivity and specificity rates are taken into account, the prevalence rate would be slightly higher but still very close to the main estimate presented in Section 2: Number of people in England who had COVID-19.

For this reason, we do not produce prevalence estimates for every analysis, but we will continue to monitor the impacts of sensitivity and specificity in future.

You can find more information on sensitivity and specificity in a paper written by the Office for National Statistics' (ONS') academic partners and in our methods article.

# 6 . COVID-19 Infection Survey data

COVID-19 Infection Survey
Dataset | Released 24 July 2020
Latest findings from the pilot phase of the Coronavirus (COVID-19) Infection Survey.

# 7 . 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, Public Health England (PHE) 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

# 8 . Glossary

## Community

In this bulletin, we refer to the number of coronavirus (COVID-19) infections within the community. Community in this instance refers to private households, and it excludes those in hospitals, care homes or other institutional settings.

## 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. For more information, see our methodology page on statistical uncertainty.

## Credible interval

A credible interval gives an indication of the uncertainty of an estimate from data analysis. 95% credible intervals are calculated so that there is a 95% probability of the true value lying in the interval.

## False-positives and false-negatives

A false-positive result occurs when the tests suggest an individual has COVID-19 when in fact they do not. By contrast, a false-negative result occurs when the tests suggest an individual does not have COVID-19 when in fact they do. For more information on false-positives and false-negatives, see our methods article.

## Incidence rate

Incidence is the rate of occurrence of new cases of the disease over a given period of time. Incidence refers to the number of individuals who have a positive test in the study divided by the time from joining the study to their last test in everyone with more than one test. Individuals who are positive when they join the study are not included in this calculation.

# 9 . Measuring the data

Data presented in this bulletin come from the Coronavirus (COVID-19) Infection Survey, which looks to identify the percentage of the population testing positive for COVID-19 and whether they have symptoms or not. The survey helps track the current extent of infection and transmission of COVID-19 among the population as a whole.

This section of the bulletin provides a short summary of the study data and data collection methods. Our methodology article provides further information around the survey design, how we process data and how data are analysed. The study protocol specifies the research for the study.

## Response rates

Tables 1 and 2 provide information regarding responses to our survey. The current number of households invited to participate in the survey is 57,776, of which 23,811 have enrolled. In responding households, there are 50,793 eligible individuals.

At the start of the pilot study, around 20,000 households were invited to take part, with the aim of achieving data from around 10,000 households. Since the end of May, additional households have been invited to take part in the survey each week (roughly 5,000 a week). This impacts the response rate as it takes time for those invited to respond and enrol.

The response rates cannot be regarded as final response rates to the survey since those who are invited are not given a time limit in which to respond. However, as the likelihood of enrolment decreases over time, we have provided response rate information for those initially asked to take part at the start of the survey (Table 1) where response rates can be considered as relatively final. Separately, we provide response rates for those invited from 31 May (Table 2), where enrolment is still continuing.

Table 1: Responses to the COVID-19 Infection Survey (initial invitation, from 26 April 2020)

| | Households | Individuals |
|---|---|---|
| | % of total | % of total |
| Households invited to take part (total) | 20,276 100% | |
| Households enrolled | 10,329 51% | |
| Completed households (provided at least one swab) | 10,221 50% | |
| Eligible individuals in responding households (total) | | 22,254 100% |
| Individuals who provided first swab | | 21,159 95% |
| Individuals who agreed to continue | | 20,122 90% |

Source: Office for National Statistics – COVID-19 Infection Survey

Notes

1. The set sample for this study is based on the achieved sample from a previous social survey who agreed to take part in future studies. Back to table

Table 2: Responses to the COVID-19 Infection Survey (extension weeks, from 31 May 2020)

| | Households | Individuals |
|---|---|---|
| | % of total | % of total |
| Households invited to take part (total) | 37,500 100% | |
| Households enrolled | 13,482 36% | |
| Completed households (provided at least one swab) | 11,169 30% | |
| Eligible individuals in responding households (total) | | 28,539 100% |
| Individuals who provided first swab | | 22,471 79% |
| Individuals who agreed to continue | | 21,370 75% |

Source: Office for National Statistics – COVID-19 Infection Survey

Notes

1. The set sample for this study is based on the achieved sample from a previous social survey who agreed to take part in future studies. Back to table

## Coverage

Only England is included in this pilot phase of the study. We intend for the full survey to expand the size of the sample over the next 12 months and look to cover people across all four UK nations. Only private residential households, otherwise known as the target population in this bulletin, are included in the sample. People in hospitals, care homes and other institutional settings are not included.

The overall target population used in this study is 54,628,600.

## Analysing the data

All estimates presented in this bulletin are provisional results. As swabs are not necessarily analysed in date order by the laboratory, we have not yet received test results for all swabs taken on the dates included in this analysis. Estimates may therefore be revised as more test results are included.

This is a pilot study where the analysis is developed at pace, and these quality enhancements may lead to minor changes in estimates, for example, the positive test counts across the study period.

## Changes to the analysis in this bulletin

In this bulletin, we have made several changes to the way that we report results compared with previous bulletins.

When analysing incidence, we have changed our modelling approach to estimating incidence. The model now aligns with the Bayesian modelling used in this bulletin for the positive rate estimates. The model looks at the incidence rate estimates for each different type of respondent (by age, sex and region) weighted by the percentage of each type in the actual overall population. To note, the weighting in the model takes into account the entire population, though not everyone is at risk (for example, if someone has recently tested positive for COVID-19). The impact of this on the estimates is negligible and therefore does not change the overall messaging of the modelling.

We have also changed the method we use to produce our 14-day estimates. We now count any positive tests in that time period, whereas previously if an individual had a positive test followed by a negative test, they would have been counted as a negative result (that is, we used the result of the most recent test). This is to align these data with those used in the modelling, which is our headline result.

More information about the methods used in the modelling can be found in our methodology article.

## Other studies

This study is one of a number of studies that look to provide information around the coronavirus pandemic within the UK.

### Department for Health and Social Care (DHSC) data

Public Health England (PHE) present data on the total number of laboratory-confirmed cases in England, which capture the cumulative number of people in England who have tested positive for COVID-19. Equivalent data for Wales, Scotland and Northern Ireland are also available. These statistics present all known cases of COVID-19, both current and historical. The large sample size means it is possible to present known cases at local authority level.

The NHS Test and Trace scheme was launched on 28 May. The Test and Trace service ensures that anyone who develops symptoms of COVID-19 can quickly be tested to find out if they have the virus. It includes targeted asymptomatic testing of NHS and social care staff and care home residents. Additionally, it helps trace close recent contacts of anyone who tests positive for COVID-19 and, if necessary, notify them that they must self-isolate.

In comparison to PHE data and NHS Test and Trace data, the statistics presented in this bulletin take a representative sample of the community population (those in private residential households) in England, including people who are not otherwise prioritised for testing. This means that we can estimate the number of people in the community population in England with COVID-19 who do not report symptoms. This is something that is currently missing from PHE and Test and Trace data.

### COVID Symptom Study (ZOE app and King's College London)

The COVID Symptom Study app allows users to log their health each day, including whether or not they have symptoms of COVID-19. The study aims to predict which combination of symptoms indicate that someone is likely to test positive for COVID-19. The app was developed by the health science company ZOE with data analysis conducted by King's College London. Anyone over the age of 18 years can download the app and take part in the study. Respondents can report symptoms of children.

The study estimates the total number of people with symptomatic COVID-19 and the daily number of new cases of COVID-19 based on app data and swab tests taken in conjunction with the Department for Health and Social Care (DHSC). The study investigates the "predictive power of symptoms", and so the data do not capture people who are infected with COVID-19 but who do not display symptoms.

Unlike the data presented in this bulletin, the COVID Symptom Study is not a representative sample of the population. It is reliant on app users and so captures only some cases in hospitals, care homes and other communities where few people use the app. To account for this, the model adjusts for age and deprivation when producing UK estimates. The larger sample size allows for detailed geographic breakdown.

### Real-time Assessment of Community Transmission-1 and -2 (REACT-1 and -2)

Like our study, the Real-time Assessment of Community Transmission-1 (REACT-1) survey involves taking swab samples to test for COVID-19 antigens to estimate the prevalence and transmission of the virus that causes COVID-19 in the community. The study currently involves around 120,000 participants aged 5 years and above, selected from a random cross-section sample of the general public from GP registration data, which allows for more detailed geographic breakdowns of infection rates than are currently possible within our study. Trends in infection by characteristics, such as age, sex, ethnicity, symptoms and key worker status, are also possible through the study. The REACT-2 study uses a finger prick test to generate data for antibody analysis.

One of the main differences from our COVID-19 Infection Survey is that the REACT surveys do not require follow-up visits, as the study is interested primarily in prevalence at a given time point. Consequently, the incidence rate cannot be calculated from the REACT studies. It is also important to note that blood samples in the REACT-2 study are self-administered, rather than taken by a trained nurse, phlebotomist or healthcare assistant.

### PHE antibody data

PHE also publish an estimate of the prevalence of antibodies in the blood in England using blood samples from healthy adult blood donors. PHE provide estimates by region and currently do not scale up to England. Estimates in this bulletin and those published by PHE are based on different tests; PHE estimates are based on testing using the Euroimmun assay method, while blood samples in our survey are tested for antibodies by research staff at the University of Oxford using a novel ELISA. For more information about the antibody test used in this bulletin, see the COVID-19 Infection Survey protocol.

### Next steps

This edition of the bulletin presents headline analysis of the overall number of people infected with COVID-19, the regional positivity rate, incidence rate and antibodies. We provide headline figures once a week, to give regular, concise and high-quality information on COVID-19 within the community.

Our recent release, Coronavirus (COVID-19) infections in the community, offers more detailed analysis, which includes further exploration of the characteristics of those with COVID-19, such as age, sex, working location and occupation. We will also include further exploration of ethnicity when we have a large enough sample size to provide reliable analysis.

## 10 . Strengths and limitations

These statistics have been produced quickly in response to developing world events. The Office for Statistics Regulation, on behalf of the UK Statistics Authority, has reviewed them against several important aspects of the Code of Practice for Statistics and regards them as consistent with the Code's pillars of trustworthiness, quality and value.

The estimates presented in this bulletin contain uncertainty. There are many sources of uncertainty, including uncertainty in the test, in the estimates and in the quality of data collected in the questionnaire. Information on the main sources of uncertainty are presented in our methodology article.

## 11 . Related links

COVID-19 Infection Survey (Pilot): methods and further information
Methodology article | Updated 23 July 2020
Information on the methods used to collect and process the data and calculate the statistics produced from the Coronavirus (COVID-19) Infection Survey (pilot).

Coronavirus (COVID-19) infections in the community in England: July 2020
Article | Updated monthly
Analysis on the latest data about the characteristics of those who test positive for COVID-19 in England, from the COVID-19 Infection Survey.

Coronavirus (COVID-19) latest data and analysis
Web page | Updated as and when data become available
Latest data and analysis on the coronavirus in the UK and its effect on the economy and society.

Coronavirus (COVID-19) roundup
Article | Updated as and when data become available
Catch up on the latest data and analysis related to the coronavirus pandemic and its impact on our economy and society.

Deaths registered weekly in England and Wales, provisional
Bulletin | Released 21 July 2020
Provisional counts of the number of deaths registered in England and Wales, including deaths involving COVID-19, by age, sex and region, in the latest weeks for which data are available.

New survey results provide first snapshot of the current number of COVID-19 infections in England
Blog | Released 14 May 2020
A large study jointly led by the Office for National Statistics (ONS) in partnership with the Universities of Oxford and Manchester, Public Health England (PHE), and Wellcome Trust is tracking infections within a representative sample of people of all ages across England. This blog explains what these mean, why they are important and how to compare this survey with other COVID-19 estimates.

COVID-19 Infection Survey
Article | Updated 14 May 2020
Whether you have been invited to take part, or are just curious, find out more about our COVID-19 Infection Survey and what is involved.