Office for National Statistics

Statistical bulletin

Coronavirus (COVID-19) Infection Survey pilot: England, 17 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 Philippa Haughton
infection.survey.analysis@ons.gov.uk
+44 (0)203 973 4761

Release date:
17 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

- 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.
- To improve the stability of the estimates, we now base our headline estimates on our exploratory modelling results, resulting in slightly higher figures than in the previous bulletin. This is not a change in the rate of people testing positive.
- We estimate around 1 in 2,300 individuals within the community population in England had COVID-19 within the most recent week, from 6 July to 12 July 2020.
- This equates to an estimated 24,000 people (95% credible interval: 15,000 to 34,000).
- During the most recent week (6 July to 12 July), we estimate there were around two new COVID-19 infections for every 10,000 individuals in the community population in England, equating to around 1,700 new cases per day (95% confidence interval: 700 to 4,200).
- Between 26 April and 8 July, 6.3% of people tested positive for antibodies against SARS-CoV-2 on a blood test, suggesting they had the infection in the past.

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 now based on the result of our trend modelling. This means that these figures cannot be directly compared to those provided in previous bulletins.

During the most recent week of the study[1], we estimate that 24,000 people in England had the coronavirus (COVID-19) (95% credible interval: 15,000 to 34,000)[2]. This equates to 0.04% (95% credible interval: 0.03% to 0.06%) of the population in England or around 1 in 2,300 individuals (95% credible interval: 1 in 3,600 to 1 in 1,600).

We have moved to modelled estimates from non-overlapping 14-day periods for our headline figure for people testing positive for COVID-19. This is because the modelling includes all available data from participants in the modelled time period rather than just the most recent positive swab and allows us to estimate a smooth change day-by-day. In contrast, when we group the data into non-overlapping 14-day periods, as in Figure 2, our estimates jump between the periods that we create. This produces artificial steps between time periods.

The modelled estimates are slightly higher than the 14-day estimate as presented in the previous bulletin, but are still within the credible intervals. This change should not be interpreted as a change in the rate of people testing positive.

These estimates are based on 112,776 swab tests collected over the past six weeks – rather than all swab tests collected over the whole study period as provided in previous publications – of which 39 individuals from 37 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 is included in the model each week, reducing the time period to six weeks enables us to improve 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 Survey.

Figure 1: Modelling shows evidence that the decrease 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 1 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 side the most recent 6 week estimates. Further context on these modelled estimates is provided in Section 10, see also Figure 7.

Download this chart

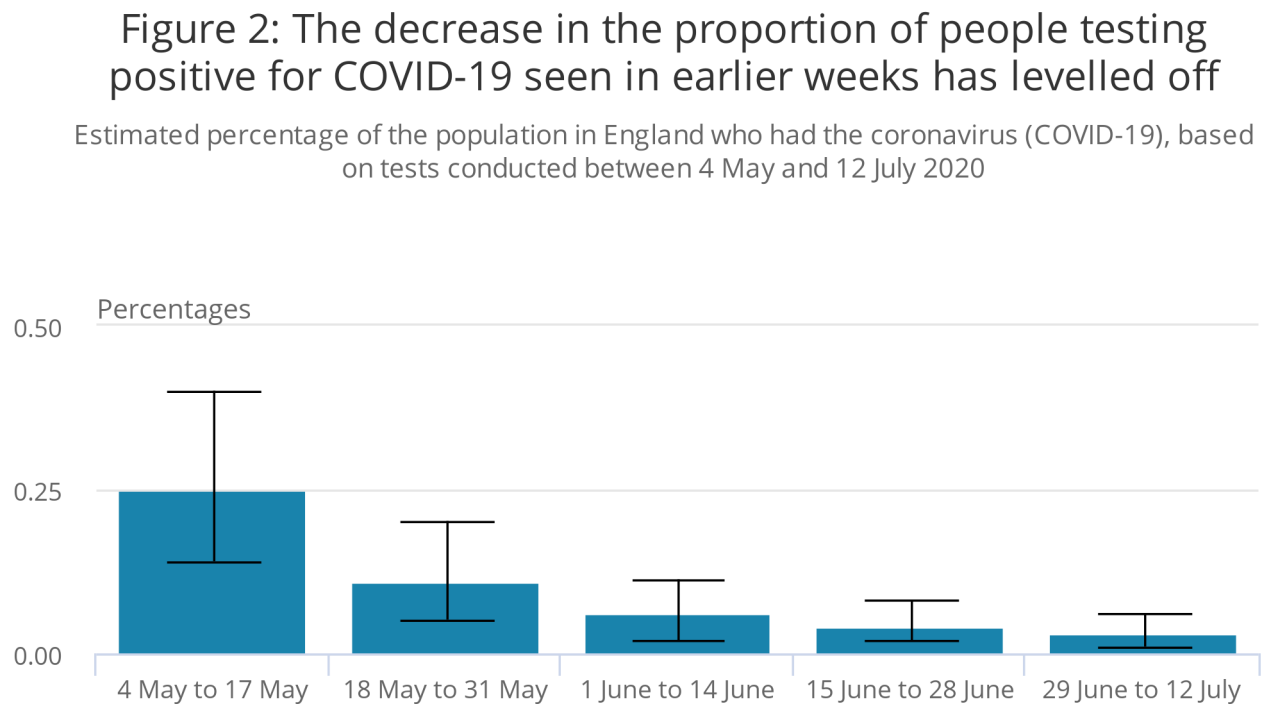
.XLSX

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. Our previous non-overlapping 14-day estimates are based on the most recent swab test result for each individual in each period. 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. The estimate for the latest non-overlapping fortnight is based on the most recent positive swabs, of which there were nine.

Figure 2: The decrease in the proportion of people testing positive for COVID-19 seen in earlier weeks has levelled off

Estimated percentage of the population in England who had the coronavirus (COVID-19), based on tests conducted between 4 May and 12 July 2020



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's 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.

3. The 14-day time periods presented in Figure 2 overlap with those presented in our previous publication, so direct comparisons are not possible.

4. All estimates are subject to uncertainty, given that a sample is only part of the wider population. The 95% confidence intervals are calculated so that, if we were to repeat this study many times, with many different samples of households, then 95% of the time the confidence intervals would contain the true value that we are seeking to estimate.

Information about the characteristics of people testing positive for COVID-19 can be found in our recent release, Coronavirus (COVID-19) infections in the community in England.

In addition, information on how our estimates compare with other sources is available in Section 10: Measuring the data.

Footnotes for Number of people in England who had COVID-19:

1. This is based on model estimates from the week's midpoint, Thursday 9 July.

2. 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.

3 . Regional analysis

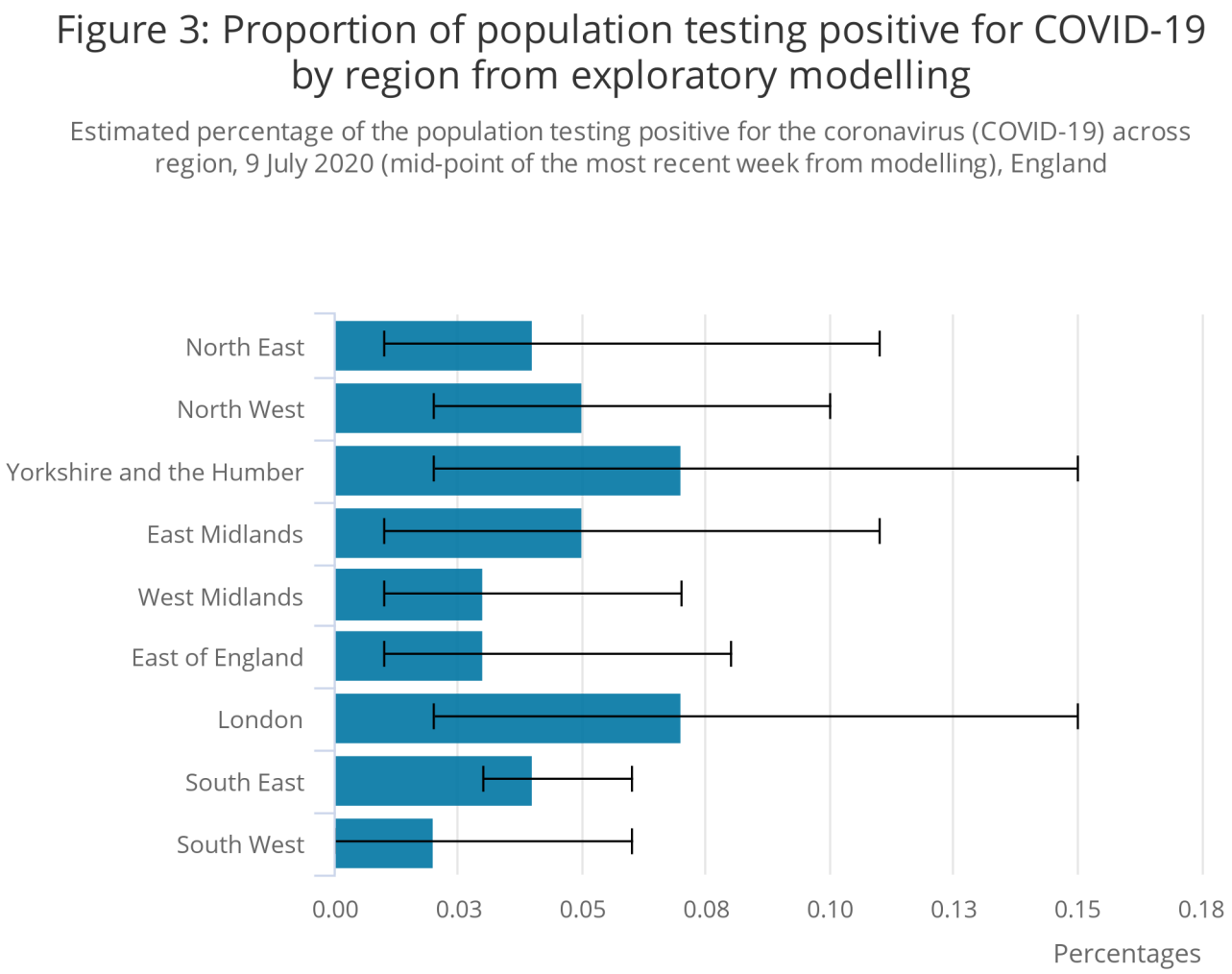
Based on nose and throat swabs, regional modelling indicates there is no evidence of a difference in the proportion of people testing positive for COVID-19 between regions

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 3: 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 region, 9 July 2020 (mid-point of the most recent week from modelling), England



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's 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 most areas appears to have 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.

Figure 4: Exploratory modelling shows that most regions appear level in recent weeks

Estimated percentage of the population testing positive for coronavirus (COVID-19) on nose and throat swabs daily between regions since 1 June 2020, England

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's 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.

Download this chart

.XLSX

4 . Incidence rate

During the most recent week (6 July to 12 July), we estimate that around 1,700 people became newly infected with COVID-19 per day (95% confidence interval: 700 to 4,200)

Using an exploratory modelling approach, we estimate that there were two new infections per 10,000 people followed for one week, or 1,700 new infections per day (95% confidence interval 700 to 4,200). This equates to an incidence rate of 0.02 (95% confidence interval: 0.01 to 0.05) of people followed for one week.

We have moved to reporting modelled estimates as our headline figures, rather than non-overlapping 14-day periods to improve the stability of our estimates and remove artificial steps between time periods. This is discussed more in Section 2.

The modelling shows that there appears to be a decreasing rate of new infections since mid-May, which has now levelled off (Figure 5).

More details on the methodology of the incidence model can be found in our methodology paper.

Figure 5: The latest exploratory modelling shows a decreasing rate of new infections since mid-May, which has now levelled off

Estimated numbers of new infections with the coronavirus (COVID-19), England, based on tests conducted daily since 11 May 2020

Notes:

1. Confidence intervals are large at the end 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 confidence intervals.

2. The model only partially controls for within household clustering by increasing the estimate of variability.

3. We have updated the modelling slightly this week, so that each new infection is counted as occurring at the midpoint of between the first positive test and the most recent previous negative test, rather than at the time of the positive test itself. This is because we do not know exactly when an individual became newly infected, only that it was sometime between these two most recent tests.

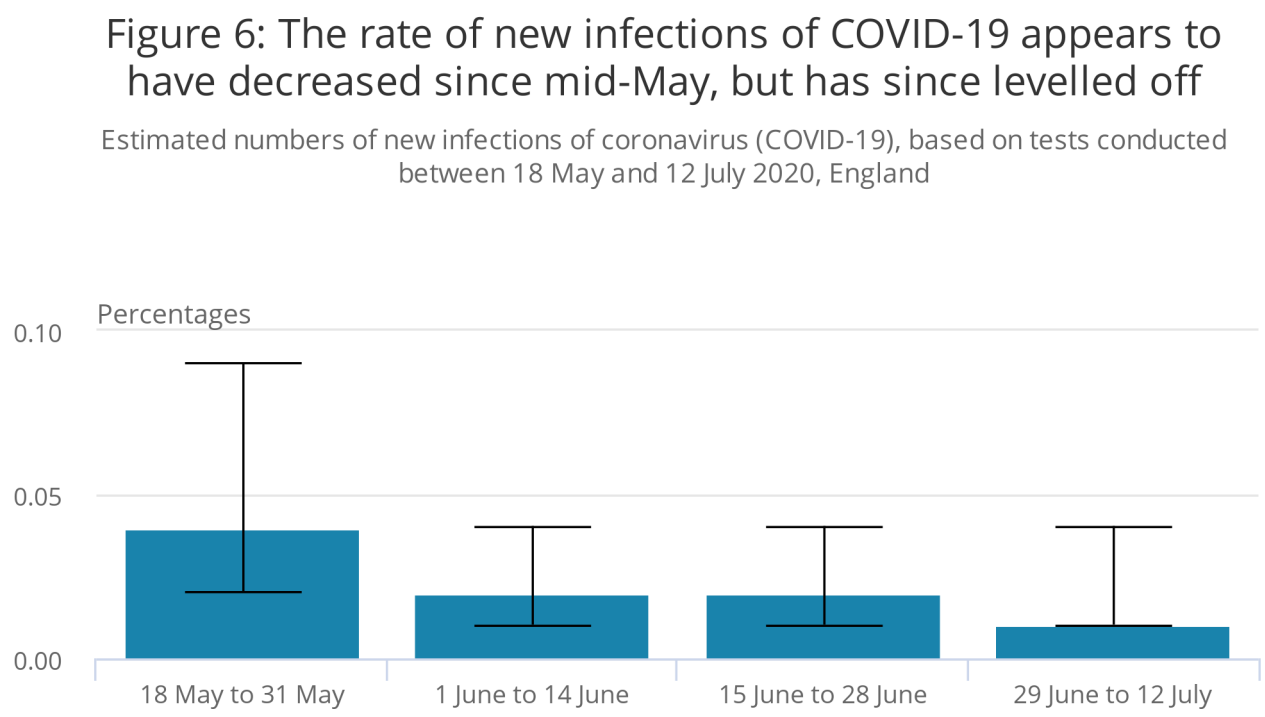
Download this chart

.XLSX

The model uses all swab test results. Our previous non-overlapping 14-day estimates are based on only the most recent swab test result for each individual in each period. These estimates show a similar trend: that decrease in the proportion of people testing positive for COVID-19 in previous weeks has levelled off.

Figure 6: The rate of new infections of COVID-19 appears to have decreased since mid-May, but has since levelled off

Estimated numbers of new infections of coronavirus (COVID-19), based on tests conducted between 18 May and 12 July 2020, England



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's 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.

We found that the infection rates within households follow a similar trend as for individuals. To avoid the impact of household clustering, our household analysis uses only the first new infection case in any given household.

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 week, we multiply the incidence rate per week by the community population (54,628,600, see Coverage) then dividing by seven to get the daily rate. We use unrounded numbers to do this, so results might differ if the published rounded numbers are used.

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

Unlike the analysis in Section 2: Number of people in England who had COVID-19, these estimates have not been weighted to be representative of the target population in England because of the relatively small numbers of new infections in the sample.

5 . Antibody data

Around 6.3% of people who provided blood samples tested positive for antibodies to COVID-19

As of 29 June 2020, 6.3% (95% confidence interval: 5.0% to 7.8%) of individuals from whom blood samples were taken tested positive for antibodies to the coronavirus (COVID-19). This estimate is weighted and equates to 1 in 16 people, or 2.8 million people, in England. Note that blood samples are taken only from those aged 16 years and over.[1]

The analysis in this bulletin is based on the most recent test result from 4,309 blood samples received since the start of the study on 26 April 2020. Of those who have provided blood samples, 4,309 providing blood samples, 211 tested positive for antibodies, and 696 individuals have provided more than one sample.

One way the body fights infections like COVID-19 is by producing small particles in the blood called antibodies. It takes between two and three weeks for the body to make enough antibodies to fight the infection but once a person recovers, antibodies remain in the blood at low levels. This is what helps to prevent individuals from getting the same infection again. We try to measure the presence of antibodies to work out who has had COVID-19 in the past.

More information on how our estimates compare with other studies can be found in Section 10: Measuring the data.

Notes for Antibody data:

1. Changes in the rate of people testing positive for antibodies between bulletins should not be interpreted as a trend over time. This is because it relates to a change in the number of individuals whose blood has now been tested for antibodies and the introduction of the weighting method in the bulletin published on 9 July.

6 . 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 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).

Whilst 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 we have estimated what prevalence would be in two scenarios using different possible test sensitivity and specificity rates.

What is test sensitivity and specificity?

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 85% and 98%.

Our study involves participants self-swabbing under the supervision of a study healthcare worker. It is possible that some participants may take the swab incorrectly, which could lead to more false-negative results. However, research suggests that self-swabbing under supervision is likely to be as accurate as swabs collected directly by healthcare workers.

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.

We know the specificity of our test must be very close to 100% as the low number of positive tests in our study means that specificity would be very high even if all positives were false. For example, in our most recent six weeks of data, 50 of the 112,776 total samples tested positive. Even if all these positives were false, specificity would still be 99.96%.

We know that the virus is still circulating so it is extremely unlikely that all these positives are false. However, it is important to consider whether many of the small number of positive tests we do have might be false. There are a couple of main reasons we do not think that is the case.

Symptoms are an indication that someone has the virus, therefore if there are many false-positives we would expect to see more false-positives occurring among those not reporting symptoms. If that were the case, then risk factors such as working in healthcare would be more strongly associated with symptomatic infections than with asymptomatic infections. However, in our data the risk factors for testing positive are equally strong for both symptomatic and asymptomatic infections.

The percentage of individuals reporting no symptoms among those testing positive has remained stable over time despite substantial declines in the overall number of positives. If false positives were high, the percentage of individuals not reporting symptoms among those testing positive would increase when the true prevalence is declining.

More information on sensitivity and specificity is included in a paper written by ONS' academic partners.

What is the impact on our estimates?

We have used Bayesian analysis to calculate what prevalence would be in two different scenarios, one with medium sensitivity and the other with low sensitivity. The table below shows these results alongside the weighted estimate of percent testing positive from the most recent two weeks of data.

Scenario 1 (medium sensitivity, high specificity)

Based on similar studies, the sensitivity of the test used is plausibly between 85% and 95% (with around 95% probability) and as noted above the specificity of the test is above 99.9%.

Scenario 2 (low sensitivity, high specificity)

To allow for the fact that individuals are self-swabbing, Scenario 2 assumes a lower overall sensitivity rate of 60% (or between 45% and 75% with 95% probability), incorporating the performance of both the test itself and the self-swabbing. This is lower than we expect the true value to be for overall performance but provides a lower bound.

Table 1: The effects of test sensitivity on estimates

Reference period: 29 June to 12 July		95% Credible Interval	
		Lower	Upper
Estimated average % of the population that had COVID-19 (weighted)	0.03	0.01	0.06
Prevalence rate in scenario 1 (medium sensitivity, high specificity)	0.04	0.02	0.08
Prevalence rate in scenario 2 (low sensitivity, high specificity)	0.06	0.03	0.12

Source: Office for National Statistics

The results 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. This is the case even in Scenario 2, where we use a sensitivity estimate that is lower than we expect the true value to be. 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.

7 . COVID-19 Infection Survey data

COVID-19 Infection Survey
Dataset | Released 17 July 2020
Latest findings from the pilot phase of the Coronavirus (COVID-19) Infection Survey.

8 . 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 and Wellcome Trust.

9 . 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 methodology page on sensitivity and specificity analysis

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.

10 . Measuring the data

Data presented in this bulletin comes 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 will help 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 2 and 3 provide information regarding responses to our survey. The current number of households invited to participate in the survey is 53,320, of which 22,124 have enrolled. In responding households, there are 47,205 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 enroll.

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 2) where response rates can be considered as relatively final. Separately we provide response rates for those invited from 31 May (Table 3), where enrolment is still continuing.

Table 2: Responses to the COVID-19 Infection Survey (initial invitation, from 26 April)

	Households		Individuals	
		% of Total		% of Total
Households invited to take part (total)	20,276	100%		
Households enrolled	10,331	51%		
Completed households (provided at least one swab)	10,215	50%		
Eligible individuals in responding households (total)			22,237	100%
Individuals who provided first swab			21,776	98%
Individuals who agreed to continue			18,510	83%

Source: Office for National Statistics

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

	Households		Individuals	
		% of Total		% of Total
Households invited to take part (total)	33,044	100%		
Households enrolled	11,793	36%		
Completed households (provided at least one swab)	9,801	30%		
Eligible individuals in responding households (total)			24,968	100%
Individuals who provided first swab			20,002	80%
Individuals who agreed to continue			17,002	68%

Source: Office for National Statistics

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

More about coronavirus

- Find the latest on coronavirus (COVID-19) in the UK.
- All ONS analysis, summarised in our coronavirus roundup.
- View all coronavirus data.
- Find out how we are working safely in our studies and surveys.

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:

- we use modelled estimates for the headline number of people testing positive rather than 14-day non-overlapping time periods
- we use modelled estimates for the headline number for incidence rather than 14-day non-overlapping time periods
- we base our incidence calculation on new infections occurring at the mid-point between the last negative and first positive test, rather than the point of the positive test
- we update figures for the most recent six weeks of the study for people testing positive for the virus, rather than since the study began

We now present modelled estimates as our headline estimates for positivity rate because the modelling includes all available data from participants in the modelled time period and allows us to estimate a smooth change day-by-day compared to our non-overlapping 14-day periods.

It is important to note that our two sets of estimates (modelled and non-modelled) give very similar results (as discussed in the bulletin). Both estimates indicate the same trend of infections having levelled off in recent weeks.

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

We now base our incidence calculation on new infections occurring at the mid-point between the last negative and first positive test result rather than the point of the positive test result because this gives more stable estimates towards the end of the study period. We do not know when the new infection occurred, only that it is was at some point between the last negative and first positive. At the end of the study period, we are still waiting for swab test results, so do not have many individuals included in the analysis. Counting positive tests as new infections at the time of the visit artificially inflates the incidence rates in the last few days.

We now provide estimates of current infections based on the latest six weeks of data for people testing positive for the virus, rather than the whole study period because this gives the most timely estimates. Timeliness is an important aspect of statistical quality. The models that we use are very complex and so take a long time to process data. As our survey continues to grow, using the whole study period would affect our ability to report latest results quickly.

Figure 7 shows the latest six-week estimates overlaid on the historic time series of the whole study period from last week's bulletin (updated results are not yet available due to time to process). The smoothing effect of the model is apparent, but the results are similar. The latest six-week estimates are lower than the historic data, because they do not take into account the higher number of positive test results that were registered at the start of the period.

Figure 7: Comparison between latest six-week modelling and previously reported modelling

Estimated percentage of the population in England testing positive on nose and throat swabs for the coronavirus (COVID-19) daily from 26 April to 5 July 2020 and since 1 June 2020 (weighted)

Download this chart

.XLSX

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/KCL)

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 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 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 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 and above, selected from a random cross section sample of the general public from GP registration data, which allows for more detailed geographical 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 to 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 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.

11 . 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 the quality of data collected in the questionnaire. Information on the main sources of uncertainty are presented in our methodology article.

12 . Related links

COVID-19 Infection Survey (Pilot): methods and further information
Methodology article | Updated 6 July 2020
Information on the methods used to collect the data, process it, 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 | 14 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.