Office for National Statistics

Article

Coronavirus (COVID-19) Infection Survey technical article: analysis of reinfections of COVID-19: June 2021

This release provides data about reinfections of COVID-19 from the COVID-19 Infection Survey. This analysis has been produced in partnership with University of Oxford.

Contact:
Elizabeth Pereira, Steffi Giji and Victoria Gabb
infection.survey.analysis@ons.gov.uk
+44 (0)1329 447683

Release date:
29 June 2021

Next release:
To be announced

Table of contents

1 . Main points

- Instances of reinfection with coronavirus (COVID-19) were rare in the survey.
- There is evidence that COVID-19 reinfections are milder than initial infections.
- Viral load was lower in reinfections than initial infections.
- Self-reported symptoms were less common in episodes of reinfection than in initial infections.

2 . Overview

This article presents analysis on the number of COVID-19 reinfections identified among participants of the Coronavirus (COVID-19) Infection Survey and the characteristics of reinfections in relation to viral load (as inferred from a cycle threshold (Ct) value) and presence of self-reported symptoms.

To identify reinfections, we have included all data from across the UK collected since the survey commenced on 26 April 2020 up to 5 June 2021.

The risk of reinfection varies from person to person, depending on when they were first infected. People who were first infected in the early part of the survey have had more opportunity to become reinfected compared with someone who has experienced their first infection more recently. Therefore, this analysis uses "participant days at risk" to determine the number of reinfections.

Overall, there were an estimated 15.2 reinfections for every 100,000 participant days at risk (95% confidence interval: 12.7 to 18.0), averaged for the entire at-risk period. This means that once participants had accumulated 100,000 days at risk of reinfection between them, we would expect there to have been around 15 reinfections in that time.

The estimated rate for reinfections with a strong positive test (with Ct less than 30) was 2.7 per 100,000 participant days at risk (95% confidence interval: 1.8 to 4.1) up to 5 June 2021. This means that once participants had accumulated 100,000 days at risk of reinfection between them, we would expect there to have been around three reinfections with a strong positive test (indicating a high viral load).

The risk of reinfection peaked around the first month after becoming at risk for reinfection and declined afterwards. The estimated rate for all reinfections was 5.0 per 100,000 participant days at risk (95% confidence interval: 3.1 to 8.1) 90 days after becoming at risk for a reinfection.

Reinfection episodes, on average, had a lower viral load (median: 32.8, interquartile range: 30.6 to 34.0) than initial infection episodes (median: 24.5, interquartile range: 19.7 to 31.2).

More individuals reported symptoms in their initial infection episode within 35 days after the first observed positive test (57.1%, 95% confidence interval: 48.6% to 65.2%) than in their reinfection episode (21.8%, 95% confidence interval: 15.6% to 29.6%).

3 . Number of reinfections identified

Key definitions relating to reinfection, at-risk period and initial episodes

The sample for this analysis is individuals in the survey who have had at least one positive test recorded in the survey and meet our criteria for being "at risk" of reinfection where:

- 90 days has elapsed since an individual's first positive test in thesurveyand their most recent test result was negative
- if 90 days has not passed since their first positive test in the survey, the individual's last positive test has been followed by four consecutive negative tests

For the purposes of this analysis, we refer to these individuals as "at risk" of reinfection. An individual being classified as "at risk" reflects that it is possible for a positive test of theirs to be considered a reinfection. The "at-risk period" refers to the period following the first time we could have defined a reinfection.

A reinfection is therefore defined as when an individual who meets these criteria has a positive test. The 90-day threshold is arbitrary but follows recommendations from Public Health England and the Centre for Disease Control and Prevention. Our definition of reinfection accounts for intermittent reverse transcriptase polymerase chain reaction (RT-PCR) positivity and reduces the likelihood that we are misinterpreting a longer initial infection as a reinfection, whilst also attempting to capture earlier reinfections.

Standard definitions of reinfection rely on whole genome sequencing of both initial and subsequent positive swab tests, or on testing positive for antibodies at a previous blood test, neither of which are currently available for most participants in the survey. It is important to note that no definition of reinfection using positive swab tests alone can differentiate reinfections as distinct from ongoing detection of virus, including after intermittent negative tests. SARS-CoV-2 RT-PCR positivity may persist for prolonged periods (up to 90 days in a small minority of individuals) following initial infection without indicating a viable virus. RT-PCR tests may also be intermittently negative despite continued detection of the virus.

Participants will have had their first positive swab in the survey at varying times and will subsequently have become at risk of reinfection at varying times.

For the purposes of this analysis, "initial episode" refers to an individual's first episode of COVID-19 recorded in the survey and "reinfection episode" refers to a subsequent infection episode meeting our reinfection criteria. Further definitions can be found in the glossary.

Reinfection rates

Reinfections, as identified in the survey, have been rare up to 5 June 2021. The rate of reinfections is provided per 100,000 participant days of risk. This is calculated by dividing the number of reinfections identified by the cumulative number of participants days from the first time a participant became at risk for reinfection to their last test in the survey to date. The estimated rate for all reinfections was 15.2 per 100,000 participant days at risk (95% confidence interval: 12.7 to 18.0) over the entire at-risk period. The estimated rate for reinfections with a strong positive test (with Ct less than 30) was 2.7 per 100,000 participant days at risk (95% confidence interval: 1.8 to 4.1) over the entire at-risk period (accompanying dataset 1a).

Table 1: Rate of reinfections per 100,000 participant days at risk
Estimated rate of COVID-19 reinfections per 100,000 participant days at risk, averaged for entire at-risk period, 26 April 2020 to 5 June 2021, UK

Definition	Number of participants at risk	Number of identified reinfections	Estimated rate of reinfections (per 100,000 participant days at risk)	Lower 95% confidence interval	Upper 95% confidence interval
All reinfections definition	17,434	133	15.2	12.7	18.0
Reinfections with Ct less than 30	17,434	24	2.7	1.8	4.1

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

Notes:

1. For the purposes of this analysis, we define reinfection as a new positive test 90 days or more after an initial first positive test which was preceded by at least one negative test, or where an individual has had a subsequent positive test following four consecutive negative tests regardless of the time since the first positive.

2. 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.

Flexible parametric models were used to investigate how the overall rate of reinfection varied over time after an individual first became at risk. The rate of reinfections was highest in the month after individuals first became at risk for reinfection (accompanying dataset 1b). Increased risk at the start of the at-risk period may reflect the fact that some early reinfections could still be the same initial infection in individuals who had intermittent positive and negative RT-PCR tests over a long period of time. After the first month, the reinfection rate declined. After 90 days at risk, the estimated rate for all reinfections was 5.0 per 100,000 participant days at risk (95% confidence interval: 3.1 to 8.1). There were not enough strong positive tests (Ct less than 30) after 90 days of risk to provide an estimated rate of infections for this sub-group.

Figure 1: Risk of reinfection peaked around the first month after becoming "at risk" for reinfection and declined afterwards

Estimated rate of COVID-19 reinfection according to time at risk per 100,000 participant days, 26 April 2020 to 5 June 2021, UK

Notes:

1. For the purposes of this analysis, we define reinfection as a new positive test 90 days or more after an initial first positive test which was preceded by at least one negative test, or where an individual has had a subsequent positive test following four consecutive negative tests regardless of the time since the first positive test.

2. 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.

Download the data

Time between infection episodes

In those with reinfections, the median time between the first positive test in the initial episode and first positive test in the reinfection episode was 115 days (interquartile range: 108 to 139) (accompanying dataset 1c).

4 . Viral load

The strength of a positive test is determined by how quickly the virus is detected and is measured by a Ct value. Positive results with a low Ct value (30 or less) indicate episodes of infection with a higher viral load and tests with a high Ct value (more than 30) indicate a lower viral load. In unadjusted analyses, viral loads were higher (Ct values were lower) in initial episodes of COVID-19 infection (median Ct value: 24.5, interquartile range: 19.7 to 31.2) than for reinfection episodes (median Ct value: 32.8, interquartile range: 30.6 to 34.0) (accompanying dataset 1d).

Figure 2 shows that most people had a high viral load (low Ct value) in the initial episode and a lower viral load (high Ct value) in the reinfection episode, which suggests these individuals may have had a stronger immune response to the initial infection which helped them respond to the reinfection more effectively. Some individuals had low viral load (high Ct values) at both episodes, which could be because of monthly tests occurring when participants are nearing the end of infection in both episodes, meaning their viral load is lower because they have had the infection for longer (accompanying dataset 1e). Very few participants had a low viral load (high Ct value) in their initial infection and a high viral load (low Ct value) at reinfection; these individuals may have had a weaker immune response to their initial infection.

Figure 2: Viral loads were typically lower (higher Ct values) in reinfection episodes than in initial infection episodes

Minimum Ct value and self-reported symptoms by infection episode, 26 April 2020 to 5 June 2021, UK

Notes:

1. This analysis illustrates minimum reported Ct values and self-reported symptoms in individuals with reinfections, split by initial and reinfection episodes of COVID-19.

2. Symptoms were self-reported and not clinically diagnosed.

3. Differences between Ct values observed at initial infection and reinfection were statistically significant (paired Wilcoxon test, $p < 0.001$).

Download the data

5 . Symptoms

This analysis considers whether any symptoms were reported at visits within 35 days following the first observed positive test corresponding to the episode of initial infection or reinfection. At each visit we ask about symptoms in the last seven days, including any specific self-reported symptoms (cough, fever, shortness of breath, loss of taste, loss of smell, myalgia, fatigue, sore throat, headache, abdominal pain, diarrhoea, nausea or vomiting) and any general symptoms participants have that they think may be related to COVID-19.

In unadjusted analyses, individuals were less likely to self-report symptoms within 35 days following the first observed positive test in their reinfection episode than in their initial episode. Figure 3 shows the proportion of individuals who reported symptoms within 35 days following the first observed positive test associated with the initial infection episode and the reinfection episode. People were more likely to report symptoms within 35 days following the first observed positive test in their initial episode (57.1%, 95% confidence interval: 48.6% to 65.2%) than in their reinfection episode (21.8%, 95% confidence interval: 15.6% to 29.6%), suggesting that reinfections may be more likely to be asymptomatic.

Figure 3: Fewer people reported symptoms within 35 days following the first observed positive test in their reinfection episode than in their initial infection episode, suggesting reinfections are more likely to be asymptomatic

Percentage of people reporting symptoms within 35 days following their first observed positive test by episode, 26 April 2020 to 5 June 2021, UK

Notes:

1. A confidence interval gives an indication of the degree of uncertainty of an estimate, showing the precision of a sample estimate. The 95% confidence intervals are calculated so that if we repeated the study many times, 95% of the time the true unknown value would lie between the lower and upper confidence limits. A wider interval indicates more uncertainty in the estimate. Overlapping confidence intervals indicate that there may not be a true difference between two estimates.

2. Symptoms are self-reported and were not clinically diagnosed.

3. This data shows unweighted percentages of individuals who met our definition of COVID-19 reinfection, split by whether the individual reported symptoms at any visit within 35 days following the first observed positive test of their initial or reinfection episode respectively.

4. Differences between those reporting any symptoms within 35 days of initial infection and reinfection were statistically significant (McNemar test, $p < 0.001$).

Download the data

Further symptom analysis can be found in the latest edition of Coronavirus (COVID-19) Infection Survey: characteristics of people testing positive for COVID-19 in countries of the UK.

6 . Coronavirus (COVID-19) Infection Survey technical data

Coronavirus (COVID-19) Infection Survey reinfections technical data
Dataset | Released 29 June 2021
Findings from the Coronavirus (COVID-19) Infection Survey technical data on reinfections.

7 . Collaboration

This analysis was produced by Owen Gethings - Office for National Statistics (ONS) Senior Statistical Officer, 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
- Thomas House - University of Manchester, Department of Mathematics: Reader in mathematical statistics

8 . Glossary

At-risk period

The "at-risk" period consists of any individual in the sample that meets either of the following criteria: 90 days has elapsed since an individual's first positive test in thesurveyand their most recent test result was negative or, if 90 days has not passed since their first positive test in the survey, the individual's last positive has been followed by four consecutive negative tests.

COVID-19 infection

We define a coronavirus (COVID-19) infection as an individual testing positive for SARS-CoV-2, with or without symptoms, on a swab taken from the nose and throat using reverse transcriptase polymerase chain reaction (RT-PCR).

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 multiple times, 95% of the time the true unknown value would lie between the lower and upper confidence limits. A wider interval indicates more uncertainty in the estimate. Overlapping confidence intervals indicate that there may not be a true difference between two estimates. For more information, see our methodology page on statistical uncertainty.

Cycle threshold (Ct) value

A cycle threshold (Ct) value is the cycle number at which the PCR product crosses a threshold of detection and can give an indication of the viral load, or strength of a positive SARS-CoV-2 test. The lower the Ct value, the higher the viral load and stronger the positive test.

Participant days at risk

The rate per 100,000 participant days at risk allows us to understand the rate of reinfection depending on how long participants have been "at risk" of reinfection and accounts for participants having initial infection episodes and entering the at-risk period at different times. It also accounts for individuals having different amounts of time until their last test in the survey to date. The rate is calculated by dividing the number of reinfections identified by the cumulative number of participants days from the first time a participant became at risk to their last test in the survey to date. For example, if there were 1,000 participants who had each spent 100 days in the at-risk period, and we identified one reinfection in that time, the rate would be one reinfection per every 100,000 at-risk days.

Viral load

Viral load refers to the quantity of virus in a bodily fluid such as saliva or blood. High viral load is associated with lower Ct values (indicating a strong positive test), and low viral load is associated with higher Ct values.

9 . Data sources and quality

In this article, we refer to the number of coronavirus (COVID-19) reinfections within the Coronavirus (COVID-19) Infection Survey. The survey includes individuals from the community population; community in this instance refers to private residential households, and it excludes those in hospitals, care homes and/or other institutional settings in the UK. For the purposes of this analysis, all estimates of COVID-19 reinfections are unweighted.

More information on measuring the data and its strengths and limitations is available in the Coronavirus (COVID-19) Infection Survey statistical bulletin.

Our methodology article provides further information around the survey design, how we process data and how data are analysed.

Methods and technical information

All estimates of COVID-19 reinfections presented in this article are unweighted, as the sample for the analysis includes only those who meet our criteria for reinfection and as such there is not a nationally representative population to weight this to. As we currently have low counts of participants reinfected with COVID-19, confidence intervals associated with this analysis may be wide, and where applicable measures of uncertainty are presented.

The strength of a positive test is determined by how quickly the virus is detected by reverse transcriptase polymerase chain reaction (RT-PCR) and is measured by a cycle threshold (Ct) value. Positive tests with a Ct value of less than 30 are considered a strong positive test and indicate a higher viral load, with higher Ct values indicating a weaker positive test. High Ct value can be seen in the early stages of infection when virus levels are rising, or late in the infection when the risk of transmission is low.

This analysis of reinfections considers individuals with a new positive test and within this, a subgroup is considered where any positive in the reinfection episode has a Ct value less than 30. We do not restrict to this group alone as one question addressed by the analysis was whether reinfections tended to have higher Ct values indicating a lower viral load. You can find more information on Ct values in a paper written by our academic partners at the University of Oxford.

Time-to-event data is commonly analysed using the Cox proportional hazards model. The Cox model makes no assumptions about the shape of the underlying hazard function and does not produce smoothly varying estimates of rates over time. In contrast, a flexible parametric model uses restricted natural cubic spline functions to model the baseline cumulative hazard (used here), baseline cumulative odds of survival, or some more general baseline distribution in survival analysis models. This provides a smoothly varying estimate of the rates of events over time, as shown in Figure 1.

For this article, we present the percentage of people reporting symptoms at any visit within 35 days following a first observed positive test corresponding to an episode of COVID-19. Previous analysis of symptoms has considered only those individuals with a strong positive test (Ct value less than 30) to exclude the possibility that symptoms are not identified because we pick up individuals very early or late in their infection. However, this analysis of symptoms considers symptoms reported by individuals regardless of the Ct value as this allowed us to examine whether reinfections were milder in terms of both symptom burden and viral load.

10 . Related links

Coronavirus (COVID-19) Infection Survey, UK
Bulletin | Updated weekly
Estimates for England, Wales, Northern Ireland and Scotland. This survey is being delivered in partnership with University of Oxford, University of Manchester, Public Health England and Wellcome Trust.

Coronavirus (COVID-19) Infection Survey: characteristics of people testing positive for COVID-19 in England
Article | Updated fortnightly
Characteristics of people testing positive for COVID-19 from the Coronavirus (COVID-19) Infection Survey, including antibody data by UK country, and region and occupation for England. Antibodies data published before 3 February 2021 are available in this series.

Ct threshold values, a proxy for viral load in community SARS-CoV-2 cases, demonstrate wide variation across populations and over time
Academic article | 4 April 2021
Data from the Coronavirus (COVID-19) Infection Survey were used by academic partners from the University of Oxford to examine Ct threshold values as a measure of viral load.

Investigation and management of suspected SARS-Cov-2 reinfections: a guide for clinicians and infection specialists
Guidance article | 15 March 2021
Information on investigating and managing SARS-Cov-2 reinfections for clinicians from Public Health England.

Common Investigation Protocol for Investigating Suspected SARS-CoV-2 Reinfection
Webpage| 27 October 2020
Protocol from the Centre for Disease Control and Prevention to support investigations into COVID-19 reinfection cases.

Interim Guidance on Ending Isolation and Precautions for Adults with COVID-19
Webpage| 16 March 2021
Information on isolation and precautions (including data regarding reinfections of COVID-19 variant viruses) from the Centre for Disease Control and Prevention.

New National Surveillance of possible COVID-19 Reinfection
Webpage | 17 June 2021
Press release from Public Heath Engand on reinfection of COVID-19

COVID-19 Infection Survey: methods and further information
Methods article | Updated 26 March 2021
Information on the methods used to collect the data, process it, and calculate the statistics produced from the COVID-19 Infection Survey pilot.

Coronavirus (COVID-19) latest insights
Interactive tool | Updated as and when data become available
Explore the latest data and trends about the coronavirus (COVID-19) pandemic from the ONS and other official sources.

Coronavirus (COVID-19) roundup
Web page | 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.

COVID-19 Infection Survey (CIS)
Article | Updated regularly
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.