

1 **Ct threshold values, a proxy for viral load in community SARS-CoV-2 cases, demonstrate wide**  
2 **variation across populations and over time**

3

4 A. Sarah Walker<sup>1,2,3,4</sup>, Emma Pritchard<sup>1,2</sup>, Thomas House<sup>5,6</sup>, Julie V Robotham<sup>2,7</sup>, Paul J Birrell<sup>7,8</sup>, Iain  
5 Bell<sup>9</sup>, John I Bell<sup>10</sup>, John N Newton<sup>11</sup>, Jeremy Farrar<sup>12</sup>, Ian Diamond<sup>9</sup>, Ruth Studley<sup>9</sup>, Jodie Hay<sup>13,14</sup>,  
6 Karina-Doris Vihta<sup>1,2</sup>, Tim Peto<sup>1,2,3,15</sup>, Nicole Stoesser<sup>\*1,2,3,15</sup>, Philippa C. Matthews<sup>\*1,15</sup>, David W.  
7 Eyre<sup>\*1,2,14,16</sup>, Koen B. Pouwels<sup>1,2,17</sup> and the COVID-19 Infection Survey team

8

9 \* contribution considered equal

10

11 <sup>1</sup> Nuffield Department of Medicine, University of Oxford, Oxford, UK

12 <sup>2</sup> The National Institute for Health Research Health Protection Research Unit in Healthcare Associated  
13 Infections and Antimicrobial Resistance at the University of Oxford, Oxford, UK.

14 <sup>3</sup> The National Institute for Health Research Oxford Biomedical Research Centre, University of Oxford,  
15 Oxford, UK

16 <sup>4</sup> MRC Clinical Trials Unit at UCL, UCL, London, UK

17 <sup>5</sup> Department of Mathematics, University of Manchester, Manchester, UK

18 <sup>6</sup> IBM Research, Hartree Centre, Sci-Tech Daresbury, UK

19 <sup>7</sup> National Infection Service, Public Health England, London, UK

20 <sup>8</sup> MRC Biostatistics Unit, University of Cambridge, Cambridge Institute of Public Health, Cambridge,  
21 UK

22 <sup>9</sup> Office for National Statistics, Newport, UK

23 <sup>10</sup> Office of the Regius Professor of Medicine, University of Oxford, Oxford, UK

24 <sup>11</sup> Health Improvement Directorate, Public Health England, London, UK

25 <sup>12</sup> Wellcome Trust, London, UK

26 <sup>13</sup> University of Glasgow, Glasgow, UK.

27 <sup>14</sup> Lighthouse Laboratory in Glasgow, Queen Elizabeth University Hospital, Glasgow

28 <sup>15</sup> Department of Infectious Diseases and Microbiology, Oxford University Hospitals NHS Foundation

29 Trust, John Radcliffe Hospital, Oxford, UK

30 <sup>16</sup> Big Data Institute, Nuffield Department of Population Health, University of Oxford, Oxford, UK

31 <sup>17</sup> Health Economics Research Centre, Nuffield Department of Population Health, University of

32 Oxford, Oxford, UK

33

34 See Acknowledgements for the Coronavirus Infection Survey team

35

36 Keywords: SARS-CoV-2, viral load, symptoms, community

37

38 Running title: Community SARS-CoV-2 positives

39

40 Corresponding author

41 A. Sarah Walker, [sarah.walker@ndm.ox.ac.uk](mailto:sarah.walker@ndm.ox.ac.uk)

42

43

44 **ABSTRACT** (151 words)

45 Information on SARS-CoV-2 in representative community surveillance is limited, particularly cycle  
46 threshold (Ct) values (a proxy for viral load). Of 3,312,159 nose and throat swabs taken 26-April-2020  
47 to 13-March-2021 in the UK's national COVID-19 Infection Survey, 27,902(0.83%) were RT-PCR-  
48 positive, 10,317(37%), 11,012(40%) and 6,550(23%) for 3, 2 or 1 of the N, S and ORF1ab genes  
49 respectively, with median Ct=29.2 (~215 copies/ml; IQR Ct=21.9-32.8, 14-56,400 copies/ml).  
50 Independent predictors of lower Cts (i.e. higher viral load) included self-reported symptoms and  
51 more genes detected, with at most small effects of sex, ethnicity and age. Single-gene positives  
52 almost invariably had Ct>30, but Cts varied widely in triple-gene positives, including without  
53 symptoms. Population-level Cts changed over time, with declining Ct preceding increasing SARS-CoV-  
54 2 positivity. Of 6,189 participants with IgG S-antibody tests post-first RT-PCR-positive, 4,808(78%)  
55 were ever antibody-positive; Cts were significantly higher in those remaining antibody-negative.  
56 Community SARS-CoV-2 Ct values could be a useful epidemiological early-warning indicator.

57

58 **IMPACT STATEMENT**

59 Ct values from SARS-CoV-2 RT-PCR tests vary widely and over calendar time. They have the potential  
60 to be used more broadly in public testing programmes as an “early-warning” system for shifts in  
61 infectious load and hence transmission.

62 **INTRODUCTION**

63 After initial reductions in SARS-CoV-2 cases in mid-2020, following release of large-scale lockdowns  
64 (Flaxman et al., 2020), infection rates have undergone waves of resurgence and suppression in many  
65 countries worldwide. Proposed control strategies include new local or national lockdowns of varying  
66 intensity and mass testing, but these have major economic and practical limitations. In particular,  
67 mass testing of large numbers without symptoms (Yokota et al., 2020), and hence low pre-test  
68 probability of positivity, can mean most positives are false-positives depending on test specificity. For  
69 example, with 0.1% true prevalence, testing 100,000 individuals with a 99.9% specific test with  
70 perfect sensitivity gives 100 true-positives, but also 100 false-positives (positive predictive value  
71 (PPV) 50%), whereas specificity of 99.5% increases false-positives to 500 (PPV=17%), and of 99.0% to  
72 999 (PPV=9%), with even lower PPV with imperfect sensitivity (Adams, Ainsworth, Anand, & et al.,  
73 2020).

74

75 Mathematical models are powerful tools for evaluating the potential effectiveness of different  
76 control strategies, but rely on population-level estimates of infectivity and other parameters.  
77 However, there are few unbiased community-based surveillance studies, including individuals both  
78 with and without symptoms. Estimates of asymptomatic infection rates vary, being 17-41% overall in  
79 recent reviews (Buitrago-Garcia et al., 2020; Byambasuren et al., 2020), but these included many  
80 studies of contacts of confirmed cases. Higher prevalence of asymptomatic infection has been  
81 reported in screening of defined populations (30% (Buitrago-Garcia et al., 2020)) and community  
82 surveillance (e.g. 42% (Lavezzo et al., 2020), 72% (Riley, Ainslie, Eales, Walters, Wang, Atchison,  
83 Fronterre, et al., 2020)). Studies have generally indicated lower rates of transmission from  
84 asymptomatic infection (Buitrago-Garcia et al., 2020; Byambasuren et al., 2020); this may be a proxy  
85 for SARS-CoV-2 viral load as a key determinant of transmission. Finally, most studies rely on  
86 “average” estimates of the asymptomatic infection percentage, independent of characteristics and

87 viral load, and have not quantified temporal variation in these key parameters for mathematical  
88 models across the community.

89

90 Here we therefore characterise variation in SARS-CoV-2 positive tests in the first eleven months of  
91 the UK's national COVID-19 Infection Survey. In brief (details in Methods), the survey randomly  
92 selects private households to provide a representative UK sample, recruiting all consenting  
93 individuals aged 2 years or older currently resident in each household to provide information on  
94 demographics, symptoms, contacts and relevant behaviours and self-taken nose and throat swabs  
95 for RT-PCR testing (Pouwels et al., 2020). A randomly selected subset are approached for additional  
96 consent to provide blood samples for IgG S-antibody testing if aged 16 years or older. At the first  
97 visit, participants can provide additional consent for longitudinal follow-up (visits every week for the  
98 next month, then monthly for 12 months from enrolment). We estimate predictors of RT-PCR cycle  
99 threshold (Ct) values (as a proxy for viral load), propose a classification for the strength of evidence  
100 supporting positive RT-PCR test results in the community, and demonstrate how this has changed  
101 over time. We also provide a preliminary assessment of seroconversion rates for community cases.

102

## 103 **RESULTS**

### 104 *Number and percentage of positive swabs*

105 From 26 April 2020 to 13 March 2021, 440,479 participants from 217,887 households in the COVID-  
106 19 Infection Survey had one or more RT-PCR results from nose and throat swabs (median 8 results  
107 per participant (IQR 6-9, range 1-19)). Participants were recruited between April 2020 and March  
108 2021 (**Supplementary File 1**). Of 3,312,159 RT-PCR test results, 27,902 (0.84%, 95% CI 0.83-0.85%)  
109 were positive, in 21,831 individuals from 16,214 households. 2,966 (14%) of these individuals were  
110 positive at their first test in the study and 18,865 (86%) subsequently, after median 5 negative tests  
111 (IQR 3-6, range 1-14).

112

113 *Viral characteristics*

114 Overall, 10,317 (37%), 11,012 (40%) and 6,550 (23%) swabs were positive for three, two or one of the  
115 three SARS-CoV-2 genes (N protein, S protein and ORF1ab) respectively (**Table 1**; 23 positives with  
116 missing Ct and gene detection excluded from this and all subsequent analysis; samples with only the  
117 S-gene detected generally not called positive, see Methods). The majority of two-gene positives  
118 (9,513 (86%)) were ORF1ab+N positive from 16 November 2020 onwards, reflecting the emergence  
119 and expansion of B.1.1.7 in the UK (Walker et al., 2021). B.1.1.7 leads to S-gene target failure (SGTF)  
120 and was estimated to account for 88% of SGTF from this time (Public Health England, 2020). Where  
121 multiple genes were detected, the Cts were highly correlated (Spearman  $\rho=0.98$ ,  $p<0.0001$ ). Taking  
122 the per-swab mean Ct across positive genes, the overall median Ct was 29.2 (IQR 21.9-32.8; range  
123 9.2-38.7), reflecting the study's surveillance design testing individuals in the community at fixed  
124 timepoints regardless of symptoms. Based on calibration data (**Supplementary Figure 1**), this  
125 corresponds to a median viral load of  $\sim 215$  copies/ml (IQR 14-56,400). Ct varied strongly by number  
126 of genes detected (Kruskal-Wallis  $p=0.0001$ ), but not by their specific pattern after adjusting for  
127 number ( $p=0.08$ ). There is no fixed Ct threshold for determining positivity (see Methods); however,  
128 only 38 (0.1%) Ct values  $>37$  were recorded (5 positive on ORF1ab+N).

129

130 Of note, whilst single-gene positives almost invariably had  $Ct>30$ , with or without reported  
131 symptoms, triple-gene positives without reported symptoms had widely varying Ct, as did ORF1ab+N  
132 positives after 16 November 2020 (SGTF, compatible with B.1.1.7) (**Figure 1**). Ct values were slightly  
133 but significantly lower in other double-gene positives vs single-gene positives, with a small number of  
134 low Ct values in ORF1ab+N positives before 16 November likely reflecting early B.1.1.7 cases.  
135 Further, whilst the percentage reporting symptoms increased linearly as Ct values dropped from 35  
136 ( $\sim 30\%$  reporting symptoms around the positive test) to 28 ( $\sim 60\%$  reporting symptoms), below 28 the  
137 percentages reporting symptoms increased only slightly (to  $\sim 70\%$  at  $Ct=10$ ) (**Figure 2**).

138

139 Evidence supporting positive results

140 Combining information on Ct values, symptoms and pre-test probability of being positive, 21,329  
141 (77%), 4,741 (17%) and 1,809 (6%) positive tests had “higher”, “moderate” or “lower” evidence  
142 supporting genuine presence of viral RNA (**Table 2**; definitions in Methods). Even though “higher”  
143 evidence was based only on number of genes detected (two or three), “higher” evidence positives  
144 were more likely to be symptomatic than “moderate” evidence positives ( $p < 0.0001$ ), but were  
145 similarly likely to have occupational risk factors ( $p = 0.48$ ). “Higher” evidence positives were more  
146 likely to occur in households with other positives ( $p < 0.0001$ ).

147

148 Predictors of Ct values

149 In multivariable regression models, Ct values were independently lower (i.e. viral loads higher) with  
150 more genes detected (8.2 lower in triple-gene vs single-gene positives (95% CI 7.9-8.5)), if symptoms  
151 were reported around the test (2.0 lower (1.8-2.2)), and at the first positive identified per-participant  
152 (2.2 lower than subsequent positives (2.2-2.5)), and if the positive was not the participant’s first test  
153 in the study (0.6 lower (0.2-0.9)) (all  $p < 0.0001$ ; **Supplementary File 2A**; see Methods for details of  
154 collection of symptoms). By far the strongest effect was associated with triple-gene positives. Men  
155 had slightly lower Ct values than women (0.3 lower (0.1-0.5)  $p = 0.001$ ), and there was marginal  
156 evidence of lower values in those reporting non-white ethnicity (0.3 lower (0-0.6)  $p = 0.08$ ). Compared  
157 with those not reporting symptoms, Ct values were lower in those reporting cough/fever/anosmia  
158 (2.5 lower (2.3-2.8)) than other symptoms only (0.9 lower (0.7-1.2); heterogeneity  $p < 0.0001$ ).

159 Associations were similar for symptoms at the positive test. After adjusting for these factors, there  
160 was no evidence of independent effects of age ( $p = 0.33$ ) or deprivation ( $p = 0.67$ , **Supplementary File**  
161 **2A**). Even after adjusting for these factors, Ct values were 1.4 (1.2-1.6) lower in individuals where  
162 another household member was positive at any point in the study ( $p < 0.0001$ ; other effects similar).

163

164 However, number of genes detected and symptoms are both potential mediators of effects of  
165 demographic factors (**Supplementary Figure 2**). Excluding these potential mediators (number of  
166 genes detected, symptoms), Ct values remained independently lower (i.e. viral loads higher) at the  
167 first positive identified per-participant, where the positive was not the participant's first test in the  
168 study, and in men, but were also slightly lower with increasing deprivation ( $p=0.0005$ ; Ct 1.0 lower in  
169 the most vs least deprived (95% CI 0.6-1.5)) and in younger adults ( $p=0.0001$ ; those aged 17-24 1.0  
170 lower (0.3-1.7) than those under 12, and 1.4 lower (0.8-2.0) than those aged 70+) (**Supplementary**  
171 **File 2B**). Results were similar adjusting for date of the positive test.

172

### 173 Temporal changes in Ct values, evidence and symptomatic percentages

174 There were strong effects of calendar time on the distribution of Ct values (**Figures 3A&B**), the  
175 percentages self-reporting symptoms, or cough/fever/anosmia (**Figure 3C**), and strength of evidence  
176 supporting each positive result (**Figure 3D**; all  $p<0.0001$ ). In particular, Ct values were markedly  
177 higher in July-August 2020 when population positivity rates were low, with correspondingly very low  
178 percentages with symptoms at/around positive tests, and more "lower" evidence positives.  
179 Decreases in Ct values in late August/early September and December 2020 coincided with increases  
180 in percentages reporting symptoms and of "higher" evidence positives, and, in England (**Figure 3B**),  
181 with initial rises in official estimates of positivity rates (Office for National Statistics, 2021) after very  
182 low rates in July/early August 2020, and with much stronger rises in December 2020 (expansion of  
183 B.1.1.7). Ct levels rose, and correspondingly percentages reporting symptoms and of "higher"  
184 evidence positives declined, as positivity peaked during November 2020 and January 2021  
185 lockdowns.

186

187 However, even within "higher" evidence positives, median Ct varied strongly over time being higher  
188 in July/early August 2020 and after November 2020 and January 2021 lockdowns (**Figure 4A**).  
189 "Lower" evidence positives also formed a larger percentage of all tests during July/early August 2020,



190 despite overall positivity rates being very low (e.g. 0.022% in the 3 weeks starting 20 July 2020;  
191 **Figure 4B**). However, interestingly, from September 2020, the percentage of “lower” evidence  
192 positives increased proportionately with “moderate” and “higher” evidence positives (**Figure 4B**).  
193 The lowest non-zero observed rate of “low evidence” positives was 0.005% (both in early June and  
194 late August), providing an upper bound on the rate of false-positives as defined by identifying virus  
195 when none present.

196

### 197 Relationship with serostatus

198 One or more IgG S-antibody results were available for 6,540 (30%) participants with positive swabs.  
199 Less than 5% of antibody tests taken >30 days before the first positive swab (not necessarily the  
200 onset of infection) were positive (**Figure 5**), rising to 12% in the 30 days before the first swab positive  
201 (likely reflecting late detection of infection), 47% in the following 14 days and then 72-81%  
202 thereafter. Overall, of 6,189 participants with one or more antibody tests after their first positive  
203 swab, 4,808 (78%) were ever antibody-positive; with higher rates in those reporting symptoms  
204 around their first positive swab (2,945/3,315 (89%) vs 1,863/2,874 (65%) of those not reporting  
205 symptoms,  $p < 0.0001$ ). Median (IQR) Ct values were also significantly lower in those ever antibody-  
206 positive to date (24.9 (18.5-31.0) vs 33.0 (29.9-34.3) in those not antibody-positive,  $p < 0.0001$ ).  
207 Results were similar restricting to 1,477 (24%) with a negative antibody result within [-120,+21] days  
208 of their first positive swab. A small number of participants appeared to have become infected despite  
209 antecedent high anti-spike antibody titres, one case in particular which had “higher evidence”  
210 positive swab tests separated by four consecutive negative swabs with 65 days between positive  
211 swabs.

212

## 213 **DISCUSSION**

214 In this large community surveillance study, we found wide variation in Ct values (a proxy for viral  
215 load). Whilst Ct values were independently associated with several factors, including symptoms

216 at/around the test as previously reported (Edwards et al., 2020; S. Lee et al., 2020), their effects were  
217 small compared with population-level variability. Notably both triple-gene positives and S-gene  
218 target failures compatible with the B.1.1.7 variant without reported symptoms had widely varying Ct,  
219 including many with low values (**Figure 3A**), potentially explaining variation in dispersion (“k”) and  
220 super-spreading events, particularly from those without symptoms but with low Ct/high viral loads  
221 (Endo, Centre for the Mathematical Modelling of Infectious Diseases, Abbott, Kucharski, & Funk,  
222 2020; Rasmussen & Popescu, 2021).

223

224 Compared with other single/double positives, Ct values were significantly lower in triple-gene  
225 positives and S-gene target failures compatible with the B.1.1.7 variant (after mid-November 2020).  
226 However, direct comparisons of viral load with B.1.1.7 vs other variants are not possible within this  
227 analysis, given lack of knowledge as to the true underlying variant over the included time period. We  
228 found lower Ct in those reporting cough/fever/anosmia/ageusia than other symptoms, and other  
229 symptoms vs no symptoms, supporting the importance of the “classic” symptoms for identifying the  
230 most infectious cases. Lower Ct values in the first positive per-participant likely reflects the natural  
231 history of viral load post-infection, and higher Ct values in those positive at their first test in the study  
232 over-representation of long-term shedders in this group. Lower Ct values in men and those reporting  
233 non-white ethnicity, although small, are consistent with poorer outcomes in these groups.  
234 Interestingly, small effects of age and deprivation were mediated by self-reported symptoms and  
235 number of positive genes. That is, when adjusting for these latter factors no association was  
236 observed with Ct, but without adjustment younger individuals (as shown in (Jones et al., 2020) but  
237 not (Jacot, Greub, Jaton, & Opota, 2020)) and those from more deprived areas had slightly lower Ct  
238 values, suggesting that these factors may affect the intrinsic level of virus present (**Supplementary**  
239 **Figure 2**). The small size of the effects mean they may variably be detected depending on study size  
240 and power.

241

242 Ct values varied strongly over time, as did symptoms and evidence supporting positives, suggesting  
243 changing viral burden in infection cases, with less severe infections during July/early August 2020.  
244 This strongly refutes hypotheses that declines in positivity during this period were due to declines in  
245 viral fitness. During this time, higher Ct values were also noted in the English point-prevalence  
246 surveillance study, REACT (Riley, Ainslie, Eales, Walters, Wang, Atchison, Fronterre, et al., 2020), and  
247 lower virus levels in Lausanne, Switzerland (Jacot et al., 2020). However, Ct values were higher even  
248 in “higher” evidence positives during this period, consistent with shifting viral burden (**Figure 4A**).  
249 Such a shift may also explain the preceding shift towards “moderate” evidence positives and the  
250 concurrent higher percentage of “lower” evidence positives, since the less virus present, the less  
251 likely it is to be detected on multiple genes. Whilst these findings are consistent with lower viral  
252 inoculum during this period (Gandhi, Beyrer, & Goosby, 2020), we cannot assess whether this is  
253 predominantly due to behaviour (e.g. increased time outdoors, face mask use (Gandhi & Rutherford,  
254 2020)) or other reasons (e.g. environmental/climatic factors, including relating to transport of swabs  
255 for testing). Whilst decreases in Ct values in July/early August 2020 preceded increases in positivity  
256 rates in England, later declines in Ct in early December coincided with, rather than preceded,  
257 increases in positivity due to B.1.1.7 expansion. This may potentially reflect faster transmission of  
258 B.1.1.7 but may also reflect greater sensitivity to changes in Ct distribution when case numbers are  
259 small. Subsequent increases in Ct reflected stabilising and then declining positivity in both periods.  
260  
261 We used laboratory, clinical and demographic evidence to classify our confidence in positive results.  
262 Around 70% had 2 or 3 genes detected (“higher” evidence), providing assurance in overall results,  
263 with only 0.1% of Ct values over 37. Whilst Ct values are not directly comparable between studies,  
264 REACT has also validated a Ct threshold of 37 for single-gene positives for their test performed in  
265 Germany (Riley, Ainslie, Eales, Walters, Wang, Atchison, Diggle, et al., 2020), and in the Public Health  
266 England (PHE) Schools study, only samples with Ct<37 were positive on repeat testing of the same  
267 swab at PHE laboratories (Ladhani et al., 2020). However, every diagnostic test has false-positives,

268 here defining a false-positive as detection of virus by RT-PCR when no virus is present in a sample, so  
269 some of our single gene “lower”, or even “moderate”, evidence positives are inevitably false.  
270 However, the false-positive rate (as defined) would generally be expected to be approximately  
271 constant over time, since it is either random or driven by external factors, although cross-  
272 contamination (which should be minimised by good laboratory practice) may theoretically be related  
273 to background prevalence/viral load. Variation in the percentage of all tests accounted for by “lower”  
274 evidence positives, and in particular the proportionate increases in “lower” evidence positives as  
275 “higher” evidence positives increased during September 2020 supports more genuinely lower-level  
276 infections occurring during the summer, and an overall false-positive rate for this test of below  
277 ~0.005% i.e. at least 99.995% specificity.

278

279 With recent expansion of antigen assays, there has been considerable debate on what “positivity”  
280 means, and hence what is a “false-positive” or a “false-negative”. First, it is clear that the detection  
281 of viral RNA is neither the same as infectiousness, although a strong relationship between Ct values  
282 and infection in contacts is observed (L. Y. W. Lee et al., 2021), nor a “disease” in its own right.  
283 However, surveillance has very distinct goals from clinical testing with its focus on isolation and  
284 contact tracing, particularly given the large percentage of asymptomatic infections. It is appropriate  
285 for surveillance to focus on detection of viral RNA, given its goal to estimate burden of  
286 current/ongoing cases that have occurred in the community. However, it is essential to recognise the  
287 difference between the RT-PCR test result (viral RNA has been detected) and the appropriate clinical  
288 action, which may legitimately differ depending on Ct value, for example if the infection is likely to  
289 have occurred sometime previously, as well as other information (e.g. preceding PCR positivity or  
290 serology). RT-PCR assays test for viral RNA presence, and hence it is more relevant to consider limits  
291 of detection, rather than “false-positives” per se. Although they were a small minority (6%), one  
292 question is whether single-gene positives with high Ct (defined as  $\geq 34$  in our study) solely represent  
293 long-term shedding of non-transmissible virus (Moraz et al., 2020), with, for example, infectious virus

294 recovered from only 8% (95% CI 3-18%) of samples with Ct>35 in a PHE study (Singanayagam et al.,  
295 2020) and studies reporting no growth of virus for Ct thresholds from >24 to >34 or higher (Jefferson,  
296 Spencer, Brassey, & Heneghan, 2020). Whilst we have not directly assessed household transmission  
297 in this analysis, it was notable that Ct values were significantly lower in positives where anyone else  
298 in the same household was ever positive, supporting a role for greater within-household  
299 transmission with lower Ct values. Ct values were 0.6 higher in positives that were a participant's first  
300 study test (where long-term shedders would be expected to be overrepresented), but these formed  
301 only 14% of the positives.

302

303 Our evaluation of serological responses is one of few in the community to our knowledge, and  
304 highlights that a significant minority (~20%) of RT-PCR-positive cases do not appear to seroconvert,  
305 particularly those with higher Ct values and not reporting symptoms. A recent systematic review  
306 estimated that 95% of adults with laboratory confirmed SARS-CoV-2 infection developed IgG  
307 antibodies (Arkipova-Jenkins et al., 2021), peaking around 25 days. However, only 23% of included  
308 studies were in outpatient settings and 14% included only participants with asymptomatic or mild  
309 disease. Our community setting, with higher percentages not reporting symptoms and higher Ct  
310 values (both associated with seroconversion), likely explains our lower overall seroconversion  
311 estimate compared with these previous studies. We observed a small number of new swab positives  
312 in antibody-positive individuals: unfortunately whole genome sequence data were not available to  
313 confirm potential re-infections. Presumed re-infections have been reported elsewhere (Tomassini et  
314 al., 2020), including in individuals without previous functional and/or durable antibody responses  
315 (Goldman et al., 2020; To et al., 2020), and may remain relevant to virus transmission, whether they  
316 occur with or without symptoms. Our data and others (Lumley et al., 2021) suggest that these may  
317 occur in the presence of anti-spike antibodies, which correlate with neutralising antibody titres.  
318 These antibody titres are unlikely to have been false-positives, given the context, persistence, and

319 known diagnostic and analytical specificity of the assay (National Sars-CoV-Serology Assay Evaluation  
320 Group, 2020), or to all reflect laboratory identifier errors, and further analyses are ongoing.

321

322 A major study strength is its design, namely being a large-scale community survey recruiting  
323 randomly selected private residential households, and testing participants regardless of symptoms.  
324 However, its size and scale is also a limitation, since we were not able to collect additional data to  
325 comprehensively characterise individual positives. We may have underestimated the initial  
326 prevalence of symptoms due to originally asking about current symptoms before July 2020  
327 (subsequently symptoms in the 7 days preceding the visit). As this was only at the earliest visits,  
328 mostly weekly, only very transient symptoms between visits would likely have been missed. Similar  
329 rates of symptom reporting in the first and last parts of the period analysed suggests that this  
330 question was likely generously interpreted in any case. We made no attempt to collect additional  
331 information on symptoms after positives were identified to minimise recall bias. This may partly  
332 explain why we observed higher rates of positive tests without reported symptoms than recent  
333 reviews (Buitrago-Garcia et al., 2020; Byambasuren et al., 2020); however, many studies in these  
334 reviews tested close contacts of index cases identified through symptoms and therefore might  
335 plausibly have higher viral loads. We compared distributions of Ct values to overall positivity rates in  
336 England, since these are the longest series of official statistics available; overall UK positivity  
337 estimates are not produced because the 4 countries making up the UK have different policies and  
338 timings regarding community restrictions including lockdowns.

339

340 Ultimately the importance of asymptomatic and low virus level infections depends on their  
341 transmissibility and their prevalence; regardless of limitations in symptom ascertainment, infection  
342 without recognition has the potential for onward transmission and unascertained infections are likely  
343 critical for avoiding resurgence after lifting lockdown (Hao et al., 2020). Our findings support the use  
344 of Ct values and genes detected more broadly in public testing programmes, predominantly testing

345 symptomatic individuals and case contacts, as an “early warning” system for shifts in potential  
346 infectious load and hence transmission, and hence the risks posed by individuals to others. This has  
347 recently also been proposed on the basis of theoretical work linking effective reproduction numbers  
348 to population level Ct (Hay, Kennedy-Shaffer, Kanjilal, Lipsitch, & Mina, 2020). In our study, declines  
349 in mean and median Ct values preceded or at least coincided with increases in office estimates of  
350 positivity rates (**Figure 3B**); given the far larger numbers that would be available in testing  
351 programmes, future research should investigate whether the greater power afforded by continuous  
352 outcomes could lead to significantly earlier detection of future positivity increases, particularly within  
353 small geographical areas. Ct data are widely available within laboratory management systems;  
354 providing comparisons across the wide variety of commercial assays were interpreted carefully, they  
355 could be used alongside available risk factor and symptom information to facilitate more informed  
356 and effective individual-level and public health responses to the SARS-CoV-2 pandemic.

357

## 358 **MATERIALS AND METHODS**

359 This study included all positive SARS-CoV-2 RT-PCR results between 26 April 2020 and 13 March 2021  
360 from nose and throat swabs taken from participants in the Office for National Statistics (ONS) CIS  
361 (ISRCTN21086382). The survey randomly selects private households on a continuous basis from  
362 address lists and from previous surveys to provide a representative UK sample (**Supplementary File**  
363 **1**). If anyone aged 2 years or older currently resident in an invited household agreed verbally to  
364 participate, a study worker visited the household to take written informed consent, which was  
365 obtained from parents/carers for those 2-15 years; those aged 10-15 years provided written assent.  
366 The study protocol is available at [https://www.ndm.ox.ac.uk/covid-19/covid-19-infection-](https://www.ndm.ox.ac.uk/covid-19/covid-19-infection-survey/protocol-and-information-sheets)  
367 [survey/protocol-and-information-sheets](https://www.ndm.ox.ac.uk/covid-19/covid-19-infection-survey/protocol-and-information-sheets). Recruitment started 26 April 2020 in England, 29 June 2020  
368 in Wales, 29 July 2020 in Northern Ireland and 21 September 2020 in Scotland.

369

370 Individuals were asked about demographics, symptoms, contacts and relevant behaviours  
371 (<https://www.ndm.ox.ac.uk/covid-19/covid-19-infection-survey/case-record-forms>). To reduce  
372 transmission risks, self-taken nose and throat swabs were obtained following study worker  
373 instructions. Parents/carers took swabs from children under 12 years. At the first visit, participants  
374 were asked for (optional) consent for follow-up visits every week for the next month, then monthly  
375 for 12 months from enrolment. In a random 10-20% households, those 16 years or older were invited  
376 to provide venous blood monthly for assays of anti-trimeric spike protein IgG using an immunoassay  
377 developed by the University of Oxford (National Sars-CoV-Serology Assay Evaluation Group, 2020).  
378 All participants in households where anyone tested positive on a swab were also invited to provide  
379 blood monthly. Venous blood was not taken at any visit where any person in the household had  
380 classic COVID-19 symptoms (fever, cough or anosmia). The study received ethical approval from the  
381 South Central Berkshire B Research Ethics Committee (20/SC/0195).  
382  
383 Swabs and blood samples were collected by study workers at household visits and couriered  
384 overnight to testing laboratories at ambient temperatures. They were analysed at the UK's national  
385 Lighthouse Laboratories at Milton Keynes (National Biocentre) (from 26 April 2020 to 11 February  
386 2021) and Glasgow (from 16 August 2020) using identical methodology, with swabs from specific  
387 regions sent consistently to one laboratory. RT-PCR for three SARS-CoV-2 genes (N protein, S protein  
388 and ORF1ab) used the Thermo Fisher TaqPath RT-PCR COVID-19 Kit, analysed using UgenTec Fast  
389 Finder 3.300.5 (TaqMan 2019-nCoV Assay Kit V2 UK NHS ABI 7500 v2.1). The Assay Plugin contains an  
390 Assay specific algorithm and decision mechanism that allows conversion of the qualitative  
391 amplification Assay PCR raw data from the ABI 7500 Fast into test results with minimal manual  
392 intervention. Samples are called positive in the presence of at least single N gene and/or ORF1ab but  
393 may be accompanied with S gene (1, 2 or 3 gene positives). There is no specific Ct threshold for  
394 determining positivity. S gene is not considered a reliable single gene positive (as of mid-May 2020).  
395 Blood was analysed at the University of Oxford. Antibody titres were considered positive above 8



396 million units (National Sars-CoV-Serology Assay Evaluation Group, 2020) on the original fluorometric  
397 version of the assay and 42 units on the colorimetric version of the assay (used from 1 March 2021).  
398

399 Twelve specific symptoms were elicited at each visit (cough, fever, myalgia, fatigue, sore throat,  
400 shortness of breath, headache, nausea, abdominal pain, diarrhoea, loss of taste, loss of smell), as was  
401 whether participants thought they had (unspecified) symptoms compatible with COVID-19. From 26  
402 April through 22 July 2020, questions referred to current symptoms, and from 23 July 2020 to the  
403 preceding 7 days. Any positive response to any symptom question at the swab-positive visit defined  
404 the case as symptomatic “at” the test; we also separately defined any positive response at the swab-  
405 positive visit or visits either side (regardless of time between visits) as symptomatic “around” the  
406 test.

407

408 To investigate the potential increasing contribution of false-positives as population prevalence  
409 declines, from 2 August 2020 we arbitrarily classified in real-time each positive as:

- 410 • “Higher” evidence: two or three genes detected (irrespective of Ct).
- 411 • “Moderate” evidence: single-gene detected and (i) Ct below the 97.5<sup>th</sup> percentile of “higher”  
412 evidence positives (<34; supporting this threshold, whole genome sequences had been  
413 obtained from three single gene positives with Ct 30.8-33.1 by 2 August) or (ii) higher pre-  
414 test probability of infection, defined as any symptoms at/around the test or reporting  
415 working in a patient-facing healthcare or care/residential home.
- 416 • “Lower” evidence: all other positives; by definition single-gene detected at Ct≥34 in  
417 individuals not reporting symptoms/working in relevant roles.

418

419 As the Ct distribution was skewed to the left, we assessed independent predictors using median  
420 (quantile) regression. Results were broadly similar using random effects model for mean Ct with a  
421 random effect per household. We used 5 knot natural cubic splines (knots at the

422 10<sup>th</sup>/25<sup>th</sup>/50<sup>th</sup>/75<sup>th</sup>/90<sup>th</sup> percentiles of observed unique values) to assess non-linearity in the effect of  
423 calendar time, age and deprivation (index of multiple deprivation rank). Multivariable models for Ct  
424 values were constructed by first choosing the more strongly univariably predictive factor from the  
425 collinear variables (symptoms at/around the test, number of genes detected/supporting evidence for  
426 each positive) then using backwards elimination on the remaining variables. Deprivation was  
427 assessed using the index of multiple deprivation (IMD) in England, a score based on lower layer super  
428 output areas with average population of 1500 people and incorporating seven domains to produce  
429 an overall relative measure of deprivation (income, employment, education, skills and training,  
430 health and disability, crime, barriers to housing services and living environment)  
431 (<https://www.gov.uk/government/statistics/english-indices-of-deprivation-2019>) and equivalent  
432 scores in the other three countries comprising the UK. Each country's scores were converted to a  
433 within country percentile. All analyses were conducted in Stata 16.1.  
434

435 **ACKNOWLEDGEMENTS**

436 **Office for National Statistics:** Iain Bell, Ian Diamond, Alex Lambert, Pete Benton, Emma Rourke,

437 Stacey Hawkes, Sarah Henry, James Scruton, Peter Stokes, Tina Thomas.

438 **Office for National Statistics, Analysis:** John Allen, Russell Black, Heather Bovill, David Brauholtz,

439 Dominic Brown, Sarah Collyer, Megan Crees, Colin Daghish, Byron Davies, Hannah Donnarumma, Julia

440 Douglas-Mann, Antonio Felton, Hannah Finselbach, Eleanor Fordham, Alberta Ipser, Joe Jenkins, Joel

441 Jones, Katherine Kent, Geeta Kerai, Lina Lloyd, Victoria Masding, Ellie Osborn, Alpi Patel, Elizabeth

442 Pereira, Tristan Pett, Melissa Randall, Donna Reeve, Palvi Shah, Ruth Snook, Ruth Studley, Esther

443 Sutherland, Eliza Swinn, Heledd Thomas, Anna Tudor, Joshua Weston.

444 **Office for National Statistics, Secure Research Service:** Shayla Leib, James Tierney, Gabor Farkas, Raf

445 Cobb, Folkert van Galen, Lewis Compton, James Irving, John Clarke, Rachel Mullis, Lorraine Ireland,

446 Diana Airimitoiaie, Charlotte Nash, Danielle Cox, Sarah Fisher, Zoe Moore, James McLean, Matt Kerby.

447 **University of Oxford, Nuffield Department of Medicine:** Ann Sarah Walker, Derrick Crook, Philippa C

448 Matthews, Tim Peto, Emma Pritchard, Nicole Stoesser, Karina-Doris Vihta, Alison Howarth, George

449 Doherty, James Kavanagh, Kevin K Chau, Stephanie B Hatch, Daniel Ebner, Lucas Martins Ferreira,

450 Thomas Christott, Brian D Marsden, Wanwisa Dejnirattisai, Juthathip Mongkolsapaya, Sarah

451 Hoosdally, Richard Cornall, David I Stuart, Gavin Sreaton.

452 **University of Oxford, Nuffield Department of Population Health:** Koen Pouwels.

453 **University of Oxford, Big Data Institute:** David W Eyre.

454 **University of Oxford, Radcliffe Department of Medicine:** John Bell.

455 **Oxford University Hospitals NHS Foundation Trust:** Stuart Cox, Kevin Paddon, Tim James.

456 **University of Manchester:** Thomas House.

457 **Public Health England:** John Newton, Julie Robotham, Paul Birrell.

458 **IQVIA:** Helena Jordan, Tim Sheppard, Graham Athey, Dan Moody, Leigh Curry, Pamela Brereton.

459 **National Biocentre:** Ian Jarvis, Kirsty Howell, Bobby Mallick, Phil Eeles.

460 **Glasgow Lighthouse Laboratory:** Jodie Hay, Harper Vansteenhouse.

461 Department of Health: Jessica Lee.

462

#### 463 **FUNDING**

464 This study is funded by the Department of Health and Social Care. ASW, EP, JVR, TEAP, NS, DE, KBP  
465 are supported by the National Institute for Health Research Health Protection Research Unit (NIHR  
466 HPRU) in Healthcare Associated Infections and Antimicrobial Resistance at the University of Oxford in  
467 partnership with Public Health England (PHE) (NIHR200915). ASW and TEAP are also supported by  
468 the NIHR Oxford Biomedical Research Centre. EP and KBP are also supported by the Huo Family  
469 Foundation. ASW is also supported by core support from the Medical Research Council UK to the  
470 MRC Clinical Trials Unit [MC\_UU\_12023/22] and is an NIHR Senior Investigator. PCM is funded by  
471 Wellcome (intermediate fellowship, grant ref 110110/Z/15/Z) and holds an NIHR BRC Senior  
472 Fellowship award. The views expressed are those of the authors and not necessarily those of the  
473 National Health Service, NIHR, Department of Health, or PHE.

474

475 The funders had no role in study design, data collection and interpretation, or the decision to submit  
476 the work for publication.

477

#### 478 **COMPETING INTEREST**

479 DWE declares lecture fees from Gilead, outside the submitted work. No other author has a conflict of  
480 interest to declare.

481

#### 482 **DATA AVAILABILITY**

483 De-identified study data are available for access by accredited researchers in the ONS Secure  
484 Research Service (SRS) for [accredited research purposes](#) under part 5, chapter 5 of the Digital  
485 Economy Act 2017. For further information about accreditation, contact  
486 [Research.Support@ons.gov.uk](mailto:Research.Support@ons.gov.uk) or visit the [SRS website](#).

487

488 **CONTRIBUTIONS**

489 The study was designed and planned by ASW, JF, JB, JN, IB, ID, KBP and JVR, and is being conducted  
490 by ASW, IB and RS. This specific analysis was designed by ASW, KBP, DWE, PCM, NS and TEAP. ASW,  
491 EP, KBP, TH, and KDV contributed to the statistical analysis of the survey data. JH conducted analysis  
492 of the RT-PCR data. ASW drafted the manuscript and all authors contributed to interpretation of the  
493 data and results and revised the manuscript. All authors approved the final version of the  
494 manuscript.

495

496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546

## REFERENCES

- Adams, E. R., Ainsworth, M., Anand, R., & et al. . (2020). Antibody testing for COVID-19: A report from the National COVID Scientific Advisory Panel [version 1; peer review: 1 approved]. *Wellcome Open Res*, 5:139. doi:<https://doi.org/10.12688/wellcomeopenres.15927.1>
- Arkhipova-Jenkins, I., Helfand, M., Armstrong, C., Gean, E., Anderson, J., Paynter, R. A., & Mackey, K. (2021). Antibody Response After SARS-CoV-2 Infection and Implications for Immunity : A Rapid Living Review. *Ann Intern Med*. doi:10.7326/M20-7547
- Buitrago-Garcia, D., Egli-Gany, D., Counotte, M. J., Hossmann, S., Imeri, H., Ipekci, A. M., . . . Low, N. (2020). Occurrence and transmission potential of asymptomatic and presymptomatic SARS-CoV-2 infections: A living systematic review and meta-analysis. *PLoS Med*, 17(9), e1003346. doi:10.1371/journal.pmed.1003346
- Byambasuren, O., Cardona, M., Bell, K., Clark, J., McLaws, M., & Glasziou, P. (2020). Estimating the extent of asymptomatic COVID-19 and its potential for community transmission: systematic review and meta-analysis. *Official Journal of the Association of Medical Microbiology and Infectious Disease Canada*, 5, 223-234.
- Edwards, T., Santos, V. S., Wilson, A. L., Cubas-Atienzar, A. I., Kontogianni, K., Williams, C. T., . . . Cuevas, L. E. (2020). Variation of SARS-CoV-2 viral loads by sample type, disease severity and time: a systematic review. *MedRxiv*, <https://www.medrxiv.org/content/10.1101/2020.1109.1116.20195982v20195981.full.pdf>.
- Endo, A., Centre for the Mathematical Modelling of Infectious Diseases, C.-W. G., Abbott, S., Kucharski, A. J., & Funk, S. (2020). Estimating the overdispersion in COVID-19 transmission using outbreak sizes outside China. *Wellcome Open Res*, 5, 67. doi:10.12688/wellcomeopenres.15842.3
- Flaxman, S., Mishra, S., Gandy, A., Unwin, H. J. T., Mellan, T. A., Coupland, H., . . . Bhatt, S. (2020). Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. *Nature*, 584(7820), 257-261. doi:10.1038/s41586-020-2405-7
- Gandhi, M., Beyrer, C., & Goosby, E. (2020). Masks Do More Than Protect Others During COVID-19: Reducing the Inoculum of SARS-CoV-2 to Protect the Wearer. *J Gen Intern Med*. doi:10.1007/s11606-020-06067-8
- Gandhi, M., & Rutherford, G. W. (2020). Facial Masking for Covid-19 - Potential for "Variolation" as We Await a Vaccine. *N Engl J Med*. doi:10.1056/NEJMp2026913
- Goldman, J. D., Wang, K., Roltgen, K., Nielsen, S. C. A., Roach, J. C., Naccache, S. N., & et al. (2020). Reinfection with SARS-CoV-2 and Failure of Humoral Immunity: a case report. *MedRxiv*, <https://www.medrxiv.org/content/10.1101/2020.09.22.20192443v1.full.pdf+html>.
- Hao, X., Cheng, S., Wu, D., Wu, T., Lin, X., & Wang, C. (2020). Reconstruction of the full transmission dynamics of COVID-19 in Wuhan. *Nature*, 584(7821), 420-424. doi:10.1038/s41586-020-2554-8
- Hay, J. A., Kennedy-Shaffer, L., Kanjilal, S., Lipsitch, M., & Mina, M. J. (2020). Estimating epidemiologic dynamics from single crosssectional viral load distributions *MedRxiv*, <https://doi.org/10.1101/2020.10.08.20204222>.
- Jacot, D., Greub, G., Jatou, K., & Opat, O. (2020). Viral load of SARS-CoV-2 across patients and compared to other respiratory viruses. *MedRxiv*, <https://www.medrxiv.org/content/10.1101/2020.1107.1115.20154518v20154512.full.pdf>.
- Jefferson, T., Spencer, E., Brassey, J., & Heneghan, C. (2020). Viral cultures for COVID-19 infectious potential assessment – a systematic review. *Clinical Infectious Diseases*, *ciaa1764*.
- Jones, T. C., Mühlemann, B., Veith, T., Biele, G., Zuchowski, M., Hoffmann, J., . . . Drosten, C. (2020). An analysis of SARS-CoV-2 viral load by patient age. *MedRxiv*, <https://www.medrxiv.org/content/10.1101/2020.06.08.20125484v1>.
- Ladhani, S., Baawuah, F., Beckmann, J., Okike, I., Ahmad, S., Garstang, J., . . . Linley, E. (2020). *Prospective active national surveillance of preschools and primary schools for SARS-CoV-2*

- 547 *infection and transmission in England, June 2020 (sKIDs COVID-19 surveillance in school KIDs)*  
548 *Phase 1 Report*. Retrieved from
- 549 Lavezzo, E., Franchin, E., Ciavarella, C., Cuomo-Dannenburg, G., Barzon, L., Del Vecchio, C., . . .  
550 Crisanti, A. (2020). Suppression of a SARS-CoV-2 outbreak in the Italian municipality of Vo'.  
551 *Nature*, 584(7821), 425-429. doi:10.1038/s41586-020-2488-1
- 552 Lee, L. Y. W., Rozmanowski, S., Pang, M., Charlett, A., Anderson, C., Barnard, M., . . . Peto, T. E. A.  
553 (2021). An observational study of SARS-CoV-2 infectivity by viral load and demographic  
554 factors and the utility lateral flow devices to prevent transmission. *JAMA (submitted)*,  
555 [http://modmedmicro.nsms.ox.ac.uk/wp-](http://modmedmicro.nsms.ox.ac.uk/wp-content/uploads/2021/01/infectivity_manuscript_20210119_merged.pdf)  
556 [content/uploads/2021/01/infectivity\\_manuscript\\_20210119\\_merged.pdf](http://modmedmicro.nsms.ox.ac.uk/wp-content/uploads/2021/01/infectivity_manuscript_20210119_merged.pdf).
- 557 Lee, S., Kim, T., Lee, E., Lee, C., Kim, H., Rhee, H., . . . Kim, T. H. (2020). Clinical Course and Molecular  
558 Viral Shedding Among Asymptomatic and Symptomatic Patients With SARS-CoV-2 Infection  
559 in a Community Treatment Center in the Republic of Korea. *JAMA Intern Med*.  
560 doi:10.1001/jamainternmed.2020.3862
- 561 Lumley, S. F., O'Donnell, D., Stoesser, N. E., Matthews, P. C., Howarth, A., Hatch, S. B., . . . Oxford  
562 University Hospitals Staff Testing, G. (2021). Antibody Status and Incidence of SARS-CoV-2  
563 Infection in Health Care Workers. *N Engl J Med*, 384(6), 533-540.  
564 doi:10.1056/NEJMoa2034545
- 565 Moraz, M., Jacot, D., Papadimitriou-Olivgeris, M., Senn, L., Greub, G., Jatou, K., & Opota, O. (2020).  
566 Clinical importance of reporting SARS-CoV-2 viral loads across the different stages of the  
567 COVID-19 pandemic. *MedRxiv*,  
568 <https://www.medrxiv.org/content/10.1101/2020.07.10.20149773v1>.
- 569 National Sars-CoV-Serology Assay Evaluation Group. (2020). Performance characteristics of five  
570 immunoassays for SARS-CoV-2: a head-to-head benchmark comparison. *Lancet Infect Dis*.  
571 doi:10.1016/S1473-3099(20)30634-4
- 572 Office for National Statistics. (2021). Coronavirus (COVID-19) Infection Survey, UK: 26 March 2021  
573 [https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsan](https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/bulletins/coronaviruscovid19infectionsurveyspilot/26march2021)  
574 [ddiseases/bulletins/coronaviruscovid19infectionsurveyspilot/26march2021](https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/bulletins/coronaviruscovid19infectionsurveyspilot/26march2021).
- 575 Pouwels, K. B., House, T., Pritchard, E., Robotham, J. V., Birrell, P. J., Gelman, A., . . . Team, C.-I. S.  
576 (2020). Community prevalence of SARS-CoV-2 in England from April to November, 2020:  
577 results from the ONS Coronavirus Infection Survey. *Lancet Public Health*, 6, E30-38.  
578 doi:10.1016/S2468-2667(20)30282-6
- 579 Public Health England. (2020). *Investigation of novel SARS-CoV-2 variant: Variant of Concern*  
580 *202012/01; Technical briefing 2*  
581 ([https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/948152/Technical_Briefing_VOC202012-2_Briefing_2_FINAL.pdf)  
582 [data/file/948152/Technical\\_Briefing\\_VOC202012-2\\_Briefing\\_2\\_FINAL.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/948152/Technical_Briefing_VOC202012-2_Briefing_2_FINAL.pdf)) Retrieved from
- 583 Rasmussen, A. L., & Popescu, S. V. (2021). SARS-CoV-2 transmission without symptoms. *Science*,  
584 371(6535), 1206-1207. doi:10.1126/science.abf9569
- 585 Riley, S., Ainslie, K. E. C., Eales, O., Walters, C. E., Wang, H., Atchison, C. J., . . . Elliott, P. (2020).  
586 Transient dynamics of SARS-CoV-2 as England exited national lockdown. *MedRxiv*,  
587 <https://www.medrxiv.org/content/10.1101/2020.1108.1105.20169078v20169071>.
- 588 Riley, S., Ainslie, K. E. C., Eales, O., Walters, C. E., Wang, H., Atchison, C. J., . . . Elliott, P. (2020).  
589 Resurgence of SARS-CoV-2 in England: detection by community antigen surveillance.  
590 *MedRxiv*, <https://www.medrxiv.org/content/10.1101/2020.09.11.20192492v2>.
- 591 Singanayagam, A., Patel, M., Charlett, A., Lopez Bernal, J., Saliba, V., Ellis, J., . . . Gopal, R. (2020).  
592 Duration of infectiousness and correlation with RT-PCR cycle threshold values in cases of  
593 COVID-19, England, January to May 2020. *Euro Surveill*, 25(32). doi:10.2807/1560-  
594 7917.ES.2020.25.32.2001483
- 595 To, K. K., Hung, I. F., Ip, J. D., Chu, A. W., Chan, W. M., Tam, A. R., . . . Yuen, K. Y. (2020). COVID-19 re-  
596 infection by a phylogenetically distinct SARS-coronavirus-2 strain confirmed by whole  
597 genome sequencing. *Clin Infect Dis*. doi:10.1093/cid/ciaa1275

598 Tomassini, S., Kotecha, D., Bird, P. W., Folwell, A., Biju, S., & Tang, J. W. (2020). Setting the criteria for  
599 SARS-CoV-2 reinfection - six possible cases. *J Infect.* doi:10.1016/j.jinf.2020.08.011  
600 Walker, A. S., Vihta, K., Gethings, O., Pritchard, E., Jones, J., House, T., . . . team, C.-I. S. (2021).  
601 Increased infections, but not viral burden, with a new SARS-CoV-2 variant. *medRxiv*  
602 *2021.01.13.21249721*.  
603 Yokota, I., Shane, P. Y., Okada, K., Unoki, Y., Yang, Y., Inao, T., . . . Teshima, T. (2020). Mass screening  
604 of asymptomatic persons for SARS-CoV-2 using saliva. *Clinical Infectious Diseases*, ciaa1388.  
605  
606



607 **Table 1 Genes detected in positive swabs**

Number of genes detected	All positives (N=27,879)		First positive per participant (N=21,811)	
	n (%)	Median CT* (IQR) [range]	n (%)	Median CT* (IQR) [range]
1	6,550 (23%)	33.8 (32.9-34.7) [12.7-38.7]	5,102 (23%)	33.9 (32.9-34.7) [12.7-38.7]
2	1,145 (4%)	32.3 (30.9-33.4) [10.3-37.2]	773 (4%)	32.3 (30.7-33.4) [10.3-37.2]
2: ORF1ab+N 16 Nov 2020 onwards	9,867 (35%)	26.4 (19.4-31.1) [9.2-37.8]	8,184 (38%)	25.3 (18.6-30.7) [9.2-37.8]
3	10,317 (37%)	25.3 (19.8-29.5) [9.3-36.8]	7,752 (36%)	23.9 (18.8-28.8) [9.3-36.8]
Genes detected				
N only	4,479 (13%)	33.9 (33.0-34.8) [26.1-38.7]	3,419 (16%)	34.0 (33.1-34.8) [28.2-38.7]
ORF1ab only	2,044 (7%)	33.6 (32.6-34.5) [16.8-38.3]	1,656 (8%)	33.7 (32.7-34.6) [16.8-38.3]
S only**	27 (0.1%)	34.9 (33.5-36.1) [12.7-37.3]	27 (0.1%)	34.9 (33.5-36.1) [12.7-37.3]
N+ORF1ab: before 16 Nov 2020	731 (3%)	31.9 (30.3-32.9) [10.3-37.2]	497 (2%)	31.8 (29.7-33.0) [10.3-38.2]
N+ORF1ab: 16 Nov 2020 onwards	9,867 (35%)	26.4 (19.4-31.1) [9.2-37.8]	8,184 (38%)	23.9 (18.8-28.8) [9.3-36.8]
S+ORF1ab	190 (0.7%)	32.5 (31.2-33.5) [15.1-36.6]	138 (0.6%)	32.4 (31.0-33.6) [15.1-36.6]
N+S	224 (0.8%)	33.4 (32.5-34.2) [25.0-36.8]	138 (0.6%)	33.3 (32.4-34.3) [27.3-36.8]
N+S+ORF1ab	10,317 (37%)	25.3 (19.8-29.5) [9.3-36.8]	7,752 (36%)	25.3 (18.6-30.7) [9.2-37.8]

608 \* taking the mean CT per positive swab across positive gene targets (Spearman rho=0.98 for each pair of genes where both positive, p<0.0001)

609 \*\* 17/27 before mid-May only: after this samples positive for the S gene only were not called positive overall by the algorithm and therefore reflect likely  
610 recording errors.

611 Note: excluding 23 positive results without Ct values or genes detected available. Comparing first vs subsequent positives per participant, exact p<0.0001 for  
612 both number of genes detected and specific genes detected.

613

614 **Table 2 Evidence supporting positive test results indicating presence of virus and impact on other**  
 615 **factors**

	Strength of evidence for true infection			p (exact)
	Higher	Moderate	Lower	
Number (col %) (N=27,879)	21,329 (77%)	4,741 (17%)	1,809 (6%)	
<b>Factors determining classification</b>				
Number of genes detected (row %)	3: 10,317 (48%) 2: 11,012 (52%)	1: 4,741 (100%)	1: 1,809 (100%)	
CT, median	26.2	33.4	34.8	
CT, n (row %) <34*	21,070 (98.8%)	3,613 (76%)	0 (0%)	
Symptoms around test, n (row %)	12,466 (58%)	2,243 (47%)	0 (0%)	<0.0001 (exc lower)
Occupational risk**, n (row %)	1,322 (6%)	307 (6%)	0 (0%)	0.48 (exc lower)
<b>Other factors</b>				
Cough, fever, anosmia, ageusia around test, n (row %)	9,345 (44%)	1,241 (26%)	0 (0%)	<0.0001 (exc lower)
First positive test n (row %) (vs subsequent positive test)	16,709 (78%)	3,508 (74%)	1,594 (88%)	<0.0001
First test in study, n (row %) (vs follow-up ie prior negative in study)	2,281 (11%)	482 (10%)	199 (11%)	0.49
Any genome sequence obtained, confirming presence of virus†	6,621/9,022 (73%)	544/2,315 (24%)	0/836 (0%)	<0.0001
Any other household member ever positive‡	11,493/18,494 (62%)	1,513/4,004 (38%)	318/1,525 (21%)	<0.0001

616 \* approximate 97.5<sup>th</sup> percentile of CT in higher evidence positives through 2 August when  
 617 classification first applied.

618 \*\* reported working in a patient-facing healthcare role/care/residential home.

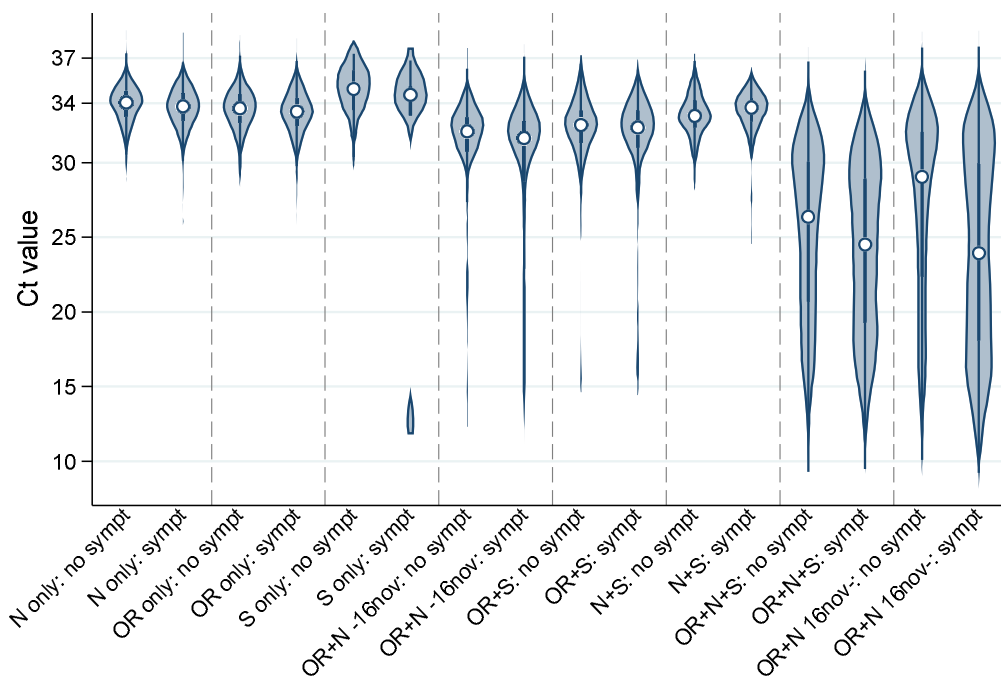
619 † any genome sequence obtained out of attempted (other positives not found or not yet attempted)

620 ‡ denominator households with 2 or more study participants.

621 Note: classification arbitrarily determined on 2 August 2020 based on the number of genes detected,  
 622 Ct values and pre-test probability (see Methods).

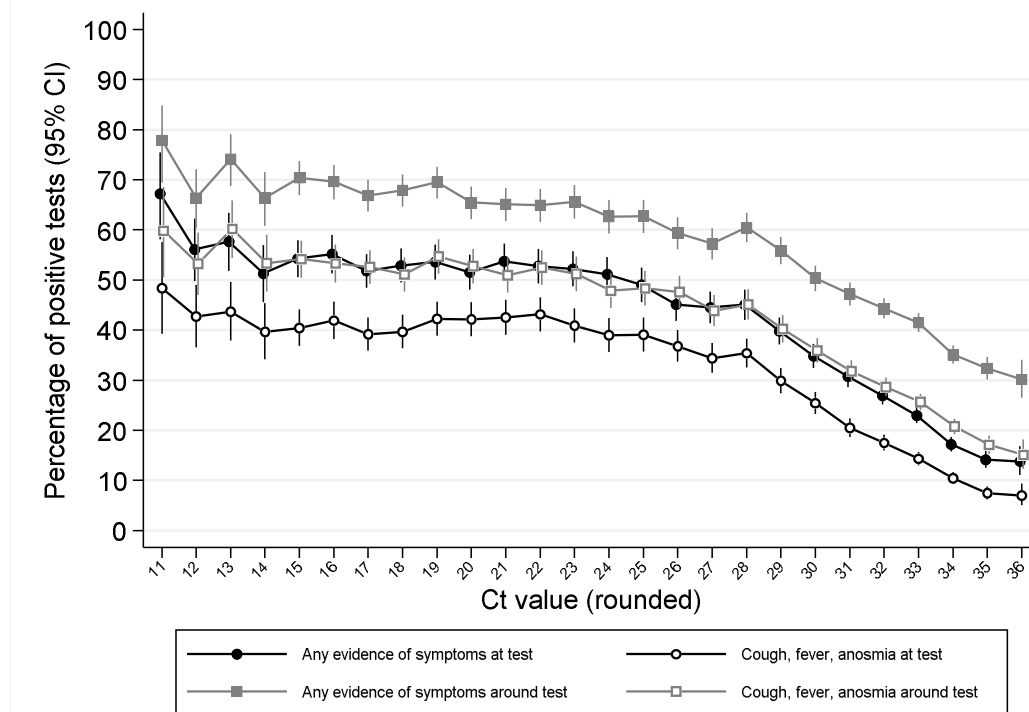
623

**Figure 1 Distribution of Ct values at each positive test by genes detected and self-reported symptoms**



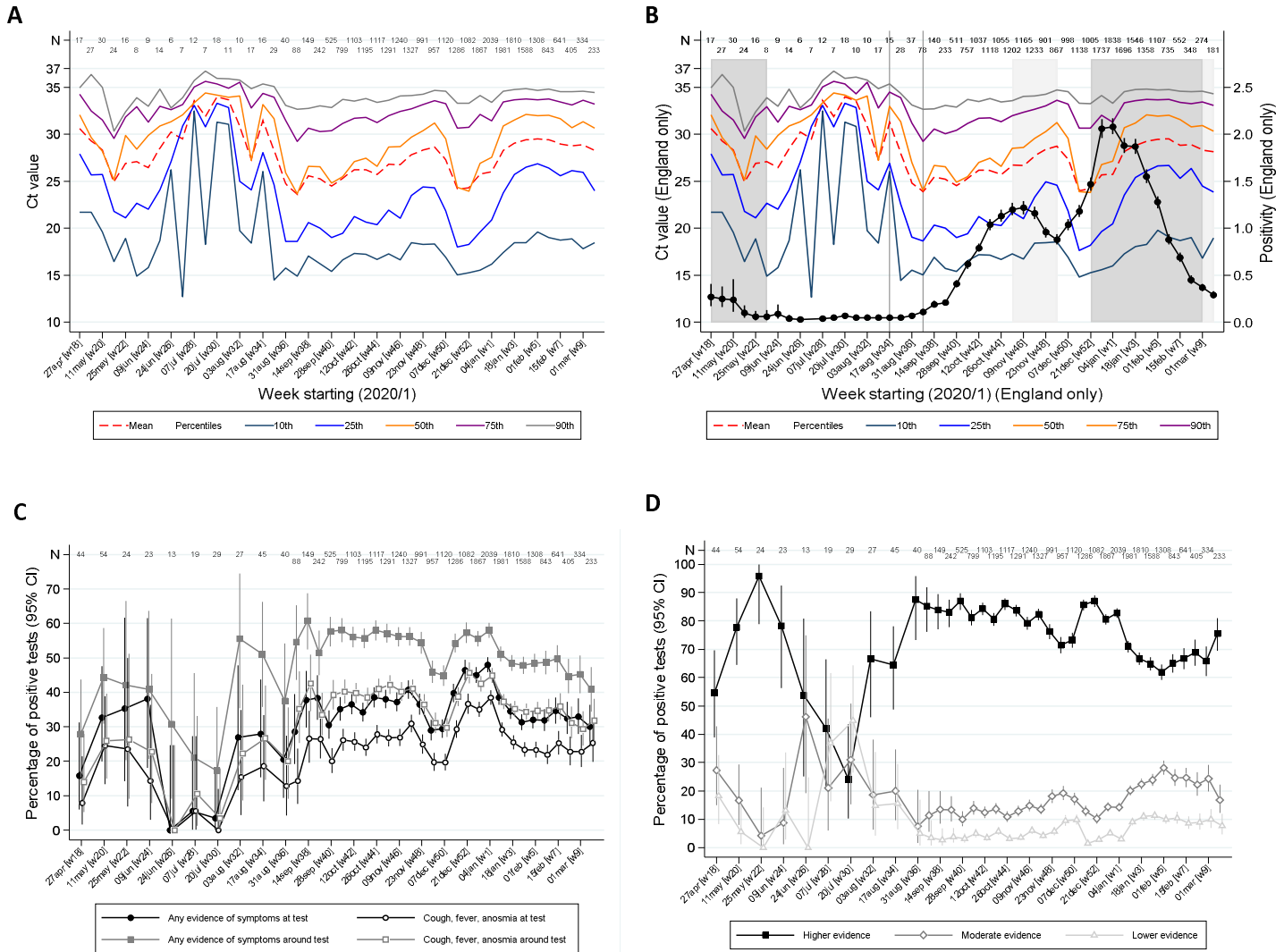
Note: points show the median and boxes the interquartile range. OR=ORF1ab.  
Positives where only the ORF1ab+N genes were detected are split by whether the swab was taken before or after 16 November 2020, reflecting the expansion of B.1.1.7 (which has S-gene target failure on the assay used in the survey).

**Figure 2 Percentage reporting symptoms by Ct value**



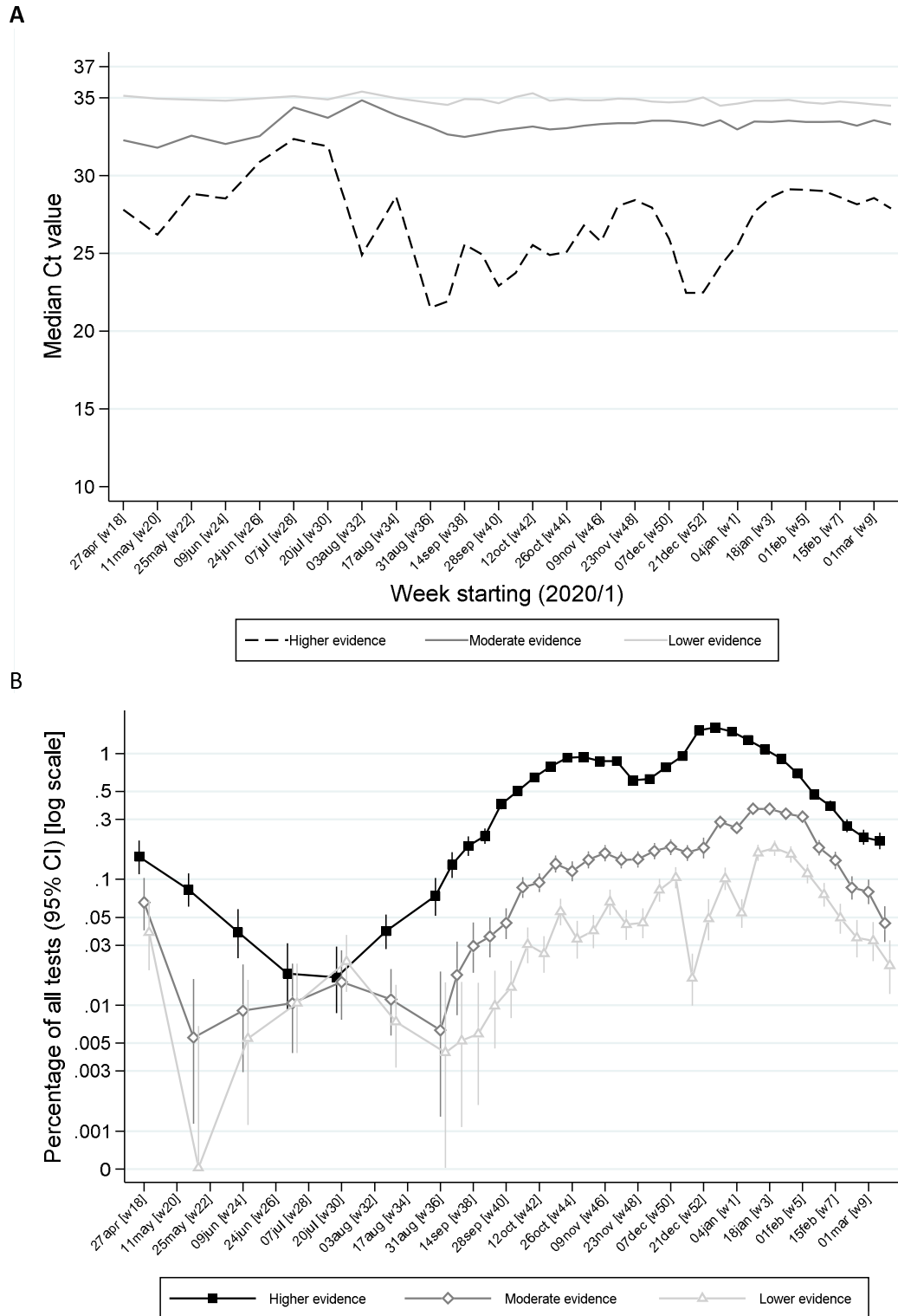
Note: Points show the percentage of positive tests with each rounded Ct value reporting any symptoms or cough, fever, anosmia/ageusia at each test or around each test (see Methods for symptoms collection and definitions). Ct values under 11 and over 36 grouped with 11 and 36 respectively.

**Figure 3** Variation over calendar time in the distribution of Ct values in the UK (A) and England (B) together with percentage positivity in England (B), and in self-reported symptoms (C) and evidence supporting positives (D)



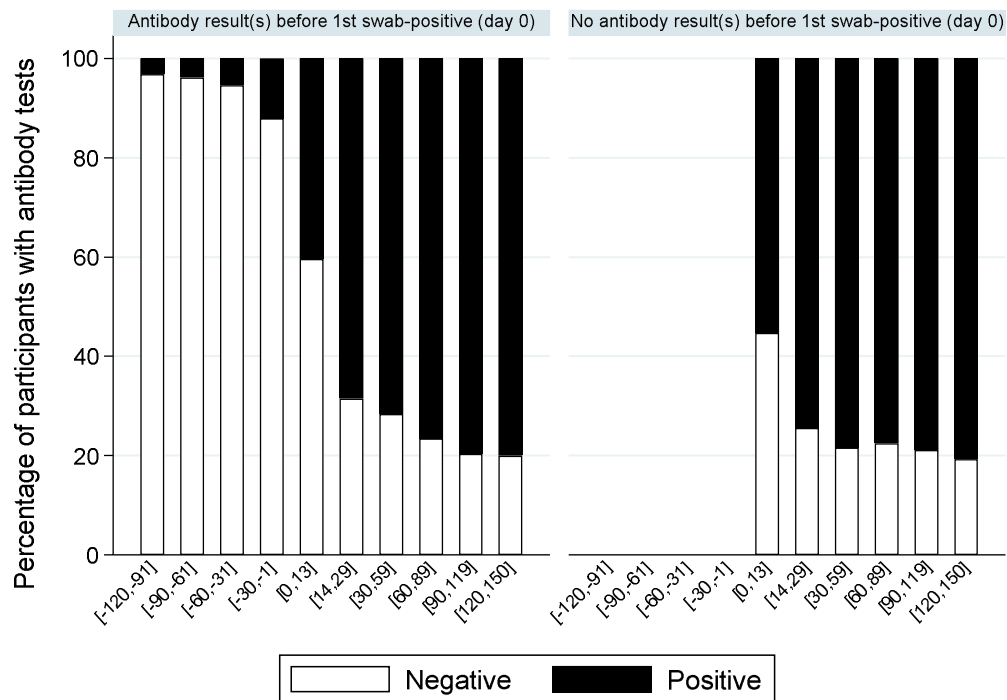
Note: panel (A) shows the distribution of Ct values each week including all positives across the UK. Panel (B) is restricted to England shown together with the official estimates of positivity as reported by the Office for National Statistics (black line) and periods of national “stay-at-home” restrictions (schools shut in dark grey, schools open in light gray). Panels (C) and (D) show the proportions reporting symptoms and with different levels of evidence supporting the positive test respectively. Variation in the width of 95% CI reflects the increase in size of the survey from mid August (**Supplementary File 1**).

Figure 4 Ct values (A) and percentage positive of all tests (B) by level of evidence and time



Note: panel (A) shows median Ct values according to level of evidence and panel (B) percentage of all swab tests positive according to level of evidence over calendar time. The early part of the study is grouped into three week periods due to lower numbers of positives.

**Figure 5 Percentage of positive antibody tests over time from first positive swab**



Note: showing the percentage of participants with S-antibody positive or negative tests according to days from their first positive swab, separately for those with and without any antibody results prior to their first positive swab.