

1 SARS-CoV-2 infectivity by viral load, S gene variants and demographic
2 factors and the utility of lateral flow devices to prevent transmission

3

4 Lennard YW Lee,^{1*} Stefan Rozmanowski,^{2*} Matthew Pang,² Andre Charlett,³ Charlotte Anderson,³
5 Gareth J Hughes,³ Matthew Barnard,² Leon Peto,¹ Richard Vipond,⁴ Alex Sienkiewicz,⁴ Susan
6 Hopkins,³ John Bell,¹ Derrick W Crook,^{1,5,6} Nick Gent,³ A Sarah Walker,^{1,5,6} Tim EA Peto^{1,5,6§}, David W
7 Eyre^{5,6,7§}

8 ¹ Nuffield Department of Medicine, University of Oxford, UK

9 ² Department of Health and Social Care, UK Government, London, UK

10 ³ Public Health England, London, UK

11 ⁴ Public Health England, Porton Down, UK

12 ⁵ NIHR Oxford Biomedical Research Centre, University of Oxford, Oxford, UK

13 ⁶ NIHR Health Protection Research Unit in in Healthcare Associated Infections and Antimicrobial
14 Resistance, University of Oxford, UK

15 ⁷ Big Data Institute, Nuffield Department of Population Health, University of Oxford, Oxford, UK

16

17 *Contributed equally

18 §Contributed equally

19

20 Corresponding author: David Eyre, John Radcliffe Hospital, Oxford, OX3 9DU, United Kingdom.
21 david.eyre@bdi.ox.ac.uk.

22 Alternative corresponding author: Tim Peto, John Radcliffe Hospital, Oxford, OX3 9DU, United
23 Kingdom. tim.peto@ndm.ox.ac.uk.

24 **Keywords:** infectivity, contact tracing, SARS-CoV-2, lateral flow device, B.1.1.7 variant

25 **Running title:** SARS-CoV-2 infectivity

26

27 Key points

28 In 2,474,066 contacts of 1,064,004 SARS-CoV-2 cases, PCR-positive tests in contacts increased with
29 higher index case viral loads, the B.1.1.7 variant and household contact. Children were less
30 infectious. Lateral flow devices can detect 83.0-89.5% of infections leading to onward transmission.

31

32 Abstract

33 **Background:** How SARS-CoV-2 infectivity varies with viral load is incompletely understood. Whether
34 rapid point-of-care antigen lateral flow devices (LFDs) detect most potential transmission sources
35 despite imperfect sensitivity is unknown.

36 **Methods:** We combined SARS-CoV-2 testing and contact tracing data from England between 01-
37 September-2020 and 28-February-2021. We used multivariable logistic regression to investigate
38 relationships between PCR-confirmed infection in contacts of community-diagnosed cases and index
39 case viral load, S gene target failure (proxy for B.1.1.7 infection), demographics, SARS-CoV-2
40 incidence, social deprivation, and contact event type. We used LFD performance to simulate the
41 proportion of cases with a PCR-positive contact expected to be detected using one of four LFDs.

42 **Results:** 231,498/2,474,066 (9%) contacts of 1,064,004 index cases tested PCR-positive. PCR-positive
43 results in contacts independently increased with higher case viral loads (lower Ct values) e.g.,
44 11.7%(95%CI 11.5-12.0%) at Ct=15 and 4.5%(4.4-4.6%) at Ct=30. B.1.1.7 infection increased PCR-
45 positive results by ~50%, (e.g. 1.55-fold, 95%CI 1.49-1.61, at Ct=20). PCR-positive results were most
46 common in household contacts (at Ct=20.1, 8.7%[95%CI 8.6-8.9%]), followed by household visitors
47 (7.1%[6.8-7.3%]), contacts at events/activities (5.2%[4.9-5.4%]), work/education (4.6%[4.4-4.8%]),
48 and least common after outdoor contact (2.9%[2.3-3.8%]). Contacts of children were the least likely
49 to test positive, particularly following contact outdoors or at work/education. The most and least
50 sensitive LFDs would detect 89.5%(89.4-89.6%) and 83.0%(82.8-83.1%) of cases with PCR-positive
51 contacts respectively.

52 **Conclusions:** SARS-CoV-2 infectivity varies by case viral load, contact event type, and age. Those with
53 high viral loads are the most infectious. B.1.1.7 increased transmission by ~50%. The best performing
54 LFDs detect most infectious cases.

55

56 Introduction

57 The global health impact of SARS-CoV-2 is profound.¹ There is widespread on-going transmission
58 despite control efforts predominantly focused on quarantining symptomatic cases and population-
59 level self-isolation.² The emergence of potentially more transmissible variants, such as B.1.1.7³ which
60 has spread widely in the UK, has hampered control. However, vaccine roll-out offers the prospect of
61 reduced disease and transmission.⁴

62 Intermittent national and regional social distancing and self-isolation measures have been imposed
63 in many countries.^{5,6} Additional self-isolation measures for “contacts” (individuals exposed to SARS-
64 CoV-2) vary by country, but generally last 7-14 days.⁷ While reducing transmission,
65 quarantine/isolation measures have indirectly had many wider effects on economic productivity,
66 well-being⁸ and non-COVID-19-related excess deaths.⁹⁻¹¹ Not all exposure to SARS-CoV-2 leads to
67 infection, e.g., in some settings only 5-7% of exposed “contacts” develop COVID-19 infection^{12,13} and
68 modelling suggests ~15% of individuals are responsible for most SARS-CoV-2 transmission.¹⁴
69 Therefore, using isolation selectively for those who are most infectious could lessen some of its
70 collateral impacts.^{12,13}

71 Our understanding of how individual infectiousness varies is limited. Several assays for infectivity
72 have been proposed. Functional assays include animal and cell culture models, whereas viral sub-
73 genomic mRNA is a nucleic acid-based measure of infectivity.¹⁵ Detection of viral protein, i.e.
74 antigen, as assessed by lateral flow devices (LFDs), has been shown to be more closely linked to viral
75 culture infectivity than PCR measurements.¹⁶ However, few of these surrogate measures of
76 infectivity have been convincingly demonstrated to predict the real-world likelihood of a SARS-CoV-2
77 infected individual infecting someone else.

78 Here we use data from the England’s national contact tracing and testing programs to explore the
79 relationship between infectivity and SARS-CoV-2 viral load, as measured by PCR cycle threshold (Ct)
80 values. We identify demographic factors associated with infectivity and assess the impact of the
81 emergence of the B.1.1.7 variant. We apply our results to a population of PCR-positive individuals to
82 estimate the proportion of infectious individuals detected by viral antigen LFDs under a range of
83 performance conditions.

84

85 Methods

86 Data from community and hospital PCR testing in England between 01-September-2020 and 28-
87 February-2021 were obtained and linked with national contact tracing data by the UK Government
88 Department of Health and Social Care. Data extracts were de-identified prior to analysis and
89 included for PCR-confirmed cases and their contacts: demographic details (age, sex, ethnicity), if
90 symptoms were present for cases and the timing of testing relative to symptom onset, and test
91 results, as well as details on the nature of the contact events.

92 Index cases and contacts

93 We defined index cases as SARS-CoV-2 PCR-positive individuals with a community-based test
94 performed by three high-throughput national testing facilities (“Lighthouse Laboratories” in Milton
95 Keynes, Alderley Park or Glasgow), which reported Ct values indicating viral load. Samples were

96 processed using the same RNA extraction and Thermo Fisher TaqPath PCR platform in each
97 laboratory (targeting S and N genes, and ORF1ab; details in Supplement). Only the first positive
98 result per person was included. Index cases without available Ct values were excluded. The B.1.1.7
99 variant contains a deletion in the S gene, resulting in S gene target failure (SGTF). Sequencing of
100 SGTF samples showed 89.5% were due to B.1.1.7 by mid-January 2021,¹⁷ so SGTF was used as a
101 proxy for B.1.1.7.

102 Contacts of index cases were defined as all individuals notified to the national contact tracing service
103 from the day of the index cases' positive test until 10 days later with whom the index case had been
104 in close proximity from 48 hours before their symptom onset to 10 days afterwards (further
105 definitions in Supplement). Contacts could be tested PCR-positive through any community or
106 hospital-based test as these were nationally reported.

107 [Statistical analysis](#)

108 We aimed to determine factors associated with PCR-positive results in contacts, including the
109 demographics, viral load and SGTF status of the index case. To identify outcomes most likely
110 representing onward transmission from the index case rather than a third party, we excluded
111 contacts named by more than one index case. We also restricted to positive test results obtained 1-
112 10 days following the index case's test date, i.e., the period when the index case may have been
113 infectious, to exclude earlier results in contacts and avoid contacts who were the source for the
114 index case's infection. Given these restrictions, the absolute proportion of contacts testing PCR-
115 positive cannot be interpreted as a secondary attack rate, because some onward transmission
116 events are excluded. Where contacts had more than one PCR test within the follow-up window, all
117 were considered to identify positive results.

118 We used multivariable logistic regression to investigate associations between PCR-confirmed
119 infection in contacts (including contacts whether or not they had PCR tests) and the index case's Ct
120 value and SGTF status (B.1.1.7 proxy), the contact event nature, the case's demographics, and
121 incidence and social deprivation index at the contact's home location. We did not adjust for
122 symptoms in the case, as these may be mediators of the effect of viral load on onward transmission.
123 We used splines to account for non-linearity in continuous variables and screened for all pairwise
124 interactions between main effects (details in Supplement).

125 We performed sensitivity analyses to test our restriction to contacts tested 1-10 days after each
126 index case and including only contacts with PCR tests. We used unadjusted linear regression to
127 investigate the proportion of the variation in Ct values in contacts that could be explained by the
128 case's Ct value.

129 [Simulations of the number of cases identified by antigen LFDs](#)

130 We used our findings to estimate the proportion of potential transmission events where the source
131 case would have been detected using an antigen LFD, using existing data on the sensitivity of four
132 LFDs: Innova, Deep Blue, Orient gene and Abbott.¹⁸ For each source case we simulated a positive or
133 negative LFD result by randomly drawing from the probability of a LFD being positive by the source
134 case's Ct value (see Supplement, Figure S1). Each simulation was repeated 1000 times. Additionally,
135 we ran simulations for a range of hypothetical LFD performances.

136 Ethics

137 The study was conducted as part of national COVID-19 surveillance under the provisions of Section
138 251 of the NHS Act 2006 and therefore did not require individual patient consent. It was approved
139 by Public Health England (PHE), the UK COVID-19 LFD oversight group and NHS Test and Trace. The
140 protocol for this work was reviewed by the PHE Research Ethics and Governance Group, which is the
141 PHE Research Ethics Committee, and was found to be fully compliant with all regulatory
142 requirements. As no regulatory or ethical issues were identified, it was agreed that a full ethical
143 review would not be needed, and the protocol was approved.

144

145 Results

146 Cases and contacts

147 3,577,246 SARS-CoV-2 PCR-positive results were available from England between 01-September-
148 2020 and 28-February-2021. Of these 1,818,456 (51%) tests were performed by three national
149 laboratories providing high-throughput community testing with a standardised PCR assay, yielding a
150 first positive test per person in 1,796,139 individuals. 27,893 (2%) were excluded as no Ct value was
151 available, 439,482 (24%) had no recorded contacts, and 264,760 (15%) had only recorded contacts
152 shared with other index cases, leaving 1,064,004 index cases in the analysis (Figure 1, Table S1).
153 487,653 (46%) cases had SGTF, consistent with B.1.1.7, increasing to near 100% by 28-February-2021
154 (Figure S2).

155 The 1,064,004 index cases had 2,974,596 contacts identified within 10 days of their test of whom
156 918,758 (31%) had a PCR test within ± 10 days of the index case and 638,456 (21%) tested PCR-
157 positive. 2,474,066 (83%) contacts were named only by a single case and are included in the analysis;
158 231,498 (9%) tested PCR positive 1-10 days after the index case's PCR-positive result, our main
159 outcome measure, i.e., consistent with possible transmission from the index case to the contact.

160 The median (IQR) age of cases and contacts was 36 (24-51) and 31 (16-49) years respectively, and
161 54% and 52% with available data were female (Table 1). Most contact events occurred within
162 households (77.4%), followed by visits to households (8.3%), workplaces or education (8.0%),
163 attending events or activities (5.3%), and outdoors (0.3%).

164 Predictors of PCR-positive results in contacts

165 On univariable analysis (Table 2, Figures S3-S6), PCR-positive tests in contacts were associated with
166 lower case Ct values (i.e. higher viral loads), SGTF in the index case, higher incidence in the local
167 population, less social deprivation, white ethnicity and male sex. Household contacts were most
168 likely to be PCR-positive. PCR-positive results were least frequent in contacts of children, with
169 highest rates in contacts of older adults.

170 Adjusted multivariable analysis showed strong evidence of effect modification (interactions) and
171 non-linear relationships, such that associations are best described graphically (Figures 2-3, S7-S11).
172 Index case Ct value was an important determinant of PCR-positive results in contacts, with an
173 approximately linear decline in positive results as Ct value increased, that was independent of the
174 nature of the contact event (Figure 2). For example, amongst household contacts, with other
175 variables set to median values/baseline categories, rates of PCR-positive tests were 11.7% (95%CI

176 11.5-12.0%) for index case Ct=15 and 4.5% (4.4-4.6%) for Ct=30. Contacts were most likely to test
177 PCR-positive after household contact (percentage of PCR-positive tests, at median Ct value=20.1,
178 8.7% [95%CI 8.6-8.9%]), followed by visitors to households (7.1% [6.8-7.3%]), contacts at
179 events/activities (5.2% [4.9-5.4%]), then work/education (4.6% [4.4-4.8%]), with outdoor contacts
180 the least likely to test positive (2.9% [2.3-3.8%]).

181 SGTF was associated with increased percentages of contacts testing PCR-positive, by 1.55-fold more
182 (95%CI 1.49-1.60) at index Ct=15, 1.55 (1.49-1.61) at Ct =20 and 1.44 (1.38-1.51) at Ct=30. At Ct
183 values near the upper limit of the assay, the relative increase in PCR-positive results fell to near 1
184 (Figure 3).

185 Contacts of children were the least likely to test positive, particularly following contact outdoors or
186 at work or in education (Figure 4). Most contact types had similar rates of PCR-positive results across
187 adult ages, except for household contact where risk increased as age increased above 35 years and
188 contact at work/education, events/activities and outdoors where risk of a PCR-positive result was
189 highest in adults in their 20s.

190 Associations between PCR-positive results in contacts and sex varied with age (Figure S8). Broadly,
191 increasing incidence increased PCR-positive contacts, likely reflecting increased acquisition from
192 third parties. There were fewer PCR-positive contacts in areas of greater social deprivation (Figure
193 S11) and amongst Black, Asian and minority ethnic groups (Figure S9-S10).

194 A sensitivity analysis supported the 1-10 day follow-up window for PCR results in contacts (Figure
195 S12). Case Ct values explained only a small proportion of the variability in contact Ct values
196 (unadjusted linear regression coefficient 0.14 [95%CI 0.13-0.14, $p < 0.001$], R-squared = 0.02).

197 Predictors of PCR-positive results in contacts attending PCR testing

198 In a sensitivity analysis restricted to contacts who had a PCR test (Table S2, Figures S13-S19), similar
199 relationships were seen between PCR-positive results and index case Ct values, contact type and
200 SGTF (Figures S19). While rates of PCR-positive results remained highest in older adult household
201 contacts, there was attenuation of the lower rates seen in children, consistent with main analysis
202 findings of less transmission from children arising from less testing being required or undertaken in
203 contacts of children (Figure S18). In contrast to the main analysis, contacts of all non-white ethnic
204 groups (Table S2) and those living in more deprived areas (Figure S17) were more likely to be PCR-
205 positive, potentially due to differences in access to and use of testing by different ethnic and
206 socioeconomic groups.

207 Proportion of cases with PCR-positive contacts detected by LFDs

208 Overall, 85.4% (197,677/231,497) of case-contact pairs with PCR-positive contacts, i.e., plausible
209 onward transmission, had case viral loads of $\geq 10,000$ RNA copies/ml (i.e. Ct ≤ 24.4) versus 75.2% of
210 all cases (800,020/1,064,004). Index cases with SGTF had lower Ct values, except for results near the
211 detection threshold (Figure S20).

212 As antigen LFD sensitivity varies by viral load, we used the distribution of viral loads in case-contact
213 pairs with a PCR-positive contact to simulate the proportion of such cases who would have been
214 detected using antigen LFDs (Figure 5). The Deep Blue LFD would have detected 85.9% (95%CI 85.8-
215 86.0%) of cases who plausibly subsequently transmitted to a contact, the Innova LFD 83.0% (82.8-

216 83.1%), the Orient Gene LFD 89.5% (89.4-89.6%) and the Abbott LFD 85.8% (85.7-86.0%).
217 Performance was very similar before and after B.1.1.7 expansion (Table S3). The performance
218 characteristics required to detect varying proportions of transmission sources by a novel LFD are
219 illustrated in Figure S21.

220

221 Discussion

222 We have performed a large-scale analysis of combined SARS-CoV-2 contact tracing and testing data
223 from England involving >2 million contacts of PCR-confirmed cases. We show SARS-CoV-2 infectivity
224 is associated with index case viral load, including after adjustment for demographic factors and type
225 of contact event. SGTF, a proxy for the B.1.1.7 variant, increased transmission by ~50% at most viral
226 loads. Onward transmission from children was relatively uncommon compared to adults, although
227 this may partly be due to less testing in their contacts. We confirm earlier findings that household
228 contact is associated with greater rates of transmission compared to workplace, educational or
229 recreational contact outside of homes.^{19,20}

230 Except SGTF, it is noteworthy that we found no evidence of significant interactions between Ct
231 values and any other variables in the analysis, i.e. the effect of viral load on infectivity is
232 generalisable across populations and settings. These results are consistent and add to a recent
233 smaller cohort study.²¹

234 Consistent with other reports³ we found that SGTF increased the proportion of contacts testing PCR-
235 positive, by around 55% at high viral loads (Ct values of 10-20), rising with moderate viral loads to a
236 maximum of 75% (Ct=25) before declining again to below 10% at low viral loads (Ct=34). SGTF also
237 affected how the likelihood of transmission varied with age, contact event type and ethnicity. The
238 higher relative infectiousness at moderate viral loads may represent increased infectiousness of
239 individual virions at viral loads where stochasticity is more important compared to higher viral loads.
240 The attenuation of the relative infectiousness at high Ct values partly arises from greater numbers of
241 wildtype strains exhibiting SGTF due to stochastic failure to detect a single gene at low viral loads. As
242 lower viral loads are less infectious, it may also reflect more PCR-positive contacts acquiring
243 infection from third parties, such that the characteristics of the index case matter less. This is
244 supported by the proportion of contacts testing PCR-positive not tending to zero at very low viral
245 loads.

246 85.4% PCR-positive contacts had an index case with a viral load of $\geq 10,000$ RNA copies/ml (Ct ≤ 24.4).
247 Hence, 85.4% of infections in contacts are potentially attributable to the 75.2% of cases overall with
248 a viral load of $\geq 10,000$ RNA copies/ml. While such data could be used to drive differential
249 interventions to prevent onward transmission with a particular focus on those with high viral loads,
250 our findings suggest that most infected individuals still have some risk of transmitting onwards
251 based on Ct values.

252 However, we show that several LFDs are sufficiently sensitive to detect most cases that led to
253 onward transmission. These tests offer potential advantages, in returning a result in 15-30 minutes,
254 not requiring laboratory infrastructure and costing significantly less than PCR tests. However post-
255 analytic infrastructure is still needed to collect results. Using the estimated sensitivity of four LFDs,

256 we estimate they would detect 83.0%-89.5% of cases leading to onward transmission. While such
257 performance is not sufficient to replace PCR for testing of all symptomatic individuals, use of LFDs in
258 addition to existing testing, particularly of those who otherwise would not be tested at all (including
259 those without symptoms), would allow many of the most infectious individuals to be identified
260 earlier, potentially preventing onward transmissions and helping to drive reproduction numbers
261 below 1, despite imperfect performance against PCR. The specificity of each LFD is another
262 important consideration, particularly as incidence falls; the false positive rate for the Innova LFD has
263 been previously reported as 0.32% (95%CI 0.20-0.48%),¹⁸ and large-scale evaluations of the other
264 LFDs are on-going. In settings where the positive predictive value of an LFD is insufficiently high,
265 confirmatory PCR testing may be required.

266 Our study has important limitations. Firstly, ascertaining infection in contacts depends on the
267 contact being reported by the case and the contact being tested. In the UK, PCR testing is only
268 recommended for those with symptoms and therefore we do not ascertain most asymptomatic
269 infections. Whilst Ct values are generally slightly lower in those without symptoms,²² they may
270 nevertheless contribute substantially to transmission.²³ Additionally, access to testing depends on
271 social and demographic factors, e.g. the relationships between PCR-positive results in contacts and
272 ethnicity varied if we conditioned on contact attendance for a PCR test (Table 2 vs. Table S2).

273 Secondly, our classification of contact events is relatively simple, e.g., we do not have any direct
274 measures of human behaviour, such as proximity or duration of contact. We also do not account for
275 the dynamic nature of viral loads over time,²⁴ relying on a single measurement at varying times post
276 infection. Despite this, the time from symptom onset to testing in the cases was relatively
277 consistent, median (IQR) 2 (1-3) days, such that measured Ct values plausibly represent similar
278 stages of the illness in cases. We use only a single assay to determine Ct values, but have calibrated
279 this to allow comparison with other platforms.

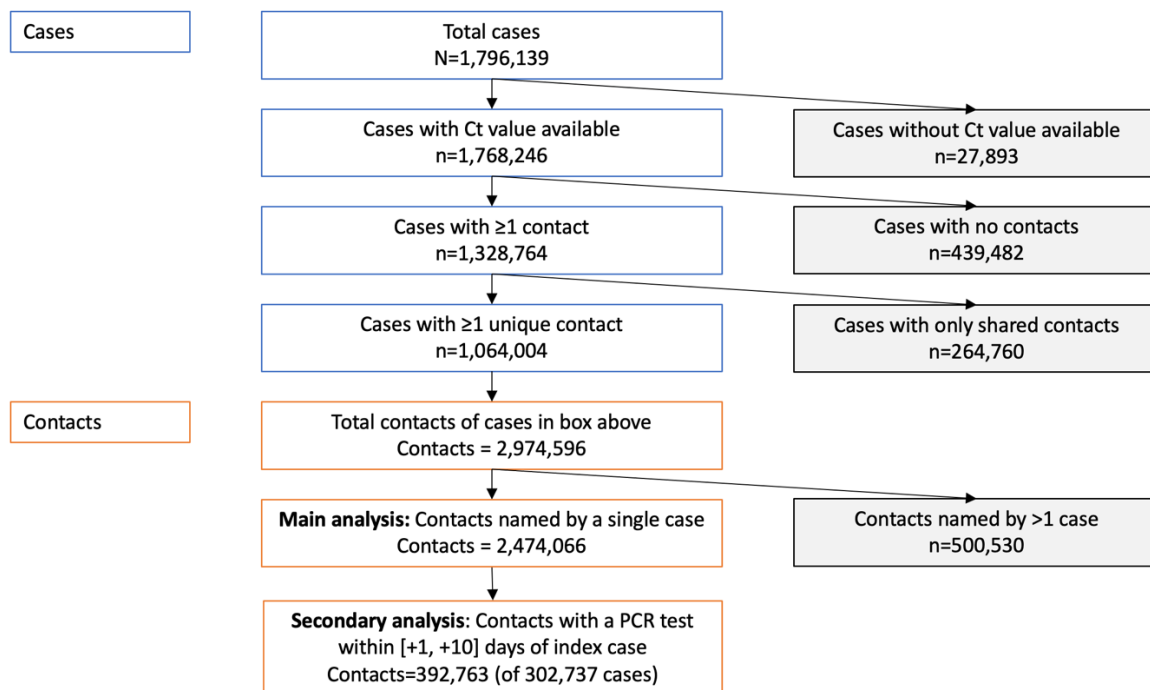
280 Finally, it was not possible to account for unobserved third-party transmission, although we
281 designed our study population to minimise this risk. This likely means that some contact events
282 identified as possible transmission events may actually not be the source of the infection in the
283 contact. It is likely that proportionally this effect is greatest at lower viral loads (higher Ct values), as
284 the likelihood of transmission rises with viral load.

285 In summary, we provide strong evidence that SGTF increases SARS-CoV-2 transmission and that
286 SARS-CoV-2 infectivity increases with increasing viral load. We show that the relative strength of the
287 effect of viral load is consistent across ages, ethnicities, and different types of contact events.
288 Despite this association, most individuals have Ct values compatible with onward transmission.²⁵
289 Nevertheless, LFDs can detect most individuals who are potential transmission sources. This
290 supports wider use of LFDs as rapid and regular screens to detect infectiousness in populations at
291 high risk of acquisition, including recent contacts of cases. Further prospective studies will be
292 required to demonstrate whether targeted isolation and/or contact tracing, together with wider use
293 of LFDs in combination with vaccination are effective in preventing ongoing SARS-CoV-2
294 transmission.

295

296 Figures

297

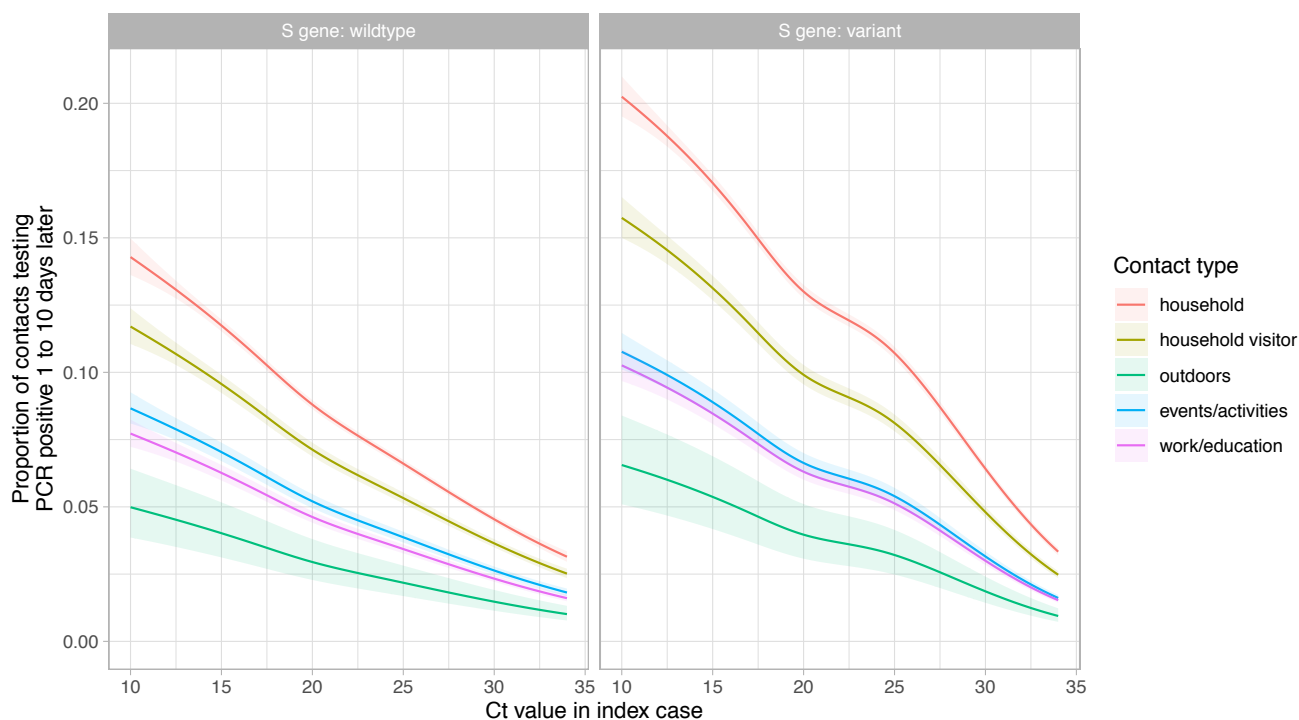


298

299

300 **Figure 1. Index cases and contacts in England, 01 September 2020 to 28 February 2021.**

301

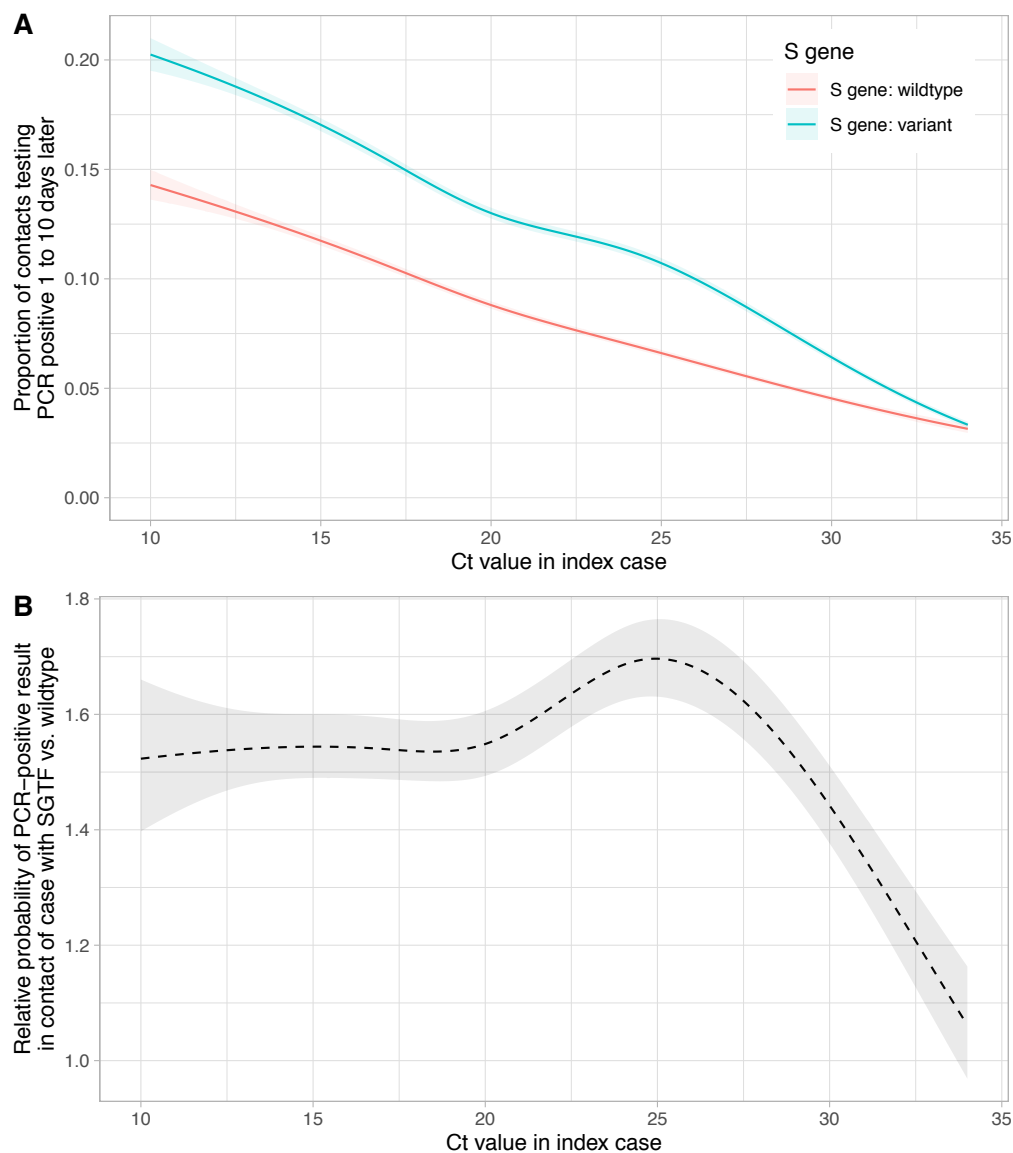


302

303

304 **Figure 2. Relationship between PCR cycle threshold (Ct) value in cases and the proportion of their**
305 **contacts with a PCR positive result, by contact type and S gene target failure.** Model predictions
306 are plotted after adjustment for index case age (set to the median value, 35 years), case ethnicity
307 (set to white), index of multiple deprivation score at contact's home address (set to median, 14,465),
308 incidence at contact's home address (set to median 350 cases per 100,000 population per week) and
309 index case sex (set to female). The shaded area indicates the 95% confidence interval.

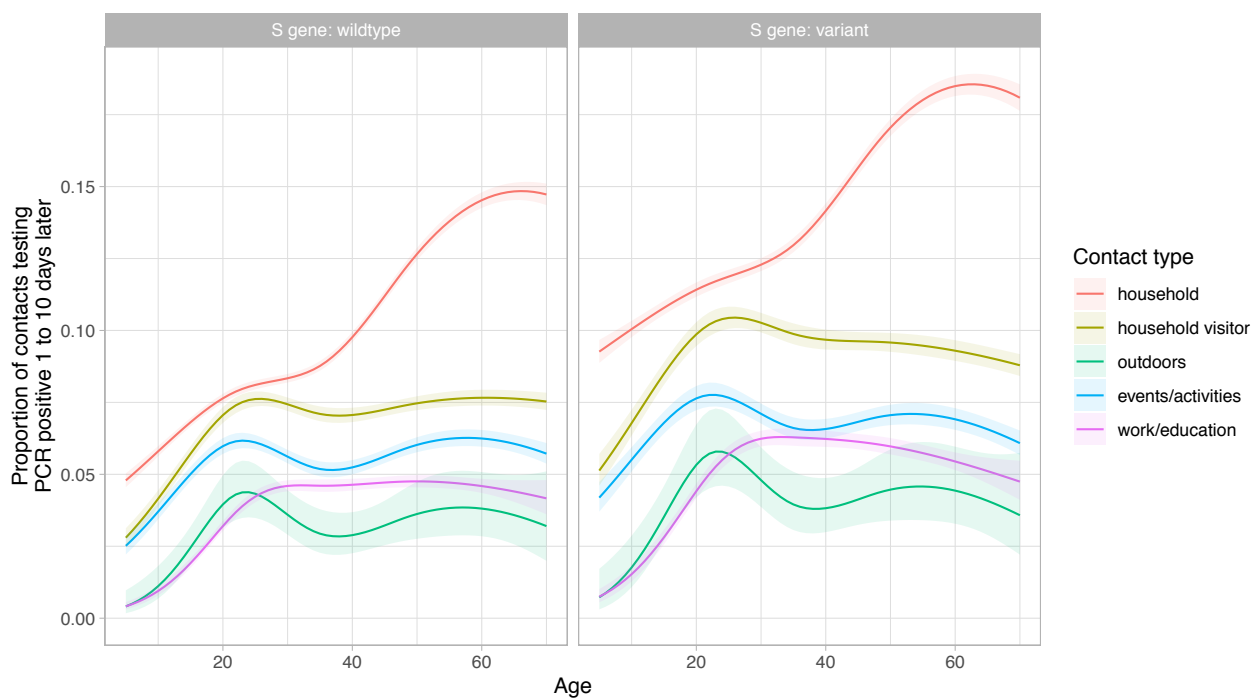
310



311

312 **Figure 3. Relationship between PCR-positive results in contacts and index case Ct value and S gene**
313 **target failure (SGTF) indicative of B.1.1.7 variant.** Panel A shows the proportion of contacts testing
314 by PCR-positive. Panel B displays the ratio of the two lines from panel A, i.e., the relative
315 infectiousness of index cases with SGTF vs. without SGTF. Model predictions are adjusted for index
316 case age, sex and ethnicity, contact index of multiple deprivation and incidence as in Figure 2.

317

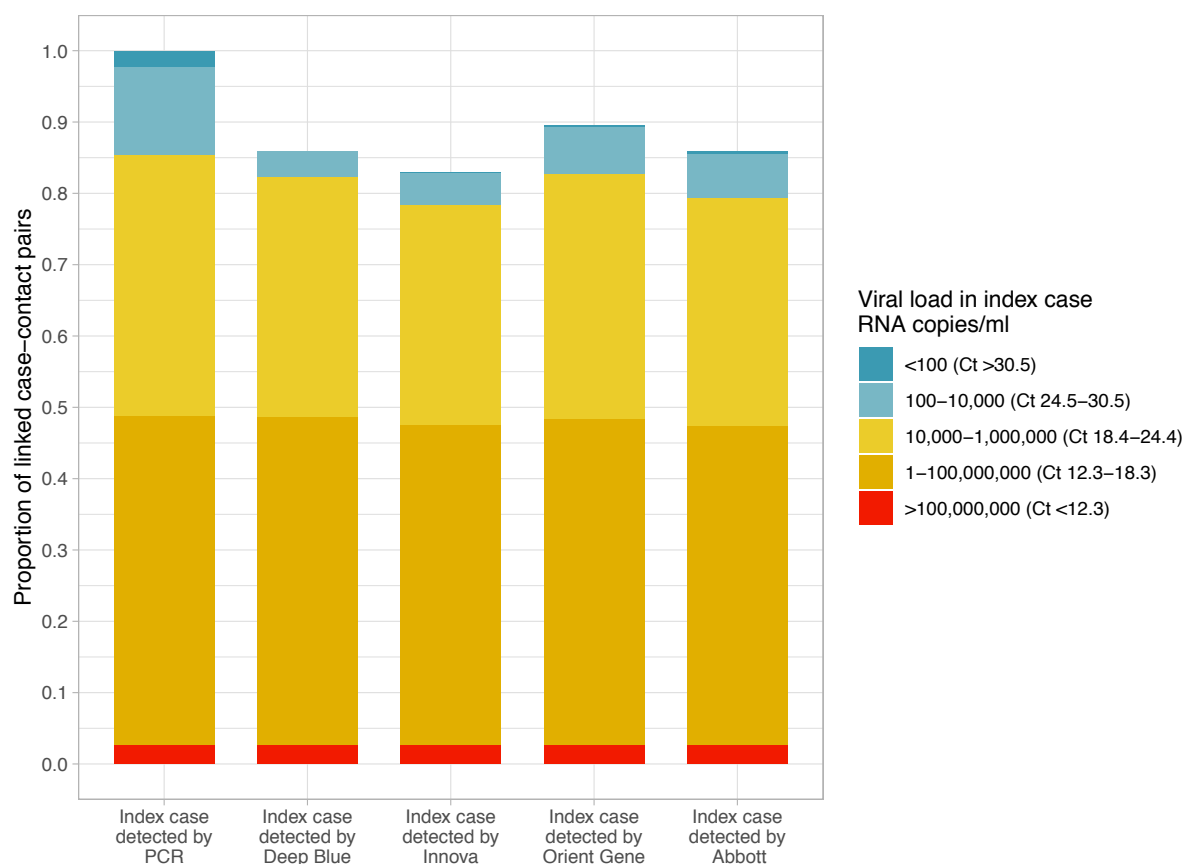


318

319 **Figure 4. Relationship between index case age and the proportion of their contacts with a PCR**
320 **positive result, by contact type and S gene target failure.** Model predictions are plotted after
321 adjustment for Ct value (set to the median Ct value, 20.1), and other variables as in Figure 2.

322

323



324

325 **Figure 5. Simulated proportion of cases with a PCR-positive contact detected using four lateral**
326 **flow devices (LFD).** The proportion of cases detected by PCR viral load group is shown in the PCR
327 column. The number of cases with a PCR-positive contact who would be detected using each LFD is
328 shown for 4 LFDs.

329 Tables

330

Variable	Case, n = 1,064,004 ¹	Contact: not PCR- positive within 1-10 days, n = 2,242,569 ¹	Contact: PCR-positive within 1-10 days, n = 231,497 ¹
Sex			
Female	560,557 (53%)	820,203 (37%)	114,837 (50%)
Male	476,967 (45%)	765,538 (34%)	99,539 (43%)
Not specified	26,480 (2.5%)	656,828 (29%)	17,121 (7.4%)
Age	36 (24 - 51)	30 (15 - 48)	37 (23 - 52)
Not available	9	720,544	17,853
Ethnic group			
Asian	128,218 (12%)	77,932 (3.5%)	9,491 (4.1%)
Black	27,658 (2.6%)	17,167 (0.8%)	1,874 (0.8%)
Mixed	27,263 (2.6%)	19,342 (0.9%)	2,297 (1.0%)
Other	15,682 (1.5%)	9,667 (0.4%)	1,170 (0.5%)
White	728,265 (68%)	585,255 (26%)	78,363 (34%)
Not available	136,918 (13%)	1,533,206 (68%)	138,302 (60%)
Incidence at home address, per 100,000 population	355 (215 - 546)	348 (207 - 524)	375 (226 - 581)
Not available	4,124	6,444	493
Deprivation index at home address (lower = more deprived, of 32,844 areas)	14,465 (11,374 - 18,704)	14,465 (11,304 - 18,649)	14,593 (11,744 - 19,165)
Not available	4,124	6,444	493
Case symptomatic	969,942 (91%)		
Days from symptom onset to test in case where symptomatic	2 (1 - 3)		
Contact type			
Events/activities		137,805 (6.1%)	8,919 (3.9%)
Household		1,718,674 (77%)	196,508 (85%)
Household visitor		189,637 (8.5%)	16,426 (7.1%)
Outdoors		8,002 (0.4%)	317 (0.1%)
Work/education		188,451 (8.4%)	9,327 (4.0%)
Days from case diagnosis to contact notification		2 (2- 3)	2 (1- 3)
Days from index case test to contact's test where tested		2 (1 - 4)	3 (2 - 5)

331

332 **Table 1. Demographics and characteristics of the study population.** ¹Frequency (%) or median (IQR).

Variable		Univariable			Multivariable		
		OR	95% CI	p-value	OR	95% CI	p-value
Incidence contact's home address, per 100,000 population*	50 (baseline)	1.00		<0.001	Interaction with SGTF, see Figure S7		
	100	1.10	1.09 - 1.11				
	200	1.25	1.24 - 1.26				
	400	1.25	1.24 - 1.26				
	600	1.42	1.41 - 1.43				
Deprivation score at contact's home address (lower = more deprived)*	7000 (baseline)	1.00		<0.001	1.00		<0.001
	14000	1.11	1.10 - 1.11		1.07	0.92 - 1.24	
	21000	1.26	1.25 - 1.27		1.16	1.00 - 1.35	
	28000	1.25	1.20 - 1.25		1.14	0.98 - 1.33	
Case Ct value (lower = higher viral load)*	10 (baseline)	1.00		<0.001	Interaction with SGTF, see Figures 2 and 3		
	15	0.81	0.80 - 0.81				
	20	0.57	0.57 - 0.57				
	25	0.44	0.43 - 0.44				
	30	0.28	0.28 - 0.29				
S gene target failure (SGTF)	Wildtype (baseline)	1.00			Multiple interactions, see other rows		
	S gene variant	1.52	1.50 - 1.53	<0.001			
Case sex	Female	1.00			Interaction with age, see Figure S8		
	Male	1.04	1.03 - 1.04	<0.001			
	Not specified	0.73	0.71 - 0.75	<0.001			
Case age*	30 years (baseline)	1.00		<0.001	Interactions between SGTF and contact type, SGTF and age, contact type and age, see Figures 2 and 4		
	10 years	0.71	0.70 - 0.72				
	50 years	1.34	1.33 - 1.34				
	70 years	1.40	1.38 - 1.41				

Contact event	Household (baseline)	1.00			Interactions between ethnicity and SGTF, ethnicity and contact type, ethnicity and age, see Figures S9 and S10
	Activities and events	0.57	0.55 - 0.58	<0.001	
	Household visitor	0.76	0.75 - 0.77	<0.001	
	Work or education	0.43	0.42 - 0.44	<0.001	
	Outside	0.35	0.31 - 0.39	<0.001	
Case ethnicity	White (baseline)	1.00			
	Asian	0.74	0.73 - 0.75	<0.001	
	Black	0.67	0.65 - 0.69	<0.001	
	Mixed	0.81	0.79 - 0.83	<0.001	
	Other	0.78	0.75 - 0.81	<0.001	
	Not available	0.76	0.75 - 0.77	<0.001	

333

334 **Table 2. Univariable and multivariable associations with the proportion of contacts testing PCR**
335 **positive.** The lower rates of PCR-positivity seen in cases without a documented sex possibly reflect
336 incomplete contact tracing or poor data quality preventing appropriate linkage of these cases.
337 *Incidence, deprivation score, index case Ct value and case age are all fitted as non-linear effects
338 with 5 default-spaced knots, example values are shown, and univariable relationships plotted in
339 Figures S3-S6. Multivariable results are presented with continuous variables set to their median
340 value and categorical variables set to baseline, figures illustrating relationships with interactions are
341 listed. See Figure S11 for the multivariable relationship for deprivation score.

342

343 References

344

345 1 Williamson EJ, Walker AJ, Bhaskaran K, *et al.* Factors associated with COVID-19-related death using
346 OpenSAFELY. *Nature* 2020; **584**: 430–6.

347 2 Nussbaumer-Streit B, Mayr V, Dobrescu AI, *et al.* Quarantine alone or in combination with other
348 public health measures to control COVID-19: a rapid review. *Cochrane Db Syst Rev* 2020.
349 DOI:10.1002/14651858.cd013574.pub2.

350 3 Davies NG, Abbott S, Barnard RC, *et al.* Estimated transmissibility and impact of SARS-CoV-2
351 lineage B.1.1.7 in England. *Science* 2021; : eabg3055.

352 4 Dagan N, Barda N, Kepten E, *et al.* BNT162b2 mRNA Covid-19 Vaccine in a Nationwide Mass
353 Vaccination Setting. *New Engl J Med* 2021. DOI:10.1056/nejmoa2101765.

354 5 Grasselli G, Pesenti A, Cecconi M. Critical Care Utilization for the COVID-19 Outbreak in Lombardy,
355 Italy. *Jama* 2020; **323**: 1545–6.

356 6 Coronavirus latest: California to issue stay-at-home orders if ICU capacity runs low – as it
357 happened. undefined. <https://www.ft.com/content/eb9a5ae9-d209-39ee-a743-5259039de9c0>
358 (accessed Jan 17, 2021).

359 7 How do measures for isolation, quarantine, and contact tracing differ among countries? – Cross-
360 Country Analysis. undefined. [https://analysis.covid19healthsystem.org/index.php/2020/05/19/how-
361 do-measures-for-isolation-quarantine-and-contact-tracing-differ-among-countries/](https://analysis.covid19healthsystem.org/index.php/2020/05/19/how-do-measures-for-isolation-quarantine-and-contact-tracing-differ-among-countries/) (accessed Jan 17,
362 2021).

363 8 Brooks SK, Webster RK, Smith LE, *et al.* The psychological impact of quarantine and how to reduce
364 it: rapid review of the evidence. *Lancet* 2020; **395**: 912–20.

365 9 Leon DA, Shkolnikov VM, Smeeth L, Magnus P, Pechholdová M, Jarvis CI. COVID-19: a need for real-
366 time monitoring of weekly excess deaths. *Lancet* 2020; **395**: e81.

367 10 Woolf SH, Chapman DA, Sabo RT, Weinberger DM, Hill L. Excess Deaths From COVID-19 and
368 Other Causes, March–April 2020. *Jama* 2020; **324**: 510–3.

369 11 Office for National Statistics. Deaths registered in England and Wales, provisional: week ending 6
370 November 2020. undefined.
371 [https://www.ons.gov.uk/releases/deathsregisteredinenglandandwalesprovisionalweekending6nove
372 mber2020](https://www.ons.gov.uk/releases/deathsregisteredinenglandandwalesprovisionalweekending6nove) (accessed Jan 17, 2021).

373 12 Bi Q, Wu Y, Mei S, *et al.* Epidemiology and transmission of COVID-19 in 391 cases and 1286 of
374 their close contacts in Shenzhen, China: a retrospective cohort study. *Lancet Infect Dis* 2020; **20**:
375 911–9.

- 376 13 Luo L, Liu D, Liao X, *et al.* Contact Settings and Risk for Transmission in 3410 Close Contacts of
377 Patients With COVID-19 in Guangzhou, China: A Prospective Cohort Study. *Ann Intern Med* 2020;
378 **173**: 879–87.
- 379 14 Sun K, Wang W, Gao L, *et al.* Transmission heterogeneities, kinetics, and controllability of SARS-
380 CoV-2. *Science* 2021; **371**: eabe2424.
- 381 15 Wölfel R, Corman VM, Guggemos W, *et al.* Virological assessment of hospitalized patients with
382 COVID-2019. *Nature* 2020; **581**: 465–9.
- 383 16 Pekosz A, Cooper CK, Parvu V, *et al.* Antigen-based testing but not real-time PCR correlates with
384 SARS-CoV-2 virus culture. *medRxiv* DOI:10.1101/2020.10.02.20205708.
- 385 17 Public Health England. Investigation of novel SARS-CoV-2 Variants of Concern. Technical briefing
386 documents on novel SARS-CoV-2 variant.
387 [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/959426/Variant_of_Concern_VOC_202012_01_Technical_Briefing_5.pdf)
388 [/959426/Variant_of_Concern_VOC_202012_01_Technical_Briefing_5.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/959426/Variant_of_Concern_VOC_202012_01_Technical_Briefing_5.pdf) (accessed March 23,
389 2021).
- 390 18 Peto T, Team UC-19 LFO. COVID-19: Rapid Antigen detection for SARS-CoV-2 by lateral flow assay:
391 a national systematic evaluation for mass-testing. *medRxiv* DOI:10.1101/2021.01.13.21249563.
- 392 19 Madewell ZJ, Yang Y, Longini IM, Halloran ME, Dean NE. Household Transmission of SARS-CoV-2.
393 *Jama Netw Open* 2020; **3**: e2031756.
- 394 20 College I. Report 38: SARS-CoV-2 setting-specific transmission rates: a systematic review and
395 meta-analysis. undefined. [https://www.imperial.ac.uk/media/imperial-college/medicine/mrc-](https://www.imperial.ac.uk/media/imperial-college/medicine/mrc-gida/2020-11-27-COVID19-Report-38.pdf)
396 [gida/2020-11-27-COVID19-Report-38.pdf](https://www.imperial.ac.uk/media/imperial-college/medicine/mrc-gida/2020-11-27-COVID19-Report-38.pdf).
- 397 21 Marks M, Millat-Martinez P, Ouchi D, *et al.* Transmission of COVID-19 in 282 clusters in Catalonia,
398 Spain: a cohort study. *medRxiv* DOI:10.1101/2020.10.27.20220277.
- 399 22 Pouwels KB, House T, Pritchard E, *et al.* Community prevalence of SARS-CoV-2 in England from
400 April to November, 2020: results from the ONS Coronavirus Infection Survey. *Lancet Public Heal*
401 2021; **6**: e30–8.
- 402 23 Buitrago-Garcia D, Egli-Gany D, Counotte MJ, *et al.* Occurrence and transmission potential of
403 asymptomatic and presymptomatic SARS-CoV-2 infections: A living systematic review and meta-
404 analysis. *Plos Med* 2020; **17**: e1003346.
- 405 24 Kissler SM, Fauver JR, Mack C, *et al.* SARS-CoV-2 viral dynamics in acute infections. *medRxiv*
406 DOI:10.1101/2020.10.21.20217042.
- 407 25 Smith LE, Potts HWW, Amlot R, Fear NT, Michie S, Rubin J. Adherence to the test, trace and
408 isolate system: results from a time series of 21 nationally representative surveys in the UK (the

409 COVID-19 Rapid Survey of Adherence to Interventions and Responses [CORSAIR] study). *medRxiv*
410 DOI:10.1101/2020.09.15.20191957.

411

412

413

414 Transparency Declarations

415 Declaration of interests

416 DWE declares lecture fees from Gilead, outside the submitted work. LYWL declares speaker fees
417 from the Merck group and Servier, outside the submitted work. No other author has a conflict of
418 interest to declare.

419 Funding Statement

420 This study was funded by the UK Government's Department of Health and Social Care. This work was
421 supported by the National Institute for Health Research Health Protection Research Unit (NIHR
422 HPRU) in Healthcare Associated Infections and Antimicrobial Resistance at Oxford University in
423 partnership with Public Health England (PHE) (NIHR200915), and the NIHR Biomedical Research
424 Centre, Oxford. The views expressed in this publication are those of the authors and not necessarily
425 those of the NHS, the National Institute for Health Research, the Department of Health or Public
426 Health England. DWE is a Robertson Foundation Fellow and an NIHR Oxford BRC Senior Fellow. ASW
427 is an NIHR Senior Investigator. LYWL is supported by the NIHR Oxford BRC.

428 Role of the funding source

429 The funder of the study provided access to the data and facilitated data linkage. The funder had no
430 role in study design, data analysis, data interpretation, or writing of the report. The corresponding
431 author had full access to all the data in the study and had final responsibility for the decision to
432 submit for publication.