An observational study of SARS-CoV-2 infectivity by viral load and demographic factors and the utility lateral flow devices to prevent transmission

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

[1] Nuffield Department of Medicine, University of Oxford, UK
[2] Department of Health and Social Care, UK Government, London, UK
[3] Public Health England, London, UK
[4] Public Health England, Porton Down, UK
[5] NIHR Oxford Biomedical Research Centre, University of Oxford, Oxford, UK
[6] NIHR Health Protection Research Unit in in Healthcare Associated Infections and Antimicrobial Resistance, University of Oxford, UK
[7] Big Data Institute, Nuffield Department of Population Health, University of Oxford, Oxford, UK

*Contributed equally
§Contributed equally

Keywords: infectivity, transmission, contact tracing, coronavirus, COVID-19, SARS-CoV-2, lateral flow device, viral antigen detection, testing

Abstract

Background:
How SARS-CoV-2 infectivity varies with viral load is incompletely understood. Antigen lateral flow devices (LFD) are rapid point-of-care tests for SARS-CoV-2 but have imperfect sensitivity compared to PCR assays. Whether LFDs can detect most potential transmission sources is unknown.

Methods:
We combined national SARS-CoV-2 PCR and contact tracing data from the United Kingdom between 21-May and 01-November-2020. We used multivariable logistic regression to investigate the relationship between PCR-confirmed infection in contacts of community-diagnosed cases and the cases' viral load, adjusting for case demographics, the prevalence of SARS-CoV-2, social deprivation, and contact event type. We used the previously reported LFD performance to simulate the proportion of cases with a PCR-positive contact that would have been detected using one of four LFDs.

Findings:
18,291/303,192 (6.0%) contacts were PCR-positive 2-7 days following their index case's PCR test. PCR-positive results in contacts decreased with lower case viral loads (adjusted OR (aOR) per unit rise in Ct value, 0.93 [95%CI 0.92-0.93, p<0.001]). Higher rates of PCR positive results were seen in household contacts, compared to household visitors (aOR 0.70 [0.63-0.78]), contacts in work/education (0.50 [0.44-0.57]) or at activities/events outside homes (0.53 [0.48-0.59]). The proportion of PCR-positive tests was independently lowest in contacts of children. The most and least sensitive LFDs would detect 90.5% (95%CI 90.1-90.8%) and 83.7% (83.2-84.1%) of cases with PCR-positive contacts respectively.

Interpretation: SARS-CoV-2 infectivity varies by case viral load, contact event type, and age. Those with high viral loads are most infectious. The best performing LFDs can detect most infectious cases.

Funding: UK Government

Research in context

Evidence before this study

A literature review was performed using PubMed, bioRxiv, medRxiv for all studies related to COVID-19 infectivity. This used the search terms, COVID-19, SARS-CoV-2, infectivity, transmission, contact testing, viral load, lateral flow devices (LFD), and LFD and was performed with no language restrictions. To date, the majority of studies have used surrogate measures of infectivity, with most studies utilising a functional assay, viral culture. One study reported in a pre-print consisted of 282 transmission events in the Catalonia region of Spain and identified an association with viral load and infectivity. This study was not powered to identify effect modification between viral load and other factors including demographics and contact type. It did not discuss the clinical implications of viral load and COVID-19 testing or potential use of LFDs. Aside from this single study, we found no other evidence regarding the determinants of infectivity (measured by contacts infected) or any papers investigating the utility of lateral flow devices to identify the most infectious individuals.

Added value of this study

This study represents the largest analysis of contacts following exposure to SARS-CoV-2 to date. It consists of a population-level sample and provides new insights into SARS-CoV-19 infectivity. We found a linear relationship between log viral load and likelihood of a contact developing PCR-positive SARS-CoV-2 infection. Higher rates of PCR-positive results were seen in household contacts, compared to household visitors, contacts in work/education or at activities/events outside homes. The proportion of PCR-positive tests was lowest in contacts of infected children. Modelling suggests that lateral flow devices, would identify individuals responsible for 84% of transmissions using the least sensitive of four kits tested, and 91% using the most sensitive.

Implication of all the available evidence

We provide strong evidence that SARS-CoV-2 infectivity varies by viral load. We show that the relative strength of this effect is consistent across ages, ethnicities and different types of contact events. This assessment could be of importance in tailoring future quarantine/isolation measures. Whilst not being a substitute for PCR, lateral flow devices, and their ability to rapidly identify individuals with the highest viral load, including those without symptoms who may not present to symptomatic testing services, are a potential tool to rapidly identify and isolate the most infectious individuals. This approach will need to be evaluated in real-world settings.

Introduction

The global health impact of SARS-CoV-2 is profound and it continues to cause significant human morbidity and mortality.[1] There is widespread on-going transmission despite control efforts focused on quarantine of predominantly symptomatic cases and population-level self-isolation.[2] Intermittent national self-isolation measures have been imposed in many countries with additional stricter self-isolation measures in specific high incidence regions and cities.[3,4] Self-isolation measures for "contacts" (individuals exposed to SARS-CoV-2) vary by country, but generally last 7-14 days.[5] While reducing transmission, quarantine/isolation measures have had many wider effects, including economic impacts, impacts on mental health and well-being,[6] as well as non-COVID-19 related excess deaths as a result of behaviour change arising from these measures and the impact of COVID-19 on healthcare systems.[7–9]

Modelling studies suggest that approximately 15% of individuals are responsible for most SARS-CoV-2 transmission.[10] Furthermore, only 5-7% of exposed "contacts" develop COVID-19 infection.[11,12] These observations imply that the majority of individuals with PCR-positive SARS-CoV-2 infection do not cause infections in close contacts, potentially as a result of reduced infectiousness, e.g. due to only short-lived or low-level production of viable virions, as well as behavioural factors. Our understanding of why many infected individuals have low infectivity is limited.

There are several proposed assays for infectivity. Functional assays to measure infectivity include animal and cell culture models. Viral subgenomic mRNA has been proposed as a nucleic acid-based measure of infectivity.[13] Viral protein, i.e. antigen, detection, as assessed by lateral flow devices (LFDs), has been shown to be more closely linked to viral culture infectivity than PCR measurements.[14] However, few of these surrogate measures of infectivity have been convincingly shown to predict the real-world likelihood of a SARS-CoV-2 infected individual infecting someone else. If a better predictor of infectivity could be identified, quarantine measures might be targeted differentially with a focus on those with the highest infectivity, potentially enhancing transmission prevention while reducing collateral impacts on economic activity and well-being. Through earlier identification of infectious individuals, it is possible that some of the population, and SARS-CoV-2 contacts, could avoid prolonged self-isolation.

Here we use data from the United Kingdom's NHS Test and Trace programme to explore the relationship between infectivity and SARS-CoV-2 viral load, as measured by PCR cycle threshold (Ct) values. We also identify other factors associated with infectivity. We apply the results to a typical population of PCR-positive individuals to show the proportion of infectious individuals detected by viral antigen LFDs under a range of performance conditions.

Methods

Infectivity datasets

Comprehensive data from community and hospital PCR testing in England between 21 May 2020 and 01 November 2020 were obtained and linked with national contact tracing data (see Supplement for contact definitions). Data linkage was undertaken by Public Health England (PHE) using patient identifiers, combining community PCR results (via the National Pathology exchange)

and hospital PCR results (via PHE's Second Generation Surveillance System) with contact tracing data collected by NHS Test and Trace. Data extracts were deidentified prior to analysis and included basic demographics and symptoms for PCR-confirmed cases and their contacts, as well as details on the nature of the contact events. Cycle threshold (Ct) values were available for community tests undertaken by the UK's high throughput Lighthouse Laboratories in Milton Keynes, Alderley Park and Glasgow using the Thermo Fisher TaqPath assay (targeting S and N genes, and ORF1ab). Assay details, including an equation for converting Ct values to viral loads, can be found in the Supplement.

Statistical analysis

We investigated the proportion of contacts of community-diagnosed cases who had a PCR-confirmed infection. Contacts could be diagnosed as PCR-positive through either community or hospital-based testing. We restricted our analysis to cases with an available Ct value. We excluded cases with no reported contacts or insufficient information for data linkage. Similarly, we excluded contacts where there was insufficient information to link Test and Trace data with PCR results.

To focus only on potential transmission pairs where there was a low likelihood of transmission from a third party, we excluded contacts with multiple potential sources of infection, i.e. contacts who had been named by multiple cases; or cases from care homes or universities. To maximise the chance of the index patient being a source for the contact, in our main analysis we only considered PCR tests taken in contacts between 2-7 days inclusive following the case's PCR result. This reduced the possibility of misidentifying the directionality of transmission between case-contact pairs and also further minimised the likelihood of acquisition from a third party. We also conducted a sensitivity analysis where we included all PCR results for contacts tested between -4 and 10 days of the case's PCR result and included the time difference between tests as a variable in the regression model.

We used multivariable logistic regression to investigate the association between PCR-confirmed infection in contacts (including named contacts whether or not they attended for PCR testing) and the Ct value in potential source cases, the nature of the contact, the case's demographics and incidence and social deprivation index at the contact's home address location (see Supplement for details). We did not adjust for symptoms in the case, as these may be mediators of the effect of viral load on onward transmission. We used backwards elimination based on the Bayesian information criterion (BIC) to select model main effects, using natural cubic splines to account for non-linearity in continuous variables (up to 5 default-spaced knots, choosing the final number of knots based on BIC). We screened for all pairwise interactions between main effects, retaining interactions that minimised BIC. We used robust standard errors to account for some contacts sharing the same source. Data analysis was performed using R, version 4.0.2.

The above analysis only captures secondary infections in contacts who attended for testing and were confirmed as PCR-positive. However, attendance rates for diagnostic testing may vary by setting or demographics, e.g. household contacts of a PCR-confirmed case with similar symptoms may be less inclined to seek confirmatory testing if they consider the pre-test probability of a positive result to be high. Therefore, we also performed an additional analysis where we considered associations with PCR-confirmed infection amongst contacts attending for PCR testing.

Finally, we used linear regression to investigate the proportion of the variation in Ct values in contacts that could be attributed to the Ct value iof the case.

Simulations of the number of cases identified by antigen LFDs

We used our findings to estimate the proportion of potential transmission events where the source case would have been detected using an antigen LFD using previous data on the sensitivity of four LFDs, including Innova, Deep Blue and Orient gene LFDs and a fourth manufacturer where permission was not available to identify the kit, denoted "LFD x".[15] For each source case we simulated a positive or negative LFD result by randomly drawing from the probability of a LFD being positive by the source case's Ct value (determined from a previously described regression model and accounting for uncertainty in the previous model's estimates, see Supplement, Figure S1). Each simulation was repeated 1000 times. Additionally, we generate simulations for a range of alternative possible lateral flow test performances.

Ethics

The study was designed as a public health surveillance analysis to support rapid clinical decision making, in accordance with the UK Policy Framework for Health and Social Care Research and was approved by Public Health England, the UK COVID-19 LFD oversight group and NHS Test and Trace. As the study was conducted as part of COVID-19 surveillance under the provisions of Section 251 of the NHS Act 2006 it did not require individual patient consent or ethical approval.

Role of the funding source

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

Results

Cases and contacts

Of 713,668 cases with PCR-confirmed SARS-CoV-2 infection, 470,144 reported ≥1 contact, of whom 219,722 also had a Ct value available. 105 cases who were identified as university students or living in care homes were excluded, leaving 219,617 cases for analysis. These cases had 721,877 contacts, of whom 159,609 were named by more than one case and therefore excluded, in a further 259,019 contact tracing was unsuccessful or incomplete, resulting in 303,192 contacts of 145,973 cases in our main analysis (Figure 1).

The median (IQR) age of cases and contacts was 35 (22-50) and 33 (19-50) years respectively, and 55% and 54% with available data were female (Table 1). Most contact events occurred within households (67.7%), followed by visits to households (12.0%), attending events or activities (11.2%), and at workplaces or education (9.1%).

Relationship between PCR-positive results in contacts and case viral load and contact event types

Of 303,192 contacts, 18,291 (6.0%) tested PCR-positive between 2 and 7 days inclusive following the date of their index case's PCR test. On univariable analysis (Table 2), PCR-positive tests in contacts were associated with lower case Ct values (i.e. higher viral loads). Household contacts were most likely to be PCR-positive, followed by visitors to households. PCR-positive results were least frequent in contacts of children, with highest rates in contacts of older adults (Figure S2).

The Ct value measured in each case remained an important determinant of PCR-positive results in contacts after adjustment for case age, contact event type, diagnostic laboratory and case ethnicity in a multivariable model (Table 2). PCR-positive results in contacts decreased with lower case viral loads; the adjusted OR (aOR) per unit increase in Ct was 0.93 (95%CI 0.92-0.93, p<0.001). Figure 2 shows the proportion of contacts testing PCR-positive by case Ct value and contact event type. There was a significant interaction between age and contact event type (heterogeneity p<0.001), and so results are shown at a median age of 33 years. Higher rates of PCR-positive results were seen in household contacts, compared to household visitors (aOR, at 33 years, 0.70 [95%CI 0.63-0.78]), contact that occurred in work/education (0.50 [95%CI 0.44-0.57]) or at activities and events outside homes (0.53 [95%CI 0.48-0.59]).

The proportion of PCR-positive tests was lowest in contacts of children, in particular in contact events in work/education, with the pattern of PCR-positive results in contacts of adults varying by contact type (Figure 3). Household contacts had the highest proportion of PCR-positive tests overall, with risk increasing with age. Within household visitors, those aged in their twenties and above 65 years had the highest proportion of PCR-positive contacts. Decreased rates of PCR-positive results were seen in contacts of Black and Asian cases, with increased rates in the north of England (Alderley Park laboratory) and decreased numbers in Scotland (Glasgow laboratory), compared to the south of England (Milton Keynes laboratory), in keeping with overall prevalence during the study period.

In a sensitivity analysis, we refitted our model considering PCR results in contacts from 4 days before to 10 days after the date of their index case's PCR test. We then analysed the adjusted odds ratio for a PCR-positive result per unit lower Ct by the time between PCR tests (Figure S3). Estimates were broadly similar from 1 to 10 days after the index cases test date but attenuated prior (consistent with acquisition from other sources for these contacts), supporting the 2 to 7 day window used in the main analysis.

Ct values in cases determined only a small proportion of the variability in Ct values in contacts (unadjusted linear regression coefficient 0.15 [95%CI 0.14-0.16, p<0.001], R-squared = 0.02).

Relationships with PCR-positive results in contacts attending PCR testing

Restricting only to contacts who had a PCR test, adjusted odds ratios for a PCR-positive test by the index case's Ct value were similar, 0.92 per unit higher (95%CI 0.91-0.92, p<0.001, Table S1, Figure S4). Similarly, within this group, household contacts were mostly likely to be PCR-positive, with the highest rates in older adults (Figure S5). In contrast to the main analysis, contacts of cases of Asian

ethnicity were more likely to be PCR-positive, potentially due to differences in access to and use of testing by ethnic group.

Proportion of cases with PCR-positive contacts detected by LFDs

Overall, 86.8% (15,883/18,291) of case-contact pairs with PCR-positive contact, i.e. plausible onward transmission, had case viral loads of ≥10,000 RNA copies/ml (i.e. Ct values of ≤24.4) and just under half of pairs, 8617 (47.1%), had case viral loads ≥1 million RNA copies/ml (Ct ≤18.3). In contrast fewer cases overall had viral loads above these levels, 76.8% of all cases (112,100/145,973) had viral loads of ≥10,000 RNA copies/ml and 35.2% (51,310) had viral loads of ≥1 million (Figure S6).

As antigen LFD sensitivity varies by viral load, we used the distribution of viral loads in case-contact pairs with a PCR-positive contact to simulate the proportion of such cases who would have been detected using antigen LFDs (Figure 4). The Deep Blue LFD would have detected 87.0% (95%CI 86.7-87.3%) of cases who plausibly subsequently transmitted to a contact, the Innova LFD 83.7% (83.2-84.1%), the Orient Gene LFD 90.5% (90.1-90.8%) and LFD x 86.6% (86.1-87.0%).

To provide an indication of the performance required by a novel LFD to detect varying proportions of transmission sources we repeated our simulation across a range of LFD performances (Figure S7). As LFD sensitivity varies by viral load (Ct value) two parameters are required to summarise performance: we use the viral loads at which 50% and 90% sensitivity are achieved, e.g. with 50% sensitivity at ~3100 RNA copies/ml (Ct ≤25.9) and 90% sensitivity at 10,000 copies/ml (Ct ≤24.4) 90.7% of cases who plausibility transmitted to one or more contact would be detected.

Discussion

We have performed a large-scale national analysis of combined contact tracing and PCR data from the United Kingdom involving over a quarter of a million contacts of PCR-confirmed cases. We show that SARS-CoV-2 infectivity is associated with index case viral load, including after adjustment for demographic factors and the nature of the contact event leading to transmission. Additional findings support onward transmission from children being relatively uncommon compared to adults, around 2% of education-based contacts of infected primary school age children (5-11 years) tested PCR-positive, compared to 6% of contacts overall. We confirm earlier findings that household contact is associated with greater rates of transmission compared to workplace, educational or recreational contact outside of homes.[16,17]

PCR-positive results in contacts decreased with lower case viral loads, after adjustment for demographic factors and the nature of contact events, the odds ratio of infectivity decreased by 0.93 for each unit increase in Ct value. It was noteworthy that we found no evidence of significant interactions between any of the other variables in the analysis and viral load, i.e. that that the effect of viral load on the infectivity is generalisable across populations and settings. These results are consistent and add to a recent smaller cohort study.[18]

If we assume that the proportion of contacts testing PCR-positive is representative of all secondary cases, whether tested or not, it is possible to estimate that proportion of onward transmission

attributable to cases with a given viral load or Ct value, e.g. 86.6% PCR-positive contacts had an index case with a viral load of ≥10,000 RNA copies/ml (Ct ≤24.4). Hence, 86.6% of infections in contacts are potentially attributable to the 76.5% (156,913/205,051) of cases overall with a viral load of ≥10,000 RNA copies/ml. While such data could be used to drive differential interventions to prevent onward transmission with a particular focus on those with high viral loads, the data suggest that most infected individuals are still at risk of transmitting onwards based on Ct measurements.

However, we are able to show that several LFDs are sufficiently sensitive to detect the majority of cases that led to onward transmission. These tests offer potential advantages, in returning a result in 15-30 minutes, not requiring laboratory or logistics infrastructure and costing significantly less than PCR tests. By applying previous estimates of the sensitivity of four LFDs and our current findings, we estimate that LFDs would detect between 83.7% and 90.5% of cases leading to onward transmission. While such performance is not sufficient to replace PCR for testing of all symptomatic individuals, use of LFDs in addition to existing testing, particularly of those who otherwise would not be tested at all (including those without symptoms), would allow many of the most infectious individuals to be identified earlier, potentially preventing onward transmissions and helping to drive reproduction numbers below 1, despite imperfect performance against PCR. We also generate simulations for a range of LFD performances (Figure S7) which provides an indicative benchmark of whether a new LFD performs sufficiently well to potentially detect the majority of infectious individuals. The specificity of each LFD is also another important consideration, the false positive rate for the Innova LFD has been previously reported as 0.32% (95%CI 0.20-0.48%),[15] however large-scale evaluations of the other LFDs described are on-going. In settings where the positive predictive value of an LFD is insufficiently high confirmatory PCR testing may be required.

Our study has several important limitations. Firstly, ascertainment of infection in contacts is dependent on the contact being reported by the case and the contact attending for testing. In the UK, PCR testing is only recommended for those with symptoms and therefore although some asymptomatic individuals were tested, we do not ascertain most of those asymptomatically infected. Whilst Ct values are generally slightly lower in those without symptoms,[19] they may nevertheless contribute substantially to onward transmission.[20] Additionally, access to testing depends on social and demographic factors, for example different relationships with PCR positive results in contacts and ethnicity were obtained if we conditioned on attendance for a PCR test by the contact (Table 2 vs. Table S1). Further, contact tracing was only successfully completed such that linkage could be achieved with PCR data for just over half of contacts. There is no obvious reason however why these should be unrepresentative in terms of estimated associations between case Ct and contact PCR-positivity.

Secondly, our classification of contact events is relatively simple, e.g., we do not have any direct measures of human behaviour, such as proximity or duration of contact beyond that definitions for contact were met, and we only have data on known contacts who were reported by cases during routine contact tracing. We also do not account for the dynamic nature of viral loads over time,[21] relying on a single measurement at varying times post infection. Despite this, the time from symptom onset to testing in the cases was relatively consistent, median (IQR) 2 (1-3) days, such that measured Ct values plausibly represent similar stages of the illness in cases. In addition, our

estimated odds ratio for the effect of Ct values was relatively invariant across a range of 1-10 days between testing of cases and contacts.

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

In summary, we provide strong evidence that SARS-CoV-2 infectivity increases with increasing viral load. We show that the relative strength of this effect is consistent across ages, ethnicities and different types of contact events. Despite this association, most individuals have Ct values compatible with onward transmission. However, there are a substantial minority of individuals who are much less likely to be a source of onward transmission. There is growing evidence of low levels of compliance with blanket self-isolation policies with less than 1 in 5 individuals adhering to isolation,[22] potentially due to perceived futility, and there is a significant cost at both an individual and community level. Potentially the finding with the most public health and policy impact is the ability of existing LFDs to detect the majority of individuals who are potential transmission sources. This supports wider use of LFDs as rapid and regular screens to detect infectiousness in populations at high risk of acquisition including those who are recent contacts of cases. Further prospective studies will be required to demonstrate whether targeted isolation and/or contact tracing, together with wider use of LFDs in combination with strategies like vaccination are an efficacious option to prevent ongoing SARS-CoV-2 transmission.

Figures

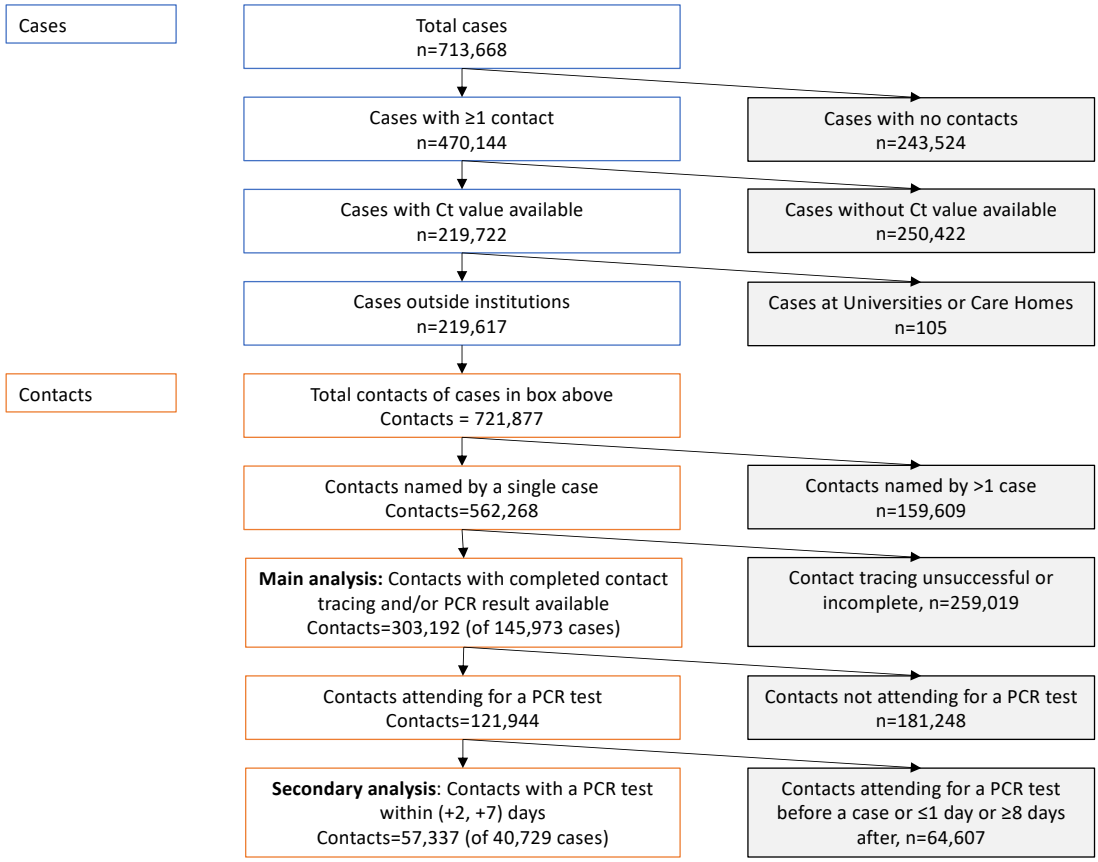

Figure 1. UK Test and Trace index cases and contacts, 21 May 2020 to 01 November 2020.

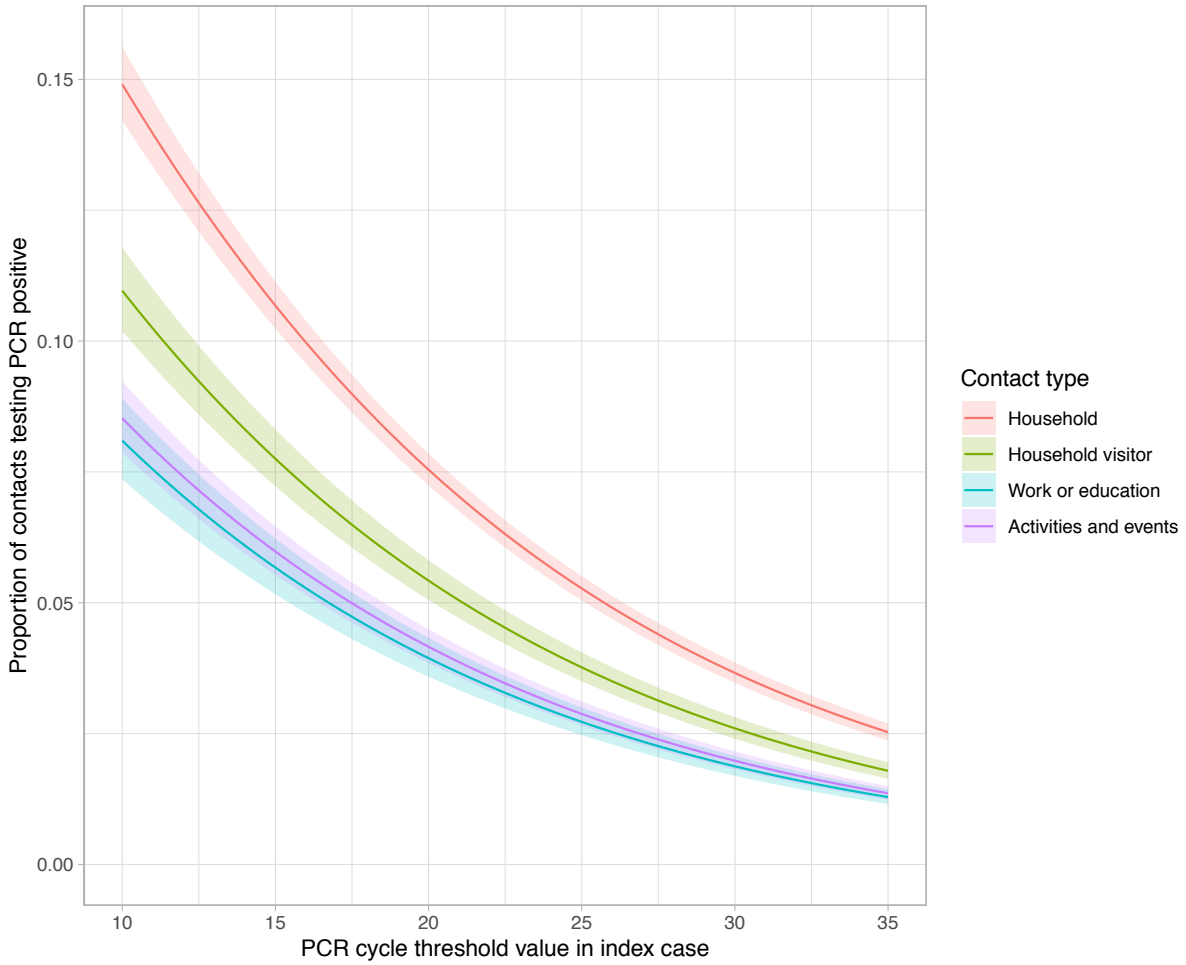

Figure 2. Relationship between PCR cycle threshold (Ct) value in cases and the proportion of their contacts with a PCR positive result, by contact event type. Model predictions are plotted after adjustment for age (set to the median value, 33 years), diagnostic laboratory (set to Milton Keynes), ethnicity (set to white). The shaded area indicates the 95% confidence interval.

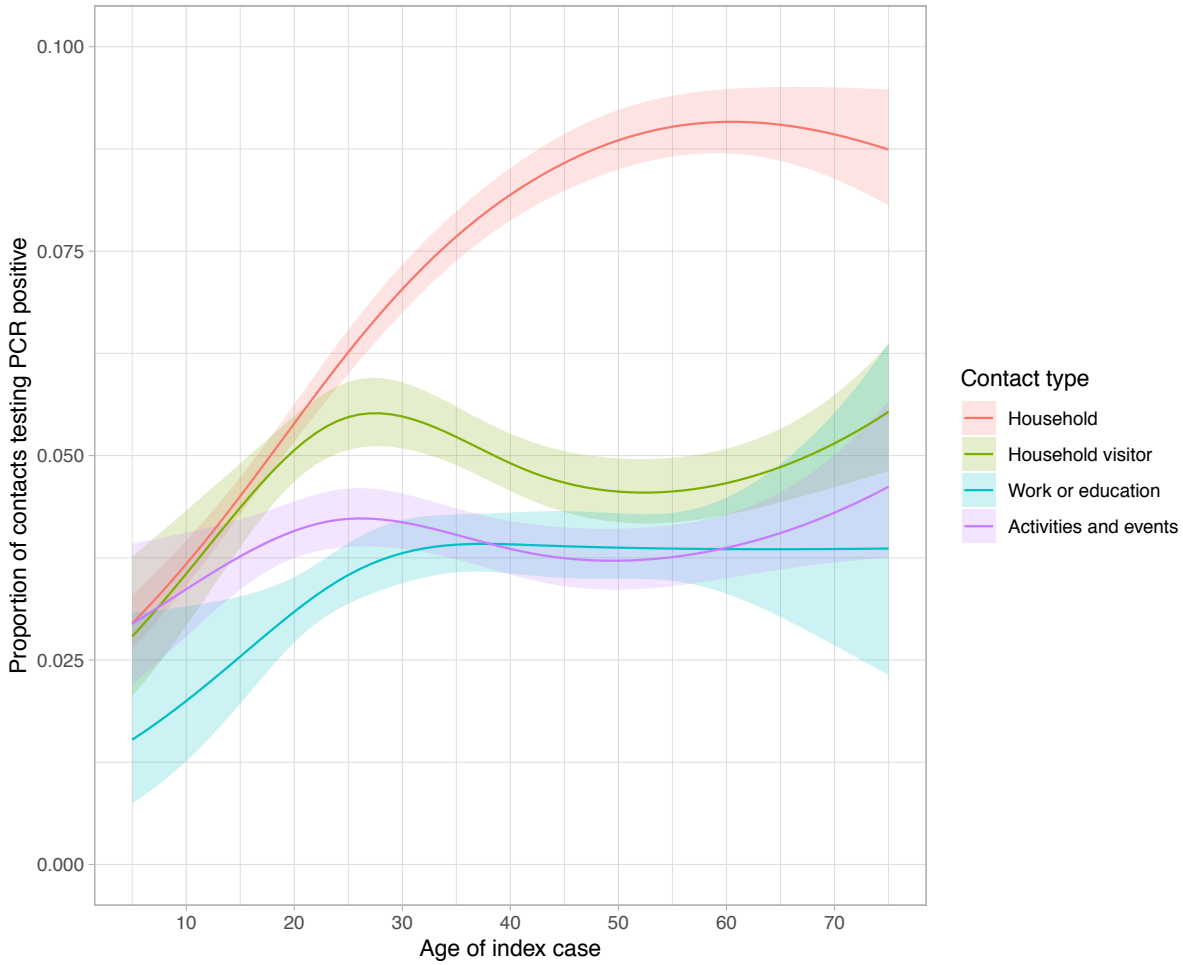

Figure 3. Relationship between case age and the proportion of their contacts with a PCR positive result, by contact event type. Model predictions are plotted after adjustment for Ct value (set to the median Ct value, 20.2), diagnostic laboratory (set to Milton Keynes), ethnicity (set to white). Age is fitted as a 4-knot spline with an interaction between age and contact event type. The shaded area indicates the 95% confidence interval.

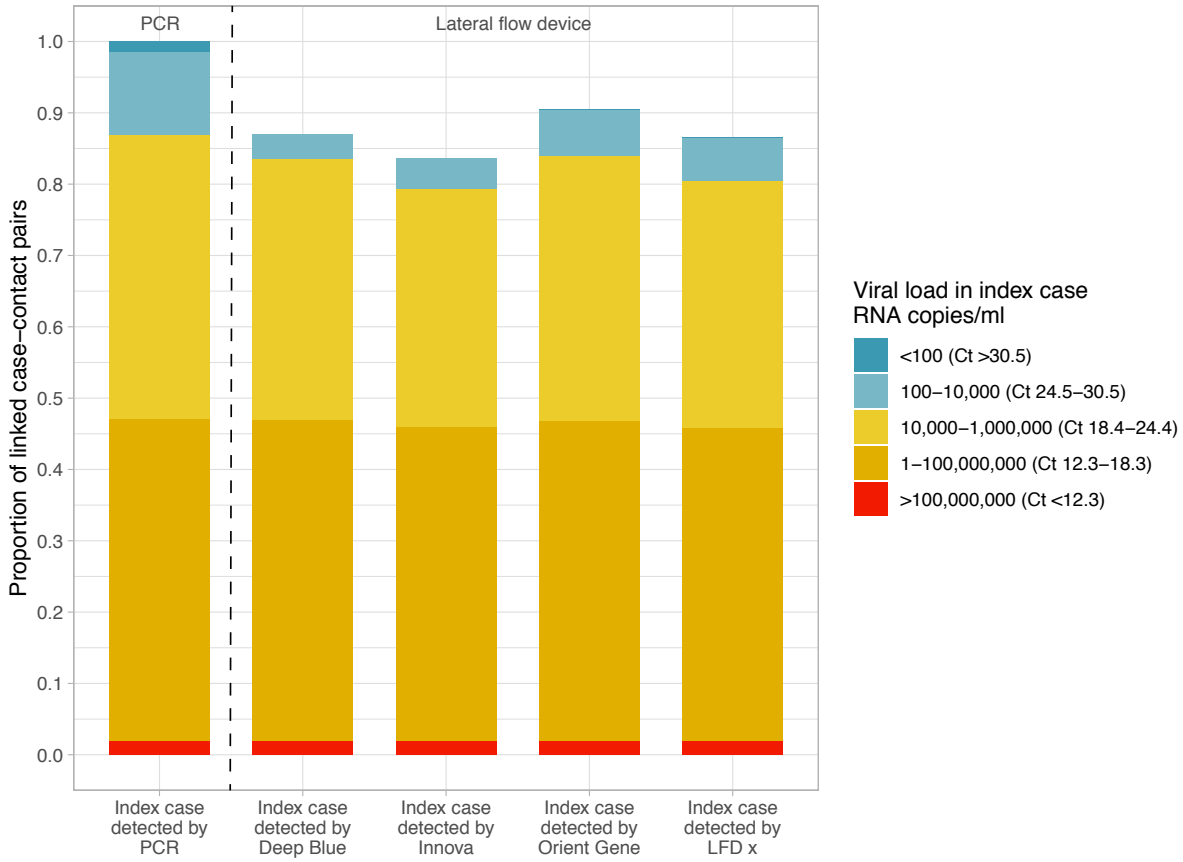

Figure 4. Simulated proportion of cases with a PCR-positive contact detected using four lateral flow devices (LFD). The proportion of cases detected by PCR viral load group is shown in the PCR column. The number of cases with a PCR-positive contact who would be detected using each LFD is shown for 4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

Tables

		Cases with ≥1 unique contact (n=145,973)	**Contacts, PCR positive within +2 to +7 days of case test** (n=18,291)	**Contacts, not PCR positive within +2 to +7 days of case test** (n=284,901)
Sex, n (%)	Male	66,548 (45.6%)	8429 (46.1%)	127,237 (44.7%)
	Female	79,425 (54.4%)	9862 (53.9%)	146,359 (51.4%)
	Not available	-	-	11,305 (4.0%)
Age, median (IQR) years		35 (22-50)	36 (23-52)	32 (19-50)
Ethnicity, n (%)	Asian or Asian British	14,241 (9.8%)	1731 (9.5%)	27,954 (9.8%)
	Black, African, Caribbean or black British	2374 (1.6%)	233 (1.3%)	4692 (1.7%)
	Mixed or multiple ethnic groups	3315 (2.3%)	400 (2.2%)	6703 (2.4%)
	Other ethnic group	1619 (1.1%)	197 (1.1%)	2986 (1.1%
	Prefer not to say / unavailable	14,007 (9.6%)	921 (5.0%)	25,552 (9.0%)
	White	110,426 (75.6%)	14,809 (81%)	217,014 (76.2%)
Prevalence, median (IQR) *per 100,000 population*		247 (67-455)	243 (70-454)	222 (57-442)
Deprivation rank, median (IQR) *higher = more deprived (of 32,844 areas)*		13,819 (11,057-17,444)	14,049 (11,057-17,800)	14,049 (11,101-17,800)
Symptomatic, n (%)		136,411 (93.4%)	9375 (51.3%)	50,192 (17.6%)
Time from symptoms to test in those symptomatic, median (IQR) days		2 (1-3)	1 (1-3)	2 (1-3)
Time from index case test to contact test, median (IQR) days		-	3 (2-4)	1 (0-3)
Contact event type, n (%)	Household	-	14,012 (76.6%)	191,171 (67.1%)
	Household visitor	-	1833 (10.0%)	34,583 (12.1%)
	Education or work	-	1055 (5.8%)	26,492 (9.3%)
	Activities and events	-	1391 (7.6%)	32,644 (11.5%)

Table 1. Demographics and characteristics of the study population.

Variable		Univariable			Multivariable		
		OR	95% CI	p	OR	95% CI	p
Prevalence at contact's home address	Prevalence, for each 1 unit change in cases per 1000	1.03	1.02 - 1.03	<0.001			
Deprivation rank at contact's home address	Per 10,000 unit change (lower = more deprived)	0.97	0.94 – 1.00	0.03			
Case viral load	Ct value, per unit change	0.93	0.92 - 0.93	<0.001	0.93	0.92 - 0.93	<0.001
Case sex	Female	1.00					
	Male	1.06	1.03 – 1.10	<0.001			
Case age*	Age, 30 years	1.00		<0.001	See Figure 3**		
	10 years	0.53	0.49 - 0.57				
	50 years	1.22	1.17 - 1.27				
	70 years	1.25	1.17 - 1.33				
Contact event	Household	1.00					
	Activities and events	0.58	0.55 - 0.62	<0.001			
	Household visitor	0.72	0.69 – 0.76	<0.001			
	Work or education	0.54	0.51 - 0.58	<0.001			
Case PCR diagnostic laboratory	Milton Keynes	1.00			1.00		
	Alderley Park	1.29	1.24 - 1.34	<0.001	1.15	1.11 - 1.20	<0.001
	Glasgow	1.00	0.95 – 1.05	0.99	0.88	0.84 – 0.93	<0.001
Case ethnicity	White	1.00			1.00		
	Asian	0.91	0.86 - 0.96	<0.001	0.84	0.80 - 0.88	<0.001
	Black	0.73	0.64 - 0.83	<0.001	0.79	0.69 - 0.91	<0.001
	Mixed	0.87	0.79 - 0.97	0.01	0.89	0.81 – 0.99	0.03
	Other	0.97	0.84 – 1.12	0.65	1.00	0.86 – 1.16	0.99
	Not available	0.53	0.49 - 0.57	<0.001	0.50	0.47 - 0.54	<0.001

Table 2. Univariable and multivariable associations with the proportion of contacts testing PCR positive. *Case age was fitted as a 4-knot spline (see Figure S1 for univariable relationship). **In the multivariable model age was fitted as a 4-knot spline, with an interaction with contact event type (heterogeneity p<0.001), see Figure 3. The multivariable model excludes 11,305 contacts without their sex recorded.

References

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

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

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

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

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

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

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

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

9 Office for National Statistics. Deaths registered in England and Wales, provisional: week ending 6 November 2020. https://www.ons.gov.uk/releases/deathsregisteredinenglandandwalesprovisionalweekending6november2020 (accessed Jan 17, 2021).

10 Sun K, Wang W, Gao L, *et al.* Transmission heterogeneities, kinetics, and controllability of SARS-CoV-2. *Science* 2021; **371**: eabe2424.

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

12 Luo L, Liu D, Liao X, *et al.* Contact Settings and Risk for Transmission in 3410 Close Contacts of Patients With COVID-19 in Guangzhou, China: A Prospective Cohort Study. *Ann Intern Med* 2020; **173**: 879–87.

13 Wölfel R, Corman VM, Guggemos W, *et al.* Virological assessment of hospitalized patients with COVID-2019. *Nature* 2020; **581**: 465–9.

14 Pekosz A, Cooper CK, Parvu V, *et al.* Antigen-based testing but not real-time PCR correlates with SARS-CoV-2 virus culture. *medRxiv* DOI:10.1101/2020.10.02.20205708.

15 Peto T, Team UC-19 LFO. COVID-19: Rapid Antigen detection for SARS-CoV-2 by lateral flow assay: a national systematic evaluation for mass-testing. *medRxiv* DOI:10.1101/2021.01.13.21249563.

16 Madewell ZJ, Yang Y, Longini IM, Halloran ME, Dean NE. Household Transmission of SARS-CoV-2. *Jama Netw Open* 2020; **3**: e2031756.

17 College I. Report 38: SARS-CoV-2 setting-specific transmission rates: a systematic review and meta-analysis. https://www.imperial.ac.uk/media/imperial-college/medicine/mrc-gida/2020-11-27-COVID19-Report-38.pdf.

18 Marks M, Millat-Martinez P, Ouchi D, *et al.* Transmission of COVID-19 in 282 clusters in Catalonia, Spain: a cohort study. *medRxiv* DOI:10.1101/2020.10.27.20220277.

19 Pouwels KB, House T, Pritchard E, *et al.* Community prevalence of SARS-CoV-2 in England from April to November, 2020: results from the ONS Coronavirus Infection Survey. *Lancet Public Heal* 2021; **6**: e30–8.

20 Buitrago-Garcia D, Egli-Gany D, Counotte MJ, *et al.* Occurrence and transmission potential of asymptomatic and presymptomatic SARS-CoV-2 infections: A living systematic review and meta-analysis. *Plos Med* 2020; **17**: e1003346.

21 Kissler SM, Fauver JR, Mack C, *et al.* SARS-CoV-2 viral dynamics in acute infections. *medRxiv* DOI:10.1101/2020.10.21.20217042.

22 Smith LE, Potts HWW, Amlot R, Fear NT, Michie S, Rubin J. Adherence to the test, trace and isolate system: results from a time series of 21 nationally representative surveys in the UK (the COVID-19 Rapid Survey of Adherence to Interventions and Responses [CORSAIR] study). *medRxiv* DOI:10.1101/2020.09.15.20191957.

Declaration of interests

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

Funding Statement

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

An observational study of SARS-CoV-2 infectivity by viral load and demographic factors and the utility lateral flow devices to prevent transmission: Supplementary materials

Supplementary methods

Data definitions

Contacts were defined as follows:[1] a person who has been close to someone who has tested PCR-positive for COVID-19 anytime from 2 days before the person was symptomatic up to 10 days from onset of symptoms. The nature of the contact could include:

- Living in the same household OR
- Face to face contact (within 1 metre for any length of time) or skin to skin contact or someone the case coughed on OR
- Within 1 metre for 1 minute of longer OR
- Within 1-2 metres for more than 15 minutes OR
- Sexual contacts OR
- Travel in the same vehicle or a plane

Ethnicity was summarised using 5 ethnic groups defined by the UK government ("white", "mixed or multiple ethnic groups", "Asian or Asian British", "Black, African, Caribbean or Black British", "Other ethnic group").

Local COVID-19 prevalence at the time of each contact event was determined using data for the lower tier local authority (LTLA) containing the contact's home address. Prevalence is calculated using the number of cases populated into CTAS per week at that LTLA per 100,000 population.

Contact deprivation indices were sourced from the latest census in 2019, averaged for each LTLA.[2]

Laboratory methods

Ct values were available for community tests undertaken by the UK's Lighthouse Laboratories in Milton Keynes, Alderley Park and Glasgow. PCR testing was undertaken using the Thermo Fisher TaqPath assay (targeting S and N genes, and ORF1ab) following extraction of 200µl of viral transport media on the Kingfisher extraction platform, yielding 60µl post-extraction of which 6µl was used in the PCR reaction. Mean Ct values were calculated across all non-missing targets.

To enable comparison of data across PCR assays the Qnostics SARS-CoV-2 Analytical Q Panel 01 (Qnostics, Glasgow, UK) was used to calibrate Ct values from the Thermo Fisher assay into equivalent synthetic RNA viral load (VL) in copies per ml. The resulting equation for converting Ct values into viral loads for the Thermo Fisher TaqPath assay was $\log_{10}(VL) = 12.0 - 0.328*Ct$.

LFD sensitivity by viral load

We used previously reported data and estimates of the sensitivity of four LFDs by viral load from a community-based evaluation.[3] Data were available for the Innova SARS-CoV-2 Antigen Rapid

Qualitative Test (Innova), Anhui Deepblue Medical Technology COVID-19 (Sars-CoV-2) Antigen Test kit (Colloidal Gold) (Deep Blue), LFD x (the manufacturer did not give consent to be named) and the Zhejiang Orient Gene Biotech Co. Coronavirus Ag Rapid Test Cassette (Orient Gene).

Data were available from 420 Innova, 177 Deep Blue, 99 LFD x and 95 Orient Gene LFD tests performed in individuals diagnosed with SARS-CoV-2 infection by PCR within the last 5 days. Self-obtained combined nasal and oropharyngeal swabs were analysed. Contemporaneous paired swabs were obtained for repeat PCR testing. Repeat PCR testing was undertaken using the Roche Cobas SARS-CoV-2 test and platform. To enable comparisons with Thermo Fisher TaqPath PCR Ct values we used data on 254 additional samples tested using both extraction and PCR assay methods to enable the following conversion to be derived using linear regression: (Cobas Ct) = 5.5 + 1.0*(Thermo Fisher Ct).

For each LFD we fitted a logistic regression model, to generate the probability of a positive LFD test for a given Ct value or viral load (Figure S1).

For each source case and LFD we simulated a positive or negative LFD result by randomly drawing from the probability of the LFD being positive by the source case's Ct value (determined from the logistic regression model and accounting for uncertainty in the model's estimates by sampling model slope and intercept terms using the mean, standard error and covariance of each parameter). Each simulation was repeated 1000 times and summary results presented. Additionally, we generated simulations for a range of alternative possible lateral flow test performances.

Supplementary figures

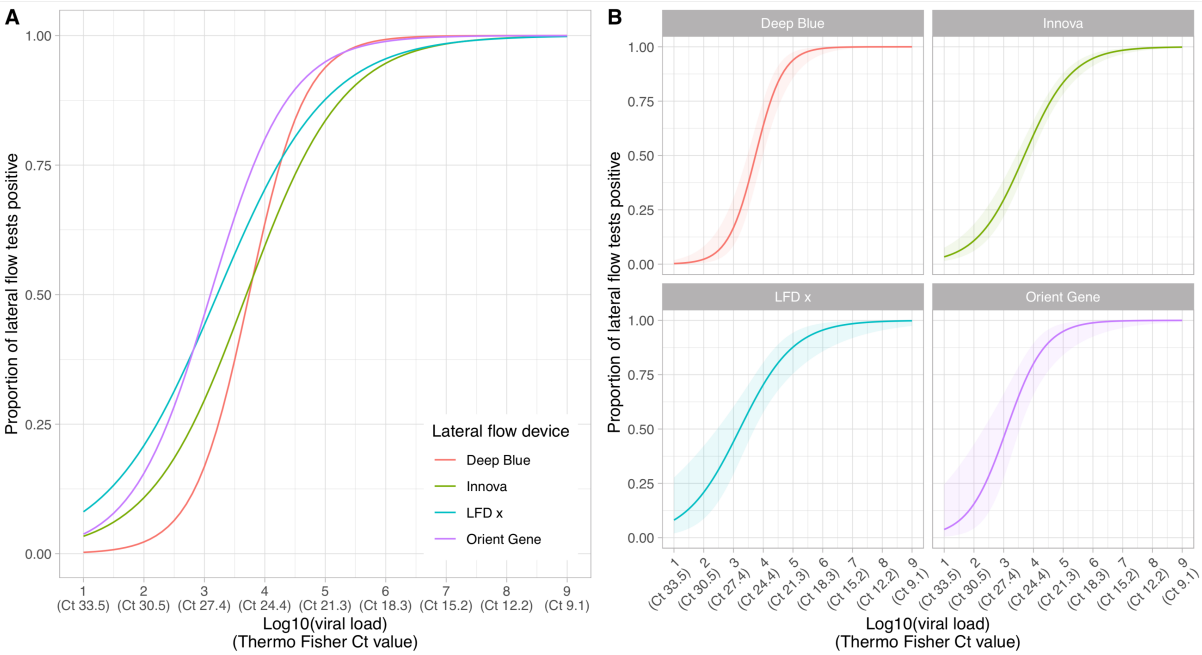

Figure S1. Sensitivity of four lateral flow devices by viral load. Thermo Fisher TaqPath assay equivalent Ct units are shown using the formula: $\log_{10}$(viral load) = 12.0 - 0.328*Ct. Panel A shows the fitted relationship, and panel B $_{\text{the}}$ 95% confidence intervals for each curve.

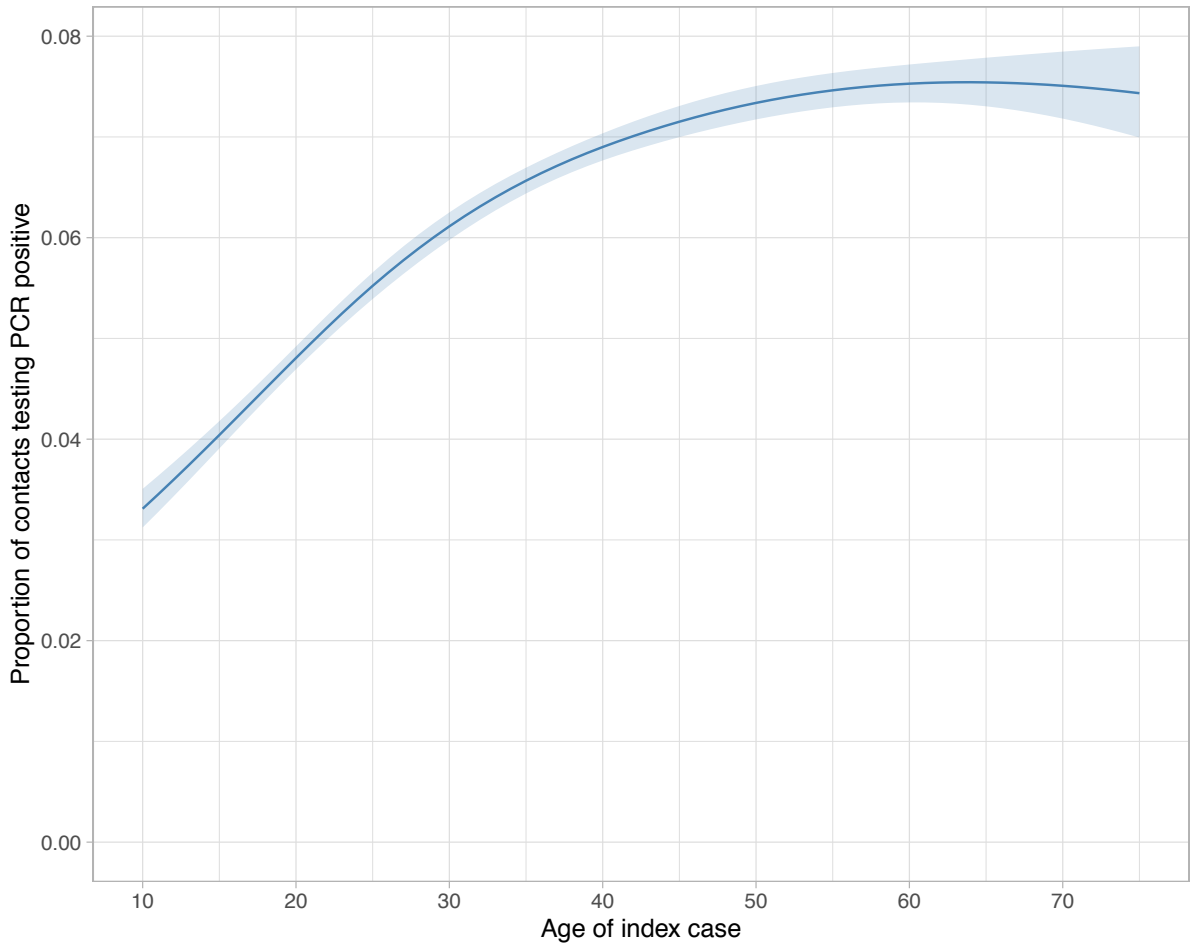

Figure S2. Univariable relationship between age of the index case and proportion of contacts testing PCR positive. Plotted based on 4-knot spline with default spaced knots, the ribbon shows the 95% confidence interval.

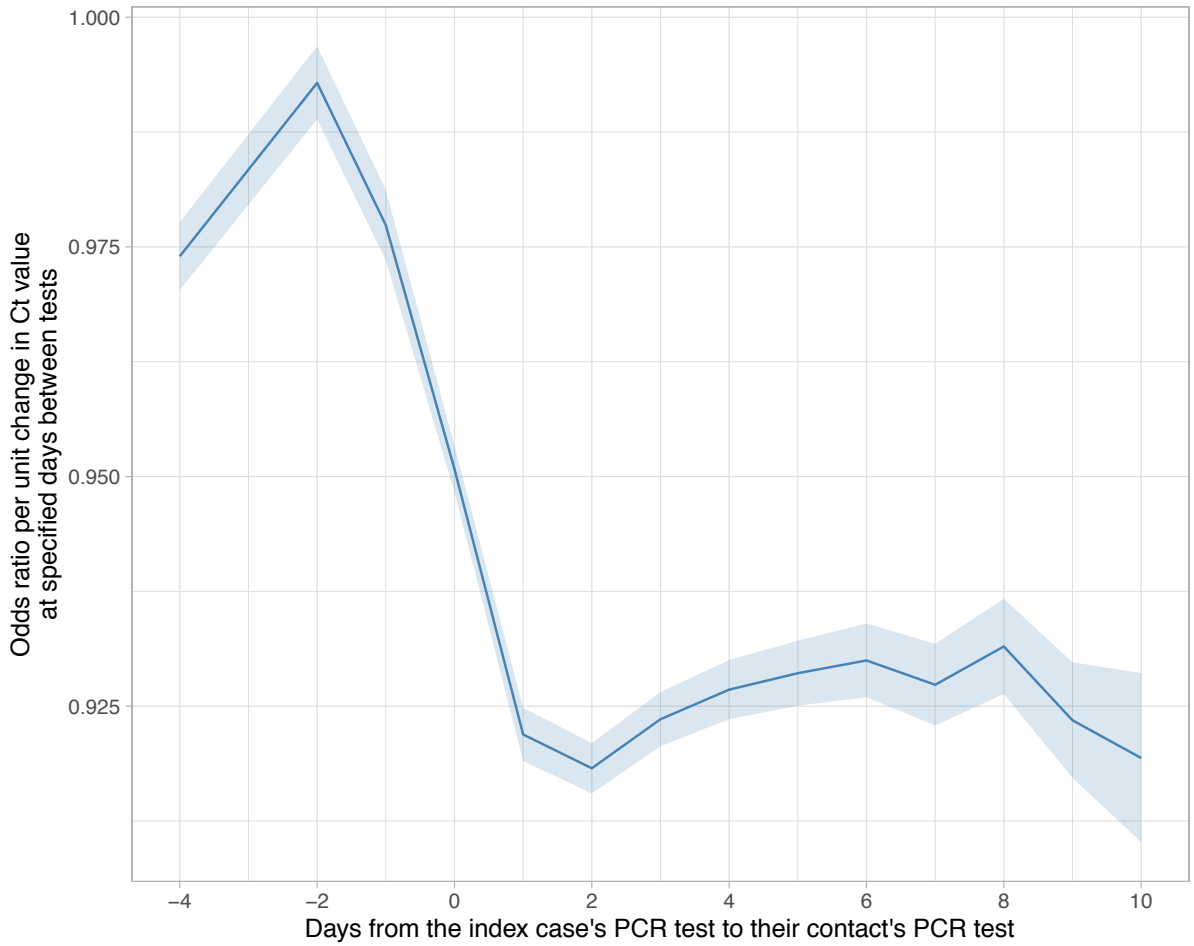

Figure S3. Sensitivity analysis to investigate the impact of the 2 to 7 day time window for positive PCR results in contacts in the main analysis. The main analysis was repeated including all positive PCR tests for contacts between 4 days before and 10 days after the index case's PCR test and including the days between tests as a main effect. Additionally, an interaction term was included between the days between PCR tests and the case's Ct value. The change in odds ratio for a positive PCR test by Ct value across differing days between tests is shown.

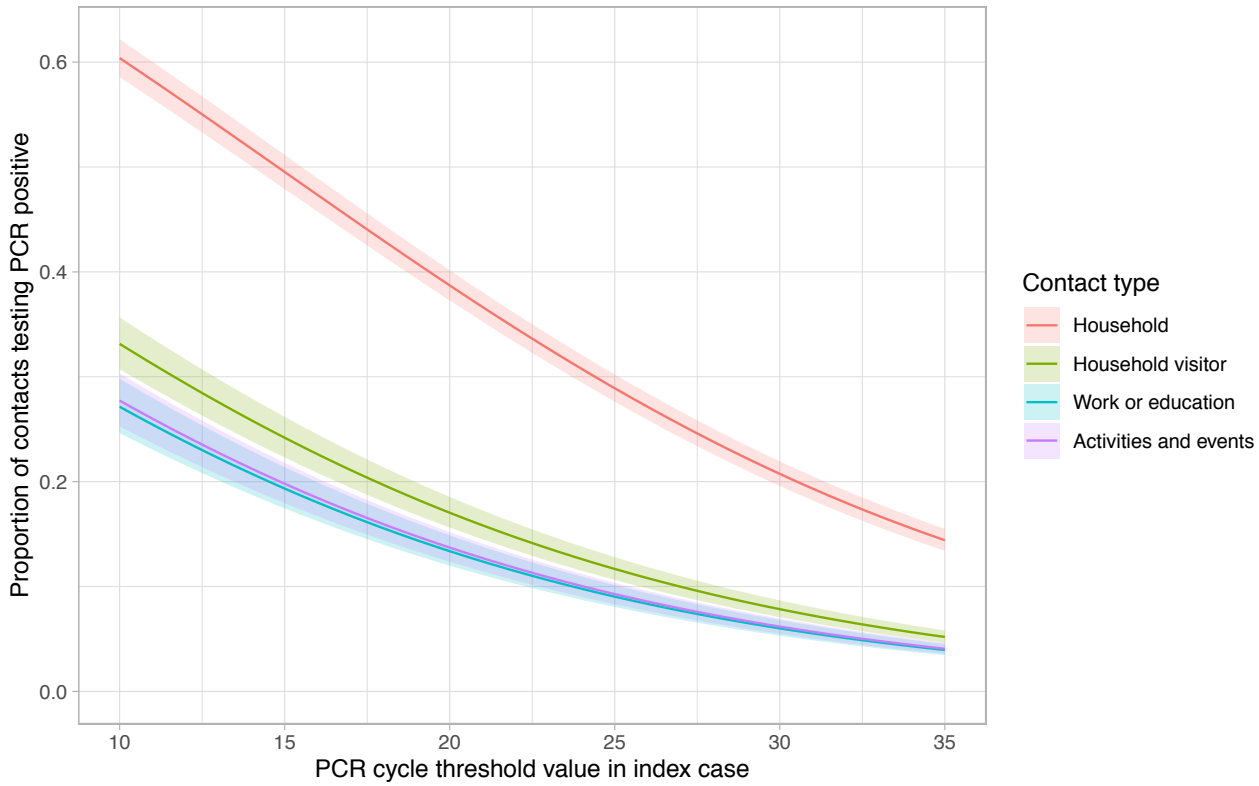

Figure S4. Relationship between PCR cycle threshold (Ct) value in cases and the proportion of their contacts who underwent PCR testing with a PCR positive result, by contact event type. Model predictions are plotted after adjustment for age (set to the median value, 31 years), diagnostic laboratory (set to Milton Keynes), ethnicity (set to white), and prevalence (set to median, 197 per 100,000 population). The shaded area indicates the 95% confidence interval.

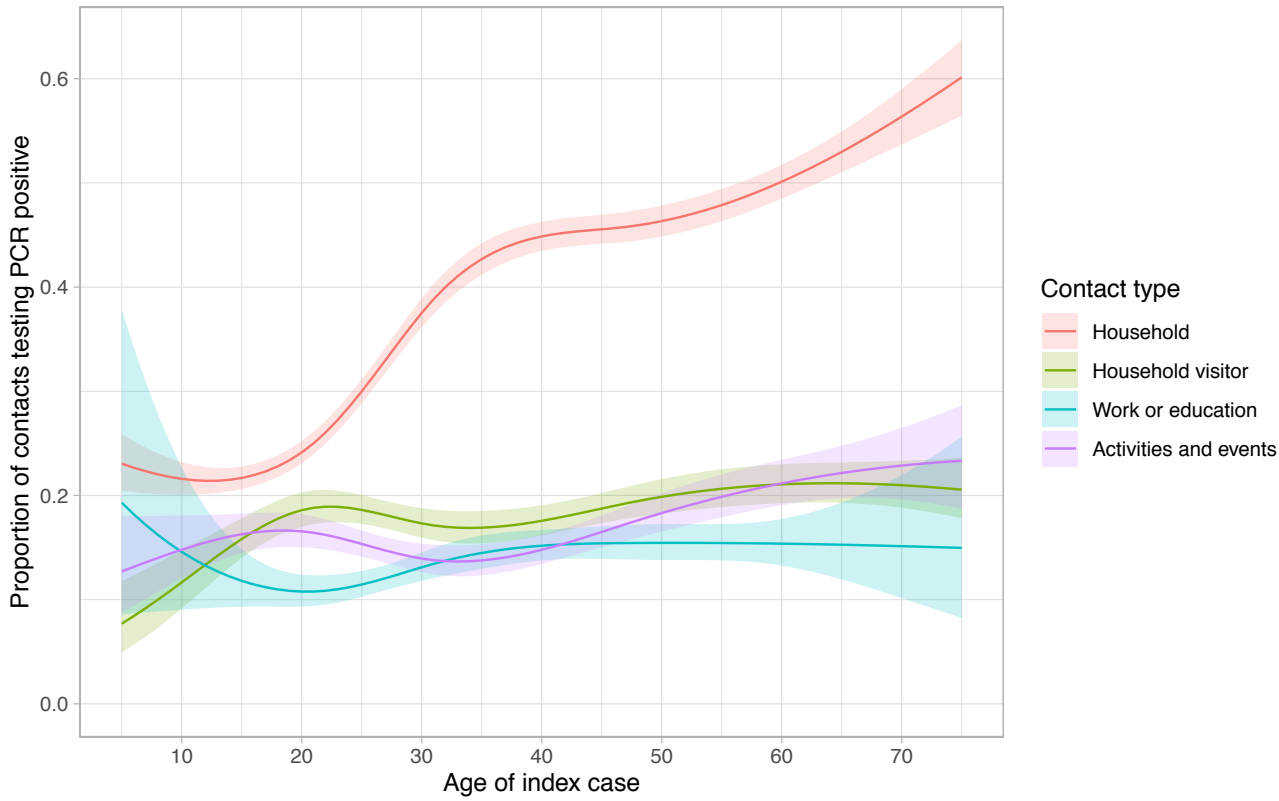

Figure S5. Relationship between case age and the proportion of their contacts who underwent PCR testing with a PCR positive result, by contact event type. Model predictions are plotted after adjustment for Ct value (set to the median Ct value, 19.9), diagnostic laboratory (set to Milton Keynes), ethnicity (set to white), and prevalence (set to median, 197 per 100,000 population). Age is fitted as a 5-knot spline with an interaction between age and contact event type (heterogeneity $p<0.001$). The shaded area indicates the 95% confidence interval.

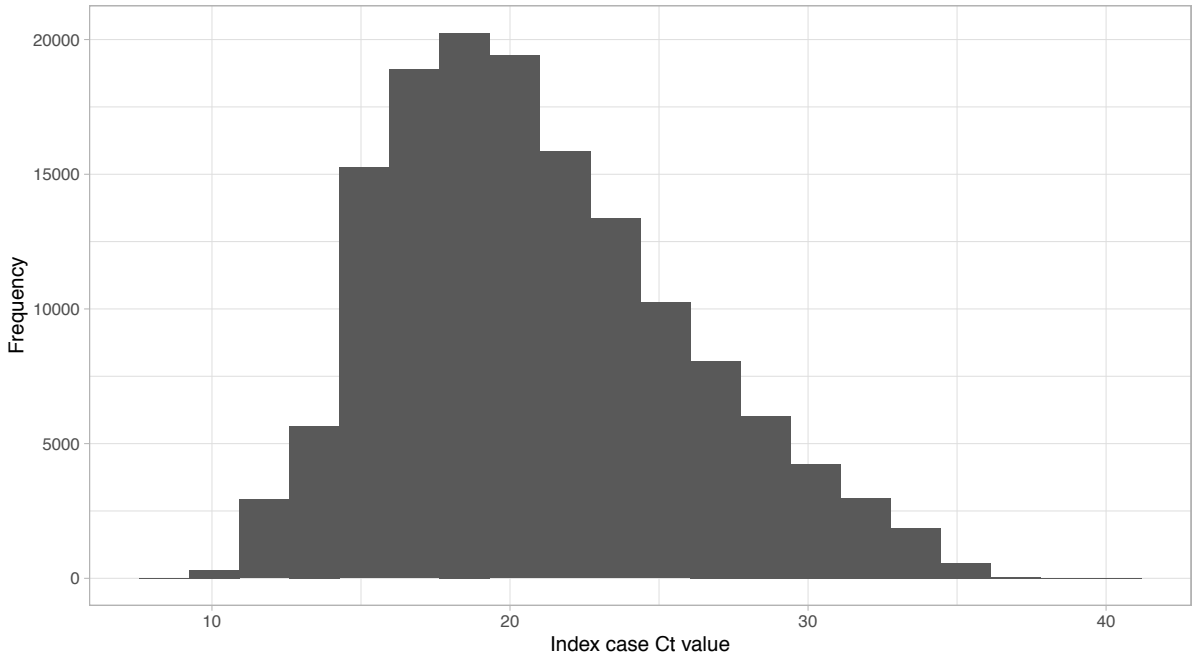

Figure S6. Distribution of mean Ct values for 145,973 index cases. Mean Ct values were calculated across all non-missing SARS-CoV-2 PCR targets.

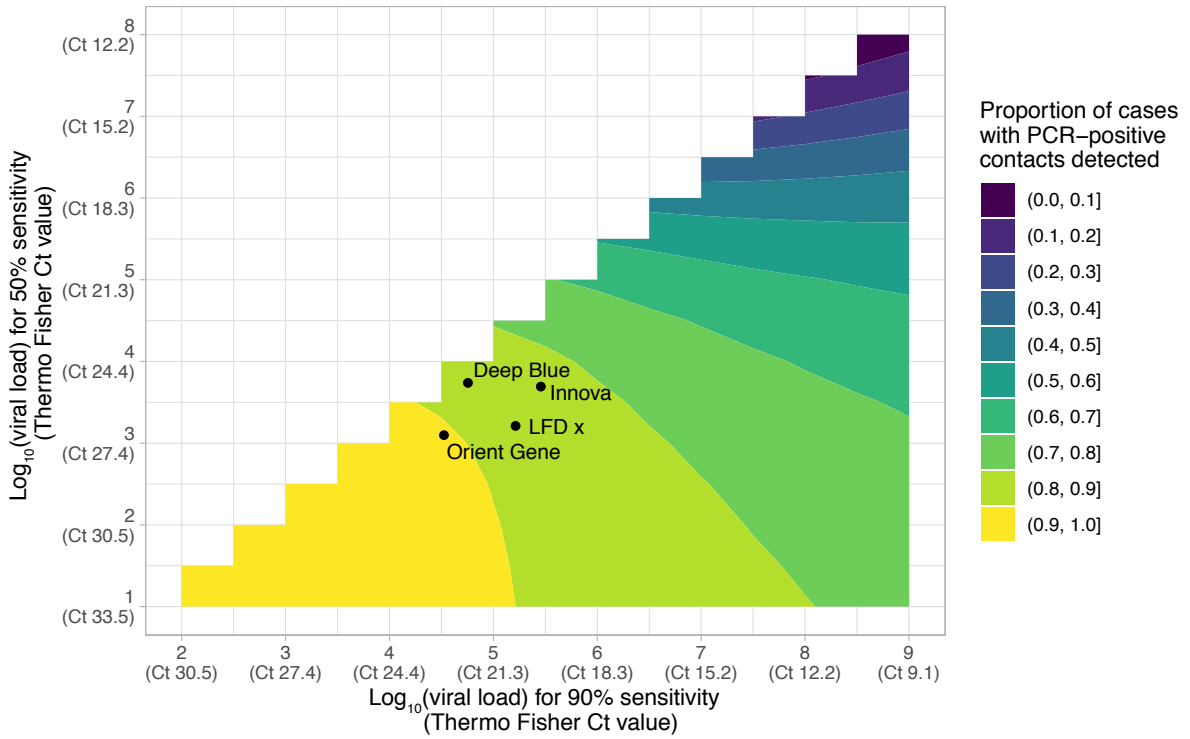

Figure S7. Simulated performance for a range of LFDs. The sensitivity of a LFD by viral load can be summarised using a logistic model with two parameters, a slope and intercept. The slope and intercept can be transformed into two alternative parameters, e.g. the viral load at which 50% of all individuals are LFD positive and the viral load at which 90% of all individuals are LFD positive. For a given combination of two parameters we simulate the proportion of cases with PCR-positive contacts that are detected by LFD, the mean result over 100 simulations is plotted. Thermo Fisher TaqPath assay equivalent Ct units are shown using the formula: $\log_{10}$(viral load) = 12.0 - 0.328*Ct. The performance of 4 LFDs (Figure S1) are overlaid on the simulation results.

Supplementary tables

Variable		Univariable			Multivariable		
		OR	95% CI	p	OR	95% CI	p
Prevalence at contact's home address	Prevalence, for each 1 unit change in cases per 1000	1.09	1.08 - 1.10	<0.001	1.05	1.46 - 1.06	<0.001
Deprivation rank at contact's home address	Per 10,000 unit change (lower = more deprived)	0.82	0.79 - 0.85	<0.001			
Case viral load	Ct value, per unit change	0.91	0.91 - 0.92	<0.001	0.92	0.91 - 0.92	<0.001
Case sex	Female	1.00					
	Male	1.10	1.06 – 1.14	<0.001			
Case age*	Age, 30 years	1.00		<0.001	**See Figure S5		
	10 years	0.61	0.56 - 0.66				
	50 years	1.63	1.55 - 1.72				
	70 years	1.78	1.65 - 1.93				
Contact event	Household	1.00					
	Activities and events	0.35	0.33 - 0.37	<0.001			
	Household visitor	0.39	0.37 - 0.41	<0.001			
	Work or education	0.29	0.27 - 0.31	<0.001			
Case PCR diagnostic laboratory	Milton Keynes	1.00			1.00		
	Alderley Park	1.54	1.47 - 1.61	<0.001	1.24	1.18 - 1.31	<0.001
	Glasgow	1.37	1.30 - 1.45	<0.001	1.16	1.09 - 1.23	<0.001
Case ethnicity	White	1.00			1.00		
	Asian	1.36	1.28 - 1.45	<0.001	1.19	1.11 - 1.27	<0.001
	Black	1.03	0.88 - 1.21	0.69	1.17	0.99 - 1.38	0.07
	Mixed	1.04	0.92 – 1.18	0.50	1.07	0.94 - 1.22	0.31
	Other	1.24	1.04 - 1.47	0.02	1.31	1.08 - 1.58	0.005
	Not available	1.05	0.96 - 1.13	0.28	1.00	0.92 - 1.09	0.99

Table S1. Univariable and multivariable associations with the proportion of contacts who underwent PCR testing having a positive result. *Case age was fitted as a 5-knot spline. **In the multivariable model age was fitted as a 5-knot spline, with an interaction with contact event type (heterogeneity p<0.001), see Figure S5.

References

1 Public Health England. Guidance for contacts of people with confirmed coronavirus (COVID-19) infection who do not live with the person. https://www.gov.uk/government/publications/guidance-for-contacts-of-people-with-possible-or-confirmed-coronavirus-covid-19-infection-who-do-not-live-with-the-person/guidance-for-contacts-of-people-with-possible-or-confirmed-coronavirus-covid-19-infection-who-do-not-live-with-the-person (accessed Jan 20, 2021).

2 gov.uk. English indices of deprivation 2019. https://www.gov.uk/government/statistics/english-indices-of-deprivation-2019 (accessed Jan 17, 2021).

3 Peto T, Team UC-19 LFO. COVID-19: Rapid Antigen detection for SARS-CoV-2 by lateral flow assay: a national systematic evaluation for mass-testing. *medRxiv* DOI:10.1101/2021.01.13.21249563.