UK Health Security Agency

The Canna model

Assessing the impact of NHS Test and Trace on COVID-19 transmission

June 2020 to April 2021

Contents

Abstract

In February 2021 NHS Test and Trace (NHSTT) provided an estimate of the impact of test, trace and self-isolation (TTI) on COVID-19 transmission in October 2020 using the Rùm model. The Canna model uses an updated framework to estimate the historical impact of TTI, in England, from June 2020 to April 2021. We estimate the reduction in transmission by considering the rate and timing of isolation among infectious individuals.

In response to comments on previous methodology, we have estimated the marginal impact directly attributable to NHSTT by comparing to a counterfactual scenario. In this counterfactual, we assume that all individuals who tested with COVID-like-symptoms, would still self-isolate without ever taking a test, together with their household contacts. We assume that isolation would be undertaken with the same level of compliance assumed for positively tested cases and their household contacts. Notably this counterfactual scenario relies on many more isolations taking place than with NHSTT, where a large proportion of this population would no longer have to isolate after a negative test.

The counterfactual has been set at the very upper limit of what is plausible without testing. In reality, we expect a positive test result will significantly increase isolation compliance; however, it is impossible to accurately determine the scale of this effect. In this study, we therefore report the full impact from TTI as well as the impact above the counterfactual. We assume that the marginal impact directly attributable to NHSTT will lie within this range.

Since August 2020, we estimate that the transmission reduction from TTI varied over time from 10 to 28% (across a 90% confidence interval). In the counterfactual this reduced to 6 to 19%. The transmission reduction from TTI, above the counterfactual varied over time from 4 to 16%. In June and July 2020, when cases remained relatively low, the transmission reduction from TTI was generally lower than for the remainder of the study period (6 to 14%).

Since August 2020, the reduction in the reproduction rate (Rt) from TTI varied over time from 0.10 to 0.44; the Rt reduction above the counterfactual varied from 0.04 to 0.22. In several periods (August 2020, November 2020, January to April 2021) our central estimates show that TTI would have been critical in reducing the reproduction rate, Rt, to below 1, thereby preventing exponential growth in infections.

We estimate that isolations occurring due to TTI over the full period of this study directly prevented 1.2 to 2.0 million secondary cases; 0.3 to 0.5 million above the counterfactual. We have not considered the impact on any onward chains of transmission; therefore, we expect that the true number of cases prevented will be significantly higher.

NHSTT notified 11 million individuals to isolate over the course of the study period (a further 21 million individuals would have been required to isolate for a short time prior to a household member receiving a negative test.) In the counterfactual scenario 25 million

individuals would have been required to isolate for the full isolation period, significantly more than with NHSTT.

Our study does not account for the impact of Pillar 1 testing in hospitals, which would have had significant additional benefits in preventing hospital outbreaks and ensuring that the right treatments were provided to those in care.

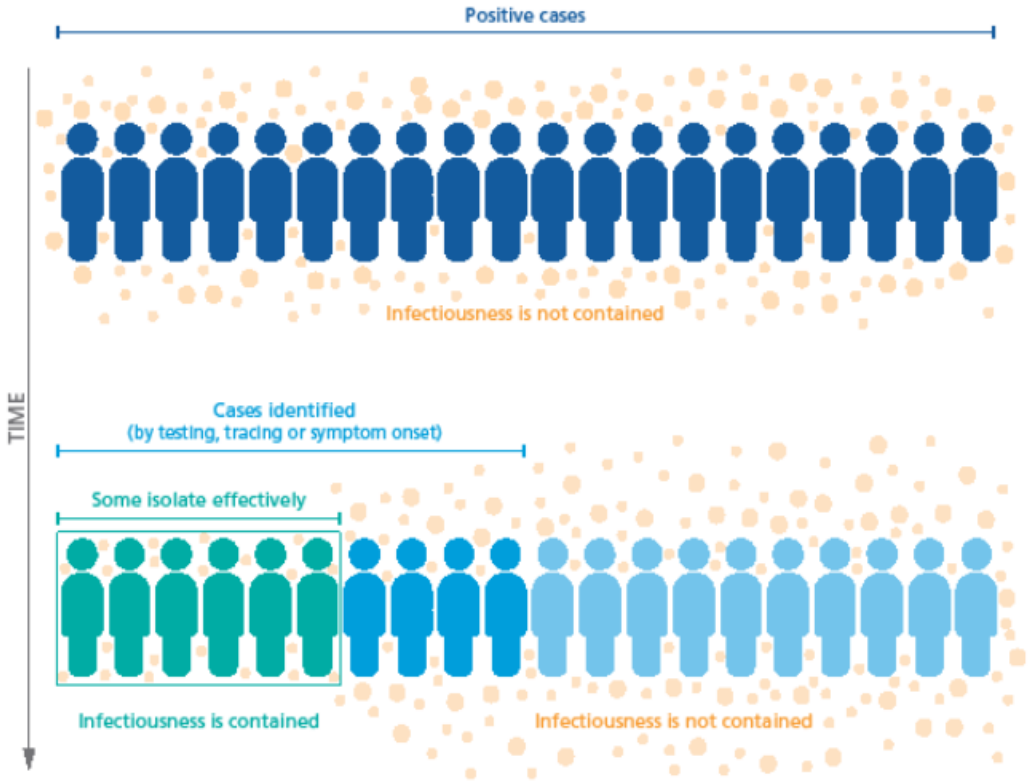
A panel of external experts from academia provided advice on the modelling throughout its development. Given the constraints, the panel regarded the core assumptions and structure as appropriate for determining the impact on transmission of test, trace and self-isolation. The panel consisted of: Prof Neil Ferguson, School of Public Health, Imperial College London; Prof Christophe Fraser, Big Data Institute, Oxford University; Dr Adam Kucharski, London School of Tropical Hygiene and Tropical Medicine; Dr James Hetherington, Director of the Centre for Advanced Research Computing, University College London; Prof Sylvia Richardson, Director of the MRC Biostatistics Unit, The University of Cambridge.

1. Model method

NHS Test and Trace (NHSTT) was set up in May 2020 to help prevent the spread of coronavirus. The combined system established rules for self-isolation and created an infrastructure to test individuals for COVID-19 and subsequently trace and notify their contacts.

The Canna1 model calculates the transmission reduction from test, trace and self-isolation (TTI) by determining the proportion of all infectious individuals undergoing isolation over the time course of their infectious period (Figure 1).

Figure 1. A simplified illustration of the Canna model



Transmission reduction occurs as a result of identification and then self-isolation of infectious individuals. The amount of transmission reduction is determined by the proportion of total infectiousness that is contained. In this study we determine this at a population level by comparing the total number of isolations from test, trace and self-isolate to the total number of infectious individuals, derived from ONS incidence estimates.

[1] Canna is a neighbour to Rùm, among the small Isles in the Inner Hebrides.

The isolation of infectious individuals is assumed to occur as a result of either; becoming a case after receipt of a positive test result, becoming a contact after being traced or symptom onset without ever engaging with NHSTT.

The timing of each isolation determines the relative amount of infectiousness that is potentially abated. The final reduction in onward infections is dependent on an individual's compliance to isolation (Equation 1).

$$\text{For each 14 day time period:}$$
$$\%\ \text{Transmission reduction from TTI} = \frac{\sum_{\text{All infectious isolations}} (\%\text{Infectiousness Abated} \times \%\text{Compliance})}{\text{Total number of infectious indviduals}}$$

Equation 1

This framework makes some notable simplifying assumptions: the relative rate of transmission and hence the reproduction rate (Rt) will scale in proportion to the number of infectious individuals not in isolation; Rt and prevalence are relatively stable over each 14-day time period; infectious and isolating individuals are evenly distributed among the population; the average rate of transmission among infectious individuals is the same, regardless of symptom expression or detectability. None of these conditions are strictly true; however, they help us to establish a tractable model system. In the concluding analyses we consider the impact of these (and other) modelling assumptions on our evaluation.

The fundamental framework described here is broadly consistent with our former publication based on the Rùm model (Department of Health and Social Care, 2021b). Here, we use a data driven approach to estimate the number of infectious cases and contacts over time. This analysis covers the period from 1 June 2020, just after NHSTT was formally established, until the end of April 2021.

Below, we review the methods and assumptions in our model and describe the Monte Carlo sampling that we used to evaluate the output uncertainty. All the parameters used in the model are summarised in Table 1.

1.1 The population of infectious individuals

We estimate the total number of infectious individuals in each discrete 14-day time period using the incidence rates provided by the ONS community infection survey (Office for National Statistics, 2021a). (Figure 11 in annex A.2 for details).

We interpolated the ONS data to provide daily estimates of incidence over the study period. We then calculated the total number of new infections falling within a 14-day window, 6 days prior to each of the 14-day study periods in which we aggregate registered cases and contacts. The 6-day delay was used to account for the average time delay between new infections (incidence) and case detection. Notably, the ONS community infection survey does not identify cases occurring in residential settings such as care

homes and prisons. For simplicity, in this study, we assume that the ONS incidence rates can be applied across the whole population.

1.2 The number of infectious cases and contacts

We used historical NHSTT data to determine the number of unique, infectious cases and contacts that were identified and reached in each 14-day time period.

Our core dataset is from CTAS (Contact Tracing and Advisory Service). In addition, we included additional test results from NPEX (National Pathology Exchange), aggregated data from the COVID-19 App and data from DFE (Department for Education) on absenteeism among school children linked to contacts with COVID-19 cases.

We ensure that there is no duplication of any individuals appearing in CTAS within a 14-day time window before and after their first registered case or contact date. We count individuals as cases or contacts falling within each 14-day study period, depending on which is registered first. Contacts need not be linked to the primary cases in the same time period. We adjusted the additional data to try to ensure that there was no double counting of cases or contacts falling outside of CTAS (see Annex A.1 for details).

1.2.1 Cases included in the study

Throughout this study we ignore the impact of NHSTT cases occurring in hospitals, associated with Pillar 1 testing, on the assumption that those cases cannot further isolate in order to prevent secondary community infections. (Although we discount the impact on community infections and Rt, we recognise that the identification of cases in hospitals is crucial for preventing hospital outbreaks). We do still count individuals traced as contacts of Pillar 1 cases, who we assume are resident in the community.

We treat all other cases, derived from Pillar 2 testing, as being equivalent in our calculations and have not attempted to differentiate the impact of TTI within any other sectors or settings (such as schools, prisons, care homes and so on, community testing and so on.). Pillar 2 cases are identified as either:

- Symptomatic PCR, where an individual has taken a polymerase chain reaction (PCR) test following symptom onset[2]

[2] We classify PCR test results as symptomatic based on self-identification of symptoms recorded at the time of test booking. The timing of symptom onset is subsequently recorded when tests are registered through CTAS.

- Asymptomatic PCR, where PCR tests are typically used 1 to 2 times a week in a variety of settings to test individuals without symptoms
- Assisted LFD, where supervised rapid Lateral Flow Device (LFD) tests were recommended for twice weekly use by a subset of the population
- Self-serve LFD, where unsupervised rapid tests were recommended for twice weekly use by a subset of the population

A proportion of LFD cases subsequently get a confirmatory PCR. When this is negative, we discount those individuals. Those with positive confirmatory PCRs are identified as LFD cases in this study.

1.2.2 Adjusting for false positive test results

To ensure that we only count positive LFD test results, without a confirmatory PCR, that represent genuine infections, we estimate the positive predictive value (PPV) of LFD tests over time and use this to scale down the number of LFD cases (see annex A.3). We carry forward this adjustment by similarly scaling down the number of contacts linked to assisted LFD tests without confirmatory PCR. For simplicity, in this study we assume that PCR cases do not include any false positives.

1.2.3 Estimating the number of infected contacts

We split CTAS contacts according to whether they are living in the same household or not. If an individual is reached twice (or more) by association with household and non-household contacts then we treat them as a household contact in order to estimate their likelihood of infection, but we use whichever notification occurs first as their isolation date.

To estimate the number of contacts that were infected we developed assumptions for the secondary attack rate (SAR) in household and non-household contacts. We based this primarily on the ATACCC study (Hakki S and ATACCC team, 2021), who estimated attack rates by conducting repeated tests on a sample of reached contacts. We also analysed NHSTT data to determine the rate at which contacts are identified as cases over time. We used this information, together with data on the penetration of different variants to convert the ATACCC study estimates into a time series (see Annex A.4).

The SAR gives us an estimate for the percentage of all secondary contacts that become cases. In this study we are only attributing isolations to contact notifications if, within a 14-day window before their contact registration, they had not previously been identified as a case (otherwise we count them as cases). Therefore we make an adjustment to account for those contacts that we have removed from our dataset (see annex A.9 for a detailed calculation; Figure 13 in annex A.5 shows the proportion of contacts that were previously cases.)

1.2.4 COVID-19 App data and school-age contacts

We uplift the final number of contacts in our dataset to account for those reached by either the COVID-19 App or school-age contacts (registered as absences in DFE schools data) that are not already accounted for in the NHSTT CTAS data (see annex A.6 and A.7).

1.3 The proportion of infectiousness abated by isolation

The timing of an infectious individual's isolation, relative to their exposure and subsequent expression of viral load, determines the proportion of (non-household) secondary infections that are potentially abated, subject to their isolation compliance (Equation 1).

1.3.1 Cases

Several academic studies have investigated the relationship between the timing of viral load, symptom onset, test sensitivity and relative levels of infectiousness (Ashcroft, and others, 2020) (Ferretti, and others, 2020) (He, and others, 2020) (Hellewell, and others, 2021). In general, they have shown that the detection of cases either through symptom onset or by asymptomatic testing is highly correlated to the expression of high viral load, which typically occurs following an incubation period of few days, during which individuals are still likely to be infectious. Overall, studies report a huge uncertainty in the timing of case detection, the potential delay before subsequent isolation, and the relative amount of infectiousness that is potentially abated. Therefore, for simplicity, in this report we have elected to use the same assumptions for all primary case isolations. We assume that on average 50 to 70% of transmission occurs prior to isolation; hence, case detection and (fully compliant) isolation will abate 30 to 50% of all secondary transmissions (central value 40%).

1.3.2 Contacts

For contacts we explicitly estimate the timing of their isolation relative to infection and use this to determine the proportion of infectiousness abated by isolation.

For CTAS contacts we define our central estimate for the timing of isolation as the difference between the registered contact notification date to the assumed exposure date[3].

[3] We note that some contact isolations could occur before CTAS notifications, particularly where contacts are in the same household as someone with symptoms or a positive test result. For consistency in this study we assume that contacts will only isolate effectively from the time of notification but recognise that the impact on transmission reduction would be higher if we brought forward the contact isolation time.

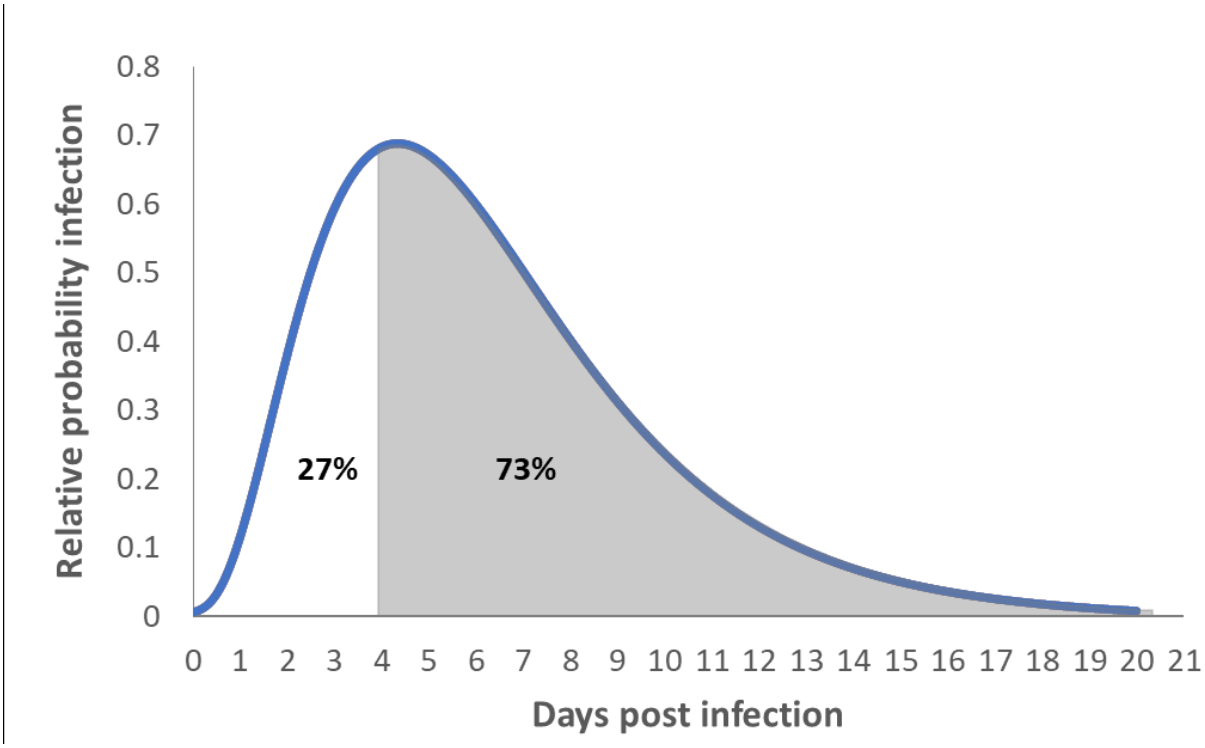
In contact tracing, positive cases are invited to list contacts that occurred over a two-day window prior to their diagnosis; this is either the time of symptom onset for symptomatic cases, or the time of the test for asymptomatic cases. The average exposure time is therefore assumed to be one-day prior to those respective timings.

For the COVID-19 App an aggregated dataset of exposure times and notification receipts are used to estimate the average timing of isolation (see annex Figure 15 for a comparison of CTAS contact and COVID-19 App notification times).

For school-age isolations the timing relative to exposure time is highly uncertain. We assume that these all occur 2 to 4 days after exposure.

We use an average infectiousness curve from the time of exposure – derived by (Ferretti, and others, 2020) – to estimate the proportion of secondary infections occurring over the time course of an infection (see illustrative example Figure 2). We use the mean day of isolation derived for all CTAS contacts or App contacts falling within each 14-day period in order to estimate the average proportion of infections potentially abated. We use a normal distribution in our sampling, which represents the uncertainty in the mean value (see parameter Table 1). We note that the real distribution of notification timings may be highly skewed; here, we are simplifying with a mean-field approximation.

Figure 2. Average infectiousness curve taken from (Ferretti, et al., 2020)



The figure illustrates the impact of a contact isolation occurring on day 4 after exposure. This would potentially prevent 73% of secondary (non-household) infections as represented by the shaded region.

1.4 The rate of compliance to isolation

In our model we modify the potential impact of each isolation to account for the rate of compliance (Equation 1). The compliance in this context represents the average reduction in the rate of (non-household) transmission occurring from the point of isolation.

Several behavioural studies and surveys have investigated people's compliance to the government's isolation rules. A recent ONS survey (Office for National Statistics, 2021b) reports that 86% of those required to self-isolate as a result of a positive test reported fully adhering to the requirements throughout their self-isolation period. The ONS also reported that 87% of reached contacts adhere to isolation requirements after being in contact with a positive case (Office for National Statistics, 2021d). Similar levels for case and contact compliance have been reported by the ONS throughout the past year.

In contrast, (Smith, and others, 2021) report that across all waves, among those with symptoms, adherence to full self-isolation was 42.5%, with 18.0% requesting a test for COVID-19.

At the upper end of the scale it is likely that results are biased by those most willing to engage with NHSTT. Conversely, the lower estimates include people that never engage at all.

For this study we adopt a central value of 80% for all cases, based on the assumption that they have engaged directly with NHSTT and will have levels of compliance closely represented by the ONS survey data. We use a central value slightly lower than the ONS results to acknowledge the fact that there may still be some bias in the survey response. We model uncertainty over a range of approximately 70 to 90%.

The value reported by (Smith, and others, 2021) comprises a mix of those who test and those who never engage with NHSTT. We therefore assume that those with symptoms who never engage with NHSTT have a much lower level of compliance of 20% (approximate range 10 to 30%).

For contacts, we assume that they will comprise a more even mix of people who responded across both surveys, with high and low levels of engagement. We therefore use a central value between (Office for National Statistics, 2021d) and (Smith, and others, 2021) of 60% (approximate range 50 to 70%).

1.5 Self-isolation from symptoms

In our model we account for the proportion of individuals with symptoms who do not engage with NHSTT (that is never take a test) but still isolate with a relatively low level of compliance. We estimate the size of this population by considering the overall proportion of all infected individuals that express symptoms (symptomatic rate) and removing the

proportion of those potentially isolating (as defined by the set of deduplicated cases or contacts in our final dataset).

$$\begin{aligned}\text{Symptomatic Only Population} \\ &= (\text{Total number of infectious indviduals} \times \text{Symptomatic Rate}) \\ &- \text{Cases detected by Symptomatic PCR not previously Contacts} \\ &- (\text{Cases detected by LFDs}^{4} \text{ not previously Contacts} \times \text{Symptomatic Rate}) \\ &- (\text{Infectious Contacts}^{5} \text{ not previouisly Cases} \times \text{Symptomatic Rate})\end{aligned}$$

Equation 2

1.6 Defining the counterfactual

In order to determine a range for the marginal impact of NHSTT, we compare the transmission reduction from TTI to an imagined counterfactual where there is no test or trace system in place. Instead we assume that there is a government policy that advises self-isolation on symptoms, as well as the isolation of all household members. We assume that all other factors remain equal over time and that there is no long-term impact on Rt or prevalence outside of each 14-day time window.

We constructed the counterfactual based on the principle that everyone who in reality tested with symptoms, would still isolate in the absence of a test (noting that a very large number of symptomatic tests are negative for COVID-19). We assume that those individuals isolate with the same level of compliance as for positive cases in our main assumptions. We further assume that symptomatic cases would also encourage household isolation at the same rate as for all tested individuals in the main dataset, also with the same rate of (contact) compliance. For consistency, we assume that household contacts in the counterfactual will isolate at the same average time after exposure as those in the main data[6]. We derive the rate of household isolation over time by taking the ratio of contacts to cases in each 14-day time period. Finally, we also allow any remaining symptomatic cases to isolate with the same low level of compliance assumed in the TTI

[4] Note that although current policy is for symptomatic individuals to take a PCR test rather than an LFD, it is unclear precisely what proportion of positive LFD results represent potentially symptomatic cases. In this calculation, we therefore assume that LFD cases have the same symptomatic rate as the wider population. PCR tests are assumed to be correctly registered as either symptomatic or asymptomatic.

[5] The number of infectious contacts is derived in Equation 12

[6] The same potential bias will occur in the counterfactual as in the main data, where it is likely that some household contacts would potentially isolate sooner, once symptoms are detected.

model. The counterfactual population is summarised in Equation 3. We assume the same SARs as for the TTI model.

$$\text{Counterfactual Cases} = \text{Cases detected by Symptomatic PCR not previously Contacts}$$

$$\text{Counterfactual Contacts} = \text{Counterfactual Cases} \times \frac{\text{Household Contacts not previously Cases}}{\text{All traced Cases}^{7} \text{ not previously Contacts}}$$

$$\begin{aligned}\text{Counterfactual Symptomatic Only Population} \\ &= (\text{Total number of infectious indviduals} \times \text{Symptomatic Rate}) \\ &- \text{Cases detected by Symptomatic PCR not previously Contacts} \\ &- (\text{Infectious}^{8} \text{ Counterfactual Contacts} \times \text{Symptomatic Rate})\end{aligned}$$

Equation 3

We also note that in this counterfactual there would be significantly more people isolating who are not infected; this could impact the transmission rate in a way that we have not modelled in this study.

1.7 Impact on the reproduction number (Rt)

We estimate the impact on the reproduction number Rt according to the following set of equations. The Rt value observed in each historical time period ($Rt_{Observed}$) is based on the ranges estimated in (Department of Health and Social Care, 2021c). We estimate Rt without TTI and then use this to calculate Rt in the counterfactual. We consider both the reduction in the observed value of Rt compared to either Rt without TTI, and Rt under the counterfactual (see Equations 4 to 7). We assume that the marginal impact of NHSTT will lie within this range.

$$Rt_{\text{Without TTI}} = \frac{Rt_{\text{Observed}}}{1 - \%\ \text{Transmission reduction from TTI}}$$

Equation 4

[7] The denominator includes all Pillar 1and2 cases not previously contacts, except self-serve LFDs without confirmatory PCR, who are not traced by CTAS.

[8] The calculation of infectious contacts is defined in Equation 12.

$$\mathrm{Rt}_{\mathrm{WithCounterfactual}} = \mathrm{Rt}_{\mathrm{Without\ TTI}}(1 - \%\ \mathrm{Transmission\ reduction\ from\ CounterFactual})$$

Equation 5

$$\mathrm{Rt}_{\mathrm{Reduction\ from\ TTI}} = \mathrm{Rt}_{\mathrm{Without\ TTI}} - \mathrm{Rt}_{\mathrm{Observed}}$$

Equation 6

$$\mathrm{Rt}_{\mathrm{Reduction\ above\ counterfactual}} = \mathrm{Rt}_{\mathrm{WithCounterfactual}} - \mathrm{Rt}_{\mathrm{Observed}}$$

Equation 7

1.8 Secondary cases prevented

We use our estimates for Rt reduction to estimate the secondary cases directly prevented by isolations over each 14-day period. This estimate is based on the simple assumption that the infectious population will approximately scale in size, over each generation of infections, with Rt (Equation 8).

$$\mathrm{Case\ reduction} = \mathrm{Total\ number\ of\ infectious\ indviduals} \times \mathrm{Rt}_{\mathrm{Reduction}}$$

Equation 8

We use the case reduction to estimate equivalent reductions in COVID-19 related hospitalisations and deaths based on the average rates observed over the time course of our study.

Importantly, these estimates do not consider any onward chains of transmission beyond each 14-day period. They should only ever be treated as a highly simplified indication of the direct impact of TTI within each discrete time window. They should not be used as a direct measure of the value-for-money of the system because they significantly underestimate the longer-term impact.

1.9 Uncertainty analysis

To determine the uncertainty in our outputs we use a simple Monte Carlo sampling method. For each parameter we define a prior distribution (see parameter table in section 2). For most parameters, we have taken a simplified approach and assumed a normal distribution that approximates the uncertainty seen in source datasets or from multiple sources. We set the standard deviation to approximately 50% of the confidence interval range above or below the central mean estimate (designed so that 2 standard deviations represent around 95% of the prior distribution).

All individual parameters are treated as uncorrelated. For each parameter that changes over time we assume that the uncertainty over the time series will be perfectly correlated;

we define the standard deviation to be a proportion of the mean so that for each sampled model run all time varying values of each parameter will be shifted the same relative distance from their central estimate.

In each model run we randomly sample from the distribution of parameters and calculate the final transmission reduction and impact on Rt. We repeat this process 10,000 times to construct an estimate of the output distribution.

In addition to the Monte Carlo sampling, we conducted a sensitivity test of each parameter, by changing the value by plus or minus 1 or 2 standard deviations whilst holding all other parameters constant at the central value. Here, we report the impact on the mean reduction in Rt from TTI, averaged over the full period of the study.

1.10 Total number of isolations

We calculate the total number of people notified to isolate in each 14-day time window and compare this to an estimate for the number of full isolations required in the counterfactual. We assume that those eligible to isolate include all deduplicated cases (including self-serve LFDs) and all contacts (including App and Schools). We do not factor in compliance or include symptomatic individuals who are not engaged with NHSTT. The total number is directly equivalent to our full deduplicated dataset (as defined in annex A.9).

We compare this figure to the counterfactual scenario where we assume that everyone who took a symptomatic PCR test for COVID-like symptoms would isolate as well as their household contacts (Equation 9). We used PCR test results (NPEX data) to identify the number of negative tests taken. To account for the estimated level of duplication among individuals taking multiple tests, we assumed the same rate as for positive tests (who we are able to deduplicate directly from their CTAS identifier). We then further scaled down the number of negative symptomatic individuals by the same proportion as for positive cases that were previously contacts.

In each 14-day window;

$$\text{Symptomatic individuals} = \text{Symptomatic PCR Cases not previously Contacts} + \text{Negative symptomatic PCR Tests} \times \left(\frac{\text{Symptomatic PCR Cases not previously Contacts}}{\text{Positive Symptomatic PCR Tests}}\right)$$

$$\text{Counterfactual Estimated Isolations} = \text{Symptomatic individuals} \times \left(1 + \frac{\text{Household Contacts not previously Cases}}{\text{All traced Cases not previously Contacts}}\right)$$

Equation 9

2. Parameter values

Table 1 summarises all the key parameters used in the model. In all cases we have indicated the distribution used for the Monte Carlo sampling and the rationale for our choices. The standard deviations are chosen to approximate the uncertainty seen in source datasets or from multiple sources. We ensure all rate parameters fall between 0 and 100% in all our sampling. For time varying parameters the standard deviation is defined as a percentage (see Monte Carlo methods). Table 2 shows the central estimate in each time period.

Table 1

Parameter Name	Description	Central Estimate	Distribution used for Monte Carlo model	Rationale and sources
LFD PPV	The rate of true positive cases among all LFD test results (without a confirmatory PCR)	See time series	No variance applied	We assume a specificity of 99.97% (Department of Health and Social Care, 2021a). This is used together with positivity rates in published LFD figures (NHS Test and Trace, 2021b) to calculate a PPV.
Rt_observed	The reproduction rate of COVID-19 in England.	See time series	Normal Distribution (mean=time series, SD=5%)	Derived from (Department of Health and Social Care, 2021c). Where weekly data does not directly align to our 14-day time periods we have adjusted based on a rolling average.
Infectious population	Estimated total population of infectious individuals in England.	See time series	Normal Distribution (mean=time series, SD=5%)	Derived from (Office for National Statistics, 2021a) incidence estimates.
Symptomatic rate	Proportion of all COVID-19 cases that express symptoms	See time series	Normal Distribution (mean=time series, SD=5%)	Based on (Office for National Statistics, 2021c). Where a monthly value is provided by ONS it is used directly. Adjacent time periods with no data use the same value as the most recent month.
Secondary attack rate	Proportion of contacts that will become infected by a case. Separate (uncorrelated) values used for contacts among the same household or non-household.	See time series	Normal Distribution (mean=time series, SD=10%)	Derived primarily from (Hakki S and ATACCC team, 2021). See Annex for methods.
Infectiousness abated: Symptomatic cases	Average proportion of infectiousness remaining after the time of isolation of symptomatic cases	40%	Normal Distribution (mean=0.4, SD=0.05)	The range used in this study is representative of several academic studies (see methods). We assume a single value in each Monte Carlo simulation. The distribution here represents the uncertainty in the mean value.

Parameter Name	Description	Central Estimate	Distribution used for Monte Carlo model	Rationale and sources
Infectiousness abated: Asymptomatic cases	Average proportion of infectiousness remaining after the time of isolation of asymptomatic cases	40%	Normal Distribution (mean=0.4, SD=0.05)	As above
Time of contact isolation	Average time of contact isolation relative to exposure time	Mean number of days derived from data	Normal Distribution (mean=time series, SD=0.5 days)	We derive the average day of contact isolation from the data in each 14-day period (see methods), which we use to sample from an average infectiousness curve to estimate transmission abated. The distribution represents uncertainty in the mean (not the individual level distribution of contact times).
Case isolation compliance	The average transmission reduction from the isolation of individuals receiving a positive test result.	80%	Normal distribution (mean = 0.80, SD=0.05)	For cases we assume a level of compliance slightly lower than the ONS survey data (Office for National Statistics, 2021b)(see methods).
Contact isolation compliance	The average transmission reduction from the isolation of individuals receiving a contact notification.	60%	Normal Distribution (mean=0.60, SD=0.05)	For contacts we assume a level of compliance falling between the (Office for National Statistics, 2021d) and (Smith, and others, 2021) (see methods).
Symptom onset isolation compliance (no engagement with NHSTT)	The average transmission reduction from the isolation of individuals who have no contact with NHSTT but express symptoms.	20%	Normal Distribution (mean=0.20, SD=0.05)	We assume a low level of compliance broadly reflecting the proportion of those who do not engage with NHSTT represented in (Smith, and others, 2021)
Additional contacts identified by the app	The number of additional exposure notifications, not already identified in CTAS, sent by the COVID-19 app as a result of a positive test being recorded in the app and contacts consenting to be traced.	See time series	Normal Distribution (mean=time series, SD=20%)	Derived from App data, assumed to overlap partially with CTAS. Additional App contacts are assumed to be non-household with equivalent SAR. See Annex A.5 for details.
Additional school-age contacts	The additional number of school-aged children identified from DFE school absentee data that are not already identified in CTAS.	See time series	No variance	Comparison of published DFE data with CTAS. School-aged contacts are assumed to have a lower SAR. See Annex A.7 for details.
School-age secondary attack rate	SAR assumed for the additional contacts identified from DFE data.	10.0%	Normal Distribution (mean=0.1, SD=0.025)	We assume a fixed SAR over the time course of our study consistent with the average non-household SAR. See Annex A.8 for details.
School-age contact timing	Timing of average school-age contact isolation relative to exposure time.	3 days	Normal Distribution (mean=3, SD=0.5)	Assumed to be towards the lower range of the CTAS timing. The distribution represents uncertainty in the mean (not the individual level distribution of contact times)

Table 2. Time series parameters

	1 Jun 20	15 Jun 20	29 Jun 20	13 Jul 20	27 Jul 20	10 Aug 20	24 Aug 20	07 Sep 20	21 Sep 20	05 Oct 20	19 Oct 20	02 Nov 20	16 Nov 20	30 Nov 20	14 Dec 20	28 Dec 20	11 Jan 21	25 Jan 21	08 Feb 21	22 Feb 21	08 Mar 21	22 Mar 21	05 Apr 21	19 Apr 21
Infectious population ('000s)	66	47	25	35	57	39	31	83	161	391	665	659	533	509	989	1288	964	747	339	155	156	195	114	46
Rt_observed	0.846	0.829	0.846	0.900	0.900	0.911	0.975	1.100	1.346	1.354	1.279	1.189	1.057	0.914	1.132	1.250	1.150	0.854	0.800	0.800	0.750	0.871	0.889	0.967
Symptomatic rate	0.547	0.547	0.547	0.547	0.547	0.547	0.547	0.547	0.547	0.547	0.547	0.547	0.547	0.547	0.547	0.547	0.609	0.609	0.592	0.592	0.498	0.498	0.526	0.526
LFD PPV	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	0.954	0.937	0.952	0.986	0.984	0.958	0.914	0.890	0.752	0.760	0.780	0.716	0.739
Household SAR	0.296	0.296	0.296	0.296	0.296	0.296	0.296	0.296	0.296	0.302	0.323	0.410	0.452	0.467	0.474	0.474	0.474	0.450	0.423	0.411	0.382	0.356	0.342	0.338
Non-household SAR	0.103	0.103	0.103	0.103	0.103	0.103	0.103	0.103	0.103	0.104	0.108	0.124	0.131	0.134	0.135	0.135	0.135	0.115	0.092	0.070	0.046	0.040	0.035	0.034
Additional App contacts ('000s)	0	0	0	0	0	0	0	0	0	43.3	118.2	304.4	102.8	110.0	406.8	307.5	108.7	NA	NA	NA	NA	NA	NA	NA
Additional school-age contacts ('000s)	0	0	0	0	0	0	0	0	0	237.5	282.5	351.7	500.6	359.1	332.5	0	0	0	0	0	102.7	138.3	0	39.9

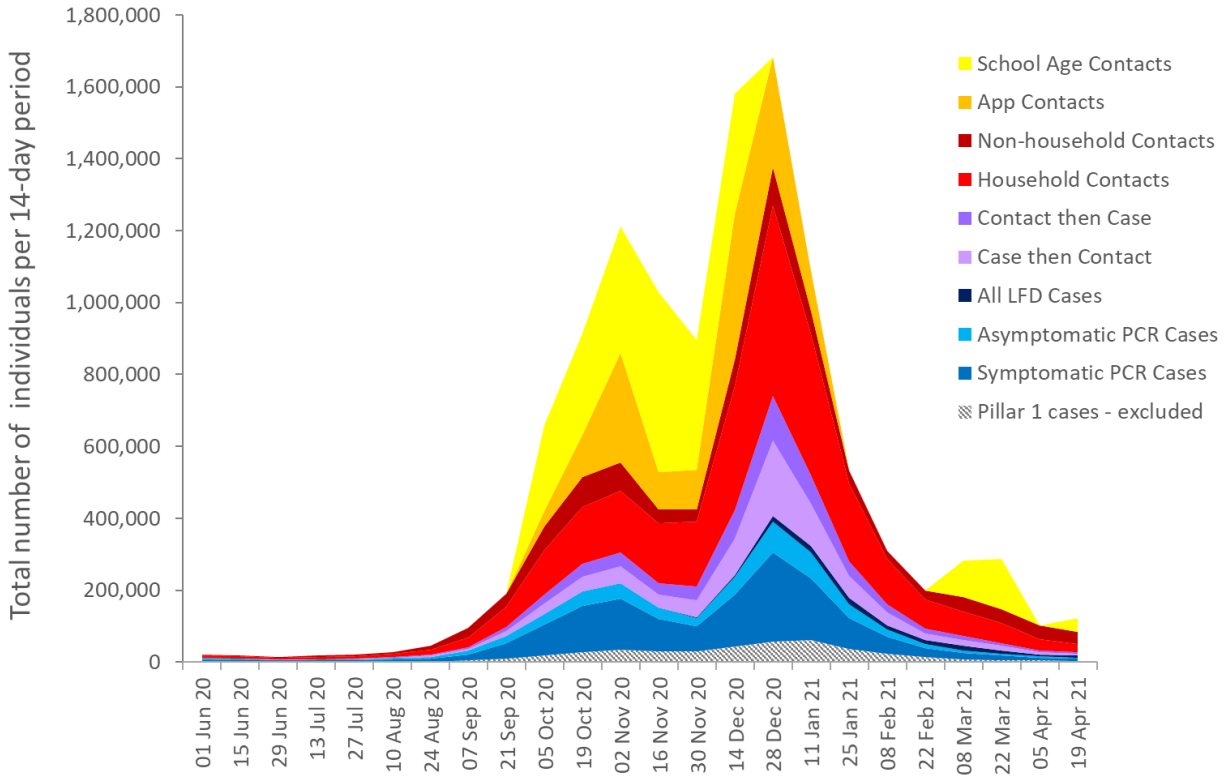
Listed are the central assumptions used for time varying parameters. The uncertainty ranges are defined in Table 1. The dates represent the start of each 14-day study period.

3. Results

3.1 Deduplicated cases and contacts

Figure 3 shows the complete data set used in this study to determine transmission reduction. This comprises all unique Pillar 2 cases (from CTAS and NPEX data) and all unique contacts (from CTAS, COVID-19 App and Schools data). This is the total number of people who we assume have been notified to isolate. The impact of Pillar 1 cases (also shown for reference in this figure) is excluded.

Figure 3. The consolidated data set of unique individuals falling into each category of case or contact in each 14-day time period extending from the dates shown

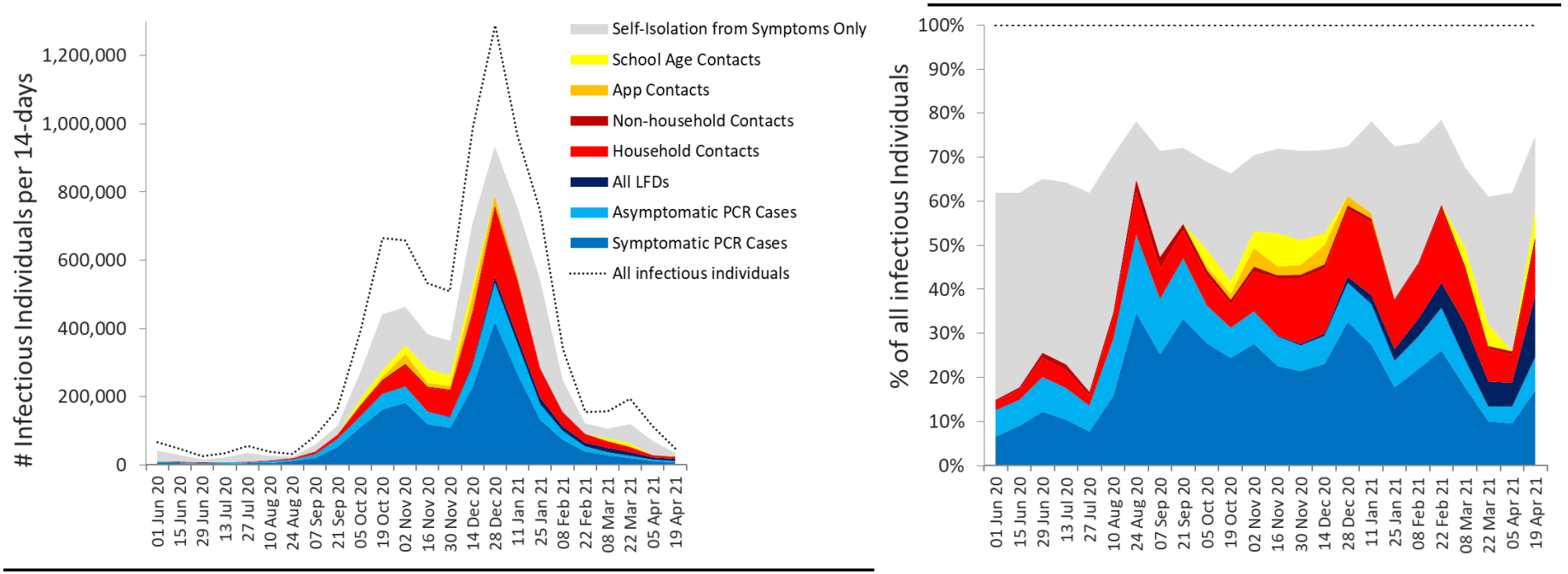


The purple regions show the overlap of cases and contacts. We ensure that all cases and contacts are unique within a 14-day window either side of their registration date. If an individual appears twice the first event is always counted, so that individuals first identified as cases that subsequently become contacts are counted as cases, and contacts that subsequently become cases are counted as contacts. Also shown are the Pillar 1 cases that we exclude on the assumption that they do not contribute towards transmission reduction. We do not include any COVID-19 App data from after January 2021.

3.2 The identification of infectious individuals

Figure 4 shows the breakdown of the population of infectious individuals (in our central assumptions). Cases identified by testing, tracing and symptom onset are compared to the total infectious population (derived directly from ONS incidence) in each 14-day period.

Figure 4. The identification of infectious individuals in the central estimate



The stack comprises the deduplicated cases identified by testing, contact tracing and the remaining population with symptoms in each 14-day time period. The cases (from Figure 3) have been adjusted to account for PPV. The proportion of infectious contacts is based on estimates of the secondary attack rates. On the righthand-side the numbers have been normalised relative to the total infectious population. Note, this figure does not consider isolation compliance; the grey region represents the entire symptomatic population not otherwise identified as a case or contact. The white region is therefore the remaining unidentified asymptomatic population.

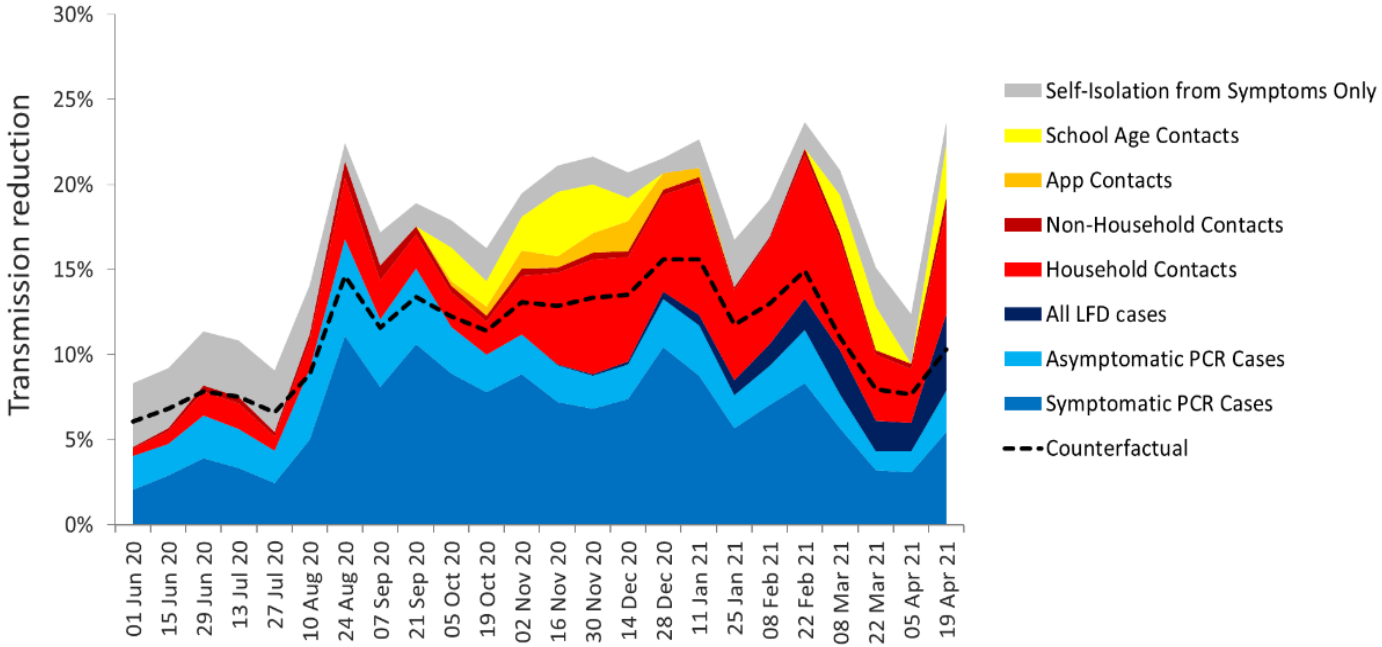
According to our central assumptions, since 10 August 2020, NHSTT identified around 25 to 65% of the total infectious population as either a case or contact. Positive tests first identified around 20 to 50% of all new cases and contact notifications around 5 to 25%.

Prior to August 2020, incidence rates were much lower, and the system was generally less impactful. There is a small rise in the incidence estimates in March 2021 that is not reflected in the number of cases detected by NHSTT; as a result, there is a lower overall identification rate in this period.

3.3 The percentage transmission reduction from TTI

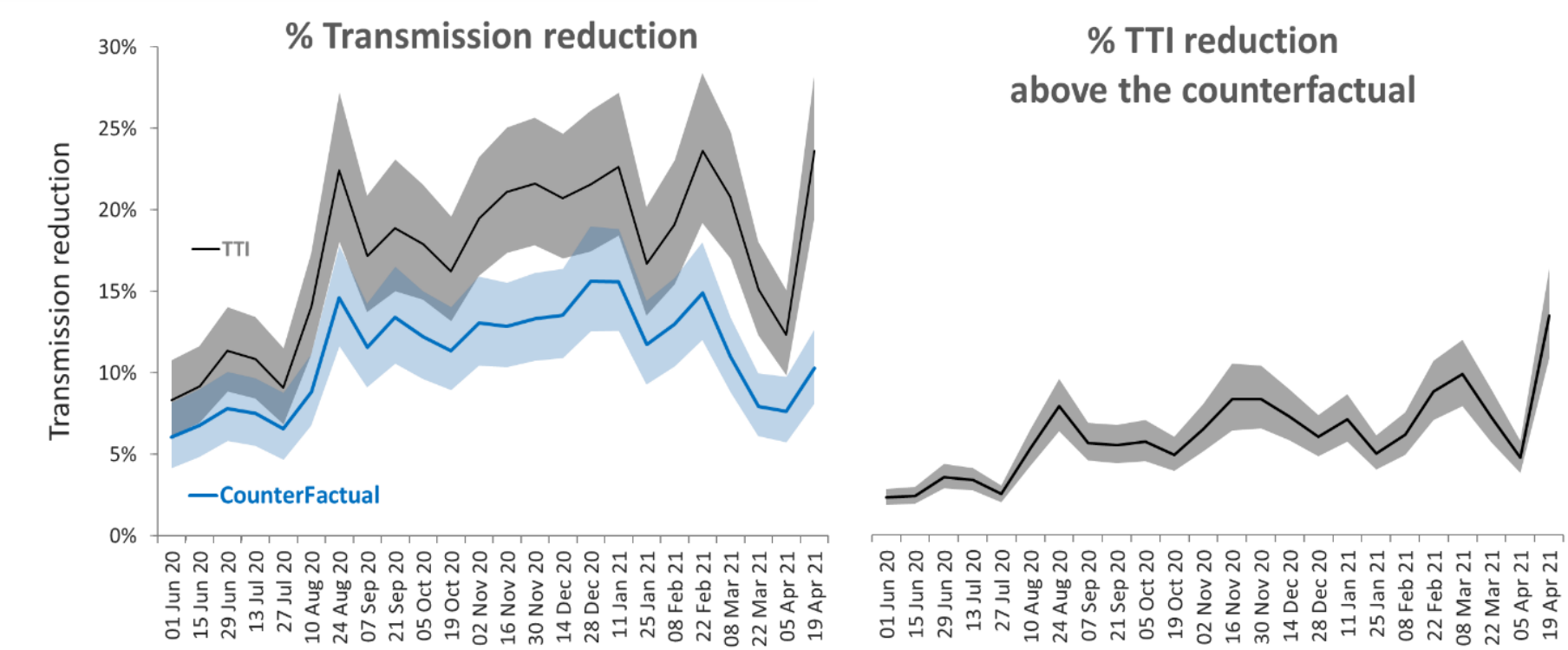
Figure 5 shows the breakdown in the total transmission reduction from TTI in our central estimate. Figure 6 compares the range in reduction from TTI and the counterfactual.

Figure 5. The central estimate for transmission reduction from TTI



The stacked bars show the contribution from cases, contacts and self-isolation (ranges are provided in annex A.12). This is compared to the counterfactual estimate (dashed line).

Figure 6. The plot on the left shows the range in the percentage transmission reduction from TTI (grey) and the counterfactual scenario (blue)



The plot on the right shows the range in the difference between them. Shaded regions show the 90% confidence interval derived from Monte Carlo sampling.

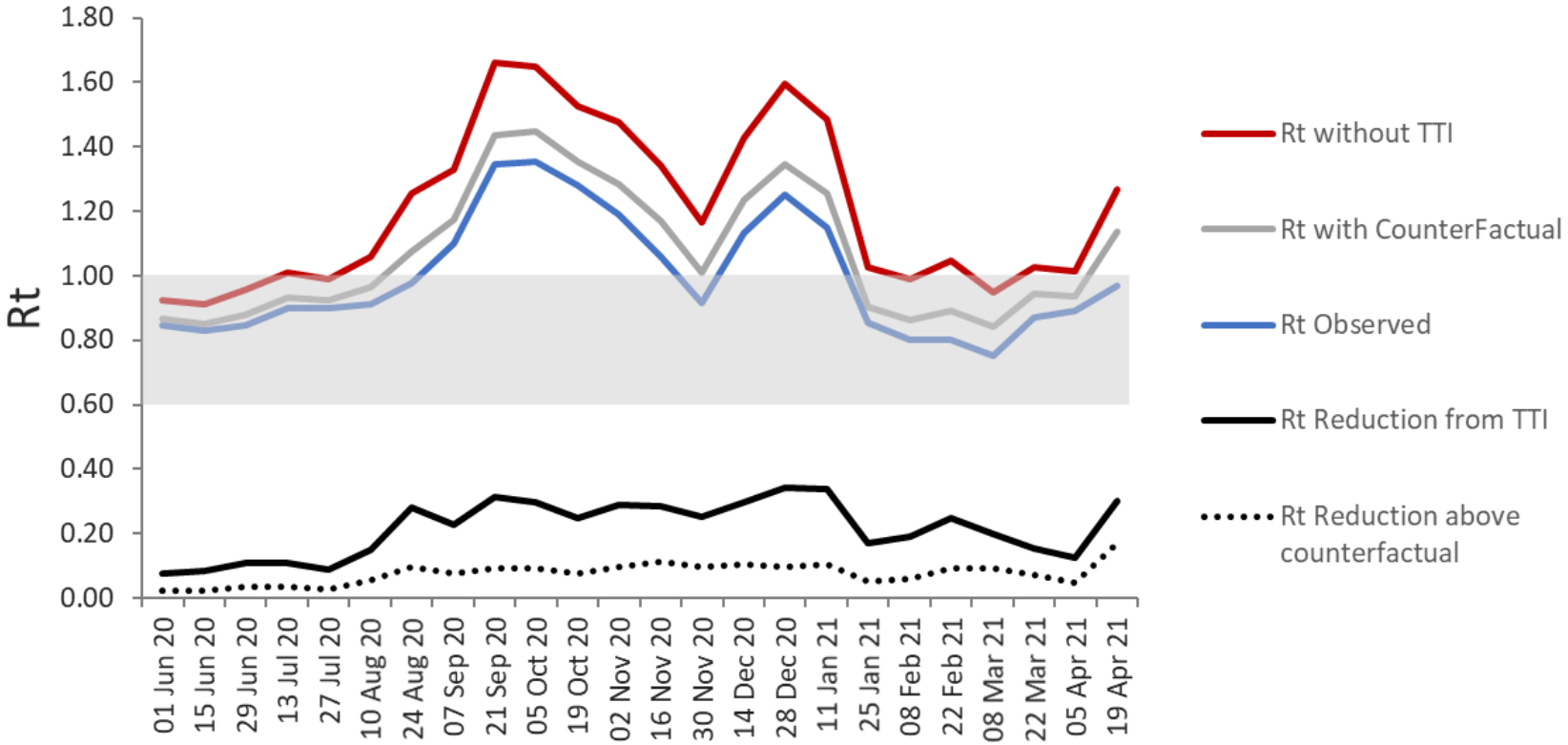
Since August 2020, the transmission reduction from TTI varied from 10 to 28% (over the 90% confidence interval derived from Monte Carlo sampling). In the counterfactual scenario, where there is no testing, but high levels of compliance to isolation with symptoms, we estimate that the transmission reduction would have varied from 6 to 19% since August 2020. In this time period, the amount of transmission reduction from TTI above the counterfactual varied from 4% to 16%.

In June and July 2020, the transmission reduction from TTI was lower (6 to 14%).

3.4 Impact on reproduction number, Rt

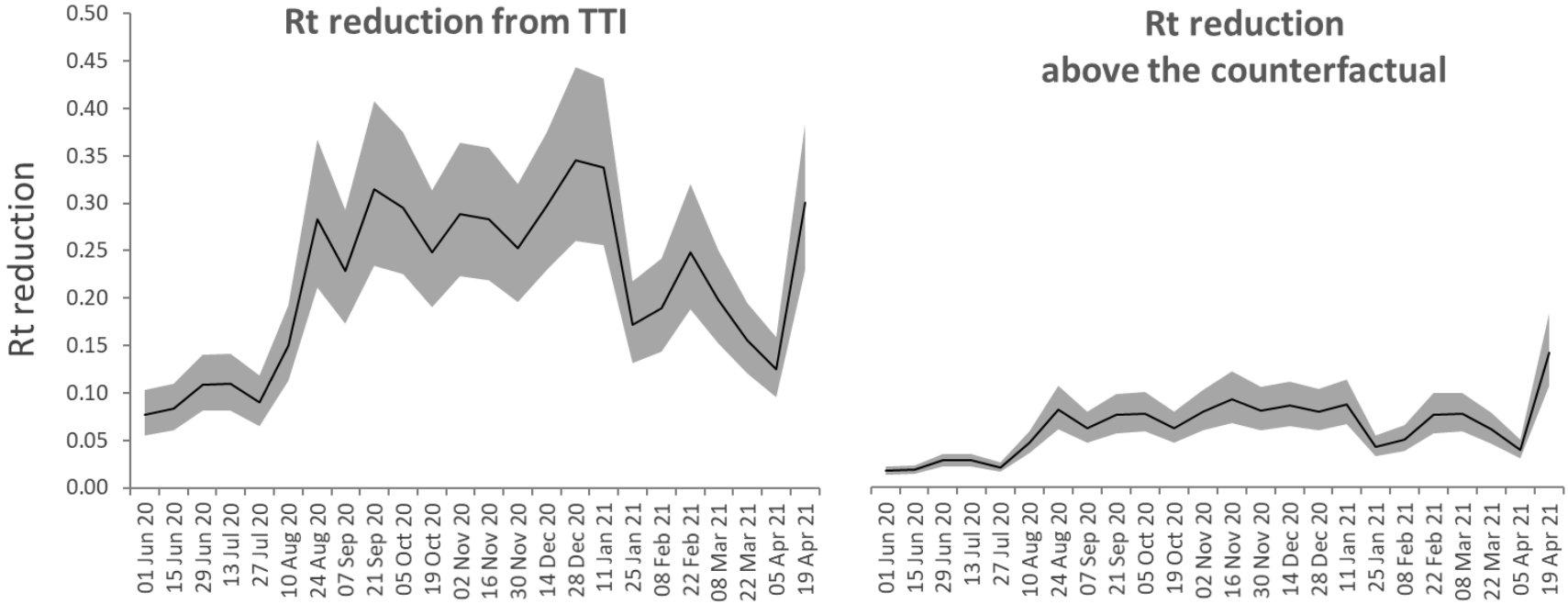
Figure 7 shows the central estimates for the impact on the reproduction number, Rt. Figure 8 shows the range in the uncertainty.

Figure 7. The central estimate for the impact on Rt in each 14-day period



The transmission reduction estimates are used to calculate Rt without TTI and then subsequently Rt under the counterfactual. The solid and dashed lines indicate the reduction in Rt that can be attributed to TTI (black line), and the level of TTI reduction that exceeds the reduction expected from the counterfactual alone (dashed line). The shaded region has been added to highlight periods where Rt has been brought below 1.

Figure 8. The range in Rt reduction from TTI and the component in excess of the counterfactual



Shaded regions show the 90% confidence interval.

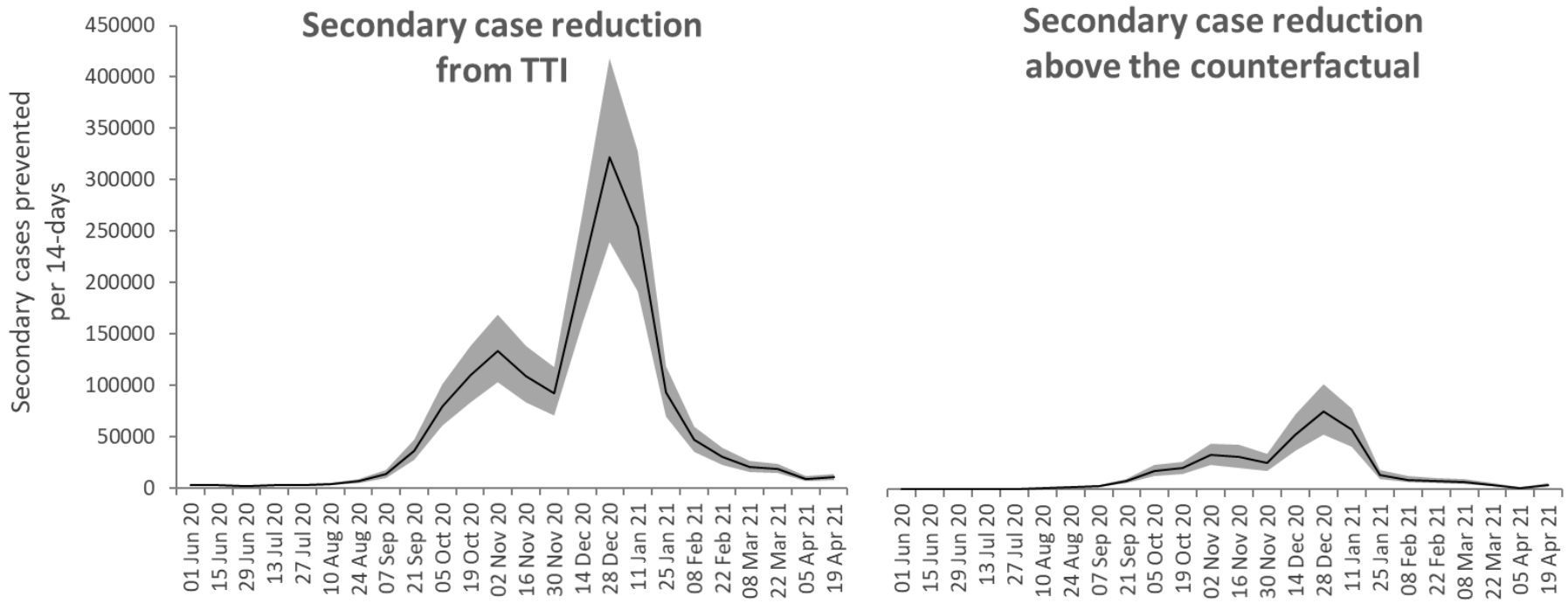
Since August 2020, we estimate that Rt reduction from TTI varied from 0.10 to 0.44. The reduction in Rt above the counterfactual varies from 0.04 to 0.22.

Our central estimate (Figure 7) shows that there are several periods where TTI has potentially brought Rt below 1 (August 2020, November 2020, January to April 2021). This could have been crucial in potentially reducing the duration and impact of lockdown.

3.5 Secondary case reduction

Figure 9 shows the secondary case reduction. These are infections directly prevented by isolations from TTI in each 14-day period. We estimate that over the study period, 1.2 to 2.0 million infections were directly prevented by TTI (0.3 to 0.5 million above the counterfactual). This does not take into account any onward chains of transmission. The biggest impact is at times of high incidence when relatively small reductions in Rt can still prevent large numbers of cases.

Figure 9. Range in secondary case reduction in each 14-day period from TTI and the number above the counterfactual



Derived from the total cases and reduction in Rt in each time window. Shaded regions show the 90% confidence interval.

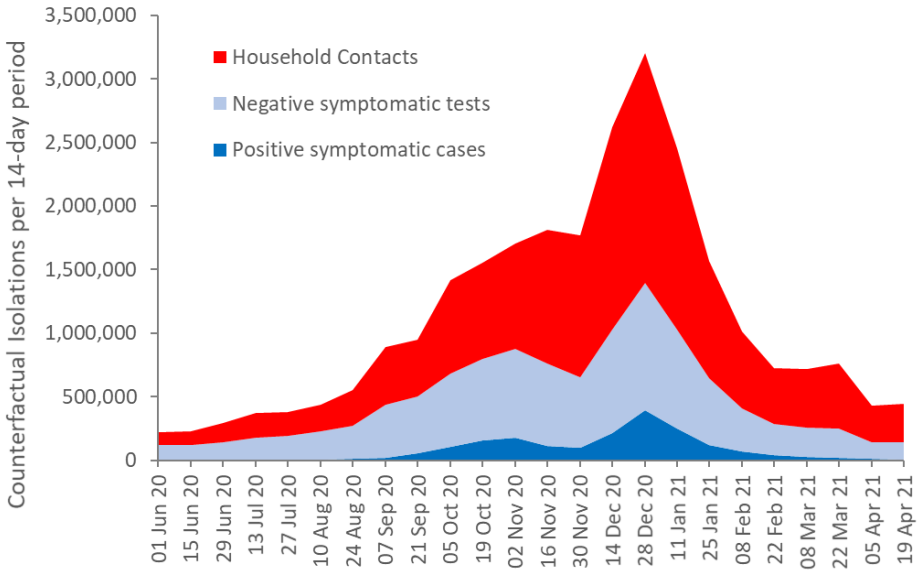
3.6 Estimating the number of isolations

The total number of isolation notifications in our main study data are plotted in Figure 3. These comprise all positive cases and their contacts. Through the whole study period (June 2020 to April 2021) these total 11 million.

Figure 10 shows the individuals isolating in the counterfactual scenario, where it is assumed that everyone who took a symptomatic PCR test would self-isolate together with their household for the full isolation period[9]. Over the course of our study the counterfactual isolations total 25 million. Therefore, an additional 16 million individuals are required to isolate in the counterfactual scenario, for the full isolation period.

Twenty-one million individuals identified in Figure 1 belong to households where individuals expressing symptoms would have received a negative test with NHSTT and therefore would have only had to isolate for a short time. In the counterfactual those individuals would have had to isolate for the full isolation period. Preventing isolations at this scale would potentially have had a huge economic benefit.

Figure 10. Isolations that are assumed to occur in the counterfactual scenario (based on central assumptions)



Isolations are based on all individuals who took a symptomatic PCR test over the course of the study period. In addition, we include an estimate of their household contacts.

[9] The full isolation period is 10 days from the day after exposure, a test or the start of symptoms. Prior to December 2020 this was 14 days.

4. Conclusion

We developed a modelling framework to evaluate the reduction in transmission from test, trace and self-isolation. This work gives a very high-level view of the impact of the whole system. It should not be used to directly evaluate the impact of specific use cases or system components. Testing and contact tracing may have prevented outbreaks in specific settings that are not considered in this analysis. More detailed work is needed for those types of evaluation.

We did not include the contribution from Pillar 1 testing in hospitals on the basis that this would not significantly impact community transmissions. However, this will have significant impact on reducing transmission within hospitals and ensuring that the correct treatments are given most effectively.

To help determine the marginal impact attributable to NHSTT we have compared the full impact of TTI to a hypothetical counterfactual scenario. In the counterfactual we have imagined that all individuals who tested with COVID-like-symptoms, self-isolated, without ever taking a test, together with their household contacts. It is extremely difficult to estimate how people would really behave without the ability to test. Our other compliance assumptions are based on behavioural survey data that is relatively robust. That showed quite high levels of isolation compliance among those who engaged with NHSTT. However, among those individuals, it is very likely that people would be much more reluctant to fully isolate without knowing they either have a positive test result or can rapidly get one.

We designed the counterfactual to maximise the transmission reduction from self-isolation in a world without TTI. The counterfactual compliance rates are set at the very upper limit of what is plausible. As such, the transmission reduction associated with TTI, above the counterfactual can be regarded as a lower limit for the marginal impact of NHSTT. The full impact of TTI sets an extreme upper limit. Notably, symptomatic case detection and household contact isolation remain in the counterfactual and are major contributors to transmission reduction. At the lower limit, any additional marginal impacts from NHSTT come exclusively from non-household contact tracing and asymptomatic testing.

Importantly NHSTT greatly reduces the number of full-term isolations that would have been required under the counterfactual, where there are no negative test results to release people from an initial period of self-isolation. This will potentially have had huge economic benefits. We note that more isolation in the counterfactual could equate to greater social distancing which itself could bring down the reproduction rate; however, we have not quantified this impact.

In this study we have only evaluated the impact on secondary infections that we can directly attribute to self-isolation. It is extremely complex to accurately model the impact on ongoing chains of transmission. This is largely because it is very hard to know how other parts of the system would have responded in the absence of NHSTT and therefore how the number of cases would have progressed over time. During phases of exponential growth, and high incidence, even very small reductions in Rt will prevent many cases. Our study has indicated several periods when TTI brought Rt below 1. This would have prevented exponential growth,

bringing incidence rates down, and will have helped to reduce the duration and economic impact of lock down and other social restrictions.

4.1 Comparison with other studies

The Rùm model previously estimated that TTI had reduced transmission by around 18 to 33% in October 2020 (Rt reduction of 0.3 to 0.6) (Department of Health and Social Care, 2021b). This is broadly consistent with our new estimate for that time period (TTI reduction in the range 13 to 22%, Rt reduction in the range 0.2 to 0.4).

A study by the Welsh government estimated that in Winter (outside of firebreaks) TTI reduced R from 1.3 to 0.8 (Welsh government, 2021). This is at the upper limit of our range. Notably, our study has slightly more pessimistic assumptions for contact isolation compliance and the timing of contact isolations (based on published surveys and NHSTT data respectively).

A previous study by (Wymant, and others, 2021) estimated the impact of the COVID-19 App using a combination of modelling and statistical techniques. Their core modelling assumptions are broadly consistent with ours; however, in addition they considered the impact on onward chains of transmission which we have not done in this study. From their different analyses they suggested that the App averted approximately 0.3 to 0.6 million future cases. In our study we estimate that the App notifications reduced transmission by an average of around 1% (see Table 4 Annex A.12) over a time period when we estimate there were approximately 6 million new infections (October to January). Under our modelling framework, we would predict that the App directly prevented approximately 0.1 million secondary infections; this estimate is lower than the range reported by (Wymant, and others, 2021) but would likely be much closer if we also accounted for onward chains of transmission in a similar way. Notably, we did not include App data from after January 2021 so this will be an underestimate of the impact over the full study period.

Other modelling work has predicted a wide range of impacts (Worden, and others, 2020) (Keeling, and others, 2020) (Kucharski, and others, 2020) (Kretzschmar, and others, 2020). This underlines the complexity and uncertainty involved in making these kinds of estimate.

The NHSTT data itself reveals that contact tracing is identifying a significant number of cases. To some extent this is direct evidence of the reach of the whole system. Figure 17 (in annex A.10) shows that since August, up to 45% of all cases were also identified as contacts. 10 to 20% were reached prior to recording a positive test; 1 to 5% of these were as non-household contacts. Importantly reaching infectious contacts before they are likely to test will help them to isolate sooner and prevent onward infections.

4.2 A review of the model framework

The modelling framework that we used makes some simplifying assumptions that were required to provide a high-level estimate of the TTI impact. Below we discuss the weaknesses and risks associated with some of these.

1. The relative rate of transmission and hence the reproduction rate (Rt) will scale in proportion to the number of infectious individuals not in isolation. The actual relationship between isolation and the reproduction rate is more complex. Perfect self-isolation would theoretically prevent all non-household transmission. It is less clear exactly how isolations affect transmissions within a household. The ultimate impact on the reproduction rate is potentially different depending on the relative rate of non-household mixing (that is the level of social distancing). We consider the approximation used in this framework to be reasonable for estimating the range in the impact of TTI on Rt. It should not be used to try to accurately predict the reproduction rate.

2. Rt and prevalence are relatively stable over each 14-day time period. This assumption enables us to compare the population of infectious individuals with cases and contacts (and hence expected isolations) over a 14-day time period. If the reproduction rate changes rapidly or is very far from 1 so that prevalence changes rapidly then this comparison becomes less valid. We note that in December 2020 to January 2021 there was a significant change in incidence and prevalence that could impact our estimates around that time.

3. Infectious and isolating individuals are evenly distributed among the population. There is strong evidence of very different rates of transmission among different individuals. Rates of social mixing are very different as are rates of compliance to social restrictions and compliance to testing and self-isolation. It is possible that those who engage with NHSTT have much lower rates of transmission than those who do not. In this case, isolations associated with NHSTT would have lower impact overall and this would potentially lead to an overestimate in our model.

4. The average rate of transmission among infectious individuals is the same, regardless of symptom expression or detectability. People who are identified as a result of symptom onset or a positive LFD may have significantly higher than average viral loads and higher levels of infectiousness. This would potentially bias our calculation for the impact of TTI towards a lower estimate.

Many of the parameters that we have used are highly uncertain and will vary considerably among different sections of the population. Our model uses mean-field approximations instead of explicitly considering the distributions among the population (for example for compliance

rates or isolation times). We have used central estimates around which we consider uncertainty. This could impact our results, particularly when distributions are highly skewed; however, it is very hard to accurately define precise distributions in most cases. We have tried to use the most robust, evidence-based central estimates that were available, with uncertainty factored into our modelling. Future investigations will aim to refine our understanding of these.

Where there was limited evidence, we have tried to make our assumptions deliberately conservative to avoid overestimating the impact of TTI. For example, we have assumed that household isolations are delayed until contact notifications occur (we maintained this delay in the counterfactual for consistency). It is likely that many household isolations could occur sooner, which would potentially increase our estimate for the overall impact.

We tested the sensitivity of our outputs to each of our assumptions. Table 3 in the annex summarises this analysis.

4.3 Future work

This report has presented a data driven framework to estimate the impact of NHSTT. In the future we hope to develop this model to further understand the impact of the system. Notably it will be particularly important to determine levels of isolation and different rates of transmission among different age groups, to account for vaccination and the Delta variant. We also plan to develop this work to consider the impact of the system in terms of reducing chains of transmission over the longer term.

5. References

Ashcroft P and others, 2020. COVID-19 infectivity profile correction Swiss medical Weekly

Department of Health and Social Care, 2021a. Lateral flow device specificity in phase 4 (post marketing) surveillance (viewed on 1 July 2021)

Department of Health and Social Care, 2021b. RUM model technical annex (viewed on 1 July 2021)

Department of Health and Social Care, 2021c. The R value and growth rate (viewed on 18 June 2021)

Explore education statistics, 2021. Attendance in education and early years settings during the coronavirus (COVID-19) outbreak; 2020 data; 2021 data (viewed on 23 June 2021)

Ferretti L and others, 2020. The timing of COVID-19 transmission medRxiv 2020.09.04.20188516

Hellewell J and others, 2021. Estimating the effectiveness of routine asymptomatic PCR testing at different frequencies for the detection of SARS-CoV-2 infections BioMed Central Medicine

He X and others, 2020. Temporal Dynamics in Viral Shedding and Transmissibility of COVID-19 Nature Medicine

Hakki S and ATACCC team[10], 2021. 'The Assessment of Transmission and Contagiousness of COVID-19 in Contacts (ATACCC) study' NIHR Health Protection Research Unit in Respiratory Infections, PHE and Imperial College London

Jarrom D and others, 2020. Effectiveness of tests to detect the presence of SARS-CoV-2 virus, and antibodies to SARS-CoV-2, to inform COVID-19 diagnosis: a rapid systematic review British Medical Journal Evidence-Based Medicine

Keeling M and others, 2020. Efficacy of contact tracing for the containment of the 2019 novel coronavirus (COVID-19). Journal of Epidemiology and Community Health

Kretzschmar M and others, 2020. Impact of delays on effectiveness of contact tracing strategies for COVID-19: a modelling study The Lancet: Public Health.

Kucharski, A and others, 2020. Effectiveness of isolation, testing, contact tracing, and physical distancing on reducing transmission of SARS-CoV-2 in different settings: a mathematical modelling study The Lancet: Infectious Diseases

NHS Test and Trace, 2021a. NHS Test and trace statistics (viewed on 1 July 2021)

NHS Test and Trace, 2021b. Coronavirus (COVID-19) Tests Conducted (viewed on 17 June 2021)

[10] S Hakki, ND Fernandez, A Koycheva, E Conibear, JL Barnett, J Fenn, KJ Madon, R Kundu, R Varro, SN Janakan, C Anderson, A Lackenby, M Zambon, J Dunning, A Lalvani

Office for National Statistics, 2021a. Coronavirus (COVID-19) Infection Survey, UK (viewed on 1 July 2021)

Office for National Statistics, 2021b. Coronavirus and self-isolation after testing positive in England: 10 May to 15 May 2021 (viewed on 2 June 2021)

Office for National Statistics, 2021c. ONS symptomatic rate estimate Coronavirus (COVID-19) Infection Survey (viewed on 1 July 2021)

Office for National Statistics, 2021d. Coronavirus and self-isolation after being in contact with a positive case in England: 1 to 5 June 2021 (viewed on 18 June 2021)

Public Health England, 2020. Investigation of novel SARS-CoV-2 (viewed on 1 July 2021)

Smith L and others, 2021. Adherence to the test, trace, and isolate system in the UK: results from 37 nationally representative surveys British Medical Journal

UK government, 2021. Coronavirus data dashboard (viewed on 1 July 2021)

Welsh government, 2021. Modelling the Current Welsh TTP (Test, Trace, Protect) System (viewed on 1 July 2021)

Worden L and others, 2020. Estimation of effects of contact tracing and mask adoption on COVID-19 transmission in San Francisco: a modelling study medRxiv 2020.06.09.20125831

Wymant C and others, 2021. The epidemiological impact of the NHS COVID-19 App Nature

Appendix

All the data underlying figures shown in this report can be found in the accompanying spreadsheet.

A.1 NHSTT data used in the study

The primary data set used in this study is provided by CTAS (Contact Tracing and Advisory Service). The CTAS data contains information for cases, registered with positive test results, and contacts whose details are provided by each case. CTAS includes all registered positive cases from PCR tests, supervised LFDs and self-served LFDs with a confirmatory PCR. Some remaining self-served LFD test results were derived from the NPEX (National Pathology Exchange) where we identified any unique specimen IDs not appearing in CTAS. We used negative symptomatic PCR test results directly from NPEX to determine the expected rates of isolation in the counterfactual scenario. We identify symptomatic PCR tests through a self-reported field recorded at the time a test is booked. The timing of symptom onset is recorded when cases are registered in CTAS.

We do not include any Pillar 1 cases in our transmission reduction calculations. We do include the contacts of Pillar 1 cases and Pillar 1 cases are also used to determine the overall proportion of contacts that become cases and the average number of contacts per case. Contacts that were previously identified as Pillar 1 cases are excluded.

We count individuals in CTAS as cases or contacts depending on which is registered first. We only count contacts that are reached, and discount those that have only been named by a case. We ensure that there is no duplication of individuals within a 14-day time window from the first registered case or contact date. We adjusted the number of additional self-served LFDs from NPEX by assuming that there will be the same rate of duplication (that is the percentage of cases that are first contacts) as we observed in the CTAS data in each 14-day time period.

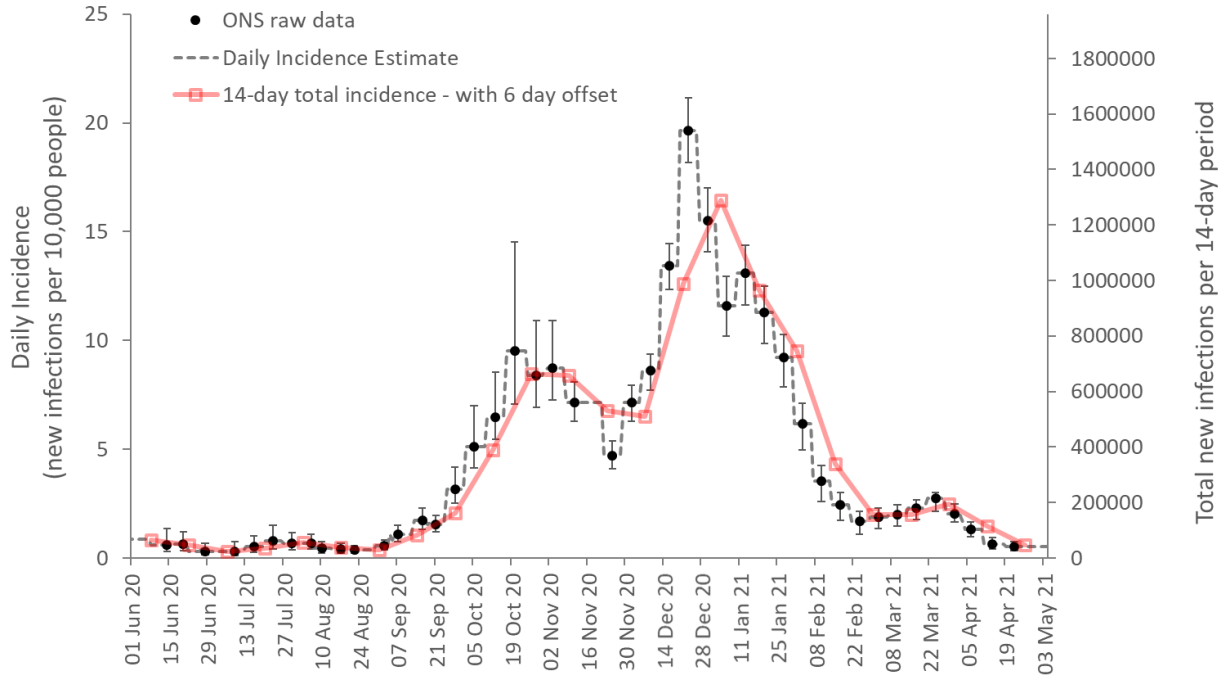
The consolidated data that we use is consistent with publicly available figures comprising all cases and contacts (NHS Test and Trace, 2021a). Our outputs (Figure 3) show slightly lower total numbers because of the stringent deduplication that we have applied.

We used data from the COVID-19 App and DFE Schools absence records to estimate the number of additional isolation notifications that were not accounted for in CTAS (see details below).

A.2 Calculating total incidence

We estimate the total number of infectious individuals in each discrete 14-day time period using the incidence rates provided by the ONS community infection survey (Office for National Statistics, 2021a) (Figure 11). We interpolated the ONS data to provide daily estimates of incidence over the study period; we assume a fixed value for each day over the ONS time periods with defined incidence rates, and where these overlapped, we took an average value. We then calculated the total number of new daily infections falling within a 14-day window, 6 days prior to each of the 14-day study periods in which we aggregate registered cases and contacts, to account for the average time delay between new infections (incidence) and case detection.

Figure 11. ONS incidence data is used to estimate the infectious population over time



Shown are the ONS incidence rate estimates for England – there are plotted with central estimates at the centre of each time period they represent. The vertical bars show the confidence intervals in that dataset. These are compared to our derived daily figures which are either constant over the ONS period or use an average where there was any overlap in ONS dates. The daily figures were then aggregated into a total estimate for each 14-day time period in our study, which also accounted for a 6-day offset (the study values are plotted here as red squares against the central date in each 14-day study period). The left y-axis shows the ONS rate per 10,000. The axis on the right shows the incidence rate converted to infections per fortnight assuming a population of 56 million. The daily rates are simply multiplied by 14 to align the 2 axes and show a direct comparison with our final study assumptions. Our final data point extends 2 days beyond the range of estimates from ONS; here we have assumed a constant rate will persist.

Notably, the ONS community infection survey does not identify cases occurring in residential settings such as care homes and prisons. For this study, we assume that the ONS incidence rates can be applied across the whole population (56 million).

A.3 Adjusting for false positive among LFDs

To ensure that we only count positive LFD cases, without a confirmatory PCR, that are genuinely infected, we estimated the positive predictive value (PPV) of LFD tests over time (Equation 10). We based this on our published data that shows the rate of LFD positivity from all recorded test results; Table 3 in (NHS Test and Trace, 2021b), and the LFD specificity; 99.97% as determined in (Department of Health and Social Care, 2021a). Results are shown in parameter Table 2.

$$\text{LFD postivity} = (\text{LFD positves})/(\text{All LFD results})$$

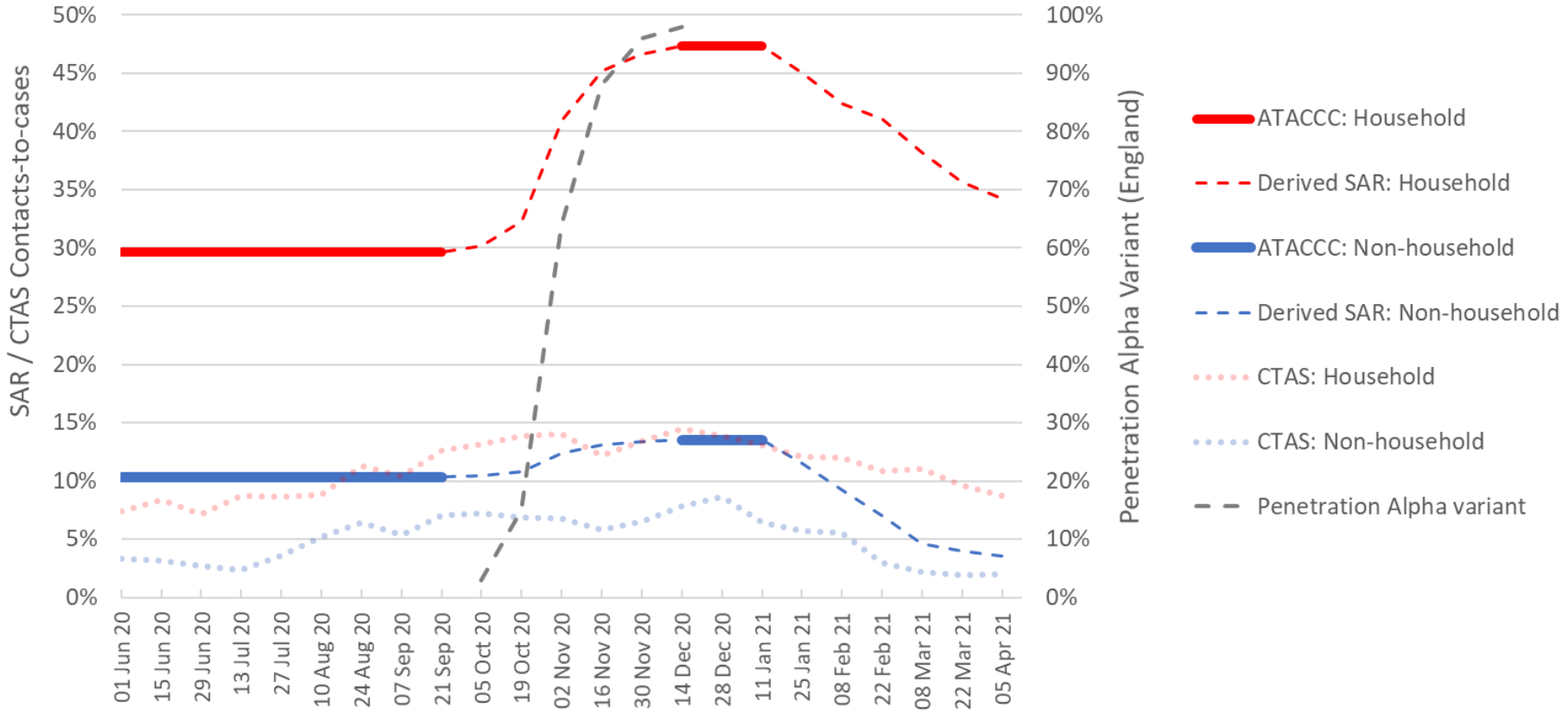
$$\text{PPV} = \frac{\text{LFD positvity} - (1 - \text{specificity})}{\text{LFD positivity}}$$

Equation 10

A.4 Household and non-household secondary attack rates (SAR)

In our modelling we rely on the secondary attack rate (SAR) to estimate the proportion of contacts that are infected. We use a combination of sources to estimate the time variation in the SAR. We start with the ATACCC study estimates for the household and non-household SAR (Hakki S and ATACCC team, 2021). These estimates are based on a sample of cases and their respective contacts who were tested over time. The SAR represents the percentage of a case's contacts that became infectious after exposure. The estimates from the study do not vary in time but are distinguished for the wild type (WT) and the B.1.1.7 (Alpha) variant. We interpolated over time between the estimate relating to the 2 variants in line with the England wide rate of penetration of B.1.1.7 (Public Health England, 2020). We further assume that there is a reduction over time as a result of vaccination since January 2021. We derive the rate of reduction in the overall SAR by scaling between January to April 2021 in proportion to the time series, derived from NHSTT CTAS data, for the number of contacts becoming case (see Figure 12).

Figure 12. Estimating the SAR for household and non-household contacts over time



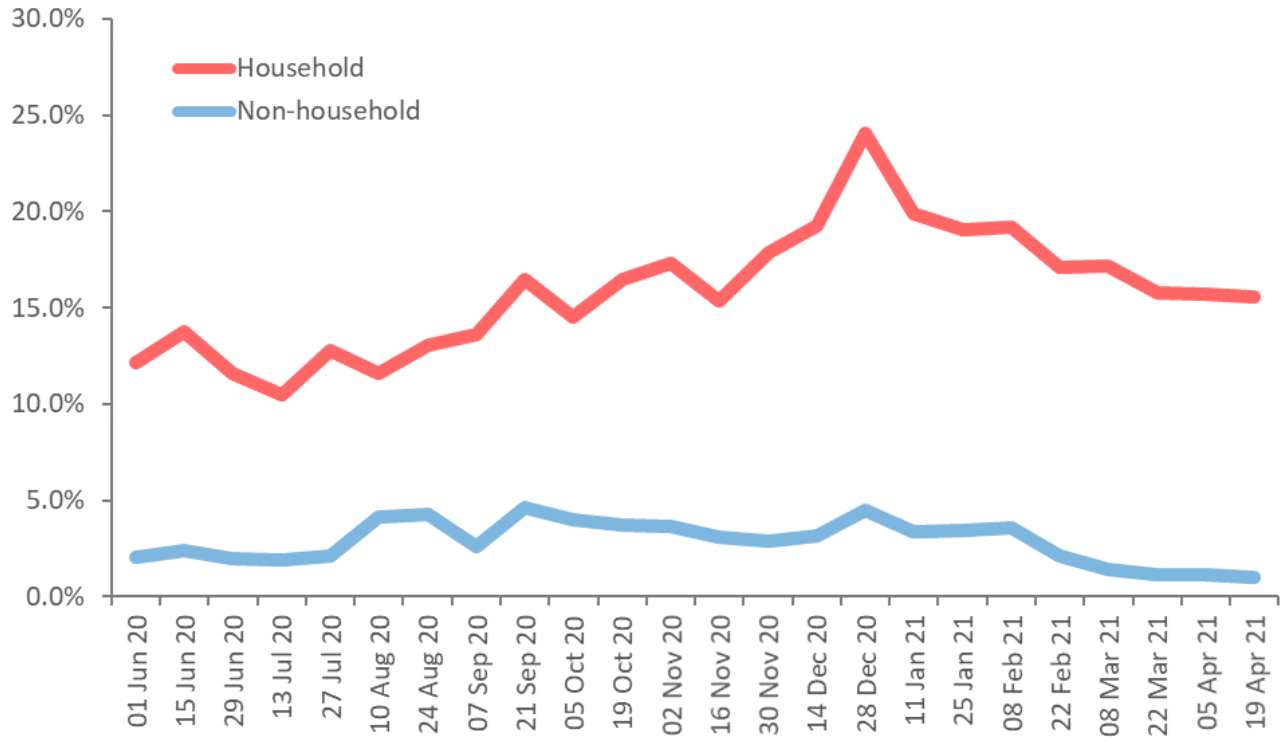
The solid parts of the lines indicate the ATACCC study central estimates for wild type and Alpha COVID-19 variants, for household (red) and non-household (blue) contacts. The interpolation between October 2020 and December 2021 is based on the proportion of cases among the England population for each variant (dashed grey line). The decrease after January 2021 is assumed to occur in proportion to the dotted red and blue lines, which show the percentage of contacts who become cases in the NHSTT CTAS data. CTAS cases are counted here if they occur within 14-days of the contact registration date.

Notably the ATACCC study estimates for the SAR are around 2 to 3 times higher than the contact-to-case rates in CTAS data. This ratio is broadly consistent with the rate of case detection that we observe in our study.

A.5 Adjusting the SAR to estimate the number of infectious contacts

The SAR gives us an estimate for the percentage of all secondary contacts that become infected per primary case. In this study we are only attributing isolations to contact notifications if, within a 14-day window before their contact registration, they had not previously been identified as a case (otherwise we count them as cases). Therefore, we make an adjustment to account for the contacts that we have removed from our dataset (see alpha term in Equation 12). Figure 13 shows the proportion of contacts that were previously cases, as is used in this adjustment.

Figure 13. The percentage of all reached household and non-household contacts that are notified after previously being registered as a case (up to 14-days prior to their contact registration time)



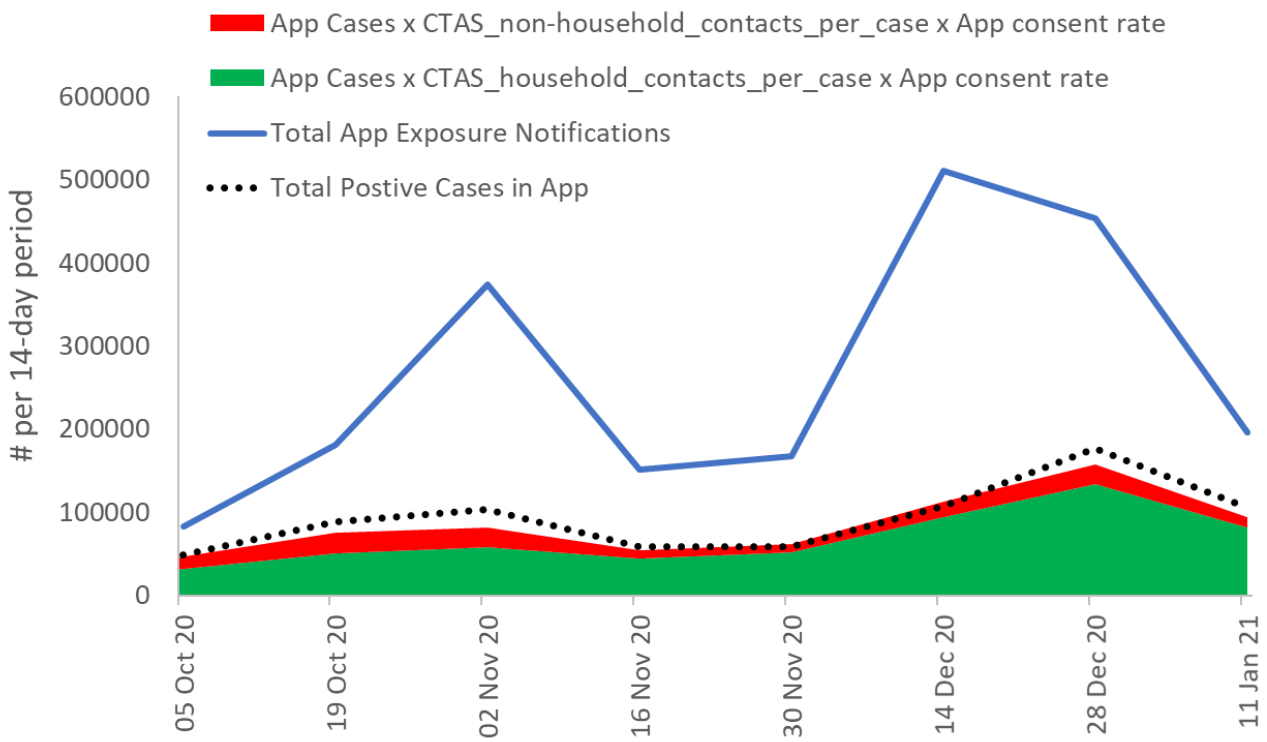
Data plotted is the percentage for total figures in each 14-day study period. We use these rates to adjust the SAR, in order to determine the rate of infectiousness among contacts in our study data.

A.6 Adding contacts from the COVID-19 App

We increase the final number of contacts to account for those reached by the COVID-19 App that are not already accounted for in the NHSTT CTAS data. We used data consistent with (Wymant, and others, 2021) which extends until Jan17th. After that time, we have not accounted for the impact of the App so the output presented here will be an underestimate.

The App data reports a total number of exposure notifications and a total number of positive cases (which represent a subset of all CTAS or NPEX registered cases). To ensure that we were not double counting contacts in CTAS and in the App, we calculated an *expected* number of household and non-household contacts, relating to the App cases, based on the average rates (contacts-per-case) derived from CTAS (as set out in Equation 3). Here we also account for the consent rate (the number of App cases who consent to have contacts notified). We assume that App exposure notifications overlap entirely with CTAS household contacts. We further assume the range of additional contacts will fall somewhere between those that are in excess of the expected household contacts, and those in excess of all expected contacts (household and non-household) (Figure 14). We take the mid-point of this range and reflect the full range in our uncertainty estimate (see parameter table for final estimates). Additional App contacts are assumed to be non-household with the same SAR as for CTAS contacts (we assume the same adjustment for the case detection rate prior to the contact notification in Equation 12).

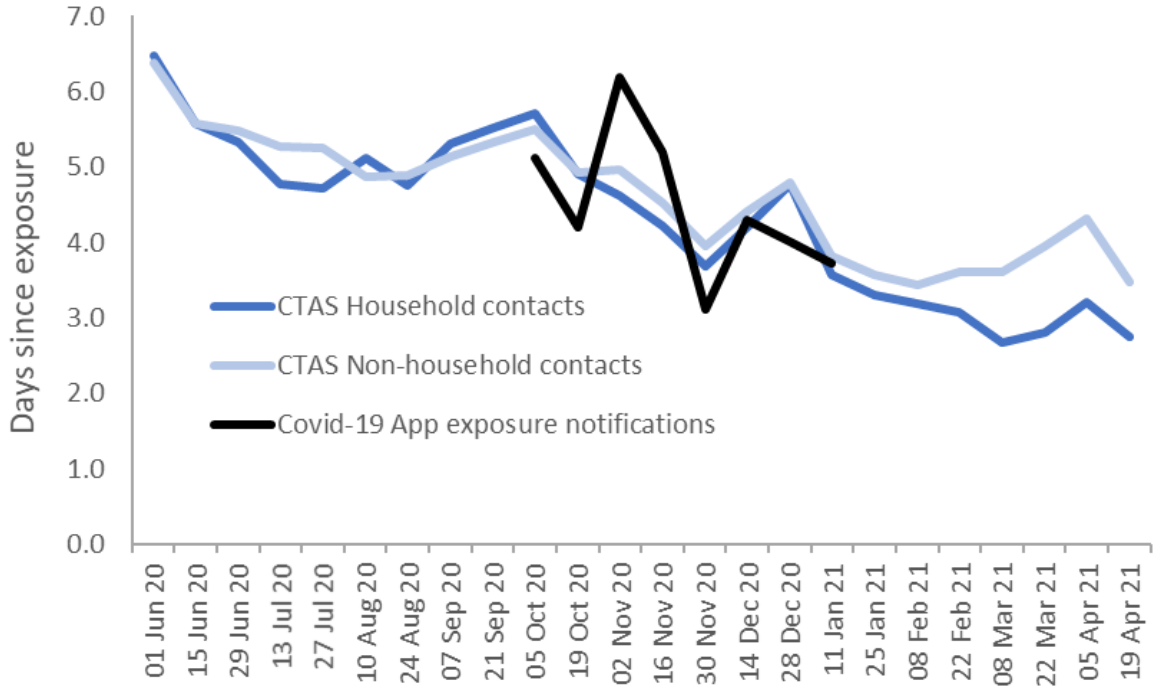
Figure 14. Estimating the rate of additional COVID-19 App exposures



The blue line is the total number of App exposure notifications (in each 14-day time period extending after the date shown). The positive cases identified are indicated by the dashed line. We calculated the expected number of household (green) and non-household (red) contacts that would have been notified assuming the same rate as among CTAS de-duplicated cases and contacts in each period. We assume 100% overlap in the household contacts reached. Our central estimate is that the additional contacts from the App lie in the middle of the range between all non-household contacts (blue – height of green) and only additional non-household contacts in excess of expected CTAS (blue – height of green and red). The variance was chosen to approximate the average range over time.

We use the average timing of App notifications for the additional contacts (as detailed in the main methods). We do not adjust the CTAS derived timing of notifications in the assumed overlap as the average time delays were very close (Figure 15).

Figure 15. The average timing of contact notifications relative to exposure time

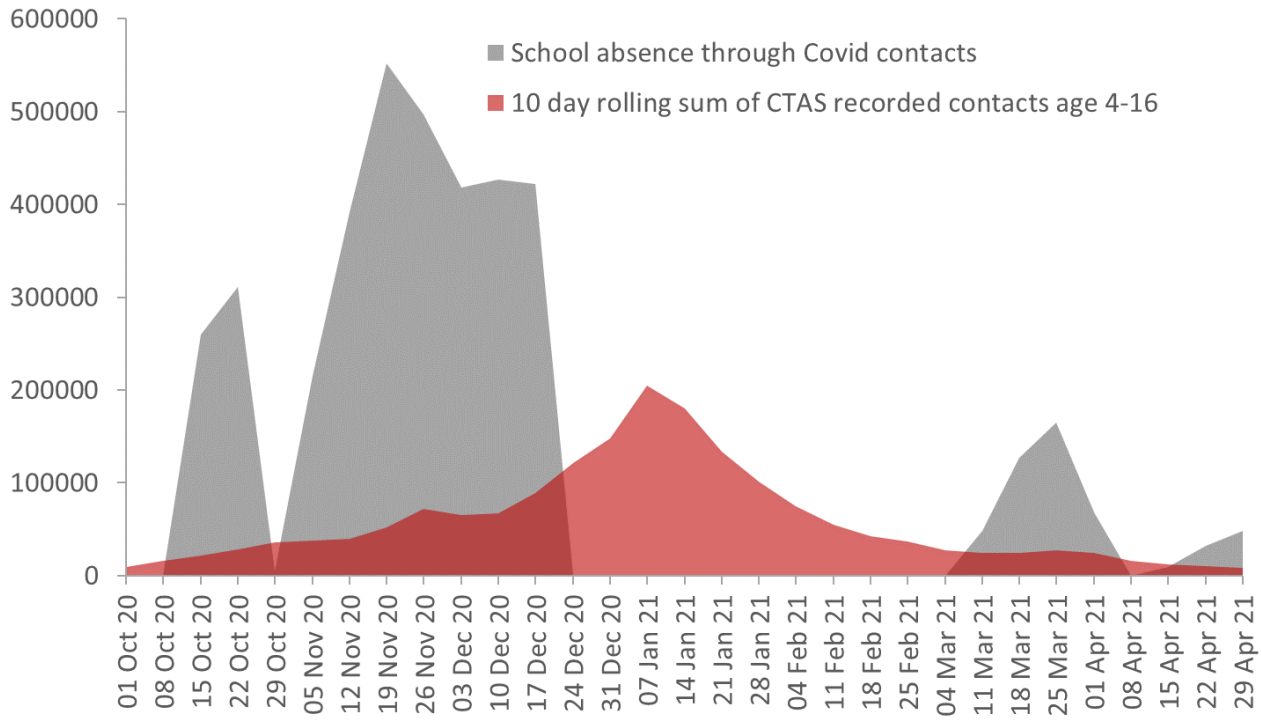


The blue lines are estimated from CTAS data (as defined in the main method) and the black line is an estimate derived from App data.

A7. Adding contacts of school-age not recorded in CTAS

Many school-aged children do not pass directly through the contact tracing system and are therefore underrepresented in CTAS. We used the figures published by DFE (Explore education statistics, 2021) to uplift the number of school age contacts in our data. We calculated the number of state-school absences linked to COVID contacts (inside or outside the school setting) registered on a single day, reported at weekly intervals. We subtracted the total number of CTAS contacts aged 4 to 16 registered in the 10-days before the date for which each DFE total was recorded (Figure 16). We define this as the weekly uplift. To further avoid double counting any isolation, we assume that the additional number of contact isolations in each 14-day period in our study will be equal to the maximum weekly uplift in that period (see parameter table for final estimates).

Figure 16. Additional school-age contact isolations



DFE data reports all recorded school absences directly linked to COVID-19 contacts (in or out of school). The grey region is the data representing a daily snapshot, reported weekly from all England state schools. At peak this is equivalent to around 6% of school-age children. We subtract the total CTAS contacts (household and non-household) in this age-group (4 to 16) in the preceding 10 days to avoid any duplication. We assume that the total additional school-age isolations in any 14-day window will be equal to the maximum uplift within that period. There are no recorded figures prior to October 2020 and these figures do not include any contact isolations that will have occurred in the holiday periods.

A.8 Estimating the SAR for school-age contacts

It is very difficult to know the precise rate of infections among the additional school-age contacts. The DFE absentee figures prior to 2020 do not identify the setting for contacts. From 2021 the figures differentiate between contacts made inside and outside the school setting. The average proportion of absences listed from contacts inside the school setting in this period is around 70%.

School-age transmissions are likely to be lower than for older age groups. However, it is not possible to determine what proportion of school-age transmissions are taking place in educational settings. Moreover, there is lots of variation in the size of bubbles that are isolated within schools. Contacts made outside schools are likely to be a mix of household and non-household.

For simplicity, here we assume a SAR of 5 to 15%, which is representative of the average non-household SAR (Figure 12). We adjust this attack rate to account for the fact that some contacts

will previously have been identified as cases, assuming the same rate of adjustment as for non-household CTAS contacts (see Equation 12).

A.9 Determining the final number of cases and contacts

To summarise our final dataset: CTAS cases and contacts are only counted once within a 14-day time window either side of their registration date. Contacts are counted if they are registered before becoming a case, and cases if they are registered before becoming a contact. Contacts are always counted as household if there are one or more household contacts recorded but the first notification date is assumed to be the isolation date. We include records for self-serve LFDs from NPEX and adjust these on the basis that there would be the same previous contact rates as for all cases. A proportion of the unconfirmed LFD cases and their associated contacts are removed in each time window to reflect the rate of false positive LFD cases (1-PPV).

We upscale the total number of contacts based on information from the COVID-19 App data and DFE records. Estimates of the SAR are used together with measurements taken directly from the data to estimate the number of infectious contacts. We effectively scale down the SAR to account for the contacts we have removed from our data who were previously cases. The final number of unique infectious cases and contacts, in each 14-day time period can be defined as:

$$\begin{aligned}\text{Total infectious cases} &= \text{Symptomatic PCR Cases not previously Contacts} \\ &+ \text{Asymptomatic PCR Cases not previously Contacts} \\ &+ \text{LFD Cases with confirmatory PCR not previously Contacts} \\ &+ \text{PPV}\left(\text{LFD Cases without confirmatory PCR not previously Contacts}\right. \\ &+ \text{NPEX selfserve LFD Cases} \times \left.\frac{\text{All other Cases not previously Contacts}}{\text{All other Cases}}\right)\end{aligned}$$

Equation 11

$$\begin{aligned}\text{Total infectious contacts}^{11} \\ &= (\text{Household Contacts not previously Cases} \times \alpha_{\text{Household}}) \\ &+ (\text{Non Household Contacts not previously Cases} \times \alpha_{\text{NonHousehold}}) \\ &+ (\text{Additional App Contacts} \times \alpha_{\text{NonHousehold}}) \\ &+ \left(\text{Additional SchoolAge Contacts} \times \text{SchoolAge SAR} \times \frac{\alpha_{\text{NonHousehold}}}{\text{NonHousehold SAR}}\right)\end{aligned}$$

Where,

$$\propto_{Household} = \frac{Household\ SAR - \ proportion\ of\ Household\ Contacts\ previously\ Cases}{1 - proportion\ of\ Household\ Contacts\ previously\ Cases}$$

$$\propto_{NonHousehold} = \frac{NonHousehold\ SAR - \ proportion\ of\ Non\ Household\ Contacts\ previously\ Cases}{1 - proportion\ of\ Non\ Household\ Contacts\ previously\ Cases}$$

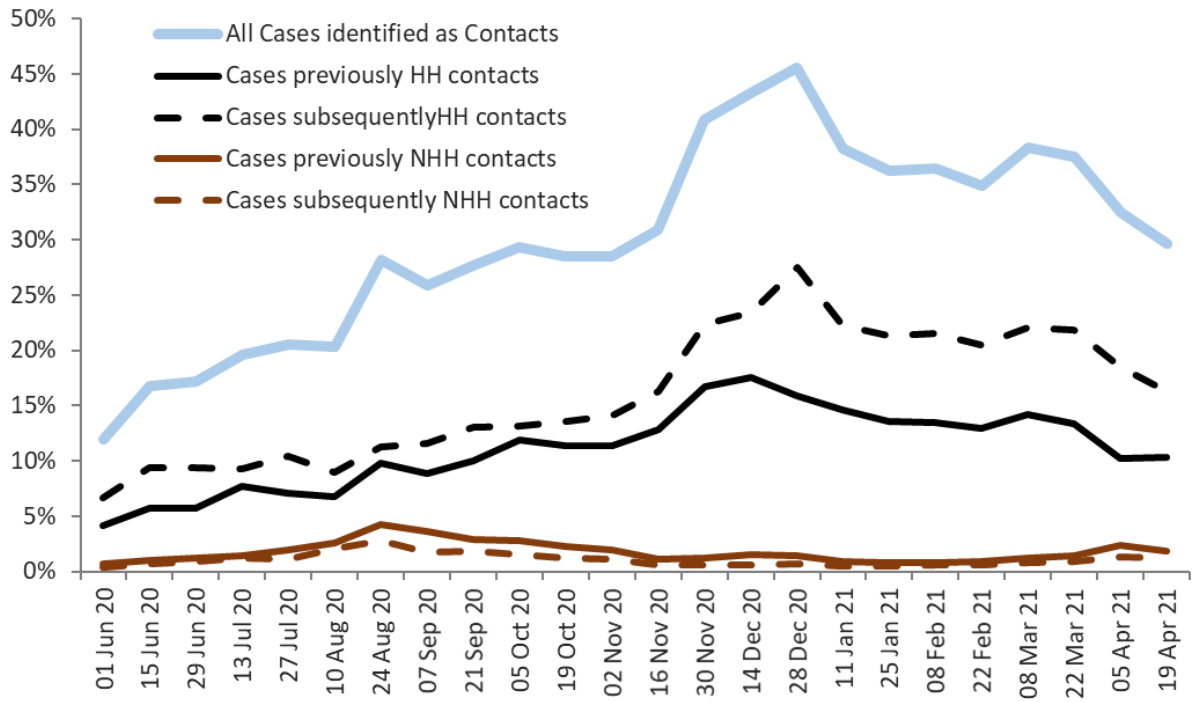
Equation 12

A.10 Measuring the reach of contact tracing among cases

Figure 2 shows the number of cases who tested positive that were also named and reached by NHSTT as contacts. The figure differentiates those who were reached before testing (who in our study would be counted as contacts), those reached after (who would be counted as cases) and the combined total. The data shows that since August 2020, between around 20 to 45% of all cases were also reached as contacts. 10 to 20% of cases were identified and reached as household contacts prior to testing positive; a further 1 to 5% were identified and reached as non-household contacts (this does not include App contacts or those appearing in DFE school records data). Prior identification by tracing would potentially bring forward isolation times and reduce secondary infections.

[11]CTAS contacts in this equation are scaled down to reflect the expected proportion of false positive LFD cases that were contact traced. The total reduction equates to around 2% of all contacts in March to April 2021 when LFD cases were proportionately highest.

Figure 17. The proportion of cases that are also contacts



The figure shows the number of cases (Pillar 1 and 2) reached in a 14-day period before, after, or either before or after, their case registration date.

A.11 Sensitivity testing individual parameters

Table 3 shows the impact when each model parameter is varied by 1 or 2 standard deviations, whilst holding all others constant at their central value. The output shown is the impact on the average Rt-reduction from TTI over the full study period. The column on the right shows the size of the standard deviation for each input parameter relative to its central value; this provides a measure of the relative uncertainty in each assumption. The red-to-blue colouring illustrates the relative output sensitivity.

Table 3

	Standard Deviation (SD)					
	-2	-1	0	1	2	SD as % of central value
Infectiousness abated for cases	0.172	0.193	0.215	0.239	0.263	13%
Rt_observed	0.194	0.205	0.215	0.226	0.237	5%
Household SAR	0.197	0.206	0.215	0.225	0.235	10%
Case Isolation Compliance	0.197	0.206	0.215	0.225	0.235	6%
Symptoms Only Isolation Compliance	0.201	0.208	0.215	0.223	0.230	25%
Contact Isolation Compliance	0.201	0.208	0.215	0.223	0.230	8%
School Age SAR	0.209	0.212	0.215	0.219	0.222	25%
Symptomatic rate	0.210	0.213	0.215	0.218	0.221	5%
Non-household SAR	0.211	0.213	0.215	0.217	0.219	10%
Additional contacts identified by the app	0.214	0.215	0.215	0.216	0.217	20%
Days to Isolate (App)	0.217	0.216	0.215	0.215	0.214	11%
Days to Isolate (Schools)	0.217	0.216	0.215	0.214	0.213	17%
Days to Isolate (CTAS)	0.227	0.221	0.215	0.210	0.205	11%
Total infectious population	0.238	0.226	0.215	0.206	0.198	5%

A.12 Impact on transmission reduction from TTI components

Table 4 shows the range of impact on transmission reduction from the different TTI system components. The lower-upper range represents the 90% confidence interval derived from the Monte Carlo analysis.

Table 4

Start of 14-day period	Symptomatic PCR cases		Asymptomatic PCR cases		LFD cases		Household Contacts		Non-Household Contacts		Covid-19 App contacts		School-age contacts		Self Isolation Symptoms Only	
	Lower CI	Upper CI	Lower CI	Upper CI	Lower CI	Upper CI	Lower CI	Upper CI	Lower CI	Upper CI	Lower CI	Upper CI	Lower CI	Upper CI	Lower CI	Upper CI
01 Jun 20	1.6%	2.6%	1.5%	2.4%	0.0%	0.0%	0.3%	0.7%	0.1%	0.1%	0.0%	0.0%	0.0%	0.0%	2.1%	5.6%
15 Jun 20	2.2%	3.6%	1.4%	2.3%	0.0%	0.0%	0.5%	1.1%	0.1%	0.2%	0.0%	0.0%	0.0%	0.0%	2.0%	5.3%
29 Jun 20	3.0%	4.9%	2.0%	3.2%	0.0%	0.0%	0.9%	2.0%	0.2%	0.5%	0.0%	0.0%	0.0%	0.0%	1.8%	4.7%
13 Jul 20	2.6%	4.2%	1.8%	2.9%	0.0%	0.0%	1.0%	2.1%	0.3%	0.6%	0.0%	0.0%	0.0%	0.0%	1.8%	4.9%
27 Jul 20	1.9%	3.1%	1.4%	2.3%	0.0%	0.0%	0.6%	1.2%	0.2%	0.4%	0.0%	0.0%	0.0%	0.0%	2.0%	5.4%
10 Aug 20	3.9%	6.3%	3.2%	5.2%	0.0%	0.0%	1.1%	2.4%	0.3%	0.6%	0.0%	0.0%	0.0%	0.0%	1.6%	4.3%
24 Aug 20	8.6%	13.9%	4.4%	7.1%	0.0%	0.0%	2.3%	5.1%	0.6%	1.4%	0.0%	0.0%	0.0%	0.0%	0.5%	1.8%
07 Sep 20	6.2%	10.1%	3.1%	5.0%	0.0%	0.0%	1.4%	3.3%	0.6%	1.2%	0.0%	0.0%	0.0%	0.0%	1.0%	3.0%
21 Sep 20	8.2%	13.3%	3.4%	5.5%	0.0%	0.0%	1.1%	3.0%	0.3%	0.7%	0.0%	0.0%	0.0%	0.0%	0.7%	2.2%
05 Oct 20	6.9%	11.2%	2.1%	3.4%	0.0%	0.0%	1.2%	3.0%	0.3%	0.6%	0.1%	0.4%	1.1%	3.0%	0.8%	2.5%
19 Oct 20	6.0%	9.8%	1.7%	2.7%	0.0%	0.0%	1.2%	2.8%	0.2%	0.5%	0.3%	0.8%	0.8%	2.2%	1.0%	3.0%
02 Nov 20	6.8%	11.1%	1.8%	3.0%	0.0%	0.0%	2.2%	4.8%	0.3%	0.6%	0.6%	1.8%	1.1%	2.9%	0.7%	2.2%
16 Nov 20	5.6%	9.0%	1.7%	2.7%	0.0%	0.1%	3.7%	7.3%	0.2%	0.4%	0.4%	1.0%	2.1%	5.6%	0.8%	2.4%
30 Nov 20	5.3%	8.6%	1.5%	2.4%	0.1%	0.1%	4.7%	9.1%	0.3%	0.5%	0.7%	1.7%	1.6%	4.3%	0.8%	2.5%
14 Dec 20	5.7%	9.2%	1.6%	2.6%	0.1%	0.2%	4.1%	8.4%	0.3%	0.5%	1.1%	2.7%	0.8%	2.0%	0.8%	2.4%
28 Dec 20	8.0%	13.1%	2.2%	3.5%	0.3%	0.5%	3.5%	8.2%	0.2%	0.4%	0.6%	1.5%	0.0%	0.0%	0.3%	1.6%
11 Jan 21	6.7%	11.0%	2.3%	3.7%	0.5%	0.8%	5.2%	10.5%	0.2%	0.5%	0.3%	0.8%	0.0%	0.0%	0.9%	2.6%
25 Jan 21	4.4%	7.1%	1.5%	2.5%	0.7%	1.1%	3.6%	7.1%	0.1%	0.3%	Data not included		0.0%	0.0%	1.5%	4.2%
08 Feb 21	5.4%	8.8%	1.8%	2.9%	1.0%	1.6%	4.1%	8.3%	0.1%	0.3%			0.0%	0.0%	1.2%	3.3%
22 Feb 21	6.4%	10.4%	2.4%	3.9%	1.4%	2.3%	5.8%	11.3%	0.3%	0.5%			0.0%	0.0%	0.8%	2.5%
08 Mar 21	4.4%	7.1%	1.5%	2.5%	2.0%	3.3%	4.4%	8.8%	0.3%	0.5%			1.3%	3.3%	0.8%	2.2%
22 Mar 21	2.5%	4.0%	0.8%	1.3%	1.4%	2.3%	2.6%	5.3%	0.2%	0.3%			1.4%	3.8%	1.3%	3.5%
05 Apr 21	2.4%	3.8%	1.0%	1.6%	1.3%	2.1%	2.1%	4.3%	0.2%	0.5%			0.0%	0.0%	1.6%	4.3%
19 Apr 21	4.2%	6.9%	1.8%	3.0%	3.5%	5.7%	4.0%	8.3%	0.6%	1.1%			1.7%	4.5%	0.7%	2.1%

About the UK Health Security Agency

The UK Health Security Agency is an executive agency, sponsored by the Department of Health and Social Care.

Website: www.ukhsa.gov.uk/
Twitter: @UKHSA_uk
Facebook: www.facebook.com/UKHSA

Published: September 2021

Gateway number: GOV-8809

UKHSA supports the UN Sustainable Development Goals