UK Health Security Agency

Testing for SARS-CoV-2 using antigen-detecting lateral flow devices

Evidence from performance assessment and evaluation for the period February 2021 to March 2022

Public Health and Clinical Oversight (PHCO)
Testing Operations, UKHSA

Contents

Background

The UK Health Security Agency (UKHSA) operates SARS-CoV-2 testing nationally. In the latter half of 2020, testing of asymptomatic individuals with antigen testing Lateral Flow Devices (LFDs) to diagnose coronavirus (COVID-19) was introduced and supported identification of positive individuals. LFDs are point of care tests based on a colloidal gold immunochromatography assay. For the test, a sample is applied directly to the test strip. If there is viral antigen in the sample, this will bind to an antibody to create a molecular complex and move along the test strip. When passing the test line, the complex is captured by another antibody resulting in a coloured line for a positive result.

LFD antigen testing was initially offered through assisted testing at designated test sites, Asymptomatic Test Sites (ATS) and was followed by the introduction of LFD self-test for at-home testing. This meant that individuals who were infectious, and may not otherwise have been found, could self-isolate and therefore reduce community transmission. Asymptomatic testing at scale allowed more people with transmissible virus to be found, with the potential to break chains of transmission.

The structure of testing for coronavirus within the UK is separated under different operational 'Pillars'. Testing of the wider population and the majority of care homes, in alignment with government policy, is operated under Pillar 2, with LFD antigen tests predominantly used for asymptomatic testing. Subsequently, this was extended to symptomatic individuals in specific circumstances. Pillar 2 has operated LFD testing through assisted-test or self-test delivery channels (for example, ATS at schools and workplaces) and home delivery. The asymptomatic testing programme, through these channels, has operated in 4 main testing groups:

- Group 1: repeat testing to detect positive cases amongst asymptomatic individuals (and remove them from circulation)
- Group 2: testing prior to an activity to reduce risk (this may be one or more tests)
- Group 3: asymptomatic testing where there is a signal of a potential outbreak (or where there has been an outbreak) to control infections, or where there is perceived to be a higher risk
- Group 4: daily testing of contacts to identify positive cases early[1]

In 2021, guidance extended the use of LFD antigen tests to include the symptomatic population in specific circumstances. The groups included within the symptomatic population were:

- concurrent testing with an extracted molecular test, for example, polymerase chain reaction (PCR), for the purposes of dispensing antiviral medication to eligible

[1] Includes contacts who were not required to self-isolate: fully vaccinated (2 vaccines), aged under 18 years old, have taken or taking part of an approved vaccine trial or not able to get vaccinated for medical reasons. Testing of vaccinated contacts was strongly recommended, in order to reduce risk. Source: Guidance for contacts of people with confirmed coronavirus (COVID-19) infection who do not live with the person (Withdrawn on 24 February 2022)

individuals with COVID-19. Subsequently diagnostic testing for antiviral medication has been solely by use of repeated LFDs
- ending of self-isolation early for individuals with COVID-19 in England testing negative on day 5 and day 6 of their self-isolation period

The national testing programme has developed and evolved in a changing landscape. Testing services undergo ongoing review to ensure they support the government strategy in response to COVID-19. Further information on the approach to testing and how it aligns with government policy is outlined in 'COVID-19 response: Living with COVID-19'.

Context

An essential component of the UKHSA national testing programme is performance evaluation of LFD testing prior to deployment and ongoing performance assessment after deployment. This provides assurance to the programme that LFDs used in different settings and populations, continue to perform as expected once deployed. To support assessment, UKHSA has implemented a series of service evaluations prior to deployment and ongoing evaluations after deployment. Service evaluations are field tests to assess the suitability of devices in the settings in which they are to be used prior to deployment, whereas ongoing evaluation monitors the performance of devices in the settings in which they are being used.

The first report on real-world LFD antigen test performance evaluation, as part of the national testing programme assessment of test devices, was published on 7 July 2021.[2] The report contained an overview of asymptomatic and symptomatic testing for COVID-19, considerations for using LFDs in wide-scale testing and available evidence at the time from LFD performance evaluations conducted and fully completed by the programme from October 2020 to May 2021.

This report builds on the first report and covers outcomes from real-world evaluations conducted by the programme, not previously published and initiated in the period February 2021 to March 2022. The focus is on 4 key areas:

- pre-deployment device evaluation
- pre-deployment testing regime evaluation
- post-deployment evaluations of LFD antigen test performance due to specific changes that may affect implementation
- evaluation of innovations and improvements

Other UKHSA reports are available which cover ongoing evaluation in post deployment.[3]

[2] Asymptomatic testing for SARS-CoV-2 using antigen-detecting lateral flow devices. Evidence from performance data October 2020 to May 2021. Published 7 July 2021. Lateral flow device (LFD) performance data.
[3] Ongoing evaluation reports can be found at: Lateral flow device (LFD) performance data

Pre-deployment device evaluation

Introduction

There is a robust process in place to ensure device performance is validated for all LFD antigen tests. Validation and performance assurance is performed prior to deployment. Before being approved for use in the testing programme, all types of LFD proposed for use in the national testing programme have been assessed at UKHSA Porton Down public health laboratories as part of an initial validation process. This work is overseen by the Oversight Group chaired by Professor Sir John Bell and the New Testing Technologies Governance Group. The full protocol is available at Protocol for evaluation of rapid diagnostic assays for specific SARS-CoV-2 antigens (lateral flow devices).

If the LFD antigen test passes Phase 3 within the UKHSA Porton Down evaluation protocol, it becomes eligible to bid under the Dynamic Purchasing System[4] for LFD procurement. For those LFD antigen tests successful in the bidding process, the next stage is regulatory review and evaluation. Over the course of the pandemic, there was an extensive programme of large-scale pre-deployment service evaluations at regional and local test sites with a focus on the performance of new device types, sampling techniques and assessment of evidence around device performance where there may be a need for 'off-label' use. LFD antigen tests where there was deemed a reasonable need to use them, in extremis, 'off-label', and those with appropriate regulatory Exceptional Use Authorisation rather than CE marking, have been evaluated by this programme in order to provide appropriate assurance and information for decisions on use as part of the response to the pandemic. The service evaluations are formal prospective evaluations designed to assess device performance in comparison to quantitative real-time polymerase chain reaction (qRT-PCR) in real-world usage and to inform operational programme management decisions. This was not required where an LFD antigen test was CE marked for the intended use. However as described above, all LFD test kits were subject to ongoing evaluation.

Between March and August 2021, 2 pre-deployment service evaluations were conducted to investigate the performance of LFD devices to be procured for the national testing programme.

The following evaluations and results relate to the service evaluations outlined in Table 1.

[4] Suppliers have to register online to access the DPS requirements.

Table 1. Summary of the 2 pre-deployment service evaluations conducted between March 2021 and August 2021 of LFD antigen tests procured for the national testing programme

	SureScreen SARS-CoV-2 antigen rapid test cassette V2	SureScreen SARS-CoV-2 antigen rapid test cassette (nasal swab) (Gold)
Total subjects	3,853	1,898
Swab collection technique	Double anterior nares	Double anterior nares
Setting	Assisted test	Self-test
Personnel using the test	Trained testing operative	End user

A further service evaluation reported below was undertaken for Innova LFD which had already successfully completed validation and been procured for use. The aim was to compare real-world performance of the COVID-19 self-test LFD by first time lay users using a dual mid-turbinate nasal swabbing technique compared to previously demonstrated performance by the same user group using throat and single mid-turbinate nasal swabbing technique.

Limitations

For all service evaluations using this methodology, it was noted that participant recruitment through regional or local test sites (RTS/LTS) was an inherent limitation. Participants attended sites for the purpose of receiving a diagnostic (PCR) test leading to a selection bias that may impact results observed, with consent occurring at time of arrival. Not all RTS/LTS sites were used for service evaluation which could have led to bias secondary to local variation, for example, local prevalence and circulating variants. Previous service evaluations for device performance of LFD antigen tests, used as comparators in this evaluation, were performed at different time periods since the start of the national testing programme. Variation in circulating variants and prevalence of SARS-CoV-2 for comparator service evaluations was present secondary to the time periods over which evaluations have been performed.

Performance evaluation of the SureScreen LFD V2 antigen test using double anterior nares swab collection in an assisted test setting

As part of pre-deployment device performance evaluation, a service evaluation was conducted between 4 March to 14 April 2021 to assess the performance of the SureScreen SARS-CoV-2 antigen rapid test cassette V2 (SureScreen LFD V2). At the time, there were limited regulatory approved LFD antigen tests on the market for testing for coronavirus.

In order to meet the increasing demand for testing, and to increase device supply chain resilience, a greater number of regulatory approved LFD antigen tests were required. The

SureScreen LFD V2 had passed Phase 3 of the Public Health England (PHE) Porton Down laboratory-based validation process and was considered suitable for further performance evaluation.

The primary objective of the evaluation was to evaluate performance of the SureScreen LFD V2 using double anterior nares swabbing, in an assisted test scenario at Regional Testing Sites (RTS). The reference standard was qRT-PCR analysis of a throat and single mid-turbinate nasal swab sample.

A total of 3,853 subjects were recruited for the service evaluation by convenience sampling from individuals attending an RTS for testing. Assisted swab collection and processing of the SureScreen LFD V2 antigen test was performed at RTS sites. A double anterior nares sample collection swabbing technique was performed by a trained operative, followed by a throat and single mid-turbinate nasal sample swab collection for a qRT-PCR test. The sensitivity and specificity of the LFD was then determined by comparing the LFD antigen test results with the qRT-PCR reference standard.

Results

The results showed an overall sensitivity of 73.23% and a specificity of 99.85% for the SureScreen LFD V2 antigen test (Table 2). All 95% confidence intervals (CI) were calculated using Clopper-Pearson method.

Table 2. Overall sensitivity and specificity for the SureScreen LFD V2 antigen assisted test[5]

Stratification	Sensitivity % (TP/TP+FN; 95% CI)	Specificity % (TN/TN+FP; 95% CI)
Overall	73.23 (320437; 68.81-77.32)	99.85 (3,411/3,416; 99.66-99.95)
Viral concentration under 10,000 RNA copies per mL	19.32 (17/88; 11.68-29.12)	-
Viral concentration = 10,000 to 1 million RNA copies per mL	70.40 (88/125; 61.58-78.23)	-
Viral concentration over 1 million RNA copies per mL	95.98 (215/224; 92.51-98.15)	-
Symptoms: yes	76.2(291/382; 71.6-80.4)	
Symptoms: no	52.7 (29/55; 38.8-66.4)	

[5] Data does not include samples that were missing, void or participants who had dropped out of PCR and LFD.

The SureScreen LFD V2 antigen test results were compared with previous device service evaluation outcomes. The results comparing SureScreen LFD V2 with combined service evaluation LFD results (non-SureScreen) are included in Table 3. The results of the combined previous service evaluations are published as part of the initial report of performance data October 2020 to May 2021.[6]

Table 3. Sensitivity and specificity of the SureScreen LFD V2 device in comparison to evaluated LFD antigen tests

Measure	SureScreen LFD V2 % (n, 95% CI)	Combined service evaluation LFD results (non-SureScreen LFD V2)[7] % (n, 95% CI)
Specificity	99.85 (3,410/3,415; 99.66-99.95)	99.65 (5698/5718; 99.46-99.79)
Sensitivity (viral concentration under 10,000 copies/mL)	19.32 (17/88; 11.68-29.12)	11.51 (32/278; 8.01-15.86)
Sensitivity (viral concentration 10,000 to 1 million copies/mL)	70.40 (88/125; 61.58-78.23)	57.21 (234/409; 52.26-62.06)
Sensitivity (viral concentration over 1 million copies/mL)	95.98 (215/224; 92.51-98.15)	87.27 (336/385; 83.52-90.43)

The service evaluation showed sensitivity of the device increased with viral concentration, as with other LFDs. Results for device sensitivity and specificity, when compared to non-SureScreen LFD antigen test previously evaluated at the time, provided assurance to the programme that the SureScreen LFD V2 had appropriate performance for use in the testing programme.

Performance evaluation of the SureScreen LFD V3 using double anterior nares swab collection in a self-test setting

The service evaluation was conducted between July to August 2021 to evaluate the performance of the SureScreen SARS-CoV-2 antigen rapid test cassette (nasal swab) (Gold), referred in this report as SureScreen LFD V3, when used by first-time lay users in a self-test setting. The reference standard comparator was qRT-PCR analysis of a throat and single mid-turbinate nasal swab sample.

[6] Results published in Asymptomatic testing for SARSCoV-2 using antigen-detecting lateral flow devices. Evidence from performance data October 2020 to May 2021. Published 7 July 2021.
[7] The comparator group includes merged data from assisted testing using Innova LFD and Orient Gene LFD.

Participants were recruited by convenience sampling of individuals attending an RTS for a coronavirus test. A throat and single mid-turbinate nasal swab sample for qRT-PCR was collected from participants at the RTS. Participants were then provided a SureScreen LFD V3 to perform at home, after leaving the RTS. Participants reported the LFD results to the service evaluation team. Barcodes were used to match the LFD results to participants qRT-PCR results for analysis. The service evaluation had a total sample of 1,898 participants during the evaluation period (demographics are included in Appendix A).

Results

Overall sensitivity for the SureScreen LFD V3 was 74.8%. This increased to 93.4% for individuals in the highest viral concentration category (over one million copies per mL). Specificity for the SureScreen LFD V3 was 99.7%. Full summary results are included in Table 4. Results for the SureScreen LFD V3 were consistent with previous service evaluations for non-SureScreen LFD antigen tests, including the DHSC COVID-19 self-test kit. The results for sensitivity of the SureScreen LFD V3 showed the lower bound of the 95% Confidence Interval (CI) for overall sensitivity to be 70.5%, above the performance threshold of 50%.[8]

Table 4. Summary of sensitivity and specificity by viral concentration categories and symptoms

SureScreen LFD V3 self-test (N=1844)[9]	Sensitivity % (TP/TP+FN; 95% CI)	Specificity % (TN/TN+FP, 95% CI)
Overall	74.83 (327/437 70.49-78.83)	99.72 (1403/1407, 99.27-99.92)
Viral concentration under 10,000 RNA copies per mL	23.53 (16/68; 14.09-35.38)	
Viral concentration = 10,000 to 1 million RNA copies per mL	69.72 (99/142; 61.45-77.14)	
Viral concentration over 1 million RNA copies per mL	93.39 (212/227; 89.34- 96.25)	
Symptoms: yes	80.06 (257/321; 75.27-84.29)	
Symptoms: no	52.33 (45/86; 41.27-63.21)	
Vaccination status: one dose	76.74 (99/129; 68.49-83.73)	
Vaccination status: 2 doses	72.94 (159/218; 66.52-78.71)	
Vaccination status: unvaccinated	76.67 (69/90; 66.57-84.94)	

[8] The LFD performance threshold / benchmark considered (50%) is based on the lowest acceptable overall sensitivity shown by an LFD device during a previous service evaluation, conducted between 4 November and 18 December 2020 in a self-testing setting, using throat and mid-turbinate nasal swabbing for sample collection. Evaluation data published in: Asymptomatic testing for SARS-CoV-2 using antigen-detecting lateral flow devices. Evidence from performance data October 2020 to May 2021. Published 7 July 2021.

[9] Data does not include samples that were missing, void or participants who had dropped out of PCR and LFD.

For all subjects (symptomatic and asymptomatic combined), the viral concentration (VC) was a statistically significant predictor of the likelihood of an LFD returning a true positive (TP), as displayed in Figure 1 and Figure 2.

Figure 1. Probability of a true positive based on viral concentration at a 95% confidence interval. The 95% CIs marked are for the 3 viral concentration categories included in Table 4, that is, under 10,000, 10,000 to 1 million, over 1 million (copies per mL)

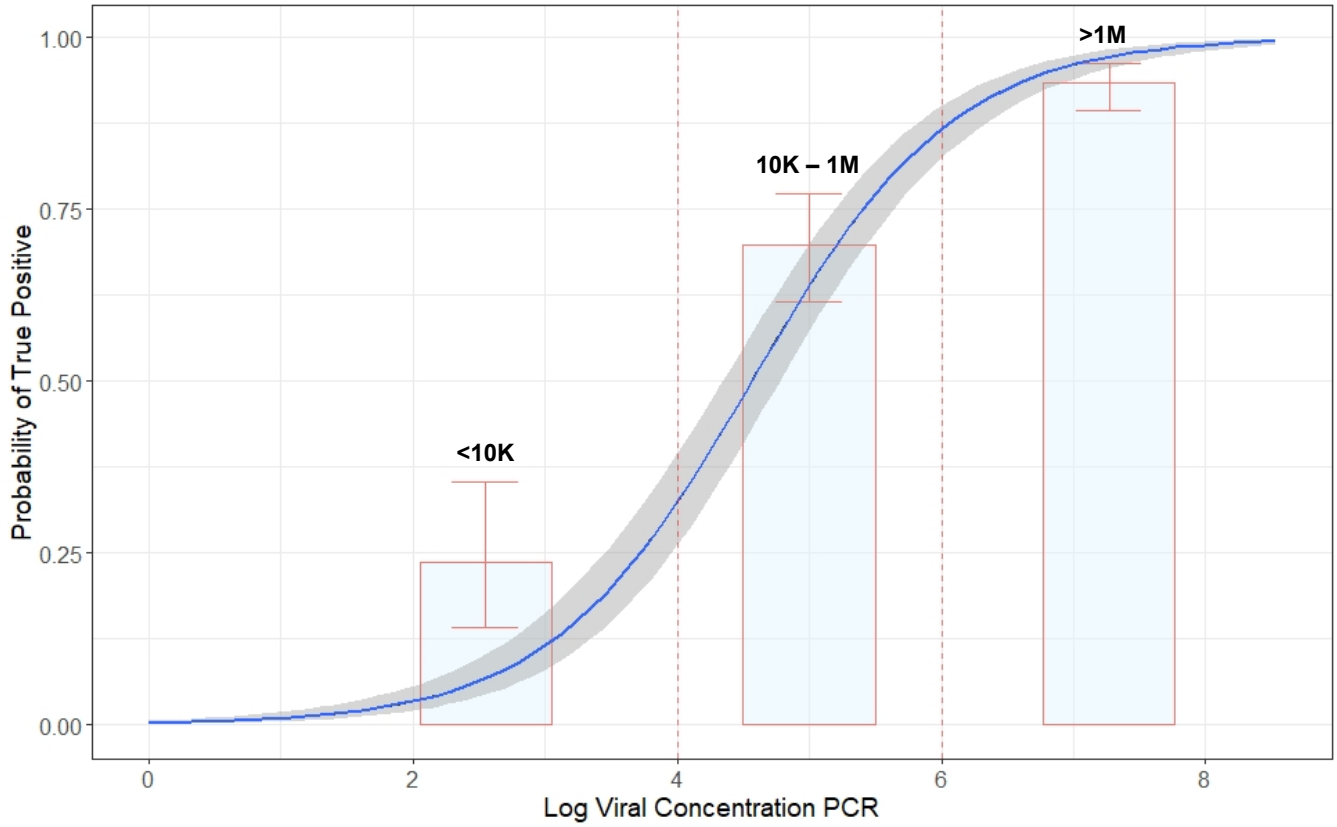
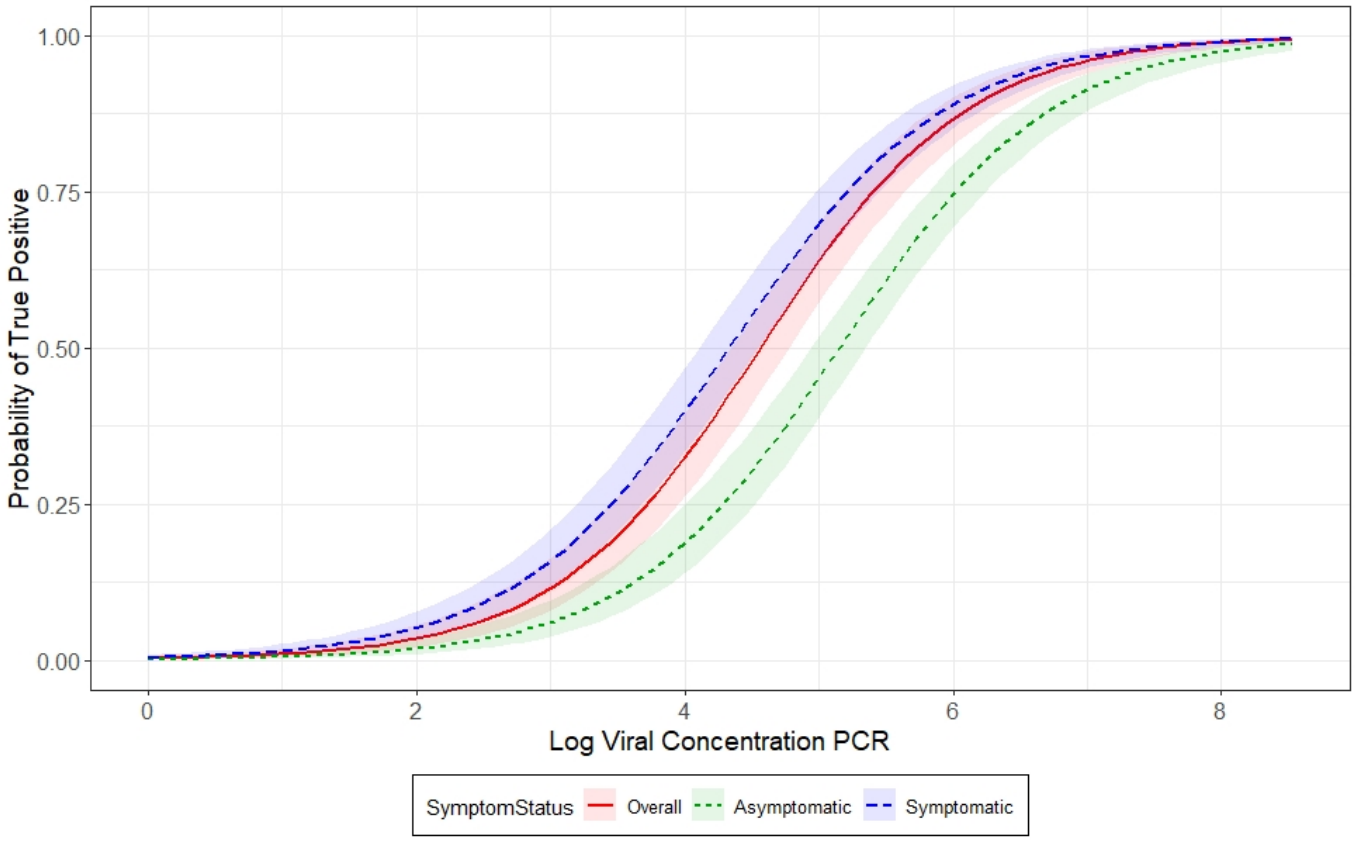
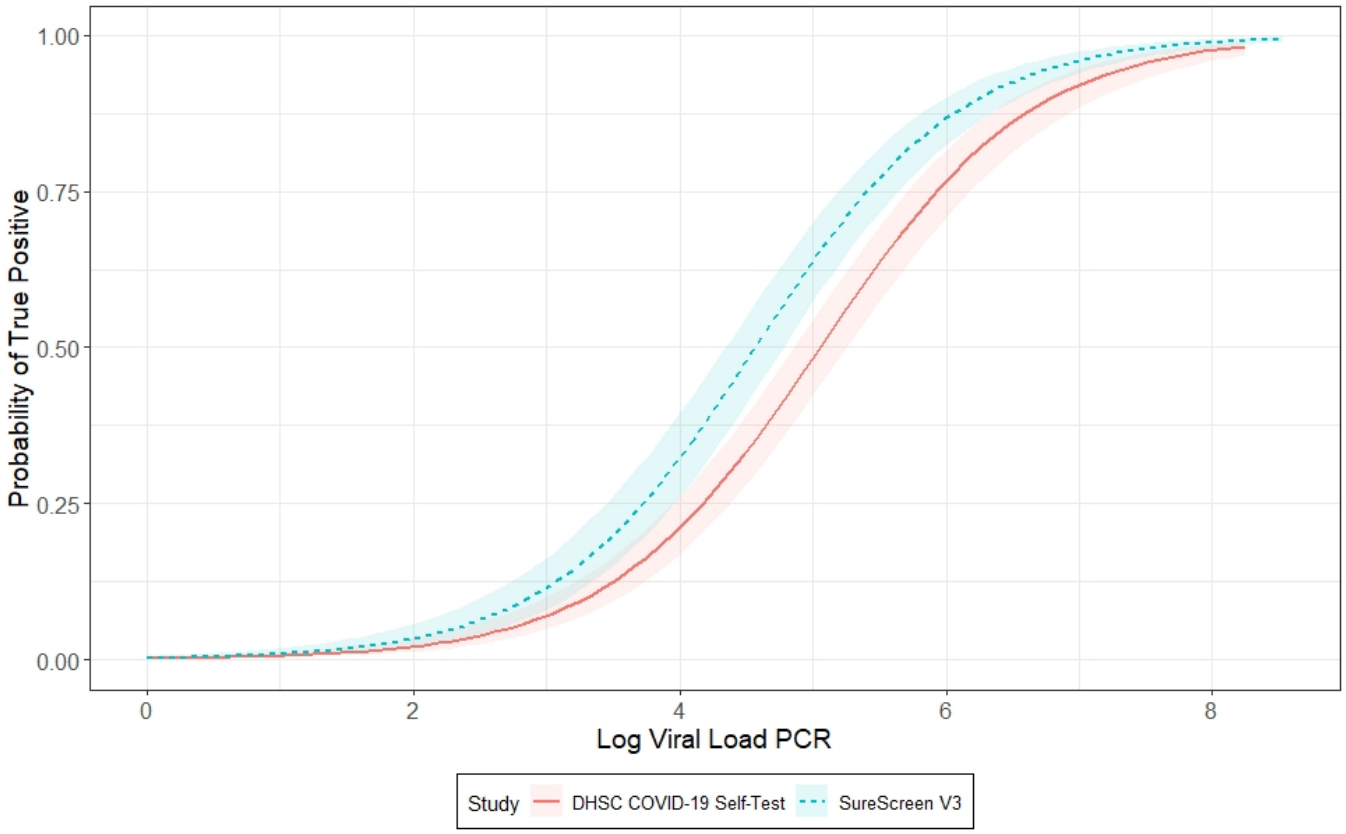


Figure 2. Probability of a true positive based on viral concentration stratified by symptom status



In comparison to DHSC COVID-19 self-test LFDs the performance of which had already been evaluated, the SureScreen LFD V3 performance was considered to be non-inferior across viral concentrations, as shown in Figure 3.

Figure 3. Comparison in probabilities of returning a TP for SureScreen V3 and DHSC Self-Test LFDs



The SureScreen LFD V3 antigen test, self-test, was considered of sufficient performance to proceed to deployment within the testing programme.

Performance evaluation of dual mid-turbinate swab collection for LFD antigen self-testing

The primary objective of this evaluation was to understand the performance of the Innova SARS-CoV-2 lateral flow antigen test (Innova LFD) when used by first-time lay users in an RTS using dual mid-turbinate swabbing for sample collection. Comparison was made to the benchmark of 50% overall sensitivity[10] and also to previous service evaluation performance data for the Innova LFD, evaluated using throat and single mid-turbinate (throat and nose) swabbing in a self-test setting. This service evaluation was conducted from 30 June to 26 July 2021 and the historical comparison service evaluation was conducted 22 May to 25 June 2021.

[10] The LFD performance threshold or benchmark considered (50%) is based on the lowest acceptable overall sensitivity shown by an LFD device during a previous service evaluation, conducted between 4 November and 18 December 2020 in a self-testing setting, using throat and mid-turbinate nasal swabbing for sample collection. Evaluation data published in: Asymptomatic testing for SARSCoV-2 using antigen-detecting lateral flow devices. Evidence from performance data October 2020 to May 2021. Published 7 July 2021.

Participants were recruited by convenience sampling of individuals attending an RTS for the purposes of receiving a qRT-PCR diagnostic test. Participants self-tested at the RTS using the Innova LFD device and using a dual mid-turbinate swab sample collection technique. This was followed by a self-collected swab sample for qRT-PCR analysis using a throat and single mid-turbinate nasal swab sample collection technique. The sensitivity and specificity of the LFD was then determined by comparing the LFD device result with the reference standard qRT-PCR result. 95% confidence intervals (CI) were calculated using Clopper-Pearson method.

The sample size of the service evaluation was 1,131 participants (demographics are included in Appendix B).

Results

Dual mid-turbinate swabbing using the Innova LFD demonstrated an overall sensitivity of 72.9%. The results were non-inferior to the benchmark overall sensitivity of 50% with a lower 95% CI of 67.7%. Summary results are included in Table 5.

When compared with the previous service evaluation using throat and single mid-turbinate swab, the evidence from this evaluation demonstrated significantly greater sensitivity for high viral concentration (more than 1 million per mL), while minimal difference was detected for lower viral concentrations (Table 5). It is possible that when this evaluation was conducted, the SARS-CoV-2 variant in circulation had a greater expression in the nose than in the throat.

The service evaluation demonstrated non-inferiority for dual mid-turbinate compared to throat and single mid-turbinate swab collection technique in a self-test setting and it appears as though performance may be better for dual mid-turbinate swabs. However, it is possible that expression of the virus in the nose and throat may differ by variant and therefore this finding may not be the case for all variants of COVID-19.

Table 5. Summary of sensitivity with an indirect comparison to previous evaluation

Stratification[11]	Innova LFD: Dual mid-turbinate self-test Total N[12]	Innova LFD: Dual mid-turbinate self-test Sensitivity % (n, 95% CI)	Innova LFD previous evaluation: throat and single mid-turbinate self-test Total N [13]	Innova LFD previous evaluation: throat and single mid-turbinate self-test Sensitivity % (n, 95% CI)	Difference[14] % (95% CI) (p value)
Overall	1,089	73.72 (230/312; 68.46-78.52)	631	65.77 (415/631; 61.92-69.47)	7.95 (1.58, 14.32) (*P = 0.02)*
VC[15] under 10,000 RNA copies per mL	48	18.75 (9/48; 8.95-32.63)	66	19.71 (13/66; 10.93-31.32)	-0.95 (-16.52, 14.63) (*P = 1)*
VC = 10,000 to 1 million RNA copies per mL	110	63.64 (70/110; 53.92-72.60)	216	54.63 (118/216; 47.73-61.42)	9.01 (-2.86, 20.87) (*P = 0.15)*
VC over 1 million RNA copies per mL	154	98.05 (151/154; 94.41-99.60)	349	81.38 (284/349; 76.89-85.32)	16.7 (11.58, 21.78) (*P < 0.005)*
Symptoms: yes	554	80.35 (184/229; 74.60-85.29)	451	68.29 (308/451; 63.78-72.57)	12.06 (5.02, 19.09) (*P = 0.001)*
Symptoms and VC under 10,000	28	21.43 (6/28; 8.30-40.95)	43	20.93 (9/43; 10.04-36.04)	0.50 (-19.46, 20.46) (*P = 1)*
Symptoms and VC = 10,000 to 1 million	69	71.01 (49/69; 58.84-81.31)	141	57.45 (81/141; 48.85-65.73)	13.57 (-0.97, 28.11) (*P = 0.08)*
Symptoms and VC over 1 million	132	97.73 (129/132; 93.50-99.53)	267	81.65 (218/267; 76.47-86.10)	16.08 (-10.22, 21.94) (*P < 0.005)*
Symptoms: no	500	52.78 (38/72; 40.65-64.67)	108	55.56 (60/108; 45.68-65.12)	-2.78 (-18.79, 13.24) (*P = 0.83)*

[11] 35 individuals did not declare their symptom status.
[12] Total positive and negative PCR results; void results removed.
[13] This data set contained only PCR positive samples.
[14] A 2-sample Chi-squared test for equality of proportions comparing the sensitivity between the current (dual mid-turbinate self-test evaluation) versus the historical evaluation (throat and single mid-turbinate self-test).
[15] VC = viral concentration.

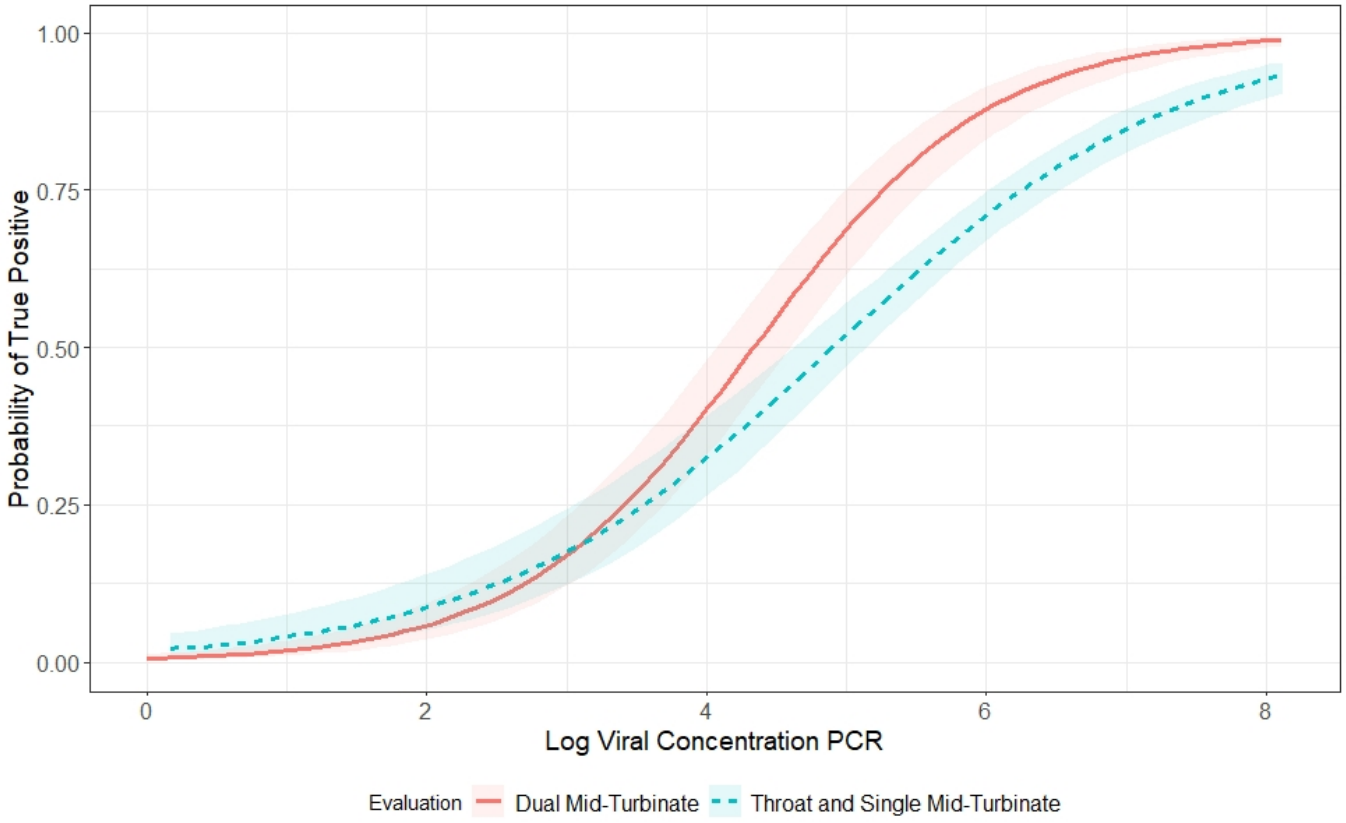
Stratification[11]	**Innova LFD: Dual mid-turbinate self-test Total N**[12]	**Innova LFD: Dual mid-turbinate self-test Sensitivity % (n, 95% CI)**	**Innova LFD previous evaluation: throat and single mid-turbinate self-test Total N** [13]	**Innova LFD previous evaluation: throat and single mid-turbinate self-test Sensitivity % (n, 95% CI)**	**Difference**[14] **% (95% CI) (p value)**
No symptoms and VC under 10,000	19	15.79 (3/19; 3.38-39.58)	19	21.05 (4/19; 6.05-45.57)	-5.26 (-35.12, 24.59) (*P = 1)*
No symptoms and VC = 10,000 to 1 million	35	48.57 (17/35; 31.38-66.01)	45	46.67 (21/45; 31.66-62.13)	1.92 (-22.06, 25.87) (*P = 1)*
No symptoms and VC over 1 million	18	100 (18/18; 81.47-100)	44	79.55 (35/44; 64.70-90.20)	20.45 (4.62, 36.29) (*P = 0.09)*
Vaccination status: one dose	394	70.97 (88/124; 61.14-78.77)	202	66.83 (135/202; 59.88-73.28)	4.1 (-6.81, 15.08) (*P = 0.51)*
Vaccination status: 2 doses	514	74.51 (76/102; 64.92-82.62)	140	67.14 (94/140; 58.70-74.84)	6.65 (-5.66, 18.96) (*P = 0.33)*
Vaccination status: unvaccinated	181	76.74 (66/86; 66.39-85.18)	288	64.24 (185/288; 58.40-69.77)	11.59 (0.47, 22.7) (*P = 0.055)*

Within this evaluation, as expected, sensitivity increased with viral concentration, with the highest sensitivity of 98% reported at the highest viral concentration category (over 1 million copies per mL). Sensitivity did not appear to depend on vaccination status when testing proportionally between population groups (equality of proportions):

- one vaccine dose versus no vaccine: Difference = -5.8; 95% CI: -18.7;7,2, P=0.44
- 2 vaccine doses versus no vaccine: Difference = -2.2; 95% CI: -15.6;11.1, P=0.85

Compared to the previous comparator service evaluation, the probability of a True Positive (TP) was higher using dual mid-turbinate swab sample collection than for throat and single mid-turbinate swab collection for viral concentrations over 100,000 copies per mL, as shown in Figure 4. At lower viral concentrations, there was less difference between swab collection techniques in terms of the likelihood to return a TP.

Figure 4. Comparison of the probability of a True Positive for 'dual mid-turbinate' and 'throat and single mid-turbinate' (by Log of VC[16])



Taking the above limitations into consideration the test was deemed to be acceptable for use in the programme.

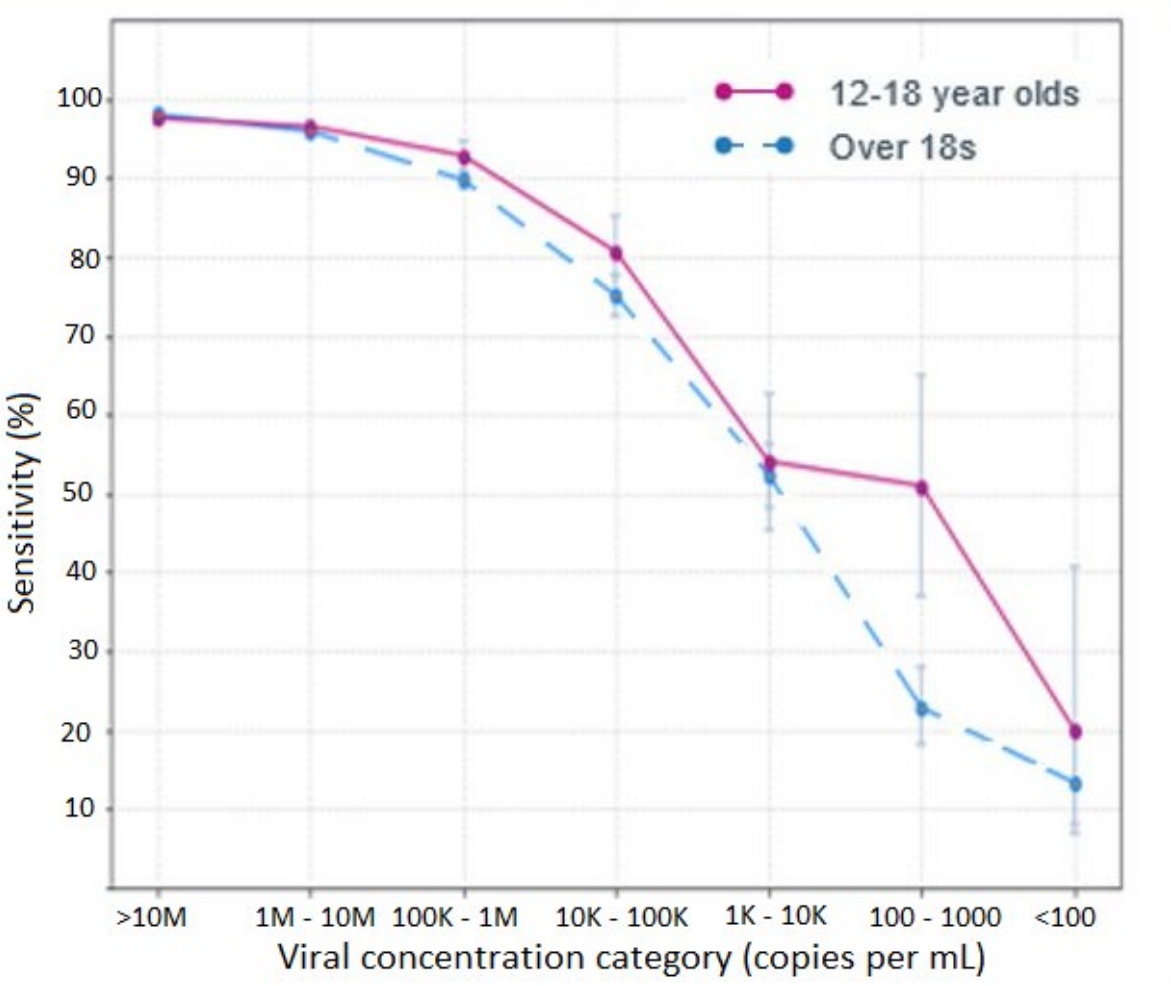
[16] The log odds of a TP event with respect to Log_{10} Viral concentration was fitted using a logistic regression model: $Log(odds(TP)) = \beta 0 + \beta 1 \bullet LogVC$

Further analysis

Further analysis of real-world data collected as part of a surge testing response between 1 April and 21 June 2021 was conducted to evaluate LFD performance by age group. PCR and LFD tests were distributed to all age groups which enabled an assessment of whether there was a difference in self-test ability to detect COVID-19 in secondary school and adult age groups,[17] using dual mid-turbinate swabbing.

The sensitivity of LFD kits within the secondary school age group compared to the adult age group showed no statistical difference in kit ability to detect SARS-CoV-2 across the range of viral concentrations (Fisher's exact test p>0.05) (Figure 5).

Figure 5. Sensitivity by estimated viral concentration in copies/mL for select surge sites in secondary school age group and adults



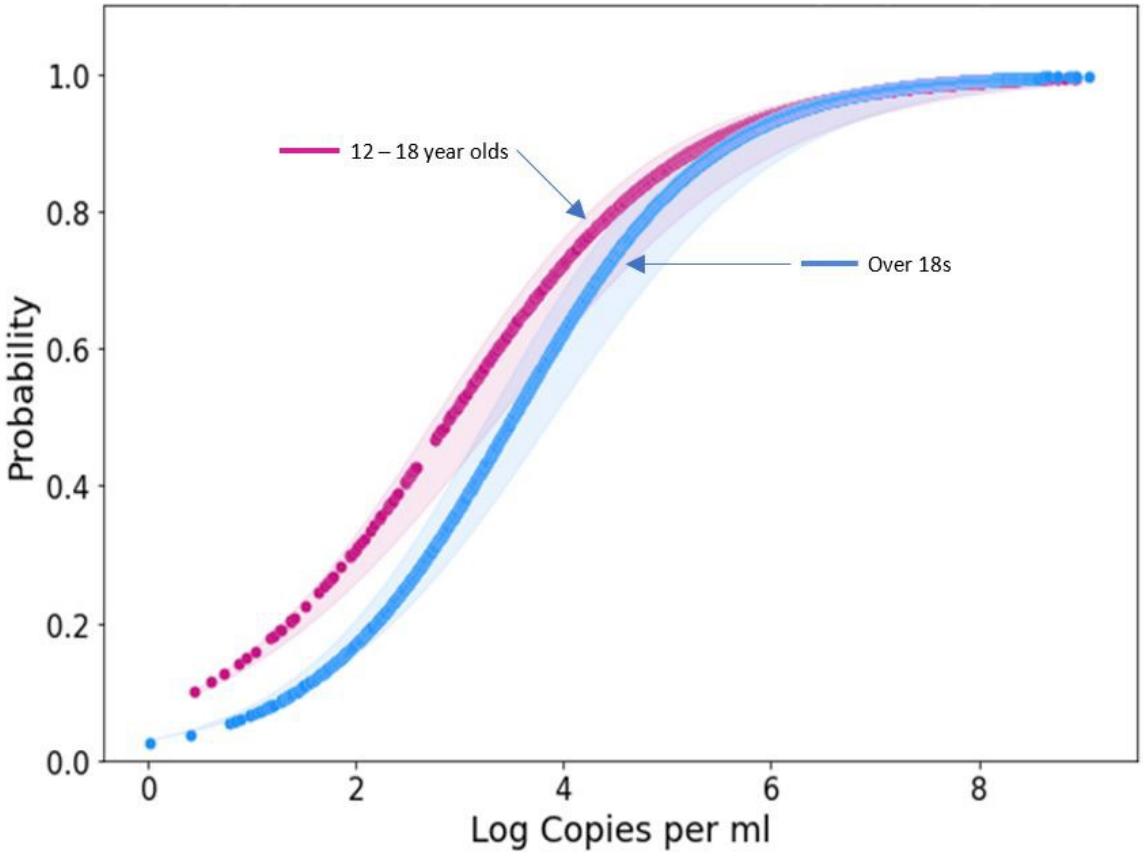
Logistic regression modelling of the paired testing data from selected surge testing sites showed no statistical difference in the clinical performance ability of the secondary school age

[17] Comparison of the DHSC COVID-19 self-test kits detection rate for SARS-CoV-2 within the secondary school age group (12 to18 year olds; see below for definition) has been compared to the detection rate within adults (over 18 years old; see below for definition). Age categories are defined using the school year cut off (current school year cut off include those who will be turning 18 before 1 September 2021)

group and adult age group to detect SARS-CoV-2 at 95% confidence interval (see Figure 6), while the ability of LFDs to detect SARS-CoV-2 in the secondary school age group is significantly greater at 90% confidence interval.

These results evidence that the use of LFDs in this age group, and in particular self-swab (under supervision), was an appropriate implementation approach in young persons.

Figure 6. Logistic regression modelling of the select surge testing site paired technology testing by age group



Pre-deployment testing regime evaluation

Introduction

Regular evaluations are conducted within the programme to assess LFD performance under different testing regimes. This also provides insight into the appropriateness of implementing the testing regime. Variables can include swab and sample collection methodology, number of tests conducted, and frequency of LFD antigen testing.

The performance of different LFD antigen testing regimes has been assessed. This includes service evaluation of dual mid-turbinate swabbing and multi-day LFD testing, retrospective analysis of same-day dual LFD testing and multi-day LFD testing, 2 randomised controlled trials (RTC) to understand the impact of a policy of daily contact testing, plus an evaluation of daily contact testing of healthcare workers in the NHS, and an evaluation of daily contact testing in the workplace.

Limitations

The service evaluations outlined below were conducted at RTS and LTS and as such have the same limitations as outlined for the service evaluations above (see section Limitations).

Evaluation of same-day dual testing LFD performance

Statistical analysis was conducted to evaluate whether 2 LFD antigen tests performed in succession gave a higher diagnostic sensitivity than one such test. In the service evaluation, participants performed 2 LFD antigen tests consecutively and a qRT-PCR test as the reference standard. The service evaluation took place between 10 February and 27 May 2021 using Orient Gene COVID-19 Ag Rapid Test Cassette LFD antigen tests (Orient Gene LFD Self-Test). Sample collection for both LFDs and PCR was self-collection by the subject. Both LFD antigen tests were collected using double anterior nares swab collection technique.

The primary objectives of the analysis were to calculate the additional sensitivity gained from performing 2 LFD antigen tests consecutively compared to a single LFD. This is based on the suggestion that undertaking multiple LFD tests at a given moment could increase sensitivity of the testing activity, considering 2 independent tests.

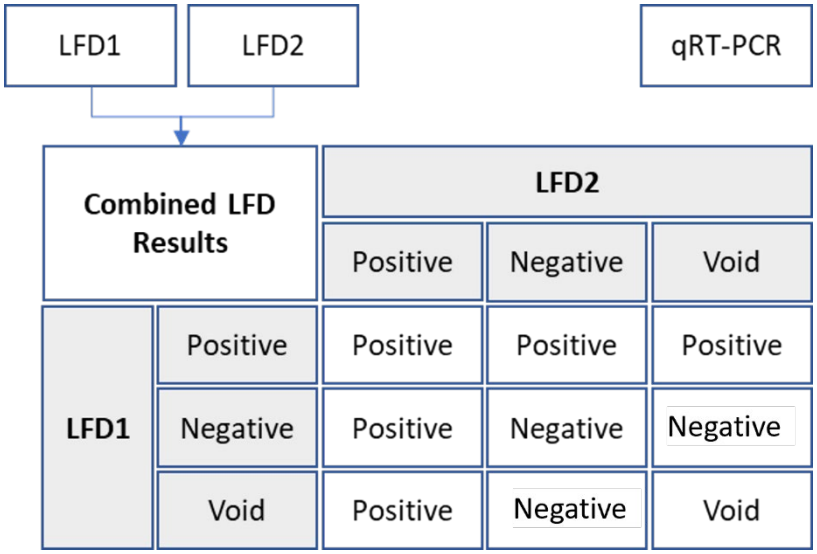
Over 2,000 subjects were suitable for evaluation, as shown in Table 6.

Table 6. Subjects suitable for evaluation as part of the analysis from same-day dual testing service evaluations

Stratification	Orient Gene LFD self-test evaluation
Subjects suitable for evaluation	2,188
PCR positive	159 (7.3%)
PCR positive with VC under 10,000 RNA copies per mL	31 (1.4%)
PCR positive with VC = 10,000 to 1 million RNA copies per mL	60 (2.7%)
PCR positive with VC over 1 million RNA copies per mL	68 (3.1%)
PCR negative	2,029 (92.7%)

Figure 7 outlines how each participant received their final result in the service evaluation, considering each individual's 2 LFD results.

Figure 7. Final result communicated to each participant considering both LFD test results



Combined sensitivity using 2 LFD antigen tests consecutively

To calculate the additional sensitivity, if any, gained from performing 2 LFD tests consecutively, the sensitivity was calculated for the results from the first LFD (LFD 1) and the second LFD (LFD 2) and for LFD 1 and 2 combined. Improvements in sensitivity by combining LFD results were minimal, as shown in Table 7.

Table 7. Individual and combined LFD results for Orient Gene LFD self-test (dual anterior nares)

Statistical category	LFD1 (Anterior nares)	LFD 2 (Anterior nares)	Combined LFD Results[18]
Total number	2,156	2,168	2,181
True positive	106	106	109
True negative	1,988	2,003	2,009
False positive	10	7	13
False negative	52	52	50
Sensitivity % (95% CI)	67.09 (59.18-74.35)	67.09 (59.18-74.35)	68.55 (60.72-75.68)
Sensitivity at VC over 1 million RNA copies per mL (%, 95%CI)	92.65 (83.67; 97.57)	94.12 (85.62; 98.37)	94.12 (85.62; 98.37)
Sensitivity at VC = 10,000 to 1 million RNA copies per mL (%, 95%CI)	67.80 (54.36; 79.38)	66.10 (52.61; 77.92)	70.00 (56.79; 81.15)
Sensitivity at VC under 10,000 RNA copies per mL (%, 95%CI)	9.68 (2.04; 25.75)	9.68 (2.04; 25.75)	9.68 (2.04; 25.75)
Specificity % (95% CI)	99.50 (99.08-99.76)	99.65 (99.28-99.86)	99.36 (98.90-99.66)

Evaluation of multi-day self-testing LFD performance

Evaluation of multi-day LFD antigen self-testing performance in symptomatic individuals

A service evaluation was conducted with the objective of evaluating whether a multi-day LFD testing regime achieved a performance that was comparable to qRT-PCR testing when applied for a low-risk symptomatic cohort. Low risk was defined as fully vaccinated (FV)[19] individuals under 50 years of age with symptoms. The LFD used in the evaluation was Acon Flowflex

[18] The total number of results for the combined LFD usage was larger than that of individual LFDs as a void result was only reported if both individual LFDs were void. If only one of the individual LFDs was void, the combined LFD usage still produced a result (see Figure 7). Therefore, fewer voids were obtained for the combined LFD usage compared to single LFD usage.

[19] Participants who, at the time of testing, had received their second vaccination dose 14 days or more prior.

SARS-CoV-2 antigen rapid test (Self-Testing) kit, referred to as the 'Acon' in this section of the report.

The evaluation recruited by convenience sampling across all ages, vaccination statuses and symptom statuses from individuals attending an RTS for the purposes of receiving a qRT-PCR diagnostic test. Day 0 was set as the day of the first PCR test. Each recruited individual had a qRT-PCR test on the day (Day 0), using a self-collected throat and single mid-turbinate sample. A further qRT-PCR sample was conducted at home on Day 1. The participant conducted 3 LFD self-tests over this period on Day 0, Day 1, and Day 2. All LFD tests were self-swabbed using a double anterior nares swab collection technique for sample collection. The participant uploaded their LFD results via an app within 4 hours of leaving the testing site (Day 0) and then daily and first thing in the morning thereafter (Days 1 and 2). Data is presented for all participants who entered the evaluation and had a valid outcome for PCR at day 0 and is also presented for the low-risk cohort.

The purpose of the 3 LFD tests was to compare 3 different multi-day LFD testing regimes to identify the optimum testing regime and to determine if any regime was non-inferior to a threshold sensitivity of 92%.[20] The evaluation did not aim at assessing compliance to different testing regimes, rather, it aimed to assess the performance of different testing regime options if subjects were tested in that manner.

Assessment of 3 different regimes was done in the evaluation:

- Regime 1: LFD test on Day 0, 1 and 2 (3 tests)
- Regime 2: LFD test on Day 0 and 1 (2 tests)
- Regime 3: LFD test on Day 0 and 2 (2 tests)

Results

A total of 2,788 participants were recruited for the evaluation of which 1,150 were associated with low risk. Table 8 shows the results for Day 0, Day 1, and Day 2 tests taken as part of the evaluation. Adherence to each testing regime for all participants was low as shown in Table 8. Out of the 2,788 subjects, 965 (34.6%) recorded results for Regime 1, 1,289 (46.2%) recorded results for Regime 2 and 995 (35.7%) recorded results for Regime 3.

[20] This margin was determined under an assumption that an estimate of the sensitivity from one of the regimens may provide a basis for demonstrating non-inferiority in terms of sensitivity to within 5% of clinical diagnostic test standards set by the Technologies Validation Group (that is, the threshold is 97%, therefore the NI margin was set to 92%: that is, if the lower 95% CI for the observed sensitivity in any regimen is greater than 92%, a conclusion of NI can be made). Technologies Validation Group standards can be found online.

Table 8. All LFD test results recorded contributing to all testing regimes

LFD day	Subjects % (missing, %)	LFD result[21]	PCR day 0 positive[22] (% of total)	PCR day 0 negative[23] (% of total)
Day 0	1,814, 65 (971, 35)	Positive	321 (82.9)	11 (0.8)
		Negative	66 (17.1)	1,375 (99.2)
Day 1	1,340, 48 (1,444, 52)	Positive	298 (90.6)	17 (1.7)
		Negative	31 (9.4)	965 (98.3)
Day 2	1,036, 37 (1,749, 63)	Positive	250 (90.6)	13 (1.8)
		Negative	26 (9.4)	724 (98.2)

The sensitivity for all subjects evaluated[24] per multi-day testing regime can be found in Table 9. To evaluate non-inferiority (NI), a threshold performance of 92% was used. As shown in Table 9, Regime 1 and Regime 3 were found to be non-inferior where the lower bound of the 95% Confidence Interval (CI) was above the sensitivity for non-inferiority.

Table 9. Sensitivity and non-inferiority for evaluable subjects per testing regime

Regimen	N (positive or negative)	Sensitivity % (95% CI)	Was NI shown? (NI margin = 92)	TP	FN
1: LFD day 0,1 and 2	965	96.51 (93.48-98.39))	Yes	249	9
2: LFD day 0 and 1	1,289	92.95 (89.52-95.53)	No	290	22
3: LFD day 0 and 2	995	95.82 (92.64-97.89)	Yes	252	11

All regimes in the low risk group,[25] as shown in Table 10, showed non-inferiority but the sample size was too small for formal non-inferiority at 90% power to be declared.

[21] Does not include void and missing results.
[22] Does not include void and missing results.
[23] Does not include void and missing results.
[24] At least a value recorded for PCR Day 0 and results for all LFD tests required.
[25] Out of the 2,788 total participants, 1,150 were in the symptomatic low-risk criteria (under 50 years of age and fully vaccinated (FV) at time of evaluation). Demographics for participants under the chosen criteria are shown on Table 17 in Appendix D. The rates of true positive (TP), false positive (FP), true negative (TN) and false negative (FN) for the low-risk symptomatic group (under 50 years of age, fully vaccinated and symptomatic) can be found on Table 18 and Table 19 in Appendix D.

Table 10. Sensitivity and non-inferiority for symptomatic low-risk subjects per regime

Regimen (under 50, FV and symptomatic)	**N** (positive or negative)	**Sensitivity %** (95% CI)	**Was NI shown** (NI margin = 92)	**TP**	**FN**
1: LFD Day 0,1 and 2	359	98.73 (95.47-99.85)	Yes	155	2
2: LFD Day 0 and 1	527	96.26 (92.44-98.48)	Yes	180	7
3: LFD Day 0 and 2	367	97.48 (93.68-99.31)	Yes	155	4

The data presented suggests that LFD testing for 3 days (Regime 1) from day of PCR or the day of seeking a test[26] (Day 0) or Day 0 and Day 2 testing (Regime 3) was comparable to the PCR conducted on Day 0.

Limitations

Participants were not randomly assigned to different testing regimes and all participants were required to follow the same 3-day testing protocol. Only one LFD type was used and therefore may not be representative of other types of LFD. Compliance data drawn from a service evaluation may not reflect real-world compliance in the event of a change to testing policy.

Evaluation of daily contact testing

Daily Contact Testing (DCT) has been proposed as an alternative to self-isolation for individuals identified as a close contact of someone who has tested positive for SARS-CoV-2. The rationale is that individuals who are identified as contacts are at a higher risk of having been infected and subsequent transmission of the virus. Contacts identified in the same household are more likely to have become infected than first-order contacts at work, school or elsewhere. Self-isolation is effective if individuals are compliant for the period required. However, early evidence suggested that compliance with self-isolation may have been as low as 11% in asymptomatic contacts.[27]

Modelling data showed daily testing of first-order contacts, using lateral flow tests, could potentially avert a similar level of onward virus transmission as self-isolation, with compliance with both self-isolation and the testing regime making significant differences to the effectiveness of each regime.

The success of DCT relies on a rapid test result. Antigen lateral flow devices (LFD) provide the quickest result turnaround of all the COVID-19 tests with results available in 15 to 30 minutes. Dependent on setting and context, DCT would involve contacts testing themselves using an

[26] 'Day of seeking a test' is not necessarily 'day of symptom onset' if applicable.
[27] Smith and colleagues. 'Adherence to the test, trace and isolate system: results from a time series of 21 nationally representative surveys in the UK.' September 2020.

LFD each day for 5 to 7 days. If the result of their test is negative, they can continue to carry out their essential day activities as usual. DCT is intended to find cases, while at the same time minimise the number of days spent in unnecessary self-isolation. It is a way to manage the risk of transmission, whilst maintaining essential services. A person who tests positive on LFD testing would follow national guidelines and self-isolate.

To understand DCT further, 2 randomised controlled trial (RCT) studies were developed and rolled out nationally along with 2 pilots of DCT, one in healthcare workers and one in the workplace. The RCTs focused on 2 different populations for inclusion: consenting English secondary schools and the wider public nationally, respectively.

A cluster randomised trial in English secondary schools comparing the impact of a policy of DCT with self-isolation

A DCT study was operated in English secondary schools over the period 18 March to 4 May 2021 and is now published in The Lancet.[28] This study sought to understand if students and staff could continue to attend school safely after they had been in close contact with a person who was positive for COVID-19. DCT would support a student's education to continue and reduce the health and wellbeing impacts of self-isolation. Consenting students and staff identified as contacts would test daily for 7 days with LFDs (intervention) or self-isolate at home for 10 days (control). Within the intervention group, if a student or staff member did not have a positive result, they were able to take part in school activities for the day on which they had tested. The study hypothesis was that the intervention arm (daily contact testing) would have increased school attendance compared to the control arm (self-isolation) (that is, superiority) and the level of transmission in the schools in the intervention arm (daily contact testing) was not inferior to (that is, not higher than in) the control arm (self-isolation). The co-primary outcomes of the study were symptomatic PCR-confirmed COVID-19 to estimate in-school transmission and COVID-19-related school absence.

There were limitations acknowledged in the study method with appropriate mitigations in place. The key limitations were that participating schools and colleges were not always able to participate due to competing pressures, the study was reliant on linkage to NHS T&T data, using incidence data meant within-school transmission was not a directly measure (this was estimated by controlling for the rate of community infections, as a proxy for the extent of introductions into the school), the trial period was over low to moderate SARS-CoV-2 incidence.

Results

In total, 201 schools were randomly assigned to either the intervention group (n=102) or the control group (n=99). Over the 10-week period, 2,432 (42.4%) of the 5,763 intervention group contacts participated in DCT.

[28] Young and colleagues. 'Daily testing for contacts of individuals with SARS-CoV-2 infection and attendance and SARS-CoV-2 transmission in English secondary schools and colleges: an open-label, cluster-randomised trial.' Lancet 2021: volume 398 (10,307) pages 1,217 to 1,229. DOI: 10.1016/S0140-6736(21)01908-5

Results for the co-primary outcomes of the study:

- symptomatic PCR-confirmed COVID-19 to estimate in-school transmission; there were 657 symptomatic PCR-confirmed infections during 7,782,537 days-at-risk in the control group and 740 during 8,379,749 days-at-risk in the intervention group
- COVID-19-related school absence; among students and staff, there were 59,422 (1·62%) COVID-19-related absences during 3,659,017 person-school days in the control group and 51,541 (1·34%) during 3,845,208 person-school days in the intervention group

The study concluded that DCT in schools was non-inferior to self-isolation for control of COVID-19 transmission. They found similar rates of symptomatic infections among students and staff when the control group and intervention groups were compared.

A non-inferiority randomised controlled trial comparing the risk of onward transmission from contacts using DCT in comparison to self-isolation

A national study of DCT has now been published in Lancet Respiratory Medicine.[29] The trial was operated as a non-inferiority study to assess if individuals and households following a DCT protocol (daily LFD testing for 7 days) would not lead to an increase of transmission in comparison to a control group of standard self-isolation guidance (for 10 days). The study prospectively followed-up on participants through NHS Test and Trace contact tracing data to assess the impact on transmission within each arm.

The study was a large RCT in a real-world setting but did have some limitations noted by the study team. Key limitations included were that the study was reliant on self-notification of close contacts, and it was not possible to assess the risk of transmission beyond named close contacts, only contacts who accessed testing could subsequently be identified as a case, and national restrictions were in place at the time with participants advised to minimise contact when taking part. There was also a skew noted in the DCT arm of the study towards individuals who were able to work from home.

Results

Results are now available in the published paper. In total, 49,623 participants consented to be recruited into the study. Of the participants, 26,123 were allocated to the DCT arm of the study and 23,500 allocated to the self-isolation control group. Results for the primary outcome of the study, non-inferiority of DCT in comparison to self-isolation, showed that the transmission rate among secondary contacts (the tertiary attack rate) was 7.49% in the self-isolation control arm

[29] Love and colleagues. ‘Daily use of lateral flow devices by contacts of confirmed COVID-19 cases to enable exemption from isolation compared with standard self-isolation to reduce onward transmission of SARS-CoV-2 in England: a randomised, controlled, non-inferiority trial.’ Lancet Respiratory Medicine 2022: volume 10, pages 1,074 to 1,085 https://doi.org/10.1016/S2213-2600(22)00267-3

and 6.40% in the DCT arm of the study. This is a difference of -1.09 % (95% CI: -2.16% to -0.03%).

The study concluded that DCT, in the form of daily LFD testing for 7 days, was non-inferior to self-isolation on assessment of transmission rates. Findings showed a protocol of DCT supported release of each participant from self-isolation for an average of 5.4 days. Qualitative assessment showed participants thought DCT was “a sensible, feasible and welcome means of avoiding unnecessary self-isolation” whilst, crucially, continuing to restrict non-essential activity as recommended.

An evaluation of the pilot of daily contact testing of healthcare workers in NHS acute hospital and ambulance trusts

DCT has been piloted in the NHS workforce to support NHS England and NHS Improvement (NHSE/I) to make informed decisions on a wider rollout. The pilot was run between 9 January and 28 February 2021 during a period when all identified contacts of a coronavirus case were required to self-isolate for 10 days. The small pilot study was conducted in 4 acute hospital trusts and one ambulance trust as a measure to understand the feasibility of implementing DCT in this workplace whilst maintaining essential services. The complete results have now been published in the journal Public Health.[30]

The pilot testing regime aimed to find cases whilst minimising the number of days spent in unwarranted self-isolation. Healthcare workers were able to participate if they had been identified as a close contact through workplace tracing or by notification through NHS Test and Trace (NHS T&T). Eligibility was extended at the midpoint of the pilot to individuals who were contacts of a positive household member if the participating individual had tested positive by PCR within the previous 90 days. Participants in the pilot trusts were asked to self-test using an LFD each day for 7 days. If the LFD result was negative, they could continue to work that day.

In total, aggregate data from 138 eligible contacts was reported with 111 individual-level data reports for each of the DCT participants. All participants were invited to complete a survey about their experience with 58 of the 138 who were eligible responded (42%). A further 18 telephone interviews were conducted with participants, site leads, administrators and union representatives. The Infection Prevention Control (IPC) leads for each trust were interviewed to further establish whether any transmission incidents related to the pilot had occurred.

Results

The pilot found that, with good engagement, the introduction of DCT was welcomed by healthcare workers within the pilot trusts. Uptake was high with 80% of eligible healthcare worker contacts participating. Participants found the acceptability to be high and IPC leads rated

[30] Bow SMA, Goddard A, Cope G, Sharp N, Schick J, Woods C, Jeffery K, Harrington D, Williams S, Rodger AJ, Finer S, Fowler T, Hopkins S and Tunkel SA (2022). ‘An evaluation of a pilot of daily testing of SARS-CoV-2 contacts in acute hospital and ambulance trusts in England’ Public Health volume 209, pages 46 to 51 https://doi.org/10.1016/j.puhe.2022.05.013

high confidence in the detection of any workplace transmission due to DCT with no reported incidents of onward transmission in this context. An initial set-up burden was noted by pilot trusts.

In total, 719 LFD tests were taken as part of the pilot with a mean of 6.5 tests per participant. Over the course of the pilot, one participant tested positive on LFD during their DCT period which was subsequently confirmed by PCR. A number of participants also received additional PCR testing through different mechanisms. These included testing of asymptomatic staff during a ward outbreak and 3 of the 5 pilot trusts chose to add PCR tests to the DCT regime. A total of 75 PCR results were reported, 59 of which could be matched to an LFD result on the same day. Of these participants, pilot trusts reported 5 participants tested positive on PCR, one outside the DCT testing period. There were 3 participants who had consistently tested negative on LFD (the PCR cycle threshold (Ct) values in 2 of these cases suggested this may have been due to low viral concentration). Sequencing of samples was conducted for 2 cases: the sequence of the sample for one did not match the index case, implying that the infection was not related to the index case exposure that led to participation in DCT.

Overall, the results outlined above show one participant tested positive on LFD and 4 participants tested positive on PCR during their respective pilot DCT testing periods. The study design did not allow for conclusions on whether this was greater than the number of cases than would have been detected in the absence of DCT (that is, self-isolation and voluntary twice weekly testing). However, there was no sign that introduction of DCT led to an increase in the number of positive LFD results in the wider workplace.

The evaluation team estimated that the pilot averted a total of 682 potential days of work absence (a mean of 136 days per trust), with 90% of these associated with clinical staff. The short time frame and devolved delivery model of the pilot enabled rapid generation of evidence for decision-making regarding wider rollout. However, the lack of a pre-defined control group precluded a direct comparison of DCT against self-isolation and the small size of the pilot meant that findings had limited statistical power. The evaluation also noted difficulty engaging with those who declined to take part and an inherent risk of bias in relying on the views of those involved in administering or participating in DCT.

Review of the Workplace DCT Pilot[31]

DCT in private institutions managed by NHS Test and Trace took place between December 2020 and March 2021. Instead of self-isolating at home, contacts of positive COVID-19 cases were asked to take LFD tests every day for up to 7 consecutive days. They were able to resume normal activity, including attending work, in the 24 hours following a negative test result. If they tested positive, they ceased daily testing and needed to self-isolate for the next 10 days.

The evaluation drew on the data provided by participating organisations and from the NHSTT data systems – online surveys with employees and test site administrators, for which over 1,400

[31] The full report - COVID-19: overview of daily contact testing (DCT) trial reports – is available online.

and 60 responses were received respectively plus interviews with 40 employees and test site administrators in 3 out of the 13 organisations that participated.

Results

In terms of operational feasibility, implementation of DCT was reported to be relatively easy. However, the evaluation noted huge financial and time investment incurred in the set up.

Suggestions for future improvements included provision of more support materials to explain how regular LFD testing can be trusted to replace self-isolation, with more reassurances to show that DCT is a safe (but not mandatory) option within the workplace. Other findings included that testing experience was easy with some difficulties in keeping up with changes to the operating procedures, barriers for employees and employers in accurately recording data on DCT participation and compliance due to the digital process for registering LFD tests for DCT being the same as that for asymptomatic testing for other reasons. There was evidence that calm and experienced test administrators played a key role in alleviating any discomfort that people had with being tested.

Those who took public transport to the workplace reported concerns about infecting others during the commute and cited it as a barrier to their participation.

It was noted that the guidance requiring people to wait for test results in a holding area before going into the workplace was inconsistently followed. Also, it was found that organisations needed more support to prove the credibility of DCT and that it was an officially sanctioned public health intervention. In terms of increased levels of infection at DCT sites, there were no apparent signals found.

The positivity rates were higher in DCT workplaces compared to non-DCT workplaces and the cause is unlikely to be DCT itself. The difference is present from the start of the pilot period. The allocation of workplaces to DCT/non-DCT groups was not randomised. Positivity rates fell at both DCT and non-DCT workplaces during the study.

The pilot found that just under 3 in every 5 (57%) of those individuals participating in DCT said that they would be more likely to name their contacts if they were to test positive and if DCT were available to their contacts. The decision to participate in DCT was primarily driven by individuals' sense of civic duty, curiosity about whether they had the virus, and the ease of the process, but also a recognition that their employer wanted them to do it. Financial factors also reportedly influenced uptake, in particular for those on pro-rata wages, those who regularly worked overtime and some agency workers.

Barriers for participation included concerns about infecting others during the 7-day testing period and being treated with some wariness by others in the workplace because they were regarded as carrying a higher risk of being infectious.

Following DCT, it was found that the majority of participants did not change either the amount they left home (75%) or the amount of contact they had with people outside their household (67%) following a negative result.

Overall, DCT reduced the number of working days lost to self-isolation and reduced anxiety around coronavirus in their workforce. A majority of test site administrators (82%) and employees (73%) reported that they would definitely participate in DCT again.

Modelling developed using pilot data indicates that the net economic impact of DCT is highly dependent on the cost to the business of individuals who are required to self-isolate. If they are just as productive while working from home, don't get sick pay, or any output loss can be covered at minimal net cost, then it is not clear that the employer would see value for money (VfM) from daily testing; where the opposite is true, they might well. Therefore, DCT will not represent VfM for all workplaces and needs to be appropriately targeted.

It is important to note that this was a relatively small pilot study, and its limitations mean that the evidence cannot be considered definitive proof that the introduction of DCT does not increase the risk of workplace transmission. The lack of high-quality data on DCT participants limited the evaluation analysis and precluded a direct comparison of the effects of DCT versus quarantining of close contacts. Therefore, any wider application of DCT would need careful evaluation and monitoring.

Post-deployment device evaluation

Impact of vaccination on LFD performance

LFD performance in vaccinated and unvaccinated individuals has been evaluated as part of a programme of ongoing service evaluation of testing regimes and devices.

Evaluation method

Individuals were identified by convenience sampling from those who had registered and attended an NHS Test and Trace Regional Test Site (RTS) for the purposes of receiving a diagnostic test and included symptomatic and asymptomatic individuals over the period 22 May to 6 July 2021. Participants agreed to take part in the evaluation, were 16 years old or older, and excluded those for whom a throat swab would not be possible or with cuts in their nose or healing nose piercings. Each participant performed an NHS Test and Trace COVID-19 (Innova) Self-Test LFD and a self-swab for a reference standard test, qRT-PCR. These were compared to give an evaluation of the LFD performance and provide genomic sequencing data. LFD results, as interpreted by the participant, were reported to the NHS Test and Trace evaluation team irrespective of the result. The reference standard test was sent to a Lighthouse Laboratory and was not available to the evaluation team at time of testing. Within this evaluation, a total of 694 subjects were included.

Comparison was made of LFD sensitivity rates between vaccinated (partially or fully) and unvaccinated individuals, taking into account adjustments for clinical and demographic factors, vaccination status and viral concentration. Logistic regression and propensity score modelling were used to model the probability of a true positive when comparing groups and exact 95% confidence intervals reported where appropriate.

Results

The evaluation showed sensitivity of the NHS Test and Trace COVID-19 self-test LFD to be generally comparable between vaccinated and unvaccinated groups regardless of the timing of testing in relation to vaccination.

Sensitivity of LFDs did not differ between those vaccinated and those unvaccinated (Table 11). The likelihood of a TP did not depend on vaccination status, after adjusting for other factors. The chance of a True Positive increased with viral concentration (OR=2.25; $p<0.001$).

Table 11. Sensitivity by vaccination status

Vaccine status	Unvaccinated[32] (n=312) % (95%CI)	Partially vaccinated 1[33] (n=71)	Partially vaccinated 2[34] (n=201)	Fully vaccinated[35] (n=110)
Sensitivity	64.19 (58.58,69.53)	68.57 (56.37, 79.15)	62.63 (55.49, 69.38)	62.73 (52.99, 71.76)
Difference (in comparison to unvaccinated)	N/A	4.38 (-8.22, 15.43; p=0.48)	1.56 (-6.84, 10.12; p=0.72)	1.46 (-8.61, 12.09; p=0.78)

[32] Unvaccinated: those at the time of testing who had not received any doses of vaccine.
[33] Partially vaccinated 1 (PV1): those who were tested within 21 days of their first dose.
[34] Partially vaccinated 2 (PV2): those who were tested over 21 days after their first dose and did not receive a second dose, or were tested within 14 days of their second dose.
[35] Fully vaccinated: those tested after 14 days post second vaccination dose.

Post-deployment innovation and improvement

Evaluation of machine learning for LFD results interpretation and reporting

LFD tests have been used extensively across the UKHSA testing programme to support rapid access to testing for SARS-CoV-2 with minimal user training required. Although intuitive, reading and interpreting an LFD result can be affected by variation between individuals reading the test. Some individuals can also find the reading and interpretation of an LFD challenging.

Research has been performed using an AI algorithm based on machine learning to interpret results. The rationale of the evaluation was to understand the performance of such a reader and if variability in interpretation of results could be reduced. Initial training of the algorithm used baseline photos from SARS-CoV-2 spiked samples and subsequently by LFD results linked to a qRT-PCR swab sample. Quality Control (QC) processes were also embedded through training of the algorithm including, for example, assessment of image quality.

Proof of concept diagnostic accuracy study

The diagnostic accuracy study compared the AI against 2 sample populations, site operatives at Asymptomatic Testing Sites (ATS) and health and social care staff self-testing as part of a routine testing regime. Participants within each sample population were invited to take part in the evaluation through convenience sampling and included if they were willing and able to participate and did not have any common COVID-19 symptoms at the time. At ATS, the trained operative who read the LFD test took a photo of the LFD and uploaded the image along with their interpretation of it to a web-based portal for the image to be read by the machine-learning algorithm. In the self-test group, the participant did this themselves and uploaded the image to the NHS Digital web service. The LFD trialled was Innova only.

AI sensitivity and specificity were calculated against ATS staff or self-testers. An expert panel resolved any discrepant results to create a ground truth against which the AI results were further compared to produce a final sensitivity and specificity.

Results

As detailed in publication[36], the team reported positive results from the machine learning based AI algorithm. The sample size included a total of 59,164 images from ATS and 58,667 from self-test health and social care staff, of which 56,776 and 50,999 were valid and suitable for analysis respectively.

[36] LFD AI Consortium. ‘Machine learning for determining lateral flow device results for testing of SARS-CoV-2 infection in asymptomatic population.’ Cell Reports Medicine October 2022: volume 3, issue 10. 100784. DOI: 10.1016/j.xcrm.2022.100784.

For ATS, the AI returned a sensitivity against ground truth of 97.6% (123 out of 126) and specificity of 99.95% (56,488 out of 56,542). In the self-test population, the AI sensitivity compared to ground truth was 100% (30 out 30) and specificity 99.28% (32,750 out of 32,986).

These sensitivities and specificities did not use PCR as the reference test and are not the sensitivity of the LFD but rather the sensitivity of the AI at accurately reading the LFD. These metrics provided proof of concept of the ML-based AI algorithm on which the further evaluation detailed below was based.

Evaluation of LFD digital reader[37]

A further evaluation of the digital reader (DR) was conducted to understand its real-world effectiveness, ease of use and value for money. Nearly 1.5 million reports and over 1.8 million images were submitted by staff from adult social care, primary care and some private sector employers who were taking part in regular asymptomatic testing.

Results

Output accuracy was assessed by comparing the results from the DR to the results from an expert panel ('ground truth') for 2,000 randomly selected images. Sensitivity was 73.3% (11 out of 15, 95%CI: 44.9, 92.2) and specificity 99.95% (1,933 out of 1,934, 95%CI: 99.71, 100).

When the LFD result was compared to the PCR result, digitally read LFDs showed a higher overall sensitivity, 72.6% (70.3, 74.9), than self-read LFDs, 57.9% (56.9, 58.8). The gain in sensitivity was marginal at high viral concentrations of above 1 million copies per mL (5.9%), 17% for medium viral concentrations (10,000 to 1 million copies per mL) and more than 100% for low viral concentrations (below 10,000 copies per mL). Specificity of the digitally read LFDs was 96.02% (95.68, 96.35) compared to self-read specificity of 99.77% (99.76, 99.78). It is worth noting that the DR algorithm is set to prioritise the detection of positive test results.

User journey performance was investigated by integrating multiple data sources which included web-portal analytics, post-marketing surveillance reports and individual level time-series data. Overall, users found that the DR process was easy to use with very high rates of successful use of the service at first attempt (indicating no user or technical issues) and no learning curve on subsequent success rates was observed. Familiarity with the service did result in users being able to complete the process faster. The median time taken by a user to complete the DR process (143 seconds) was the same as that taken to complete the current self-report process (144 seconds) indicating that there was no additional burden on users in using the service. Retention rates, defined as the proportion of people who used the service more than once, were generally high indicating again that the service was not burdensome for users. Females and those over 30 years of age (77.2%) generally had higher retention rates compared to males and the 10 to 29 age group (60%). The real-world impact of the latter is limited given that this younger age group is less likely to be susceptible to vision or cognition problems that could affect interpretation of LFT results.

[37] Full report available at COVID-19: LFD digital reader evaluation.

Value for money (VFM) modelling was undertaken to assess the benefits of deploying the DR at a population level. A worst-case-scenario approach to the modelling was used, taking the best-case performance of self-read LFDs and a reasonably optimistic outcome scenario and comparing it to a worst-case performance of the DR-LFD to give a pessimistic value for money analysis. For each context the DR data set was compared to all self-read results for ASC and primary care staff, assuming the same underlying true prevalence. Based on this, the DR demonstrated a relative improvement in sensitivity of 89.42% with the lower bound being 34.29% in ASC. The relative gain in sensitivity was less marked in primary care, but this remained significant with an average gain of 33.72% and a lower bound at 18.96%.

The model was extended further to include the observed average prevalence rate of 1% over the last 2 years and the current prevalence rate of 5%. At the current prevalence rate of 5%, DR would prevent an additional 4,353 first generation infections per million tests in adult social care, and 1,980 in primary care. Furthermore, considering a cost scenario of £0.3 per read, for 1.5 million tests and vaccination rate of at least 70% (vaccinated and boosted), a net benefit of at least £3.64 million would be observed in ASC, while the net benefit in primary care would be at least £2.34 million

Appendix A. Demographics: Performance evaluation of the SureScreen LFD V3 using double anterior nares swab collection in a self-test setting

Table 12. Data characteristics and concordance rates

Demographics and key characteristics	Statistic or category	Number (N=1,898)[38]
Age (years)	Median	31
	Minimum to maximum	(16 to 82)
Sex	Female	986 (51.9%)
	Male	910 (47.9%)
	Missing	2 (0.1%)
Symptoms (n,%)	Yes	964 (50.8%)
	No	877 (46.2%)
	Missing	57 (3%)
LFD result (n,%)	Positive	332 (17.5%)
	Negative	1,565 (82.5%)
	Void or missing	1 (0.1%)
	Drop out	0 (0.0%)
PCR result (n,%)	Positive	437 (23%)
	Negative	1,408 (74.2%)

[38] Age, sex, symptoms, LFD and PCR results include 54 voids and dropouts. Concordance, viral concentration and vaccination status has voids and dropouts removed.

Demographics and key characteristics	Statistic or category	Number (N=1,898)[38]
	Void or missing	53 (2.8%)
	Drop out	0 (0.0%)
Concordance (n,%)	Concordant	1,730 (93.8%)
	True positive	327 (17.7%)
	True negative	1,403 (76.1%)
	Discordant	114 (6.2%)
	False positive	4 (0.2%)
	False negative	110 (6.0%)
Viral concentration (copies/mL) (n,%)	Under 10,000	68 (3.7%)
	10,000 to 1 million	142 (7.7%)
	Over 1 million	227 (12.3%)
	Negative	1,407 (76.3%)
Vaccination status (n,%)	One dose	436 (23.6%)
	Two doses	1,192 (64.6%)
	None	216 (11.7%)

Appendix B. Demographics: Performance evaluation of dual mid-turbinate swab collection for LFD antigen self-testing

Table 13. Data characteristics and concordance rates

	Statistic or category	Dual mid-turbinate service evaluation (N=1,131)	Throat and single mid-turbinate evaluation (N=635)
Age (years)[39]	Median	31	32
	Minimum to maximum	(16 to 79)	(0 to 75)
Sex	Female	540 (47.7%)	300 (47.4%)
	Male	591 (52.3%)	333 (52.6%)
Symptoms (n,%)	Yes	569 (50.3%)	454 (71.5%)
	No	525 (46.4%)	109 (17.2%)
	Missing	37 (3.4%)	72 (11.3%)
LFD result (n,%)	Positive	233 (20.6%)	415 (65.5%)
	Negative	874 (77.3%)	216 (34.1%)
	Void	24 (2.1%)	3 (0.5%)
PCR result (n,%)	Positive	315 (27.9%)	635 (100%)
	Negative	798 (70.6%)	0
	Void	18 (1.6%)	0

[39] There were 8 individuals who were under 16 years or had no age information and were filtered out of analysis.

	Statistic or category	Dual mid-turbinate service evaluation (N=1,131)	Throat and single mid-turbinate evaluation (N=635)
Concordance (n,%)[40]	Concordant	1005 (88.9%)	415 (65.4%)
	True Positive	230 (20.3%)	
	True Negative	775 (68.5%)	
	Discordant	84 (7.4%)	220 (34.6%)
	False Positive	2 (0.2%)	
	False Negative	82 (7.3%)	
Viral concentration (copies/mL) (n,%)	Under 10,000	63 (5.6%)	67 (10.6%)
	10,000 to 1 million	110 (9.7%)	218 (34.3%)
	Over 1 million	155 (13.7%)	350 (55.1%)
	Negative	803 (71.0%)	
Vaccination status (n,%)	One dose	414 (36.6%)	204 (32.2%)
	Two doses	529 (46.8%)	141 (22.2%)
	None	188 (16.6%)	289 (45.6%)
	Total		

[40] All LFD or PCR void samples removed from this calculation.

Appendix C. Evaluation of same-day dual testing LFD performance

Table 14. Data characteristics and concordance results

	Statistic or category	Dual anterior nares (N=2,231)
Age (years)	Median	36
	Minimum to maximum	(16 to 90)
Sex	Female	1,072 (48.1%)
	Male	1,156 (51.8%)
	Missing	3 (0.1%)
Symptoms (n,%)	Yes	1,402 (62.8%)
	No	819 (36.7%)
	Prefer not to say	10 (0.5%)
Viral concentration (copies/mL) (n,%)	Less than 10,000	31 (1.4%)
	10,000 to 1 million	60 (2.7%)
	Over 1 million	68 (3.1%)
	Negative	2,029 (92.7%)
	Total	2,188 (100%)
Vaccination status (n,%)	One dose	573 (25.7%)
	Two doses	214 (9.6%)
	None	1,441 (64.6%)

	Statistic or category	Dual anterior nares (N=2,231)
	NA	3 (0.1%)
	Total	2,231 (100%)

Table 15. LFD and PCR results

Test result	LFD 1	LFD 2	Combined LFD	PCR
Positive	119 (5.3%)	117 (5.2%)	126 (5.6%)	159 (7.1%)
Negative	2,073 (92.9%)	2,087 (93.6%)	2,091 (93.7%)	2,029 (90.9%)
Void	31 (1.4%)	19 (0.9%)	6 (0.3%)	35 (1.6%)
Drop out or missing	8 (0.4%)	8 (0.4%)	8 (0.4%)	8 (0.4%)

Table 16. Concordance results

Test result	LFD1 result	LFD2 result	Combined LFD
Concordant	2,094 (97.1%)	2,109 (97.3%)	2,118 (97.1%)
Concordant: True positive	106 (4.9%)	106 (4.9%)	109 (5.0%)
Concordant: True negative	1,988 (92.2%)	2,003 (92.4%)	2,009 (92.1%)
Discordant	62 (2.9%)	59 (2.8%)	63 (2.9%)
Discordant: False positive	10 (0.5%)	7 (0.3%)	13 (0.6%)
Discordant: False negative	52 (2.4%)	52 (2.4%)	50 (2.3%)
Total	2,156 (100%)	2,168 (100%)	2,181 (100%)

Appendix D. Evaluation of multi-day antigen LFD self-testing performance

Table 17. Demographic data of all participants

Characteristic	All participants (N = 2,788)	Regime 1 (N = 965 (34.6%))	Regime 2 (N = 1,289 (46.2%))	Regime 3 (N = 995 (35.7%))
Mean age in years (SD)	37.45 (12.21)	40.98 (12.22)	39.55 (11.94)	41.04 (12.18)
Median age in years (range)	35.00 (0.00, 85.00)	40.00 (0.00, 78.00)	38.00 (0.00, 78.00)	40.00 (0.00, 78.00)
Age under 50	2,266 (81.3%)	704 (73.0%)	999 (77.5%)	726 (73.0%)
Age 50 and over	522 (18.7%)	261 (27.0%)	290 (22.5%)	269 (27.0%)
Sex: female	1,461 (52.4%)	528 (54.7%)	708 (54.9%)	544 (54.7%)
Sex: male	1,320 (47.3%)	435 (45.1%)	578 (44.8%)	449 (45.1%)
Sex: unknown	7 (0.3%)	2 (0.2%)	3 (0.2%)	2 (0.2%)
Asymptomatic	1,201 (43.1%)	472 (48.9%)	601 (46.6%)	490 (49.2%)
Symptomatic	1,587 (56.9%)	493 (51.1%)	688 (53.4%)	505 (50.8%)
VC under 100 copies/mL (% of positives)	71 (11.4%)	22 (8.1%)	29 (8.8%)	22 (7.9%)
VC = 100 to 1,000 copies/mL (% of positives)	37 (5.9%)	14 (5.1%)	17 (5.2%)	14 (5.0%)

Characteristic	All participants (N = 2,788)	Regime 1 (N = 965 (34.6%))	Regime 2 (N = 1,289 (46.2%))	Regime 3 (N = 995 (35.7%))
VC = 1,000 to 10,000 copies/mL (% of positives)	59 (9.5%)	22 (8.1%)	29 (8.8%)	22 (7.9%)
VC = 10,00 to 100,000 copies/mL (% of positives)	77 (12.4%)	35 (12.8%)	40 (12.2%)	37 (13.3%)
VC = 100,000 to 1 million copies/mL (% of positives)	136 (21.8%)	63 (23.1%)	76 (23.1%)	65 (23.4%)
VC = 1 million to 10 million copies/mL (% of positives)	169 (27.1%)	82 (30.0%)	94 (28.6%)	83 (29.9%)
VC = over 10 million copies/mL (% of positives)	74 (11.9%)	35 (12.8%)	44 (13.4%)	35 (12.6%)
Negative	2,165 (77.7%)	692 (71.7%)	960 (74.5%)	717 (72.1%)
Unvaccinated	217 (7.8%)	27 (2.8%)	46 (3.6%)	27 (2.7%)
Partially vaccinated	84 (3.0%)	16 (1.7%)	21 (1.6%)	16 (1.6%)
Fully vaccinated	2,487 (89.2%)	922 (95.5%)	1,222 (94.8%)	952 (95.7%)

Table 8. Demographic data of participants that fit the low-risk symptomatic criteria

Characteristic	All subjects (N = 1,150)	Regime 1 (N = 359 (31.2%))	Regime 2 (N = 527 (45.8%))	Regime 3 (N = 367 (31.9%))
Mean age in years (SD)	33.86 (7.94)	35.35 (7.58)	34.76 (7.65)	35.44 (7.55)
Median age in years (range)	34.00 (16.00, 49.00)	35.00 (18.00, 49.00)	35.00 (16.00, 49.00)	35.00 (18.00, 49.00)
Age under 50	1,150 (100%)	359 (100%)	527 (100%)	367 (100%)
Sex: female	638 (55.5%)	204 (56.8%)	298 (56.5%)	207 (56.4%)
Sex: male	510 (44.3%)	155 (43.2%)	228 (43.3%)	160 (43.6%)
Sex: unknown	2 (0.2%)		1 (0.2%)	
Symptomatic	1,150 (100%)	359 (100%)	527 (100%)	367 (100%)
VC under 100 copies/mL (% of positives)	22 (6.9%)	5 (3.1%)	9 (4.6%)	5 (3.0%)
VC = 100 to 1,000 copies/mL (% of positives)	16 (5.0%)	8 (4.9%)	9 (4.6%)	8 (4.8%)
VC = 1,000 to 10,000 copies/mL (% of positives)	27 (8.4%)	13 (8.0%)	15 (7.7%)	13 (7.9%)
VC = 10,000 to 100,000 copies/mL (% of positives)	38 (11.8%)	22 (13.5%)	25 (12.9%)	23 (13.9%)
VC = 100,000 to 1 million copies/mL (% of positives)	87 (27.1%)	41 (25.2%)	49 (25.3%)	41 (24.8%)
VC = 1 million to 10 million copies/mL (% of positives)	86 (26.8%)	51 (31.3%)	56 (28.9%)	52 (31.5%)
VC = over 10 million copies/mL (% of positives)	45 (14%)	23 (14.1%)	31 (16%)	23 (13.9%)
Negative	829 (72.1%)	196 (54.6%)	333 (63.2%)	202 (55.0%)
Fully vaccinated	1,150 (100%)	359 (100%)	527 (100%)	367 (100%)

Table 9. Number of true positive (TP), false positive (FP), true negative (TN) and false negative (FN) for the low-risk symptomatic group

Subgroup	TP	FP	TN	FN	Total[41]
All Subjects Regimen 1	249	16	669	9	943
All Subjects Regimen 2	290	18	931	22	1261
All Subjects Regimen 3	252	15	695	11	973
Age under 50 – FV: symptomatic Regimen 1	155	1	193	2	351
Age under 50 – FV: symptomatic Regimen 2	180	3	325	7	515
Age under 50 – FV: symptomatic Regimen 3	155	1	199	4	359

[41] Total not including voids.

Glossary

Term	Meaning
Case of COVID-19	An individual with COVID-19 infection identified with a positive PCR result for SARS-CoV-2 virus, irrespective of the cycle threshold.
Cycle threshold (Ct)	The number of cycles at which the virus is detectable in a qRT-PCR test.
Detection rate	Among those subjects with matched pairs of LFD/PCR, the number of subjects with positive LFD divided by the number of subjects with positive PCR. This metric can be presented as a proportion or as percentage.
False negative (FN)	An LFD negative result that corresponds with a PCR positive result.
False positive (FP)	An LFD positive result that corresponds with a PCR negative result.
Local test site (LTS)	Small local test site providing diagnostic PCR tests for those with symptoms of COVID-19.
Regional test site (RTS)	Large test site for whole region providing diagnostic PCR tests for those with symptoms of COVID-19.
True negative (TN)	An LFD negative result that is confirmed by a PCR negative result.
True positive (TP)	An LFD positive result that is confirmed by a PCR positive result.
Viral concentration (VC)	The number of SARS-CoV-2 RNA viral copies per millilitre of sample estimated by the cycle threshold in the PCR result.

About the UK Health Security Agency

UKHSA is responsible for protecting every member of every community from the impact of infectious diseases, chemical, biological, radiological and nuclear incidents and other health threats. We provide intellectual, scientific and operational leadership at national and local level, as well as on the global stage, to make the nation health secure.

UKHSA is an executive agency, sponsored by the Department of Health and Social Care

Prepared by: Public Health and Clinical Oversight (PHCO), Testing Operations, UKHSA

For queries relating to this document, please contact: ukhsa-pressoffice@ukhsa.gov.uk

Published: March 2023
Publishing reference: GOV-14281