Laboratory evaluation of 185 commercial assays for detecting SARS-CoV-2: the UK response for mass testing

Abbie Bown,[a,i,*] Tim Peto,[b,c,d,i] Angela Sweed,[a] Matthew Catton,[a] Joshua Nelthorpe-Cowne,[a] Tracy Benford,[a] LFD Laboratory Team,[g] and LFD Laboratory Support Team,[h] Tom Collinge[e] Donna Robinson,[e] Louise Oliver,[e] Susan Hopkins,[f] Derrick Crook,[c] Somya Agrawal,[e] Hannah Fordham,[e] Tom Fowler,[f] Alex Sienkiewicz,[a] Ann Sarah Walker,[b,c,d] Sir John Bell,[c] and Richard Vipond[a]

[a]UK Health Security Agency, Porton, UK
[b]Oxford University Hospitals NHS Foundation Trust, Oxford, UK
[c]Nuffield Department of Medicine, University of Oxford, Oxford, UK
[d]National Institute for Health Research Health Protection Research Unit in Healthcare Associated Infections and Antimicrobial Resistance, University of Oxford, Oxford, UK
[e]Testing Operations (COVID 19), UK Health Security Agency, UK
[f]Clinical and Public Health Group, UK Health Security Agency, UK

eClinicalMedicine 2025;■: 103416
Published Online XXX
https://doi.org/10.1016/j.eclinm.2025.103416

Summary

Background In August 2020, Public Health England and Oxford University were commissioned to design and deliver (with NHS Test and Trace, NHSTT) a rapid evaluation programme of antigen Lateral Flow Devices (LFDs) for SARS-CoV-2 for mass community testing.

Methods A three-phase evaluation process was established: 1) desktop review of kits including claimed performance and supply; 2) laboratory testing with laboratory-grown SARS-CoV-2 virus and SARS-CoV-2 virus PCR negative volunteer samples; and 3) larger-scale laboratory testing of SARS-CoV-2 PCR positive and negative clinical samples. Variant of Concern (VOC) identification in the UK (December 2020), expanded laboratory methodology. Processes also evolved to improve workflow (irradiated viral stocks, dilution matrices, sample volumes, and replicates).

Findings Overall, 1017 kits were screened at phase 1, 185 kits tested at phase 2 and 91 at phase 3. Sixteen kits failed phase 3 due to poor performance and eight more failed to detect VOC satisfactorily. Sixty-four kits were redesigns of previously failed kits. The overall pass rate for the laboratory evaluation was 35% and 5 kits were procured for the UK National Covid 19 Testing Programme.

Interpretation The evaluation results had potential, time limited commercially sensitive aspects, and public sharing was limited to kits passing phase 3. Until now, the full data set has not been published. Over 2.5 billion self-test kits were deployed by the UK government following purchasing decisions informed by this work. We offer a potential blueprint for future evaluation programmes that might be required to assess LFDs to detect cases of a pandemic novel pathogen.

Funding This study was funded by UK Department of Health and Social Care (UKHSA); UK Health Security Agency (formerly Public Health England and the National Health Service Test and Trace); and the University of Oxford NIHR Biomedical Research Centre.

Keywords: Lateral flow device; SARS-CoV-2; LFD; Mass-testing; Pandemic; Public health; Coronavirus; Laboratory evaluation

*Corresponding author. UK Health Security Agency, Porton Down, Salisbury, Wiltshire, SP4 0JG, UK.
E-mail addresses: Abbie.Bown@UKHSA.gov.uk (A. Bown), Richard.Vipond@UKHSA.gov.uk (R. Vipond).
[g]Rachael Wallis[a], Hermione Conti-Frith[a], Alexandra Alexandridou[a], Rebecca Gill[a], Clara Milroy[a], Deborah Fox McKee[a], Prem K Perumal[a], Elizabeth Truelove[a], Sian Tiley[a], N. P. McLeod[a], Cathy Rowe[a], Peter Hammond[a], Bassam Hallis[a], Raghavendran Kulasegaran-Shylini[f], Ed Blandford[f], Anna Gronert[e], Ali Vaughan[b].
[h]Ruth Elderfield[a], Deborah Wright[a], Jim Chadwick[a], Beak-San Choi[a], Daniel Bailey[a], Silvia D'Arcangelo[a], Ash Otter[a], Jodie Owen[a], Natalie E. Baker[a], Annie Richardson[a], Kit Withycombe[a], Jack Nevols[a], Becky Rayment[a], Lauren Setterfield[a], Charlie Cooper[a], Mollie Curran-French[a], Kevin R. Bewley[a], Naomi S. Coombes[a], Kevin Stephen Richards[a], Christopher Burton[a], J. E. Burton[a], Bethany Hicks[a], Hayley Bradley[a], Joseph Gernon[a], Gabrielle Harker[a], Patrick Lahert[a], Olivia Carr[a], Caoimhe Kelly[a], Anaya Ellis[a], Babak Afrough[a], John Chamberlain[a], Stephen Findlay-Wilson[a], Lucy Flett[a].
[i]Joint 1st author.

Research in context

Evidence before this study

This study was commissioned in August 2020, before any SARS-CoV-2 antigen lateral flow devices (LFDs) were available on the UK market or any evaluations of the LFD's performance characteristics had been published.

Added value of this study

This study provided the performance evidence required by the NHS test and Trace (NHSTT) to successfully introduce a testing strategy whereby individuals could test themselves for COVID-19 and reduce community transmission. A description of the evaluation method and how this evolved during the study period is detailed. The dataset presented is unique and the largest assessment of SARS-CoV-2 antigen LFDs.

Implications of all the available evidence

The study offers a standardised assurance approach for the evaluation of LFDs that might be required in the future to detect cases of a pandemic novel pathogen.

Introduction

During the early stages of the COVID-19 pandemic, diagnostic testing solely relied upon testing of symptomatic individuals via reverse transcription polymerase chain reaction (RT-PCR, henceforth referred to as PCR) technology. PCR testing requires collection and laboratory testing of samples, involving significant infrastructure, expertise and equipment, resulting in high cost and comparatively long turn-around-times. Moreover, almost half of people who have COVID-19 never show any symptoms, so may not take a PCR test, but may be infectious.[1] As a potential game changer, cheaper, quicker and more accessible testing options to identify infectious cases and reduce community transmission were considered by the UK Government.[2]

From March to June 2020, Public Health England (PHE; now UK Health Security Agency) Porton Down evaluated several high throughput commercial laboratory assays for SARS-CoV-2 IgG antibodies,[3] including a collaborative evaluation with Oxford University.[4] Following this, in August 2020, the UK Government commissioned PHE Porton Down and University of Oxford to implement a rapid evaluation pipeline of antigen LFDs suitable for SARS-CoV-2 mass community testing.[5] The evaluation of tests was to begin within one week, reflecting the urgency of the commission. For NHS test and Trace (NHSTT) to successfully introduce a testing strategy whereby individuals could test themselves and reduce community transmission, it was vital that any tests being deployed were sufficiently sensitive, specific and robust. To be eligible to be purchased by NHSTT, amongst other requirements, tests needed to have passed the evaluation process described here.

At the time of the commission, a target product profile had only just been published by the Medicines and Healthcare products Regulatory Agency (MHRA).[6] UK manufacturing capacity was limited and the majority of candidate products identified by NHSTT were from overseas suppliers, particularly in the Far East. Performance claims from suppliers were limited in nature and, based on earlier experience in the pandemic with antibody LFDs, were treated with significant caution. The LFD evaluation process was refined over the duration of testing which extended far beyond initial timelines due to the success and scale of LFD use. The initial evaluation design was based on scientific knowledge at that point in time; reasons for changes were the emergence of variants, assay manufacturing changes, the use case for LFDs and advances in laboratory reagents as described herein.

Methods

Study design

A three-phase evaluation process was established (as previously described[5]) to shortlist commercial kits for procurement. The protocol was published on the UK government website[7] and updated with the benefit of learning and experience through the process. All protocol changes, decisions and results were approved by a committee of experts.

Briefly, phase 1 was a desktop review by NHSTT to identify suitably promising available kits. Phase 2 was initial laboratory testing of LFDs with laboratory grown SARS-CoV-2 virus and PCR negative saliva samples to initially assess sensitivity, specificity, robustness, usability and cross-reactivity with seasonal coronaviruses. This allowed scarce resources in terms of trained staff, suitable high containment laboratories and sample material to be dedicated to those that performed the best. Phase 3 was larger-scale laboratory testing of negative and positive clinical swab samples in viral transport media (VTM) to further assess sensitivity, specificity and interfering substances present in clinical samples.

Phase 1: desktop review

Manufacturers were invited to submit tests for the laboratory evaluation. Very broad screening selection criteria were used, mainly the products were antigen LFDs. Initially, performance data claims made by manufacturers in their instructions were considered,

however, by the end of September 2020 this was no longer used as a selection criterion due to the inconsistency of this data.[8] Other considerations were supply capacity and costings. High numbers of applications led to a review of the selection criteria to make them more discriminatory to prioritise resources.[9]

Laboratory evaluation

Lateral flow testing

Each LFD had 200 µL (unless stated below) of sample (detailed below) added to the LFD extraction buffer and subsequent steps were performed according to manufacturer's instructions (i.e. any incubation time in the buffer, number of drops to add to the LFD cassette and timeframe to read results). Results were read by trained operators at the appropriate time under strong artificial light provided by the Microbiological Safety Cabinet. Results were interpreted by eye as positive (control band and test band present), negative (control band present and test band absent) or kit failure (control band absent). Where positive, the strength of the bands was assessed by eye and photographs taken.

Phase 2: laboratory testing

Viral strains

Wild-type ancestral SARS-CoV-2 virus (Victoria/1/2020, kindly supplied by the Peter Doherty Institute for Infection and Immunity, Melbourne) was propagated on Vero/hSLAM cells (ECACC 04091501) at PHE Porton Down (as previously described[10]) to provide a standardised stock for the laboratory evaluation. Stocks were subjected to quality control checks, including whole genome sequencing, as previously described.[11]

Analytical sensitivity testing

To evaluate LFD analytical sensitivity,[12] dilutions (1×10^4, 1×10^3 and 1×10^2 plaque-forming units per millilitre (pfu/mL) of wild type virus stock in culture medium) were made in pooled anonymised saliva donated by SARS-CoV-2 PCR-negative healthy adult volunteers and each dilution applied to 15 LFDs. Samples of the lowest dilution were also applied in triplicate to Innova (control) LFDs, providing confidence between different batches of diluted material. To allow a comparison for the LFD results, the amount of virus in each dilution was also assessed by estimating the concentration of viral RNA by PCR. For the PCR testing, samples were inactivated in AVL buffer (Qiagen) supplemented with 5% Triton X-100™ (Sigma Aldrich) prior to SARS-CoV-2 RNA detection using the Roche cobas® RT-PCR system and the associated proprietary SARS-CoV-2 assay.

Initial clinical specificity testing

To assess specificity, fresh, anonymised saliva samples were donated by healthy adult volunteers and confirmed SARS-CoV-2 PCR-negative. From this large sample set, 71 samples were tested on each LFD.

Cross-reactivity testing

Seasonal coronaviruses (viral strains 229e, NL63 and OC43) were propagated and each applied to 5 LFDs.

Acceptance testing

Upon progression to phase 3, a second batch of LFD kits were requested from the manufacturers and tested in parallel using 30 specificity samples and 15 sensitivity samples at 1×10^2 pfu/mL (as above) to ensure consistent performance.

Phase 3: clinical swab samples in VTM testing

Clinical sensitivity testing

A panel of anonymised 200 SARS-CoV-2 PCR-positive VTM respiratory swab samples obtained from patients admitted to Oxford University Hospital Trust were selected to cover a range of viral concentrations (<1000–>100 million RNA copies/ml). Samples were diluted 1:3 or 1:4 in pooled SARS-CoV-2 PCR negative saliva, aliquoted into multiple 250 µL replicates and stored frozen (–20 to –80 °C) until testing to assess clinical sensitivity. Once thawed, one aliquot was tested by PCR (to confirm the new Ct value and the viral range extrapolated following the testing of known quantitative controls). Further aliquots were thawed and 100 µL tested by LFD.

Clinical specificity and interfering substances testing

1000 anonymised fresh SARS-CoV-2 PCR-negative VTM respiratory swab samples obtained from Oxford University Hospital Trust were tested to provide robust specificity data.

Evolution of methodology

Laboratory evaluation methodology evolved over time with the process being refined due to several reasons (Table 1). Bridging studies and test controls showed no impact from these changes which led to marked improvement in efficiency and sustainably of the processes.

In December 2020, the first Variant of Concern (VOC) was identified in the UK. These were declared based on their spike protein mutation(s), independent of changes in the nucleocapsid protein (the target of most LFDs). VOCs (Table 2) were tested using the same method as phase 2 sensitivity testing but with fewer replicates at each dilution: Dilutions (1×10^4, 1×10^3 and 1×10^2 pfu/mL of virus stock in culture medium) were made in pooled anonymised saliva donated by SARS-CoV-2 PCR-negative healthy adult volunteers and each dilution applied to 3 LFDs and tested by PCR. The method later changed to applying each dilution to 5 LFDs.

	Initial protocol	Amended protocol	Reason for change
Sample volume	200 µL	100 µL (April 2021)	Smaller swabs and concerns antigen: sample buffer was too dilute
Phase 2 virus	Laboratory grown virus stock	Irradiated laboratory grown viral stock (April 2021)	Safety and increased throughput at lower containment laboratories
Specificity samples and sensitivity sample diluent	Saliva	Artificial saliva[13](January 2021) Hanks' Balanced Salt Solution with 2 mg/ml porcine gastric mucin type II (HBSS-mucin) (April 2021)	To represent nasal secretions to align with the general shift from oral/throat swab to nasal swab sampling
Phase 2 specificity samples	71 saliva samples	72 nasal swab samples in HBSS-mucin (April 2021)	
Phase 2 sensitivity samples	Sample made fresh for each evaluation, 15 replicates of all dilutions	Frozen samples and replicate numbers reduced to 5 replicates at 1×10^4, 10 replicates at 1×10^3 and 15 replicates at 1×10^2 pfu/ml (April 2021)	Improved laboratory workflow
Phase 3 clinical panel	1 panel diluted in saliva capable of testing 10 LFD kits	5 panels each capable of testing between 10 and 28 LFD kits diluted in artificial saliva (January 2021, April 2021) or HBSS (July 2021, January 2023) to a similar range of viral concentrations to the original panel, as demonstrated by PCR and confirmed by testing of replicate panels	Insufficient samples prepared.
Reading of LFD results	Positive/Negative/Fail weak, very weak	Positive with band value 1–7 (extremely weak to strong)/Negative/Fail (January 2021)	Improved data and confidence in antigen results being compared to RNA viral concentration
Variant sensitivity testing	No variants present	alpha, beta, delta, gamma and omicron testing aligned to phase 2 sensitivity testing (when variants arose)	Concerns on sensitivity performance of kits previously passing evaluation with wild type virus

Table 1: **Summary of evaluation laboratory methodology changes.**

Initial variants tested were alpha (B.1.1.7), beta (B.1.351), gamma (P.1), and delta (B.1.617.2). When Omicron variants began circulating and the number of new variants arising quickly increased, VOC investigation criteria were implemented. These principles were i) only lineages with a mutation in the nucleocapsid protein were tested; ii) mutations were only evaluated once, even if there were multiple lineages or sub-lineages with the same mutation; and iii) only if the UK prevalence of a lineage (including sub lineages) with the new mutation was above a 5% baseline and was increasing. Between December 2021 and March 2024, the only variants meeting these criteria were BA.2, BA.4, BF.7,BQ.1, BE.3 and BA.2.86. Only a subset of LFDs (i.e. those used in the UK COVID-19 testing programme) were tested on these last 3 variants. Initial testing of variant strains on prioritised LFDs was done using freshly cultured viral stocks, with other kits being tested at a later date with irradiated stocks.

WHO designation	Pangolin lineage	GISAID ID of isolation swab
Ancestral virus	B	EPI_ISL_406844[14]
Alpha	B.1.1.7	EPI_ISL_683466
Beta	B.1.351	EPI_ISL_770441
Gamma	P.1	EPI_ISL_2080492
Delta	B.1.617.2	EPI_ISL_2742236
Omicron	BA.1	EPI_ISL_7400555
Omicron	BA.2	Not available
Omicron	BA.4	EPI_ISL_13157810
Omicron	BF.7	EPI_ISL_15386825
Omicron	BQ.1.13	EPI_ISL_15715180
Omicron	BE.3	EPI_ISL_14259085
Omicron	BA.2.86	EPI_ISL_18274087

Table 2: **SARS-CoV-2 variant strains used in this study.**

Data analysis and pass criteria

Sample sizes were chosen for practical purposes (what was available and what was logistically possible).

The criteria to progress through the evaluation stages are summarised in Table 3. To pass phase 2 sensitivity testing, the LFD results from the 10^2 pfu/ml dilution were reviewed and at least 9 of the 15 LFDs (60%) need to show a positive result. To pass the phase 2 specificity, no more than 2 of the LFDs tested could give a false positive result and no false positive results were allowed with the seasonal coronavirus samples. All LFD results from the 10^2 pfu/ml dilution for VOC tested needed to show a positive result. Finally, for phase 2 testing to be passed, less than 10% of the total number of LFDs tested at phase 2 were allowed to be kit failures.

To pass Phase 3 specificity testing, no more than 7 of the LFDs tested could give a false positive result. For robustness, less than 10% of the total number of LFDs tested at phase 3 were allowed to be kit failures.

Phase 3 sensitivity data was statistically analysed to predict the proportion of infections averted when applied to a 18291 test set was analysed following the method described by Lee et al.[15]

Briefly, the infectivity of PCR-Positive cases were determined by following their contacts identified by the UK Dept of Health and Social Care. Data from 18291

	Samples tested	Analysis	Pass criteria
Phase 2 sensitivity	Wildtype virus stock diluted to approximately 1×10^2 PFU/mL (n = 15)	(Number of kits LFD positive/Number of kits tested) × 100	60%
Phase 2 specificity	PCR negative volunteer respiratory samples (n = 71–72[a])	(Number of kits LFD negative/Number of kits tested) × 100	>97%
Phase 2 robustness	All phase 2 sensitivity and specificity test samples (n = 102–116[a])	(Number of kit failures/Number of kits tested) × 100	<10%
Phase 2 seasonal coronavirus testing	229e, NL63 and OC43 (n = 15)	(Number of kits LFD negative/Number of kits tested) × 100	100%
Phase 3 sensitivity	PCR positive VTM samples spanning an range of viral concentrations (n = 200)	Regression analysis to determine sensitivity at 10 K and 100 K RNA copies/mL when adjusted for dry swab use	40% at 10 K genome copies/mL or 75% at 100 K genome copies/mL
Phase 3 specificity	PCR negative VTM samples (n = 1000)	(Number of kits LFD negative/Number of kits tested) × 100	>99.3%
Phase 3 robustness	All phase 3 sensitivity and specificity test samples (n = 1200)	(Number of kit failures/Number of kits tested) × 100	<10%
VOC sensitivity	VOC virus stock diluted to approximately 1×10^2 PFU/mL (n = 3 or 5[a])	(Number of kits LFD positive/Number of kits tested) × 100	100%

[a]Represent changes to methodology.

Table 3: **Laboratory analyses and pass and fail criteria.**

positive cases were used as a reference set against which the performance of every lateral flow kit was assessed. The observed infectivity of each source was used to determine what proportion of infectious from the total cohort of 18291 could have been averted, had the lateral test kit been used to identify and isolate the source case.

The performance of the 200 tests for each kit was analysed using logistic regression of the log transformed viral load and the logistic coefficients were used to estimate the sensitivity with 95% confidence intervals at 10,000 and 100,000 viral load. The logistic regression coefficients were then applied to the 18291 test set to estimate the probability of every sample in the source set being positive. 95% confidence intervals were constructed by bootstrapping the 18291 test (1000 runs with replacement).

The intensity of the lateral flow band was assessed by a single observer who, at the time, was blinded to the viral load concentration in the sample. The relationship between viral load and lateral flow intensity was plotted to demonstrate that the intensity was related to the viral load.

Ethics statement

Institutional ethical assessment and approval for this project was obtained from the University of Oxford. Surplus samples were used anonymously as an audit of COVID-19 tests and enabled evaluation without express consent from individuals. Informed consent was obtained from all volunteer donors of saliva and nasal swab samples.

Role of the funding source

UKHSA and its employees designed the study in conjunction with the University of Oxford, and similarly jointly collected and analysed the data, prepared the manuscript and decided to publish.

Results

A total of 1017 submissions of antigen LFD kits were made to the phase 1 desktop review process. Of these, 194 kits passed and 185 were received at PHE Porton Down and subjected to phase 2 testing between August 2020 and December 2022 (Supplementary Table S1). A variety of LFD formats were received but those that passed were all plastic cassettes with composite lateral flow strips of standard design.

The overall pass rate during the evaluation for phase 2 testing was 53% (98/185) (Fig. 1a). The reasons for individual failures were on the whole due to a single factor (sensitivity 37/185, specificity 29/185 or kit failure 12/185), but a small number failed due to multiple factors (5/185 specificity and sensitivity, 4/185 kit failure and sensitivity).

The pattern of failures changed over time (Fig. 1b). In the first six months, 79 LFDs were tested, with 68% of these failing, mostly due to poor sensitivity. As time progressed, the testing pass rate increased and during the last 18 months of the phase 2 evaluation period (from June 2021), 90% of LFDs tested passed phase 2 testing.

Sixty-four of the LFDs (from 46 different manufacturers) tested were updated versions of previously submitted tests. Three kits failed the acceptance testing process so did not progress to the phase 3 evaluation. Most of the kits demonstrated sufficient sensitivity when tested against the SARS-CoV-2 variants, with only 11 tests failing. Four of these failed sensitivity testing with the alpha variant (3/4 also failed against gamma, one of which also failed against beta and delta). One of the VOC failures was against just gamma and one against just delta. The remaining 5 kits failing VOC testing were all omicron failures (2 BA.1 and 3 B.2) (Supplementary Table S1).

Although 95 kits passed phase 2 and acceptance testing, some failed testing against emerging VOCs so

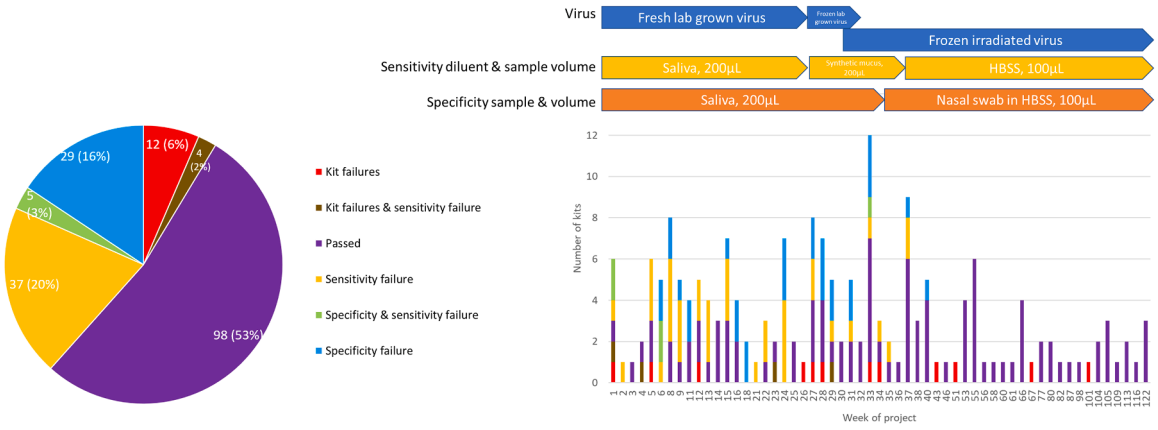

Fig. 1: Phase 2 testing results. a) Summary of overall phase 2 testing results showing number and percentage of kits passing or failing, including reason for failure (n = 185); b) Timeline of phase 2 testing results showing changes to methodology.

only 86 kits underwent phase 3 sensitivity testing. Performance, as determined by the proportion of infections averted, demonstrated a large range in sensitivity, which generally improved over the course of the evaluation period (Fig. 2). Four of the kits tested were newer versions of previously tested kits where the LFD sample buffer was changed from one containing Triton to containing Tween (a change required to meet the change in environmental toxicity regulations). All four kits showed improved sensitivity with the new buffer.

The band intensity values assigned by the trained laboratory operators aligned with the viral concentration, with an increase in band intensity as the viral concentration in samples increased. Performance, including linearity, varied between LFD kits (Fig. 3 and Supplementary Figure S1).

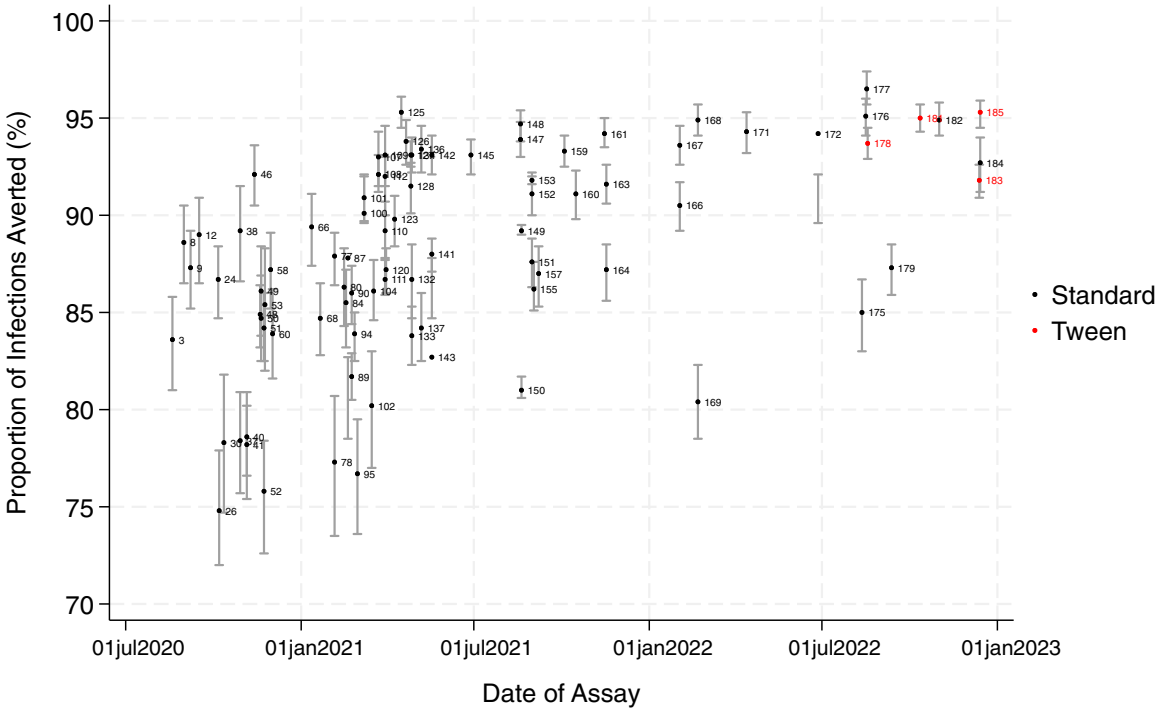

Fig. 2: Comparison of sensitivities from phase 3 testing (clinical swab samples in VTM) over the duration of the evaluation as predicted proportion of infectiousness in 18291 test set. Numbers represent different LFD listed in Supplementary Table S1, red results indicate resubmitted LFD kits with Tween buffer.

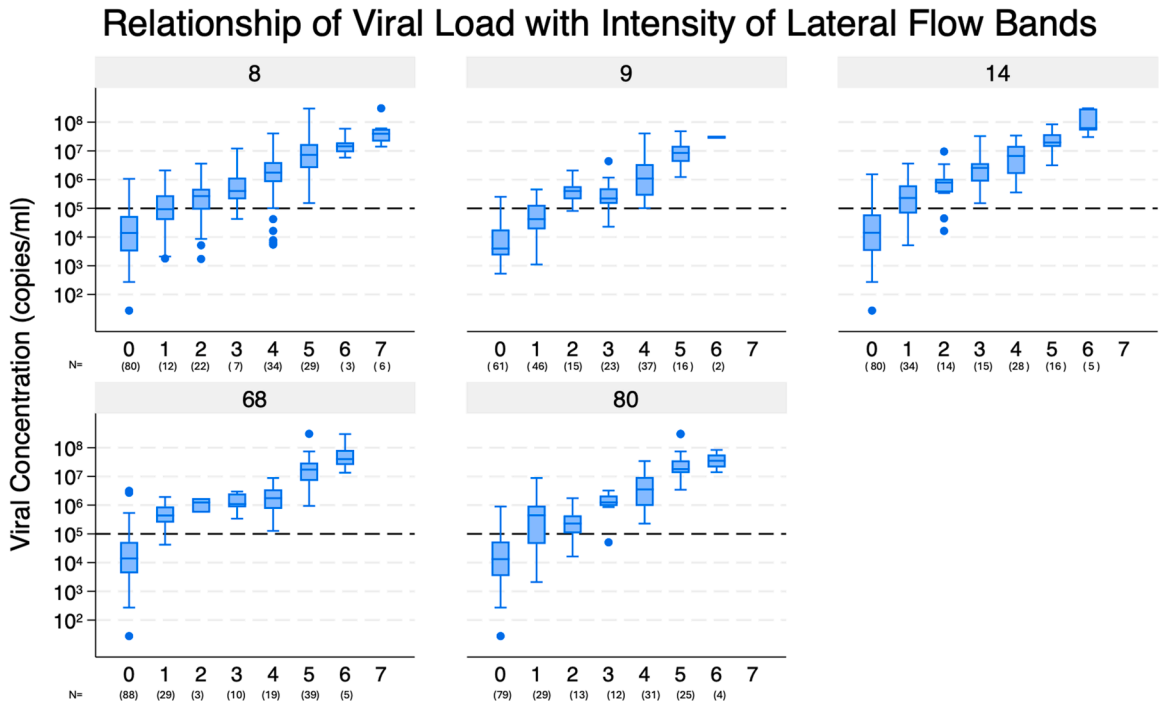

Fig. 3: Relationship of viral concentration with intensity of LFD bands for the five kits purchased by the UK government. Individual graph title number = kit reference listed in Supplementary Table S1, x-axis 0–7 = LFD band intensity values (negative to very strong) N = number of LFDs with that band value in phase 3 sensitivity testing, y-axis = viral concentration (log RNA copies/ml), dashed line represents phase 2 sensitivity pass cut-off. Higher band values with lower viral concentration indicate more sensitive tests. Box plot shows the median, IQR, the whiskers are 1.5*IQR and the outliers are outside the 1.5*IQR range.

Eighty-seven were tested for phase 3 specificity. Seventy-two (83%) of the kits tested demonstrated 100% specificity and 98% (86/88) demonstrated over 99% specificity.

The overall pass rate for the laboratory evaluation was 35%, with 64/185 kits passing Phase 3 (Fig. 4).

Discussion

Introducing new diagnostic options not requiring laboratory infrastructure was an important part of the UK Covid 19 pandemic response.[16,17] Antigen LFDs were a leading candidate to expand testing capacity and capability, so the priority of this evaluation was to develop a standardised methodology which could rapidly distinguish kits that detected infectious cases of SARS-CoV-2 with a suitable degree of sensitivity and specificity.[17]

Between August 2020 and July 2023 185 kits were studied in stages. The phase 2 assessment was not powered to identify all the suitable kits for final procurement, but to sift out those which were poor performing to focus the resources on testing the better performing kits. The phase 3 study was designed to assess performance in a 'near to field' assessment.

The 1000 SARS-CoV-2 PCR negative samples used to test specificity confirmed the chosen cutoff was low enough for clinical and community practice. Clinical samples from 200 PCR positive patients were collected from individuals tested in a hospital, rather than in field stations, and so were mainly symptomatic. At the beginning of the study the degree of heterogeneity between samples and testing was unknown and therefore a sample size of 200 was chosen essentially arbitrarily and was the maximum number that could be easily collected, reliably curated and analysed quickly. One aim was to determine whether there was any significant heterogeneity between patients including the presence of interfering substances which would distort the LFD results or factors that change the relationship between viral concentration and detectable antigen. None was observed.

The sensitivity of the tests was determined on 200 samples spanning a wide range of viral concentrations associated with infectiousness.[17,18] Assignment of band values demonstrated correlation of antigen and viral concentration, with kits showing different dynamics between protein antigen and viral RNA. There was no consensus as to the best method of summarising the overall performance of a kit. The use of a single figure to summarize the sensitivity of a kit was attractive. The most used parameter is simply giving the overall proportion of positive LFD, in a 'typical group of patients',

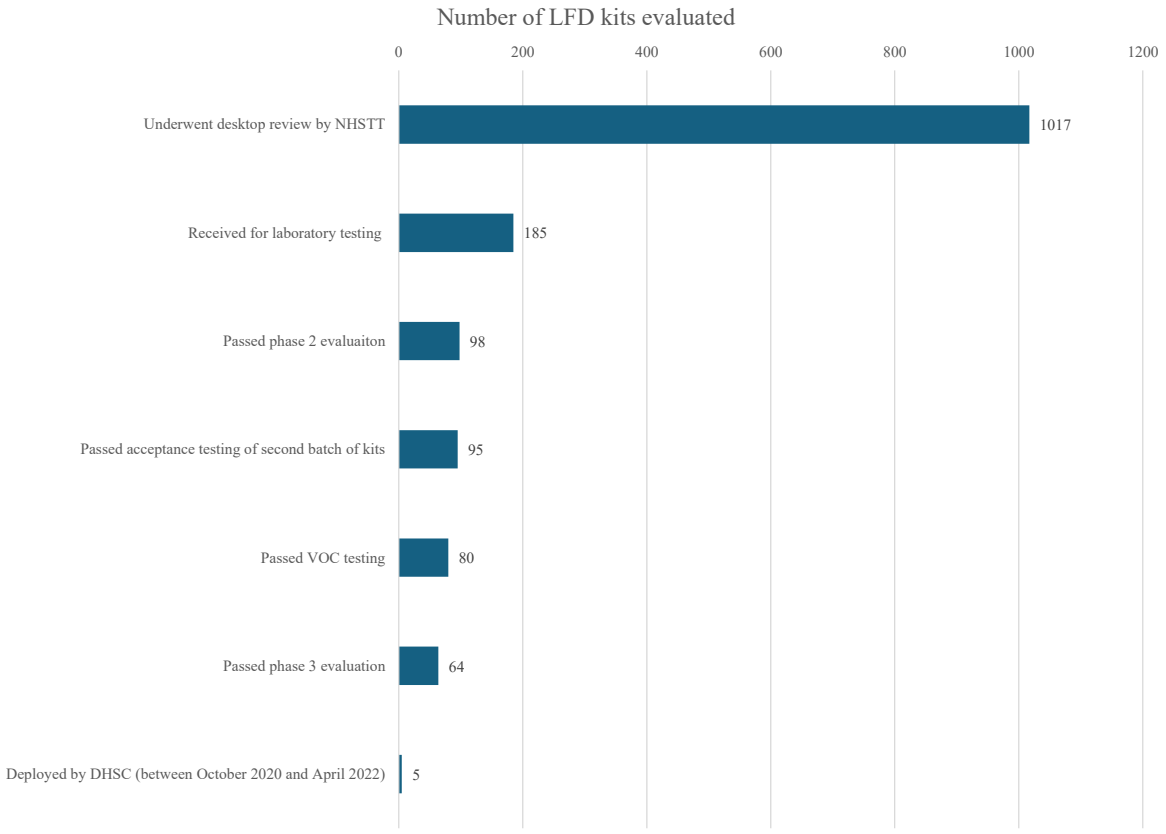

Fig. 4: Summary of number of LFD kits tested at each evaluation stage.

without regard to the range of viral concentrations. Rather than provide sensitivities of the kit at different ranges of viral concentration, we chose to use logistic regression modelling to allow a precise sensitivity to be calculated at different viral concentrations. This was designed to avoid biases when comparing results between different batches of clinical samples. The shape of the curve can be summarized by two parameters (e.g., sensitivity at 10,000 and 100,000 viral copies/ml). Differences between kits could be determined not only in terms of overall sensitivity but also in the shape of the viral concentration–sensitivity response curve. These results, in turn, allowed the overall sensitivity to detect infectiousness to be inferred using the methodology of Lee et al.[17] Using a standard set of patients PCR results identified from testing and contact tracing data, the overall sensitivity of the kit could be calculated and used to compare the performance of different kits. From this it was clear that the performance of kits, in general, improved over time.

The main weaknesses of the study were that we were not strictly following the manufacturer's instructions, this was due to the need for a standardised approach required by large scale laboratory and field studies of the 86 LFDs to enable direct comparison between all of them using matched sample panels. A mathematical adjustment was required to allow for the extra dilution made for our samples that were placed in viral transport medium and the 6 drops typically used for lateral flow tests. The 200, clinically positive samples were diluted to allow for multiple use and therefore easier comparison between different kits. Even so the number of kits tested meant that fresh batches of clinical samples were required. Comparisons between batches were potentially confounded by different viral variants or lineages prevalent at the time of the different collections, though no evidence of this was observed.

Similar data and findings have now been published (eg Pickering et al.[19]), but with far fewer LFDs being evaluated. This study has a uniquely large data set.

A critical component of the evaluation approach was engagement with industry and the Department of Health (through a forum known as the UK Rapid Antigen Test Consortium.[20] Irradiated material was provided to UK manufacturers upon request to assist with LFD product development, and suppliers were given the opportunity to discuss their results. The desired outcome was availability of high performing tests. While it cannot be attributed to the Porton evaluation process, it is notable that the sensitivity of LFDs over

time improved. This needs to be set in context internationally of similar processes and the introduction in the UK of CTDA (Coronovirus Test Device Approvals) process from July 2021 for commercially available tests, which would also have driven improvement.

Diagnostics, particularly self-tests such as LFD are likely to continue to play a globally important role. Over 2.5 billion self-test LFDs were deployed by the UK government following purchasing decisions informed by this work (personal communication, Karl Masters). The processes outlined here represent an approach that could be extended to many infectious disease self-use diagnostics to provide that standardised assurance.

Contributors

Methodology was designed by RV, TP, DC, ASi, AB, JB, ASW, HF, SH and LO. Laboratory investigations were carried out by AB, ASw, MC, JNC, TB and members of the LFD laboratory team and LFD laboratory support team. Data were accessed, verified and curated by AB, ASi, TC, DR, TP and members of the LFD laboratory team. Formal analysis was conducted by TP, AB, ASi, TC and RV. ASi, TC, AB, DR were responsible for project administration and ASi, TP and RV had supervisory responsibilities. Study resources/materials were provided by ASi, DC, SA, HF, LO, RV and TP. AB, ASi, TF, RV, TP and LO wrote the initial draft, and all authors reviewed and edited each draft during manuscript development.

Data sharing statement

Raw data collected for the study, and the study protocol, can be made available upon request to the corresponding author via email. Proposals will be reviewed by authors and collaborators on the basis of scientific merit and absence of competing interests. Following approval of a proposal, data may be shared via a secure online platform after signing a data access and confidentiality agreement. Distribution of data would require establishment and approval of an appropriate data sharing agreement.

Declaration of interests

SHo is supported by grants from MRC and NIHR. TP was Membership of National Lateral Flow Test Advisory Committee. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR, the Department of Health and Social Care (DHSC) or the UKHSA.

Acknowledgements

This study was funded by UKHSA (formerly PHE & NHSTT); the UK DHSC; and the University of Oxford NIHR Biomedical Research Centre.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.eclinm.2025.103416.

References

1 Syangtan G, Bista S, Dawadi P, et al. Asymptomatic SARS-CoV-2 carriers: a systematic review and Meta-analysis. *Front Public Health*. 2021;8:587374. https://doi.org/10.3389/fpubh.2020.587374.
2 Ramdas K, Darzi A, Jain S. 'Test, re-test, re-test': using inaccurate tests to greatly increase the accuracy of COVID-19 testing. *Nat Med*. 2020;26(6):810–811. https://doi.org/10.1038/s41591-020-0891-7.
3 Public Health England (PHE). *COVID-19:laboratory evaluations of serological assays*. London: PHE; 2021. https://www.gov.uk/government/publications/covid-19-laboratory-evaluations-of-serological-assays. Accessed December 17, 2024.
4 The National SARS-CoV-2 Serology Assay Evaluation Group. Performance characteristics of five immunoassays for SARS-CoV-2: a head-to-head benchmark comparison. *Lancet Infect Dis*. 2020;20:1390–1400.
5 Peto T, UK COVID-19 Lateral Flow Oversight Team. COVID-19: rapid antigen detection for SARS-CoV-2 by lateral flow assay: a national systematic evaluation of sensitivity and specificity for mass-testing. *eClinicalMedicine*. 2021;36:100924. https://doi.org/10.1016/j.eclinm.2021.100924.
6 Medicines & Healthcare Products regulatory Agency (MHRA). *Target product profile: point of care SARS-CoV-2 detection tests*. London: MHRA; 2020. https://www.gov.uk/government/publications/how-tests-and-testing-kits-for-coronavirus-covid-19-work/target-product-profile-point-of-care-sars-cov-2-detection-tests. Accessed December 17, 2024.
7 UK Health Security Agency (UKHSA). *Protocol for evaluation of rapid diagnostic assays for specific SARS-CoV-2 antigens (lateral flow devices*. London: UKHSA; 2022. https://www.gov.uk/government/publications/assessment-and-procurement-of-coronavirus-covid-19-tests/protocol-for-evaluation-of-rapid-diagnostic-assays-for-specific-sars-cov-2-antigens-lateral-flow-devices. Accessed April 19, 2022.
8 UK Health Security Agency (UKHSA). *LFD sensitivity: manufacturer reported versus laboratory evaluation*. London: UKHSA; 2024. https://www.gov.uk/government/publications/lfd-sensitivity-manufacturer-reported-versus-laboratory-evaluation. Accessed December 17, 2024.
9 UK Health Security Agency (UKHSA). *Lateral flow evaluation prioritisation criteria for rapid diagnostic assays for specific SARS-CoV-2 antigens*. London: UKHSA; 2022. https://www.gov.uk/guidance/lateral-flow-validation-prioritisation-criteria-for-rapid-diagnostic-assays-for-specific-sars-cov-2-antigens. Accessed December 17, 2024.
10 Ryan KA, Bewley KR, Watson RJ, et al. Syrian hamster convalescence from prototype SARS-CoV-2 confers measurable protection against the attenuated disease caused by the omicron variant. *PLoS Pathog*. 2023;19(4). https://doi.org/10.1371/journal.ppat.1011293.
11 Coombes NS, Bewley KR, Le Duff Y, et al. Assessment of the biological impact of SARS-CoV-2 genetic variation using an authentic virus neutralisation assay with convalescent plasma, vaccinee sera, and standard reagents. *Viruses*. 2023;15(3). https://doi.org/10.3390/v15030633.
12 Saah AJ, Hoover DR. "Sensitivity" and "specificity" reconsidered: the meaning of these terms in analytical and diagnostic settings. *Ann Intern Med*. 1997;126(1):91–94. https://doi.org/10.7326/0003-4819-126-1-199701010-00026.
13 Park MS, Chung JW, Kim JK, Chung SC, Kho HS. Viscosity and wettability of animal mucin solutions and human saliva. *Oral Dis*. 2007;13:181–186.
14 Caly L, Druce J, Roberts J, et al. Isolation and rapid sharing of the 2019 novel coronavirus (SARS-CoV-2) from the first patient diagnosed with COVID-19 in Australia. *Med J Aust*. 2020;212:459–462.
15 Lee LYW, Rozmanowski S, Pang M, et al. Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infectivity by viral load, S gene variants and demographic factors, and the utility of lateral flow devices to prevent transmission. *Clin Infect Dis*. 2022;74(3):407–415.
16 Department of health and Social Care. Coronavirus (COVID-19) scaling up our testing programmes. https://assets.publishing.service.gov.uk/media/5e888f05e90e0707799498b3/coronavirus-covid-19-testing-strategy.pdf; 2020. Accessed December 17, 2024.
17 Naidoo R, Lambert B, Voysey M, et al. An evaluation of the national testing response during the COVID-19 pandemic in England: a multistage mixed-methods study protocol. *BMJ Open*. 2024;14(6):e077271. https://doi.org/10.1136/bmjopen-2023-077271.
18 Eyre DW, Futschik M, Tunkel S, et al. Performance of antigen lateral flow devices in the UK during the alpha, delta, and omicron waves of the SARS-CoV-2 pandemic: a diagnostic and observational study. *Lancet Infect Dis*. 2023;23:922–932.
19 Pickering S, Batra R, Merrick B, et al. Comparative performance of SARS-CoV-2 lateral flow antigen tests and association with detection of infectious virus in clinical specimens: a single-centre laboratory evaluation study. *Lancet Microbe*. 2021;2(9):e461–e471.
20 Department of Health and Social Care (DHSC). UK government announces partnership with British rapid test manufacturers. https://www.gov.uk/government/news/uk-government-announces-partnership-with-british-rapid-test-manufacturers; 2021. Accessed December 17, 2024.