Preliminary report from the Joint PHE Porton Down & University of Oxford SARS-CoV-2 test development and validation cell: *Rapid evaluation of Lateral Flow Viral Antigen detection devices (LFDs) for mass community testing:*

Executive summary

- At the request of Ministers in the UK Department of Health and Social Care, Public Health England Porton Down and the University of Oxford developed and delivered the infrastructure required to identify the most promising LFDs with the best performance characteristics

- Extensive pre-clinical and clinical evaluation of LFDs has been completed both in the laboratory and in the field

- LFDs show acceptable viral antigen detection with high specificity, sufficient sensitivity and low kit failure rates

- One LFD, the *Innova SARS-CoV-2 Antigen Rapid Qualitative Test,* is nearing completion of the four-phase evaluation and the performance characteristics are summarised in this report.

1. Background

National governments and international organisations including the World Health Organisation (WHO) have highlighted the importance of testing and subsequent contact tracing to halt the chain of transmission of SARS-CoV-2, the virus responsible for COVID-19. The current 'gold standard' diagnostic procedure involves reverse-transcription polymerase chain reaction (RT-PCR) testing in specialised laboratories. However, there are significant challenges in expanding these testing facilities to increase capacity to identify those with asymptomatic infections or to test contacts with individuals with COVID-19, and turnaround time is typically >24 hours depending on testing location. It is widely accepted that PCR alone will not provide sufficient volumes of tests to enable mass testing at a scale that can help to identify infectious people - whether symptomatic or asymptomatic - and help break chains of transmission fast.

The development of point of care diagnostic devices for COVID-19 has formed an important part of the WHO's *"Co-ordinated global research roadmap"* since March 2020. As such, manufacturers across the world have responded to this call to align investment into this global research priority with the leading candidate being the development of Lateral Flow Devices (LFD) for COVID-19. In the summer of 2020, the NHS Test and Trace Innovation Team identified a pipeline of new products that could enable saturation testing through comprehensive and repeated testing. They concluded that these tests would need to perform with sufficient sensitivity and very high specificity so that they could be used to detect and direct responses to emerging outbreaks. This could also provide national population surveillance. In order to do so, a need was identified for evaluation of devices to be completed at pace, reliably and to a high standard so that any orders could be made with sufficient confidence. DHSC Ministers therefore commissioned PHE Porton Down to establish a time-limited SARS-CoV-2 LFD test development and validation cell in collaboration with the University of Oxford. In this document, we report on the systematic and rapid evaluation of LFDs over the past three months, which have been used by HM Government to inform decisions on increasing rapid COVID-19 testing capability in the United Kingdom.

2. Scientific Background

LFDs can be designed to test for different protein targets and are routinely used in healthcare settings as a result of their affordability, ease of use, short time to deliver a result, and high-test accuracy, e.g. pregnancy tests that detect human Chorionic Gonadotropin (hCG). In brief, a liquid sample is placed on a conjugation pad where the analyte (or antigen) of interest is bound by conjugated antibodies. The analyte-antibody mix subsequently migrates along a membrane (e.g. nitrocellulose) by capillary flow across both 'test' and 'control' strips. These strips are coated with antibodies detecting the analyte of interest and a positive test are confirmed by appearance of a coloured line; denoting successful detection of the analyte or antigen of interest.

SARS-CoV-2 antigen LFDs identify the presence of SARS-CoV-2 proteins, using conjugated antibodies to the spike, envelope, membrane or nucleocapsid proteins. As such, these tests differ from existing SARS-CoV-2 tests, that includes the first-generation LFDs that test for human antibody (IgM/IgG) against SARS-CoV-2, and RT-PCR tests that detect the presence of viral RNA. In contrast to the IgM/IgG "antibody tests", the test directly identifies

SARS-CoV-2 viral proteins and is not reliant on the host's immune response. In contrast to RT-PCR, LFDs detect viral protein rather than RNA. Results for LFDs are observed in 8-30 minutes, depending on the device, providing potential benefit through early interventions to halt the chain of transmission earlier in the disease course when individuals are most infectious

3. Aims & Objectives

The aim of the SARS-CoV-2 LFD test development and validation cell has been to design and deliver rapid systematic scientific and clinical evaluation for LFDs. Specifically, the objectives of the cell were to

- develop a high throughput pre-clinical evaluation platform focussing on:
  - Viral antigen detection
  - Specificity of the test
  - Cross-reactivity of the test to seasonal coronaviruses
  - Test kit failure rates
- establish a research and clinical trials infrastructure to establish the use of LFDs with regards to:
  - Specificity and viral antigen detection
  - Evaluation in the community and hospital
  - Pilot implementation of point of care testing in community and institutional settings

4. Methodology

Department of Health and Social Care evaluation (phase 1 evaluation)

The role of the DHSC was to identify a pipeline of manufacturers and products which had developed viral antigen LFD that could enable mass testing for SARS-CoV-2. A desktop review was performed of manufacturers' claimed performance and instructions for use to identify tests which, prima facie, may perform with high specificity and sufficient sensitivity to enable them to be used to detect SARS-CoV-2. As set out above, the DHSC were also responsible for commissioning work with Public Health England (PHE) Porton Down and the University of Oxford (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).

The work has been overseen by an LFD Oversight Group.

Pre-clinical evaluation (phase 2 evaluation)

Pre-clinical evaluation of potential LFDs was performed by PHE Porton Down with a team comprising staff from the Rare and Imported Pathogens Laboratory and the Virology and Pathogenesis Research Group. LFDs were evaluated against known PCR-negative controls consisting of saliva collected from healthy adult staff volunteers. The virus positive dilution series consisted of saliva from SARS-CoV-2- negative individuals that had been spiked with SARS-CoV-2 virus stock to give dilutions of 10^5, 10^4, 10^3 and 10^2 plaque-forming-units (pfu)/mL (n=60). An *a priori* "prioritisation" criteria was defined to evaluate LFDs and consisted of a kit failure rate of <10%, a specificity of ≥99% and a sensitivity of ≥50% at 10^2 pfu/mL, which corresponds to a PCR cycle threshold (Ct) value of approximately 25. LFDs that passed evaluation against the positive dilution series and negative controls were then evaluated against seasonal coronaviruses (229E, NL63 and OC43).

Secondary Care evaluation (phase 3a evaluation)

Evaluation against clinical samples was performed at PHE Porton Down with samples from a secondary healthcare setting. All LFDs were assessed against 1,000 known negative samples in viral transport medium (VTM) and 200 banked known positive VTM samples that had previously been frozen. These were diluted in saliva, aliquoted and frozen for later use. Analyses were performed to identify kit failure rates, specificity and viral antigen detection by LFDs in relation to viral load determined through PCR.

Community research evaluation (phase 3b evaluation)

Further evaluation against clinical samples was performed using volunteer samples from staff and patient volunteers. The clinical study of positive cases was conducted in collaboration with the UK Condor Programme "COVID-19 National Diagnostic Research and Evaluation Platform", specifically within the Falcon-C19 study (IRAS 284229). For the positive panel, this study involved the recruitment of adult individuals in the community with a known diagnosis of COVID-19 from 14 research sites around England. Participants were required to provide a paired swab sample (1 dry swab and 1 swab in VTM) and complete a study questionnaire. LFDs were evaluated according to the manufacturer's instruction using "dry swabs". For the negative panel, volunteers from PHE Porton Down and an acute hospital were recruited.

Community pilot field service evaluation (phase 4 evaluation)

Wider field service evaluation was performed within a number of institutions and settings. These institutions included secondary healthcare settings, PHE Porton Down, military establishments, schools and universities. Further evaluation was also performed at regional COVID-19 testing centres. Analyses were performed to identify kit failure rate, specificity and viral antigen detection by LFDs as a function of CT values. Further analyses were also performed in terms of assessing LFDs in relation to usability in the field, as well as uptake and feedback in terms of training.

5. Results

Over 130 suppliers of COVID-19 LFDs were identified by the DHSC for desktop review, 40 of which were sufficiently promising to be referred to PHE Porton Down for evaluation. To date, across phases 2-4 LFD of evaluations, a total of 20,545 LFD tests have been performed either directly or indirectly by the SARS-CoV-2 LFD test development and validation cell.

As part of phase 2 evaluations, 5,802 LFD tests were performed at PHE Porton Down across the 40 candidate devices (as of 31 October 2020). To date, only 9 kits (24.3%) have performed at a level in accordance with the UK lateral flow oversight group's *a priori* "proceed criteria" as published on the government website. All nine kits also passed cross-reactivity analyses against seasonal human coronaviruses. The remainder failed either due to false negative rates that did not pass the sensitivity threshold and/or false positives which did not pass the pre-defined specificity rate, and, in some cases, due to kit failures which exceeded the pre-defined rate.

A total of 7,185 tests with 6 LFDs had been completed at PHE Porton Down as part of the ongoing phase 3a evaluation. Similar to the pre-clinical testing phase, using these VTM samples, all kits significantly outperformed the pre-defined detection rate of 50% at the viral cycle threshold of 25, with an observed viral antigen detection of 83 to 97%. A total of 878 individuals with COVID-19 in the community were enrolled into the Falcon-C19 study to take part in Phase 3b evaluations. 5 kits are being evaluated and 2,678 tests have been performed to date. Taking a viral load of a cycle threshold of 25, the observed viral antigen detection of kits ranged from 95.2-100%. When all individuals in the analyses were analysed, irrespective of viral load, the viral antigen detection in the whole cohort was 77.8-93.9%.

The LFD that is currently in most advanced stages of validation is the *Innova SARS-CoV-2 Antigen Rapid Qualitative Test*, which reflects the fact that it was one of the first tests to be evaluated and successfully pass Phase 2. In total, across Phase 2-4 evaluation stages, 8,774 Innova LFD tests have been performed in the UK, including a diverse cohort of populations as part of the Phase 3b and Phase 4 testing: out-patient SARS-CoV-2 cases; healthcare staff; armed forces personnel; and school students aged 11-18 (Table 1). Due to the rapid implementation of Innova for mass testing in the United Kingdom, the purchasing and roll-out decisions which have been made by DHSC, we have focussed on the performance characteristics (kit failure rate, specificity and viral antigen detection) of the *Innova SARS-CoV-2 Antigen Rapid Qualitative Test.*

Innova LFD evaluation phase	LFD failures		LFD successes				
	fail//total	%	PCR+	PCR-	PCR-void	PCR-not done	TOTAL
Phase 2 negatives	0/72	0.0	0	72	0	0	72
Phase 2 positive dilution series	0/215	0.0	215	0	0	0	215
Phase 3a positives	12/212	6.0	199	0	1	0	200
Phase 3a negatives	50/1040	5.1	0	990	0	0	990
Phase 3b FALCON (Dry swabs-field)	28/296	10.4	252	15	1	0	268
Phase 3b FALCON (Dry swabs-lab)	9/221	4.2	204	8	0	0	212
Phase 3b FALCON (VTM swabs)	9/166	5.7	142	14	1	0	157
Phase 4 hospital staff	17/375	4.7	2	346	10	0	358
Phase 4 armed forces	6/163	3.8	46	111	0	0	157
Phase 4 PHE staff	19/231	8.9	0	212	0	0	212
Phase 4 school 1	311/2166	16.8	0	0	0	1855	1855
Phase 4 school 2 + 3 + 4	14/2146	0.65	0	0	0	2132	2132
Phase 4 COVID-19 testing centre	27/1973	1.4	139	1789	18	0	1946
TOTAL	502/9276	5.4	1199	3557	31	3987	8774

Table 1. Table illustrating the number of evaluations performed in the Innova LFD across phases 2-4 of the evaluations). The table demonstrates the kit failure rate and the where PCR results are available.

Kit Failure Rates

Table 1 show the Kit Failure Rates. There were marked differences in the kit failure rates ranging from 0.65% to 16.8% (P<0.00001; chi2(12)=530) suggesting that there might be differences between batches.

Limit of Detection

We measured the limit of detection of the antigen test with reference to plaque forming units and with RNA copies. Table shows the association between viral antigen detection and viral load as part of Phase 2 evaluations. Under these ideal concentrations, at a PFU of 100/mL, which corresponds to Ct of 25.5, the LFD identified >95% at this viral load.

PFU/ml	Ct equivalent	Positive LFD tests/total LFD tests	% positive
100000	16	20/20	100.0
10000	19	25/25	100.0
1000	23.7	65/65	100.0
390	25.2	5/5	100.0
100	25.5	63/65	95.5
40	28.5	3/5	60.0
20	29.3	0/5	0.0
10	30.2	0/5	0.0
5	31	0/5	0.0
2.5	31.7	0/5	0.0
1.2	32.5	0/5	0.0

Table 2. Table illustrating the limit of sensitivity for SARS-CoV-2 detection by the LFD for antigen detection using spiked saliva samples. Ct - cycle threshold. PFU - plaque forming units.

We also analysed the association between viral antigen detection and viral load as part of Phase 3a evaluations with clinical samples which were placed in viral transport medium allowing direct comparison of viral load and antigen tests (Fig1). This shows that samples with a CT<25.5 (calculated as a viral load >100,000 RNA copies/ml) had a 90% or greater chance of being detected.

Fig 1a

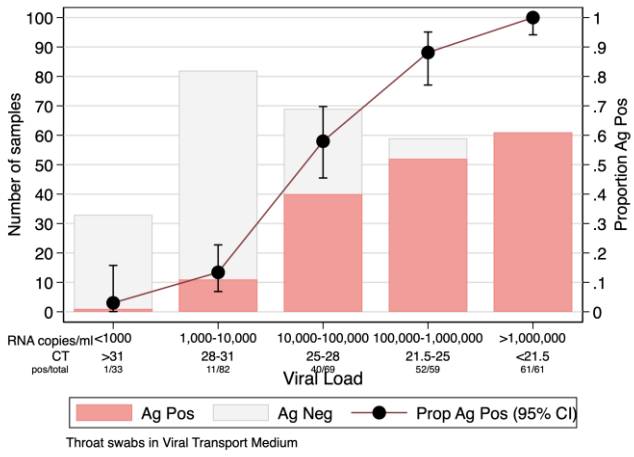

Fig 1b

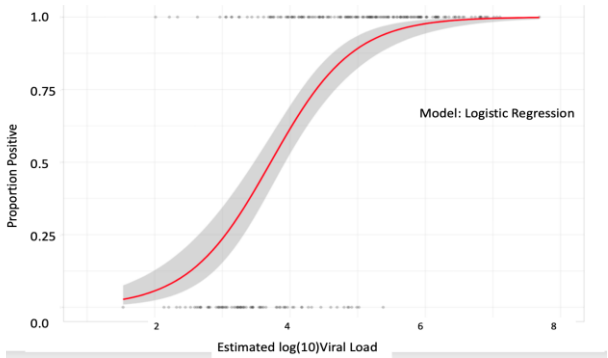

Figure 1 Proportion of Samples antigen positive according to Viral Load in Samples placed in Viral Transport Medium. 1a) Actual Numbers 1b) Proportion (with SE) estimated using logistic regression model

Specificity

Device specificity was determined through an analysis of 6,967 tests from evaluation phases 2-4. There was an overall false positive rate of 0.32% (specificity 99.68%). However, there was some indication that there was a difference in the false positive rates between laboratory-based testing (0.06%) compared to field testing (0.39%) (P=0.041 Fishers Exact Test). Our evaluations noted that where there were challenges in interpreting the results when the test result was "weak", these tests were often negative on re-testing. (Table 2).

Evaluation Phase	Testing Centre	False positives/total number	False positives and 95% CI
Phase 2 evaluation	Porton	0/72	0.00% (0.00-5.07)
Phase 3a evaluation- negative samples	Porton	0/940	0.00% (0.00-0.41)
Phase 4 evaluation- armed forces	Porton	0/105	0.00% (0.00-3.53)
Phase 4 evaluation- PHE staff	Porton	0/209	0.00% (0.00-1.80)
Phase 4 evaluation- hospital staff	Oxford	1/329*	0.30% (0.05-1.70)
Subtotal (Experienced laboratory workers)		*1/1655*	*0.06% (0.02-0.3)*
Phase 4 evaluation- school 1	Local	9/1855**	0.49% (0.26-0.92)
Phase 4 evaluation- school 2 + 3 + 4	Local	7/2130**	0.33% (0.16-0.68)
Phase 4 evaluation- COVID-19 testing centre	Local	5/1327***	0.38% (0.16-0.88)
Subtotal (Locally trained)		21/5312	0.39% (0.24-0.60)
TOTAL		22/6967	0.32% (0.21-0.47)

*This was 1 weak positive result that was also a weak positive on repeating** Weak positives result were negative on re-testing with Innova.,
*** Not photographed or repeated. Taken in setting of prevalence of 14% LFD positive results.

Table 2. Table illustrates the number of false positives in each evaluation stage and associated 95% confidence interval.

Antigen Detection in Field Studies;

Community research evaluation (phase 3b evaluation)

Viral antigen detection in individuals with confirmed SARS-CoV-2 infection was assessed in the Phase 3b evaluation as part of the FALCON-C19 research study. In particular, throat swabs were placed directly into the kit buffer solution (without using viral transport medium). Tests were performed either by laboratory scientists at PHE Porton Down or by fully trained research health care workers at the testing site. Overall 248/323 (76.8%) of the PCR positives were antigen positive. Figure 2 shows the relationship between viral load and antigen detection. There were no discernible differences in viral antigen detection in asymptomatic vs. symptomatic individuals (33/43 76.7% vs. 100/127 78.7%, p=0.78).

Phase 4 evaluation

A further series of individuals were recruited from consecutive cases from COVID19 Testing centres with tests performed by self-trained individuals and the results were compared to the Phase 3b shown above. Performance was optimal when the LFD was used by laboratory scientists (156/197 LFDs positive [79.2%, 95% CI: 72.8-84.6%])] versus trained healthcare-workers (92/126 LFDs positive [73.0%, 95% CI: 64.3-80.5%]) and self-trained members of the public given a protocol (214/372 LFDs positive [57.5%, 95% CI:52.3-62.6%]; p<0.0001 chi2(2)=30.1) (Figure 3).

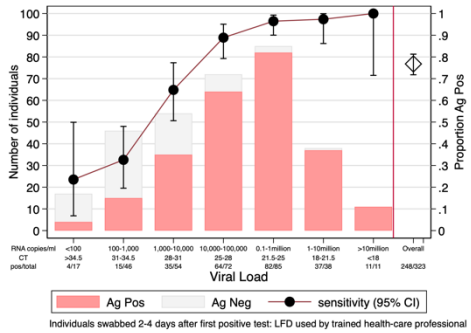

Figure 2 illustrates the association between viral antigen detection and viral load (RNA copies/ml) in phase 3b evaluation

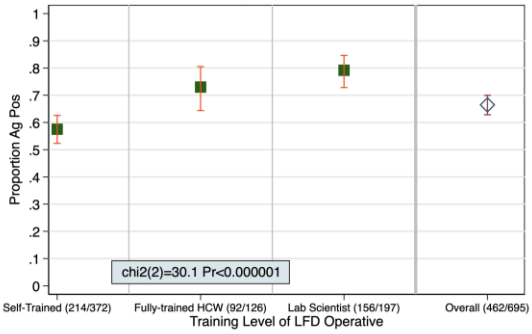

Figure 3 illustrates effect of operator and training on viral antigen detection

6. Conclusions

Enhanced identification of individuals with COVID-19 through increased availability of testing potentially offers an avenue to stop the chain of transmission for SARS-COV-2. Rapid, point of care devices for COVID-19 viral antigens offer several potential advantages over existing testing strategies.

Analyses of LFDs are at an advanced stage of evaluation in the United Kingdom. A comprehensive, systematic national pipeline has been established to rapidly evaluate the performance characteristics of LFD in laboratory and a multitude of community settings (hospitals, military establishments, schools, universities and COVID-19 testing centres). Many of the LFDs tested to date have not performed to levels established by the test and validation cell and confirmed by the LFD Oversight Group to proceed to community field service evaluations. However, a small number of LFDs have the desired performance characteristics and phase 4 evaluations have been completed for the *Innova SARS-CoV-2 Antigen rapid qualitative test.*

To date, the performance characteristics of the *Innova* LFD in the evaluations performed to date are good with a low failure rate, high specificity 99.6% and high viral antigen detection. Furthermore, issues need to be addressed to understand batch to batch variation, acceptance of the tests by the general public and the effect of operator/training effects upon performance characteristics. The delivery of appropriate training appears important to test performance. It is important to note the possibility that performance of these tests may improve with time as more research is performed within phase 4 evaluations. LFD implementation may offer advantages in national testing strategies focusing on risk reduction and warrant further testing in mass-testing scenarios. It also promises a massive increase in testing by enabling a distributed community-based use separate from the overburdened national and NHS testing laboratories.