UK Health Security Agency

LFD Digital Reader

Evaluation of real world deployment: final report

July 2022

Contents

Evaluation team

Role	Name	Institution
Project Sponsor	Susan Hopkins	UKHSA
Principal Investigator	Andrew Beggs	University of Birmingham
Co-investigators	Tom Fowler	UKHSA
	Sarah Tunkel	UKHSA
	Youssof Oskrochi	UKHSA
	Raghavendran Kulasegaran	UKHSA
Study statistician	Camila Caiado	University of Durham
Product owner	Robert Banathy	NHS Digital
Data analysts	Ashley Goddard	UKHSA
	Fegor Ichofu	UKHSA
	Jon Aveyard	NHS Digital
Project manager	Ngozi Diaku	UKHSA

Executive summary

UKHSA has worked with software developers to develop technology that can read an image of a lateral flow device (LFD) strip post testing and deliver an accurate test result. This digital reader (DR) does not require an app and can be accessed on any smart phone with a camera. When prompted, the user simply takes a photograph of their used LFD strip and the DR interprets the test result which is then uploaded directly into the national reporting system without further action from the user. The purpose of the DR is to remove subjectivity in interpretation of the LFD result, increase the accuracy of the interpreted result and improve the reporting process.

A pilot study[1] indicated that the DR was able to correctly detect approximately 25% more positive cases compared to human interpretation (self-report).

Since the pilot study, use of the DR has been expanded to include Orient Gene and Acon Flowflex lateral flow devices in addition to Innova. In total, 1,493,470 reports and 1,840,896 images were submitted by staff participating in regular asymptomatic testing across multiple different employment groups between 3 June and 31 December 2021. These groups included staff working in adult social care residential homes (ASC) and hospices, staff working in primary care, and some private sector employers.

An evaluation of the DR was carried out based on this data to understand its real-world effectiveness, and considered the following questions:

1. Output accuracy – does the DR correctly identify and report positive cases?
2. Viral accuracy compared to polymerase chain reaction (PCR) – what is the concordance between the DR and matched PCR results compared to self-read and matched PCR results?
3. User journey performance – how easy was it for users to navigate the DR process?
4. Value for money – how does the improved accuracy of the DR translate into fiscal value?

1. Output accuracy

Output accuracy was assessed by comparing the results generated by the DR for 2,000 randomly selected images with those obtained from an independent expert panel (‘ground truth'). At a sample prevalence of 0.77%, the DR had a sensitivity of 73.3% (44.9% to 92.21%) and specificity of 99.95% (99.71% to 100%). This suggested the DR was working as expected and aligns with reported real-world performance of LFD devices deployed by DHSC/UKHSA.

[1] Machine learning for determining lateral flow device results for testing of SARS-CoV-2 infection in asymptomatic populations. The LFD AI Consortium, 2022, Cell Reports Medicine 3, 100784

2. Viral accuracy compared to PCR

Viral accuracy compared to PCR was assessed by comparing the concordance between digital reader and self-read LFD results to matched PCR results taken at the same time. This was possible by virtue of the fact that the routine asymptomatic testing regime of the ASC and hospice workforce at the time required staff to undertake a PCR and LFD test at the same time once a week, and thus generated a large, matched data set which also allowed analysis of DR against viral concentrations. Concordance between DR and PCR was much greater than that of self-read and PCR, with the DR detecting 25.49% more true positive cases. The effect increased as viral loads decreased with the DR being almost twice as likely (19.67% versus 39.91%) to detect true positive cases at low viral concentrations. This was intended and desired test performance given that low viral concentrations manifest as faint bands on LFD cartridges and thus may be missed by the human eye or misinterpreted; earlier identification of positive cases represent earlier opportunities to interrupt transmission chains.

3. User journey performance

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 we observed no learning curve on subsequent success rates. 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.

4. Value for money

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.

In conclusion, compared to the self-read LFD reporting (the current status quo), the digital reader offers greater accuracy and better case-finding capability. The process is easy to use and acceptable to most users with no apparent additional burden and no learning curve. These benefits could translate into substantial financial benefits if deployed at a population level.

Commonly used abbreviations

Abbreviation	Description
Ag-RDTs	Antigen rapid diagnostic tests
AI	Artificial intelligence
ASC	Adult social care
ATS	Asymptomatic testing site
CI	Confidence interval
Ct	Cycle threshold
DA	Devolved administration
DHSC	Department of Health and Social Care
DR	Digital reader
DR-LFD	Digitally read lateral flow device
HR-LFD	Human read lateral flow device
IQR	Interquartile range
LFD	Lateral flow device
LFT	Lateral flow test
MHRA	Medicines and Healthcare Products Regulatory Agency
MVA	Moving average visits
MVS	Multi vendor strategy
NHSD	NHS Digital
PCR	Polymerase chain reaction
PMS	Post marketing surveillance

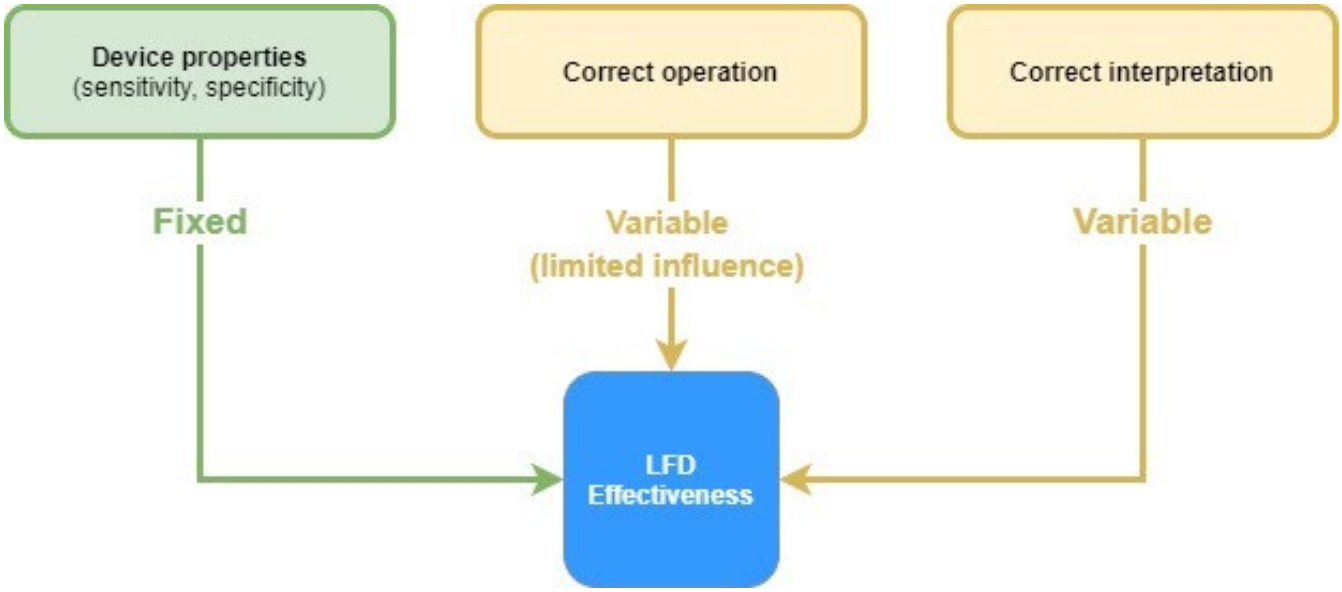
Abbreviation	Description
RDT	Rapid diagnostic test
UKHSA	UK Health Security Agency
VFM	Value for money
WCS	Worst case scenario

Background and rationale

Lateral flow devices (LFDs) are a valuable, easy to use, rapid and highly specific mass testing tool which form part of the national testing offer.

The effectiveness of an LFD is a function of the intrinsic properties of the device in terms of sensitivity and specifity, its correct operation according to the manufacturer's instructions, and the correct interpretation of the readout (Figure 1).

Figure 1. Factors affecting LFD effectiveness



All LFD devices used within the national testing offer have met the minimum standards of sensitivity or specificity as set out by Government. They are either CE marked or authorised via an exceptional use authorisation (EUA) by the MHRA. They are considered fixed variables and thus not considered as part of this evaluation.

The interpretation of a self-test LFD output is dependent on the operator being able to visually inspect the device followed by successfully interpreting the output itself. While correct execution of these steps does not generally represent an issue, under certain circumstances both visual inspection and/or interpretation of the result could reduce the performance of the lateral flow test. Examples include users with impaired vision; users in a hurry, juggling multiple activities or working under some other stress; users with cognitive impairment; and when test bands are faint and harder to see. Any of these might lead to true positive results being missed during testing, and hence to transmissions that could have been prevented.

In March 2021, a validation study was undertaken which supplemented the existing self-test LFD process with a machine learning model (digital reader) which could interpret photos of user operated LFDs from the UK's asymptomatic testing programme. The validation study

demonstrated that use of the digital reader increased sensitivity of LFDs compared to interpretation by staff trained in reading LFDs at asymptomatic testing sites (ATS) from 92.08% to 97.6% (95% CI 93.20 to 99.51%) and compared to self-reporting (by a lay person) from 6.00% to 100% (95% CI 90.2 to 100%). In the ATS group this translated to an additional 30 positive tests, and in the self-report group an additional 32 positive tests identified. The study concluded that the digital reader increased sensitivity of LFD testing and was able to detect approximately 25% more cases compared to human interpretation.

Since June 2021, the digital reader has been used by over 150 organisations across public and private sectors as part of an enhanced self-test LFD programme within the national asymptomatic testing programme. By the end of September 2021, 632,661 images had been captured and interpreted by the digital reader of which 306,265 images were processed in September alone; 89.7% of the reported images were from England and the rest were from the Devolved Administrations. The overall breakdown by use case in Table 1.

Table 1. Use case source of image submitted to the digital reader in September 2021

	Primary care (staff)	School and colleges	ASC (staff)	Private industry	Secondary care (staff)	Universities	Public industry	Not stated	Total
%	60.89	8.6	8.6	6.6	4.8	2.1	1.6	6.8	**100%**
n	186,095	26,490	26,205	20,235	14,720	6,560	4,975	20,875	**306,265**

The availability of these data and images allowed us to undertake a retrospective end-to-end analysis of the effectiveness of real-world performance of the digital reader and link it to other data which may be available (for example, PCR).

This evaluation enabled us to confirm the findings of the original validation study, determine if wider implementation represented a sound economic investment, and if there are any settings in which it would be particularly beneficial.

Aims and objectives

Aims

The aim of this evaluation was to determine whether the use of digitally read LFDs (DR-LFDs) within the national asymptomatic testing programme resulted in greater overall real-world effectiveness of LFDs compared to human read LFD (HR-LFD). If so, what would the economic impact of a wider deployment of the digital reader (DR) service be?

Objective 1. Output accuracy

Evaluate the performance of the DR in its intended use setting by the intended user population against an expert panel of reference users by:

- determining the 'ground truth' using a subset of the available data
- verifying the sensitivity and specificity of DR-LFDs versus expert user read images using the ground truth

Objective 2. Viral accuracy compared to PCR

Evaluate the concordance of DR-LFD results with those of RT-PCR by:

- matching LFD-PCR pairs from the adult social care routine dual testing (same day/day 0) regime
- calculating sensitivity and specificity of DR-LFD versus RT-PCR

Objective 3. User journey performance

To understand and identify steps in the DR-LFD process which impact user experience and reduce successful image submission by:

- collecting post-marketing surveillance data describing the user journey through the image submission and results process
- analyse data to identify points or steps in the process which have high failure or abandonment rates
- differentiate users, where possible, by use case and demographic characteristics

Objective 4. Value for money

Determine whether the digital reader would represent value for money if implemented fully or in a context-specific setting (for example, education, ASC staff, primary care staff) by using the UKHSA Value for Money tool.

Methods

This evaluation comprised 4 different analyses, each largely independent of each other.

Objective 1. Output accuracy

Output accuracy, the ability of the digital reader to correctly interpret LFD images, was assessed by comparing DR outputs to those of an expert panel of clinicians. This was done to ensure that the results of this study corroborated and aligned with those of the pilot study which allowing us to be confident that the performance of the device remained as expected. This was particularly important given that the real-world deployment of the DR, unlike the pilot, does not have a self-report arm.

When a test is presented to a digital reader, a result will be returned as positive, negative, invalid or void.

1. A positive result will be returned when a test is identifiable as an approved test type, and both the control line and test line on an LFD have detectable lines.
2. A negative result will be returned when a test is identifiable as an approved test type, and the control line has a detectable line, but the test line does not.
3. An invalid result will be returned if the test type cannot be identified or if the image quality is not good enough for a test result to be identified. In these circumstances the Digital Read journey will ask the user to take and upload another image.
4. A void result will be returned when the test type can be identified but there is no detectable control line, regardless of whether the test line is showing positive or negative. In these circumstances the results service will ask the user to take another test.

Four expert clinicians, who were previously trained in image interpretation during the pilot, were asked to analyse a defined set of images to determine the ‘ground truth’, that is, the gold standard of LFD image interpretation. Collection of these images began at 7am on 12 September 2021 and continued until 2,000 unique barcodes were accumulated at 10:04am. This time and date were selected as it represented a normal use scenario and would have reflected the prevalence and disease activity at the time.

Three of the clinicians individually assessed each image independently and blindly of each other with outputs being either positive, negative, or void. If there was consensus, no further action was taken and this was accepted as the ground truth. In cases where there was not a consensus, a fourth reviewer was requested to help resolve these discrepancies – their assessment of the image was included in final assessment and the majority view was accepted. The fourth reviewer’s views were not accepted as superior to the other 3, and the majority view was taken to be the ground truth.

The ground truth was then independently compared to the result of the DR and standard metrics of sensitivity, specificity, and estimated population prevalence were reported. The exact binomial method was used to calculate 95% confidence intervals where appropriate.

Objective 2. Viral accuracy compared to PCR

Analysis was performed to evaluate the concordance between day 0 DR-LFD and RT-PCR results. This was compared to the concordance between day 0 self-read LFD and RT-PCR results. The data was captured for defined periods from users who used either the digital reader to submit their LFD results or from users who submitted self-read LFD results in defined settings. The corresponding PCR data for the DR and self-read results were obtained where this was available.

The analysis included calculating the sensitivity[2] and specificity[3] of DR-LFD versus RT-PCR and self-read LFD versus RT-PCR and other metrics of comparison including accuracy[4], false positive rate (FPR)[5], positive predictive value (PPV)[6], and negative predictive value (NPV)[7]. Furthermore, analysis was performed to assess the range of viral concentration for both the digitally read and self-read LFD results which were PCR positives to assess how the digital reader performed compared to self-read results.

Objective 3. User journey performance

User journey performance relates to how successfully users were able to navigate the DR process end-to-end, from accessing the web portal to taking and submitting an LFD image to receiving their results. The user journey performance therefore was measured through several key metrics which are outlined in Table 2.

[2] Sensitivity = True positives / (True positives + False negatives)
[3] Specificity = True negatives / (True negatives + False positives)
[4] Accuracy: (True positives + True negatives) / Total number of tests (TP+TN+FP+FN)
[5] FPR: False positive/(True negatives + False negatives)
[6] PPV: True positives / (True positives + False positives)
[7] NPV: True negatives / (True negatives + False negatives)

Table 2. Metrics and definitions describing user journey performance

Metric	Definition	Significance
Images submitted	Number of images taken, uploaded, and submitted for DR analysis.	Represents entry into the DR analytic process. Potential outputs are valid (positive, negative, or void) or invalid results.
Valid result rate	Proportion of images successfully analysed by the DR with outcome being positive, negative or void (no control band).	Represents a normally working DR process. Repeat test on another device requested if void.
Invalid result rate	Proportion of images unable to be analysed by the DR. Further image of same device requested.	Represents a failure of the DR process – hence an invalid result. Typically, due to poor quality image, for example, out of focus, shadow, poor light, and so on
Successful submission at n^{th} try rate	Proportion of reports with successful interpretation of a submitted image at first, second or third photo attempt. User successfully completes the DR process.	Higher proportion of successful submissions at earlier attempts is a proxy measure of better and simpler user experience.
Manual report rate	Proportion of reports which, after third invalid image upload, the system asks for manual entry of result.	Represents a failure of the DR process to adequately inform the user of how best to take a suitable image. Potentially addressable.
Journey abandoned after x attempt rate	Proportion of users who choose not to complete the DR process after submission of 1 or 2 invalid images (for example, close the browser window, and so on)	Represents a failure of the DR process in that it proved too onerous for the user. Potentially not addressable.
Time to completions	Time taken for a user to fully report their result using the DR	As familiarity with the DR process increases an expected improvement in completion time is expected up to a certain limit.

Multiple data sources were utilised to allow extraction of metrics, interrogation of different dimensions of the user journey and triangulation of potential findings, these included:

- monthly post-marketing surveillance (PMS) reports provided to the MHRA by Sensyne Health
- web-based analytics provided from user interaction with the submission portal – only possible where the user had enabled cookies on their device
- the EDGE database (available for England only and does not include data from DAs) which allowed linking of unique user IDs to LFD barcodes and provided person level longitudinal data over time; demographic and contextual (use setting) data was available for each person but not to the extent which would make it identifiable

Objective 4. Value for money

The UKHSA VFM modelling tool[8] was used as the basis for this analysis. Main inputs into the VFM modelling tool were prevalence (1% or 5%), R value (1.1), estimated generic unit cost per read (£0.2 per read per 1.5 million tests or £0.3 per read per 1.5 million tests) with the related fixed costs worked into the price per test and the duration of intervention.

Further variable inputs related to the probability of outcomes in vaccinated and unvaccinated individuals (long COVID, hospitalisation, death, and so on) and these were kept at rates as estimated amongst the subgroups of interest.

A worst-case-scenario (WCS) 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.

The DR model was developed with a focus on precision, and selective sampling was used to collect all potentially positive cases (n=589) from the total images submitted on 1 and 6 September 2021. These dates were chosen as they contained the greatest number of positive cases hence were most suitable for precision calculation. An additional 500 negative samples were also chosen at random from the cohort and included in the analysis to ensure validity of the precision and sensitivity calculations, as well as verifying that no or a low number of false negatives were present.

The DR findings were compared to the self-read data over the same sampling period and the exact binomial method was used to calculate 95% confidence intervals where appropriate.

Outputs from the model included additional infections prevented, net single-vendor and multi-vendor benefits (£).

[8] Based on The Canna model: assessing the impact of NHS Test and Trace. It is worth noting that the input assumptions for this were developed in mid 2021, and these would need to be revisited for VFM considerations of new DR applications in specific settings or contexts.

Results

General

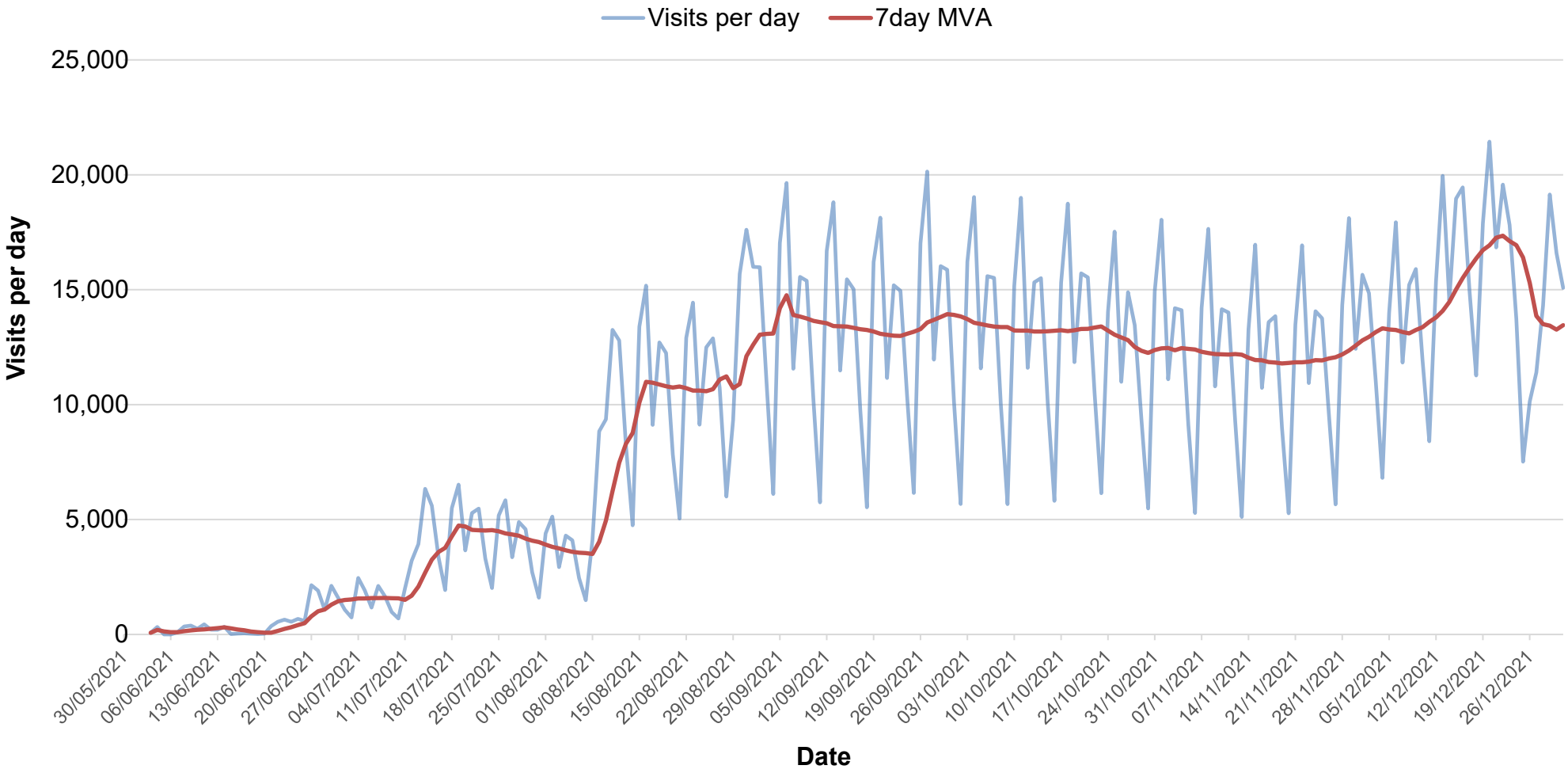
Overall, 1,840,896 images were submitted between 3 June and 31 December 2021 across all 4 nations. Summary statistics, including monthly submission volumes and outcomes are shown in Table 3.

Table 3. Digital reader usage data June to December 2021

			Outcomes (%)			
Date range	**Total images**	**Images per day (000's)**	**Positive**	**Void**	**Invalid**	**Negative**
3 to 13 June	2,104	0.19	0.9	1.4	16.0	81.8
14 June to 20 July	66,181	1.79	0.7	0.3	14.2	84.8
21 July to 31 August	313,054	7.45	0.5	0.3	14.7	83.5
1 to 24 September	295,436	12.31	0.5	0.2	14.1	85.1
1 to 31 October	371,045	11.97	0.6	0.1	13.3	85.9
1 to 30 November	327,570	10.92	0.6	0.1	13.2	86.1
1 to 31 December	465,506	15.02	1.3	0.1	15.6	83.0
Total	1,840,896					
Average			0.72	0.36	14.45	84.33
Calculated n			13,784	3,171	262,673	1,557,798

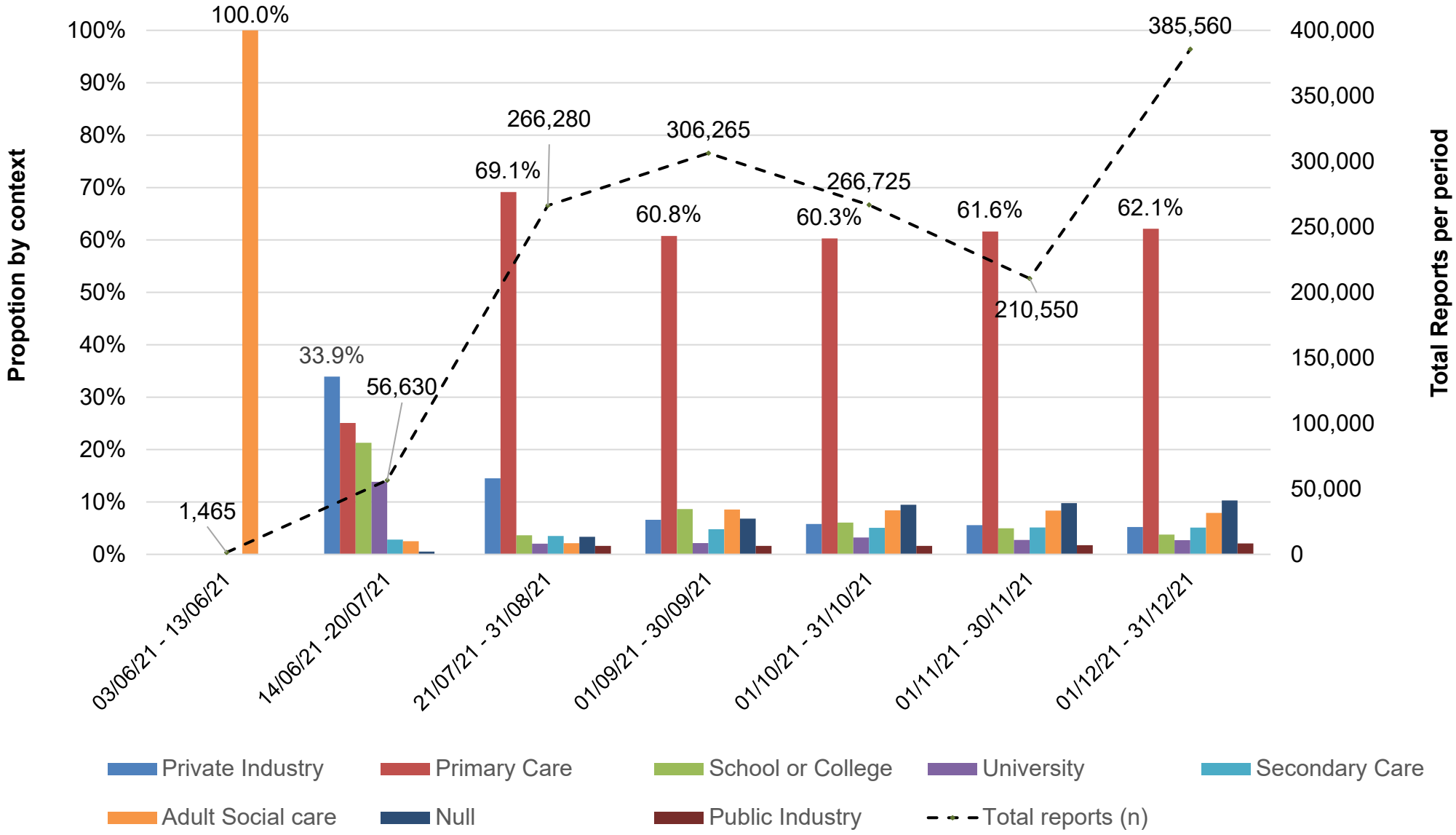
The service has grown substantially since June 2021 from approximately 200 images submitted per day to December where submission rates were above 15,000 per day. This is reflected in analytics data from the service front-end (the web portal where individuals can submit their results), demonstrating gradual and sustained increase in web-traffic (cookie-enabled users only) to the submission portal (Figure 2). The cyclical nature of the graphs is likely to represent the variation in submission days due to routine testing regimes and the influence of weekends and bank holidays (for example, submissions at the end of December).

Figure 2. Digital reader web-portal record of (cookie enabled) visitors per day



A total of 1,493,470 reports were generated during the same period which reflect the successful completion of the submission process. The number of final reports was necessarily fewer than the number of images submitted as some reports had more than one image submitted as part of a process (for example, invalid image first requiring subsequent images to be taken). Breakdown of reports by context (the setting in which the device was used) and gender across the 4 nations is presented in Figure 3 and Table 4 respectively.

Figure 1. Breakdown of submitted reports by context



Submissions from staff working in the primary care sector represent the majority of reports since inception (Figure 3), although due to the limitations of the data it was not possible to ascribe all use cases to a category and hence a minor proportion (8 to 10%) is classified as 'null'.

Table 4. Digital reader use gender breakdown

Date range	Female		Male		Unknown		Total reports
	n	%	N	%	n	%	
3 to 13 June	1,155	78.8%	310	21.2%	0	0.0%	1,465
14 June to 20 July	35,740	63.1%	20,810	36.7%	80	0.1%	56,630
21 July to 31 August	205,755	77.3%	60,360	22.7%	170	0.1%	266,285
1 to 24 September	243,690	79.6%	62,395	20.4%	180	0.1%	306,265
1 to 31 October	210,340	78.9%	56,245	21.1%	135	0.1%	266,720
1 to 30 November	167,045	79.3%	43,340	20.6%	160	0.1%	210,545
1 to 31 December	301,495	78.2%	83,760	21.7%	305	0.1%	385,560
							1,493,470

The large gender difference in reports submitted is expected and explained by the greater representation of females within the health and social care workforce. ONS estimates for 2020 suggest that 77.6% of the health and social care workforce in England and Wales were female[9], aligning with presented findings.

Objective 1. Output accuracy

A total of 2,000 images were analysed for sensitivity calculation and the 4x4 contingency table is presented in Table 5.

Of the 2,000 images, 46 were labelled as invalid either by the DR or the expert panel. The DR labelled 42 images as invalid. Of the DR invalid images, one was labelled positive by all reviewers, one had split decision between negative and positive, one could not be read by 3 of 4 reviewers. The expert panel marked 4 additional images as invalid. This included 2 images that could not be read by 2 reviewers and a further 2 that could not be read by one of the reviewers.

Two images were labelled by the DR as negative but could not be interpreted by the expert panel as the images were considered to be too blurred. Void results are defined as cases where no control line is present on the LFD device. This indicates either a failure of the device itself or user failure in correctly executing the test process. Void results are considered valid as they do not represent failure of the reader (human or digital) to correctly interpret the image. Users are asked to repeat the LFD if a void result is obtained.

[9] Number of health and care workers in England and Wales for 2019 and 2020 (ONS)

Invalid results are defined as cases where experts or the digital reader were unable to interpret the image. The cause of this could be varied but ranges from blurry images, to cropped images to poor lighting or shadows on the device. Users are given 3 attempts at providing another image for analysis – 2 additional attempts following an initial invalid result, after which they are asked to manually state their result.

Table 5. 4 x 4 contingency table of ground truth versus digital reader

		Ground truth				
		Negative	**Positive**	**Void**	**Invalid**	**Total**
Digital reader	Negative	1,933	4	0	4	**1,941**
	Positive	1	11	0	0	**12**
	Void	5	0	0	0	**5**
	Invalid	39	2	0	1	**42**
	Total	**1,978**	**17**	**0**	**5**	**2,000**

Overall, at an estimated disease prevalence of 0.77% (0.43% to 1.26%), the DR had a sensitivity of 73.3% (44.9% to 92.21%) and specificity of 99.95% (99.71% to 100%).

Limitation of interpretation: comparison was with an expert panel as opposed to general public.

Objective 2. Viral accuracy compared to PCR

The concordance of DR results and day 0 RT-PCR results was evaluated in the first instance, and this was compared to the concordance of self-read results and day 0 RT- PCR results. We limited our analysis to 3 high-performing lighthouse testing labs where the RT-PCR tests were conducted and restricted self-read LFD analysis only to those settings for which comparable DR data was available.

In total 17,898 DR LFD results had corresponding day 0 RT-PCR results; 1,359,432 self-read LFD results had day 0 RT-PCR results. The results dataset did not include any DR voids or invalids and hence these were excluded from this analysis. Summary of concordance between DR LFD results and RT-PCR results and concordance between self-read LFD results and RT-PCR results is presented in Table 6 and Table 7.

Table 6. Concordance between DR LFD results and RT-PCR

PCR	**DR LFD result**		**Grand total**
Viral load	**Negative**	**Positive**	
Negative	**14,746**	**651**	**15,397**
Low (<10,000)	2		2
Null	14,744	651	15,395

PCR	DR LFD result		Grand total
Viral load	Negative	Positive	
Positive	**524**	**1,681**	**2,205**
High (>1 million)	87	664	751
Low (<10,000)	167	119	286
Medium (10,000 to 1 million)	160	561	721
Null	110	337	447
Void	**275**	**21**	**296**
High (>1 million)		1	1
Low (<10,000)	54	6	60
Null	221	14	235
Grand total	**15,545**	**2,353**	**17,898**

Table 7. Concordance between self-read LFD results and RT-PCR

PCR	Self-read LFD result		Grand total
Viral load	Negative	Positive	
Negative	**1,168,620**	**16,070**	**1,184,690**
Low (<10,000)	103	6	109
Medium (10,000 to 1 million)	1		1
Null	1,168,516	16,064	1,184,580
Positive	**20,401**	**139,996**	**160,397**
Low (<10,000)	6,898	8,378	15,276
Medium (10,000 to 1 million)	5,564	41,074	46,638
High (>1 million)	3,789	61,840	65,629
Null	4,150	28,704	32,854
Void	**13,524**	**821**	**14,345**
Low (<10,000)	2,868	227	3,095
Medium (10,000 to 1 million)	9	39	48
High (> 1 million)	9	25	34
Null	10,638	530	11,168
Grand total	**1,202,545**	**156,887**	**1,359,432**

Further analysis was restricted to Health and Social care settings for which comparable self-read LFD and RT-PCR results were available. Summary of DR and self-read results' concordance with RT-PCR results is presented in Tables 8 and 9 respectively.

Table 8. Concordance between DR LFD results and RT-PCR across various health and social care settings

PCR	LFD negative		LFD Positive		Grand Total	
Viral load	n	%	n	%	n	%
Negative	**12,779**	**96.02**	**529**	**3.98**	**13,308**	**100.00**
Low (<10,000)	2	100.00	0	0.00	2	100.00
Null	12,777	96.02	529	3.98	13,306	100.00
Positive	**412**	**27.39 (25.15 to 29.72)**	**1,092**	**72.61 (70.28 to 74.85)**	**1,504**	**100.00**
High (>1 million)	68	14.02 (11.06 -17.43)	417	85.98 (82.57 to 88/94)	485	100.00
Low (<10,000)	128	60.09 (53.18 to 66.72)	85	39.91 (33.28 to 46.82)	213	100.00
Medium (10,000 to 1 million)	123	24.60 (20.89 to 28.62)	377	75.40 (71.38 to 79.11)	500	100.00
Null	93	30.39	213	69.61	306	100.00
Void	**242**	**94.53**	**14**	**5.47**	**256**	**100.00**
Low (<10,000)	44	89.80	5	10.20	49	100.00
Null	198	95.65	9	4.35	207	100.00
Grand total	**13,433**	**89.15%**	**1,635**	**10.85%**	**15,068**	**100.00%**

Table 9. Concordance between self-read LFD results and RT-PCR across various health and social care settings

PCR	LFD negative		LFD positive		Grand Total	
Viral load	n	%	n	%	n	%
Negative	**968,443**	**99.77**	**2,253**	**0.23**	**970,696**	**100**
Low (<10,000)	62	100.00	0	0.00	62	100
Medium (10,000 to 1 million)	1	100.00	0	0.00	1	100
Null	968,380	99.77	2,253	0.23	970,633	100
Positive	**4,780**	**42.14 (41.23 to 43.06)**	**6,563**	**57.86 (56.94 to 58.77)**	**11,343**	**100**
High (>1 million)	678	18.82 (17.55 to 20.13)	2,925	81.18 (79.87 to 82.45)	3,603	100
Low (<10,000)	2,013	80.33 (78.72 to 81.87)	493	19.67 (18.13 to 21.28)	2,506	100

PCR	LFD negative		LFD positive		Grand Total	
Viral load	n	%	n	%	n	%
Medium (10,000 to 1 million)	1,199	35.31 (33.70 to 36.94)	2,197	64.69 (63.06 to 66.30)	3,396	100
Null	890	48.42	948	51.58	1,838	100
Void	**11,201**	**98.90**	**125**	**1.10**	**11,326**	**100**
High (>1 million)	9	90.00	1	10.00	10	100
Low (<10,000)	1,739	99.03	17	0.97	1,756	100
Medium (10,000 to 1 million)	4	100.00	0	0.00	4	100
Null	9,449	98.88	107	1.12	9,556	100
Grand total	**984,424**	**99.10%**	**8,941**	**0.90%**	**993,365**	**100%**

The concordance between the PCR positives and the DR LFD results is greater than that with the self-read LFD results suggesting that the DR is significantly better at identifying positive results. This is presented in Figure 4.

Figure 4. LFD – RT- PCR concordance

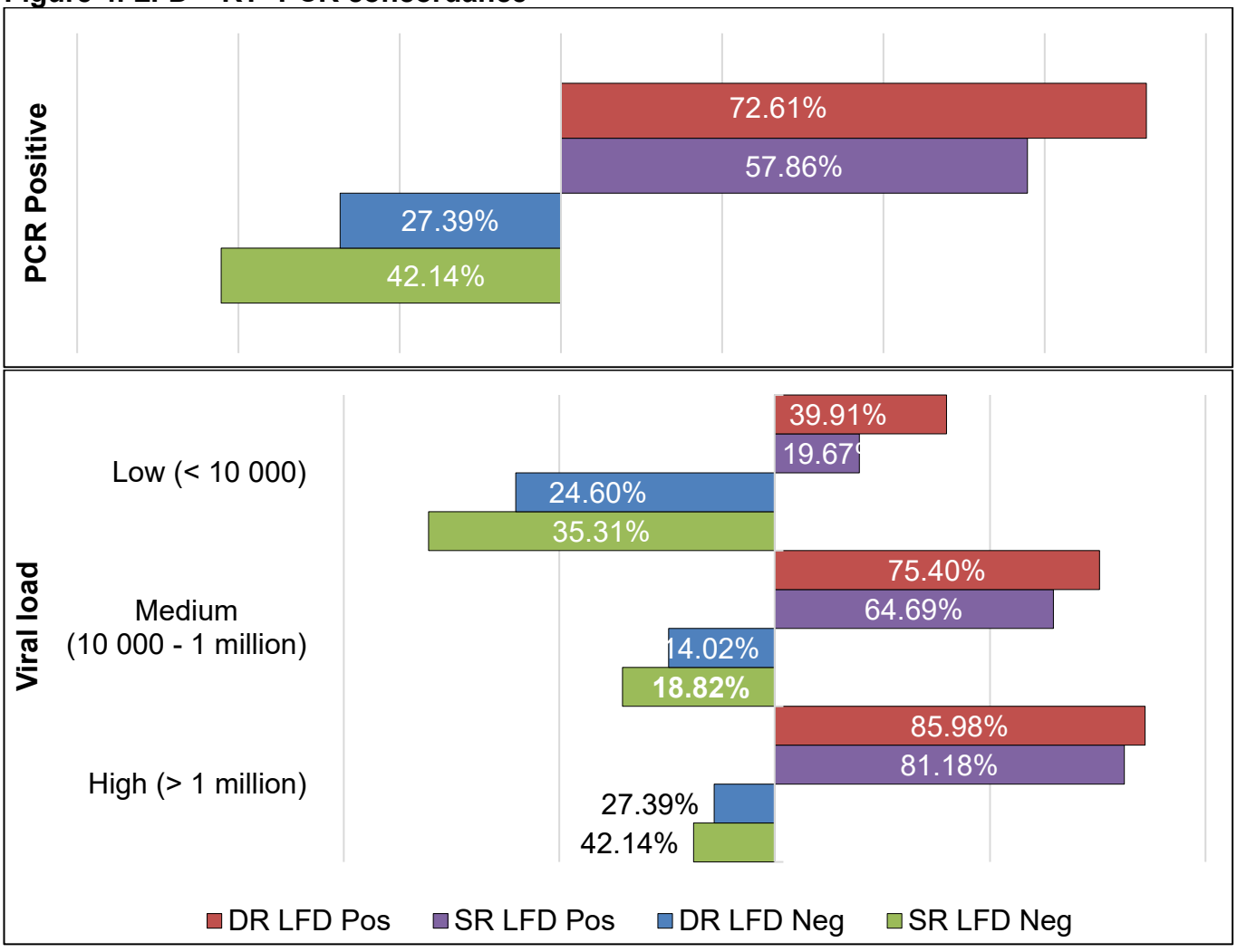


Figure 4 (top panel) shows that the rate of detection of positives by DR is 25% higher than that by self-reading. The bottom panel shows the gain made across the different viral load categories calculated using PCR Ct values. While the gain was marginal for individuals with high-viral loads (5.91%), it was nearly 17% for individuals who had a medium viral load and more than 100% for individuals with low viral loads suggesting that the DR would pick up positive cases much earlier and prevent further onward transmission.

Further analysis of the data showed that the DR had significantly higher sensitivity than self-read but marginally lower specificity (Table 10). The large sample size of the self-read LFD results, in particular the high number of reported negatives, influenced the specificity, accuracy, FPR, PPV and NPV. It is worth noting that the algorithm of the DR is set to prioritise the detection of positive test results. The outcome of this is that a higher number of false positives are returned, but fewer true positives are missed.

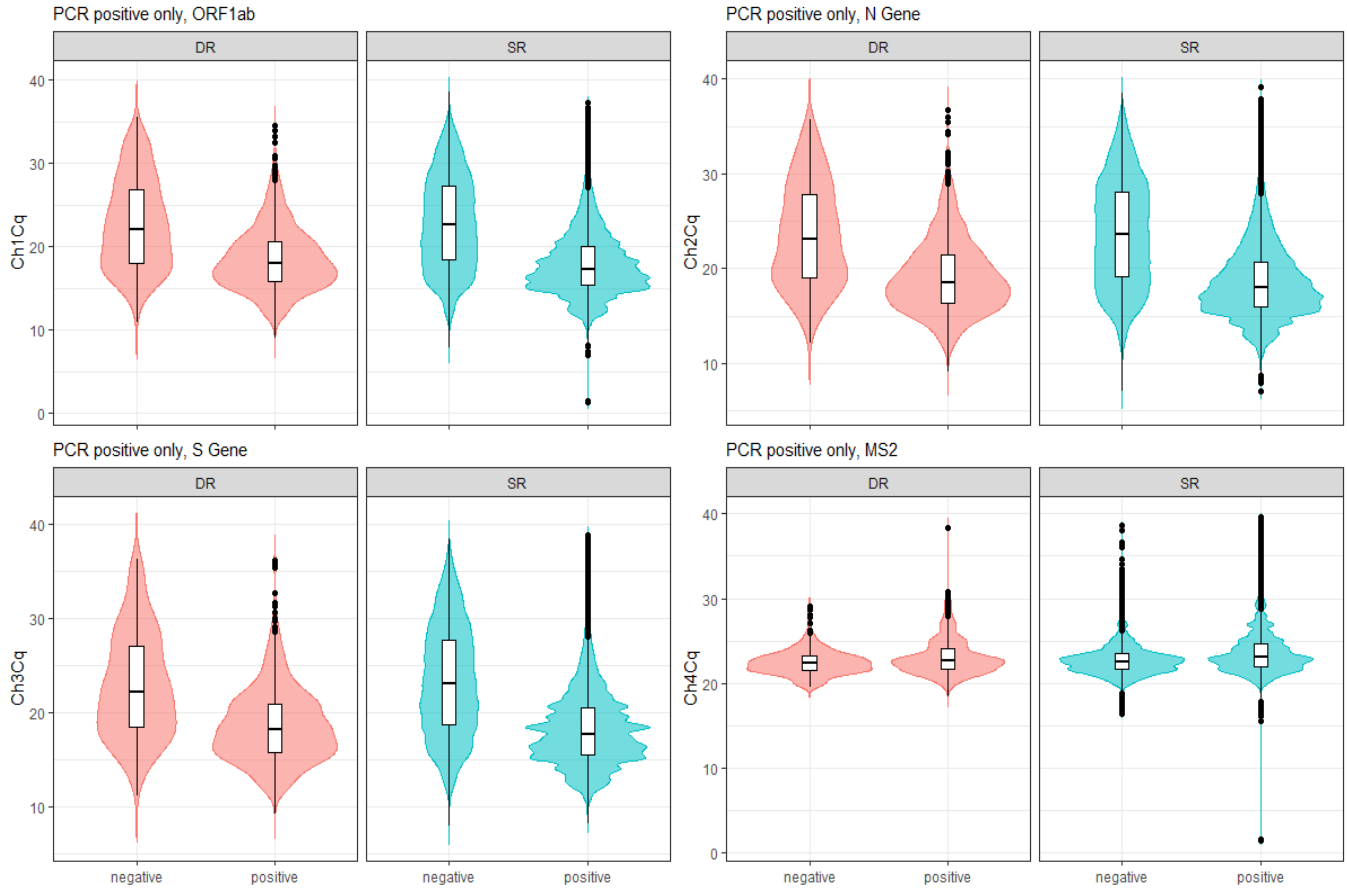
Table 10. Comparison of key metrics

	Self-read (95% CI)	**DR (95% CI)**
Sensitivity/true positive rate (TPR)	**0.5786** (0.5694 to 0.5877)	**0.7261** (0.7028 to 0.7485)
Specificity/true negative rate (TNR)	**0.9977** (0.9976 to 0.9978)	**0.9602** (0.9568 to 0.9635)
Accuracy (ACC)	**0.9928** (0.9927 to 0.9930)	**0.9365** (0.9324 to 0.9403)
False positive rate (FPR)	**0.0023** (0.0022 to 0.0024)	**0.0401** (0.0368 to 0.0436)
Positive predictive value (PPV)	**0.7444** (0.7352 to 0.7535)	**0.6737** (0.6502 to 0.6965)
Negative predictive value (NPV)	**0.9951** (0.9949 to 0.9952)	**0.9688** (0.9657 to 0.9717)

Box-Tidwell test for linearity indicates strong non-linearity for the logit transform when looking at all PCR positive tests, and for each subset by LFT results. Therefore, results of logistic regression were not fitted.

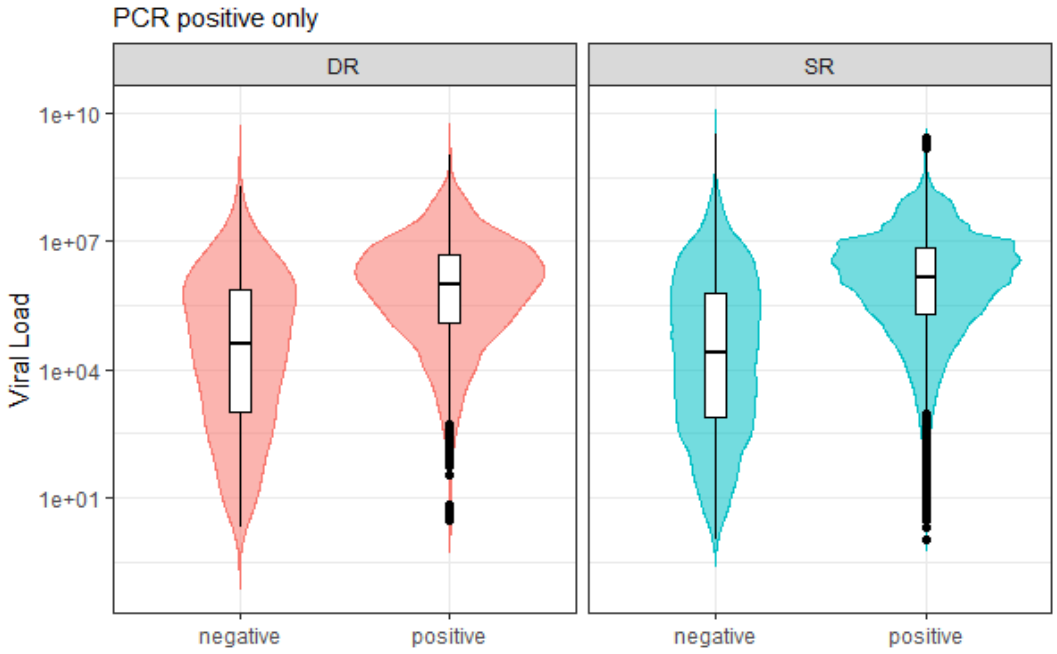
Comparison of DR and self-read RT-PCR Ct values suggests that there was very little difference in the observed median Ct values for both the result pathways and this is plotted in Figure 5.

Figure 5. Violin plots for the observed Ct values for the genes tested by RT-PCR



Viral load calculated based on the observed Ct values also suggests very little difference in the observed median viral load Figure 6.

Figure 6. Viral load calculated based on observed Ct



Objective 3. User journey performance

Demographic data was available for England only between 3 June and 19 November 2021 due to the limitations of the extract from the EDGE database (DAs do not submit to the EDGE database).

A total of 136,784 individuals submitted 1,064,234 tests results during this period. The distributions of age, gender and ethnicity are presented in Table 11 and Table 12, respectively.

Table 11. Distribution of age, gender and number of tests of LFD digital reader users

	Female		Male		Unknown		
Age	**n**	%	**n**	**%**	**n**	**%**	**Unique people**
0 to 9	75	0.1%	64	0.0%	0	0.0%	139
10 to 19	3,467	2.5%	1,673	1.2%	12	0.0%	5,152
20 to 29	16,341	11.9%	5,116	3.7%	43	0.0%	21,500
30 to 39	20,589	15.1%	6,689	4.9%	30	0.0%	27,308
40 to 49	24,892	18.2%	7,290	5.3%	20	0.0%	32,202
50 to 59	28,228	20.6%	7,175	5.2%	12	0.0%	35,415
60 to 69	10,385	7.6%	3,184	2.3%	4	0.0%	13,573
70 to 79	908	0.7%	433	0.3%	0	0.0%	1,341
80 to 89	77	0.1%	48	0.0%	0	0.0%	125
90 to 99	19	0.0%	5	0.0%	0	0.0%	24
100 to 109	1	0.0%	0	0.0%	0	0.0%	1
Unknown	2	0.0%	2	0.0%	0	0.0%	4
Total	**104,984**	76.8%	**31,679**	23.2%	**121**	0.1%	136,784

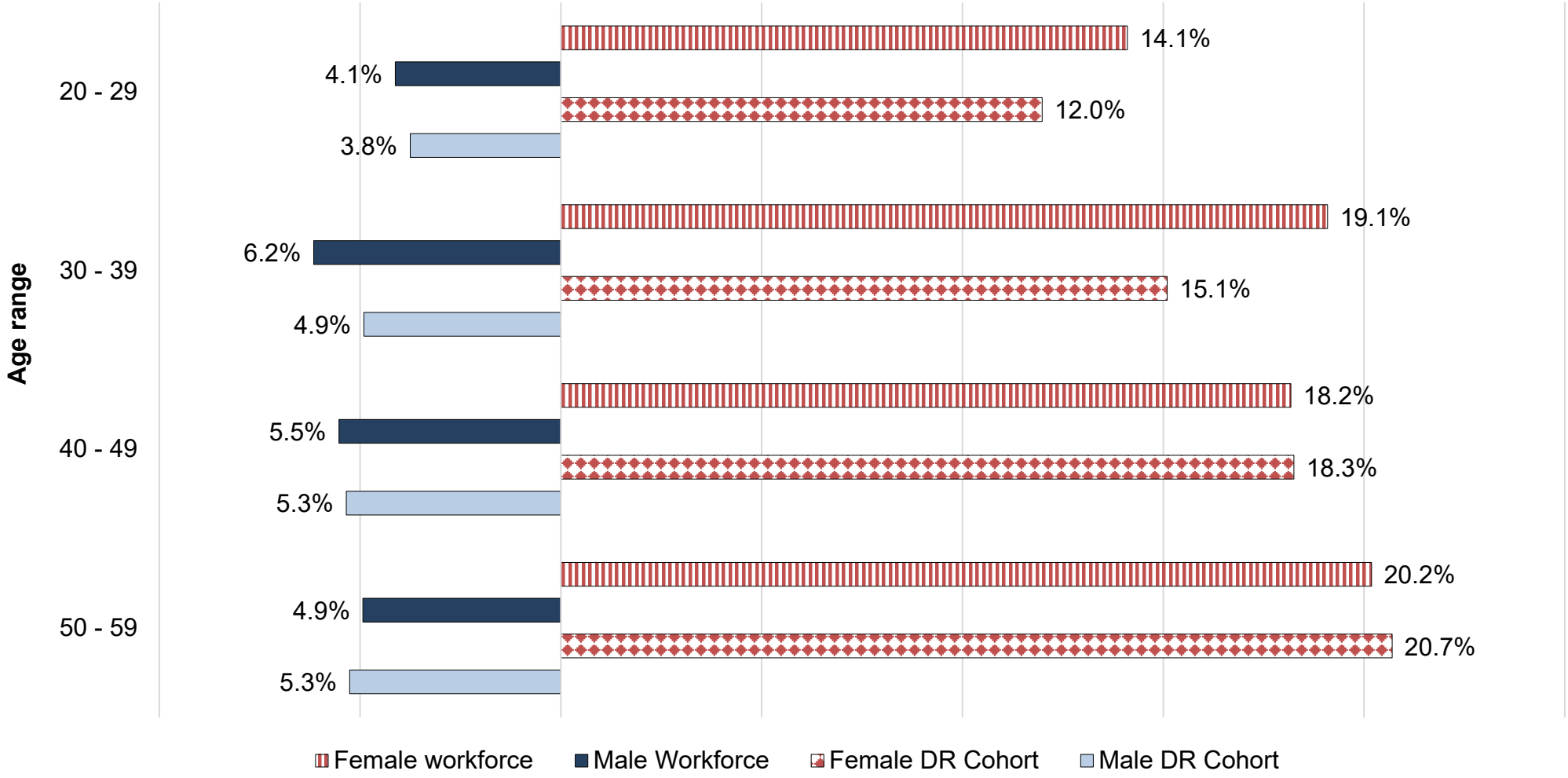
Table 12. Ethnicity distribution of LFD digital reader users

Ethnicity	**Total tests**	**Unique people**	**% of tests**	**% of users**	**Test per person**
Asian or Asian British	73,246	11,831	6.9%	8.6%	6.19
Black, African, Black British or Caribbean	18,783	3,699	1.8%	2.7%	5.08
Mixed or Multiple ethnic groups	18,544	2,871	1.7%	2.1%	6.46
White	929,372	114,234	87.3%	83.5%	8.14
Other ethnic group	2,965	576	0.3%	0.4%	5.15
Unknown or prefer not to say	21,324	3,573	2.0%	2.6%	5.97
Total	**1,064,234**	**136,784**	**100%**	**100%**	-

For the purposes of our analyses, where possible, the 0 to 9 and over 80 age groups (n=290, 0.21%) have been excluded as they do not reflect the eligible population of the original DR process. These may represent incorrect data entry (by the user) or inappropriate use of the service but, whatever the explanation, they fall outside the scope of this evaluation.

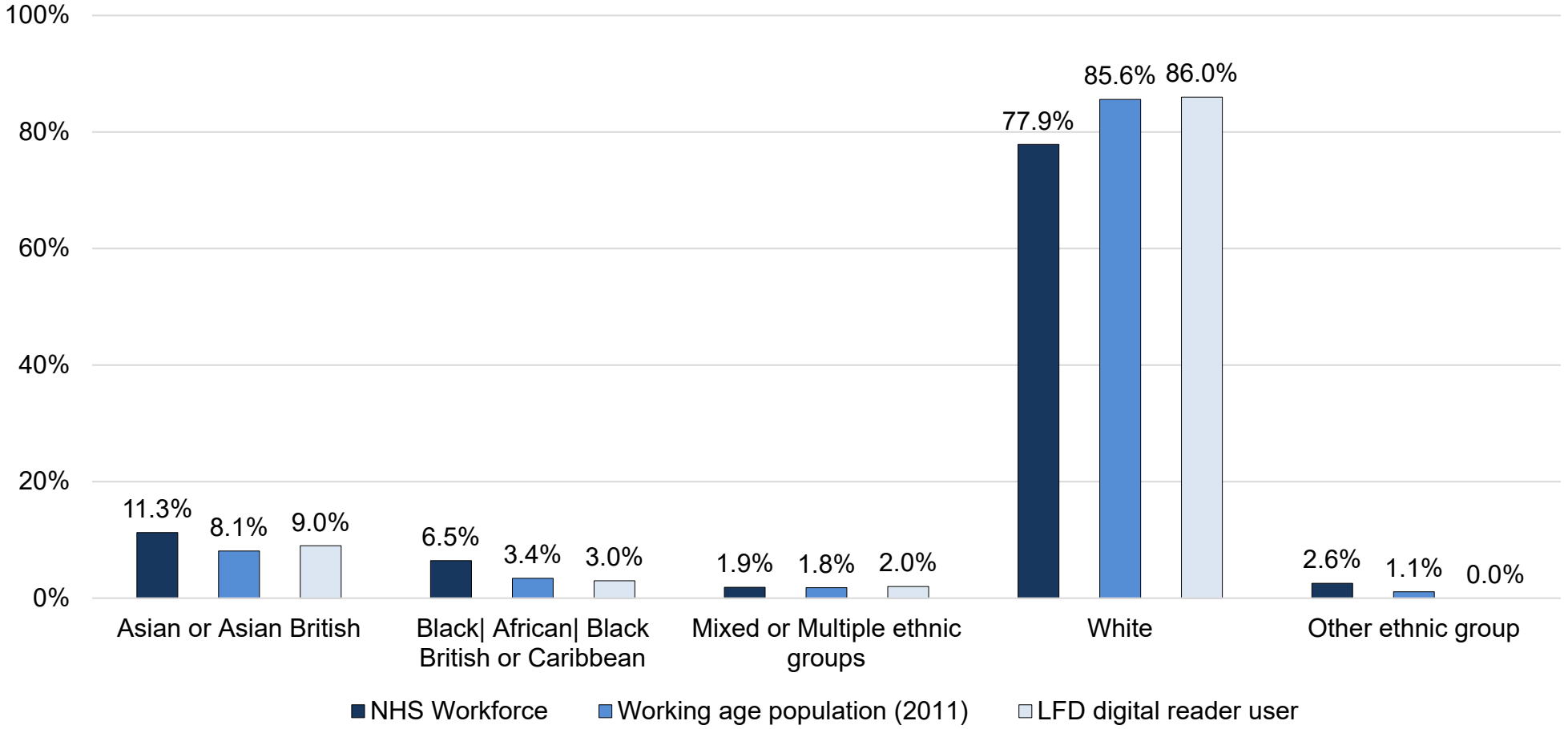
Comparison of the English DR users with the NHS workforce and the working age population (from the 2011 census) did not reveal substantial deviation from expected, thus we can be reassured that there is no impactful selection bias in terms of age (Figure 7) or ethnicity (Figure 8).

Figure 7. Age distribution of LFD DR users and health and social care staff in England 2020[10],[11]



[10] Number of health and care workers in England and Wales for 2019 and 2020 (ONS)
[11] Age ranges 0-19 (n=5,291 out of 136,663, 3.8%) and 60+ (n=15,068 out of 136,663, 11%) were excluded as these were not represented in ONS data or only had partial data in the DR set.

Figure 8. Ethnicity distribution of NHS workforce, working age population[12] and LFD digital reader users



Complete data from the monthly post-market surveillance reports were available for all 4 nations from June 2021 to 31 December 2021. Some data specifics were not available between 25 and 30 September 2021, and only partial data was available from the 1 October 2021 to 31 October 2021 due to technical issues. In total 1,840,896 images were processed up to 31 December 2021 and key performance metrics related to the user journeys are presented in Table 13.

[12] NHS Workforce Statistics October 2021, including selected provisional statistics for November 2021 (NHS Digital)

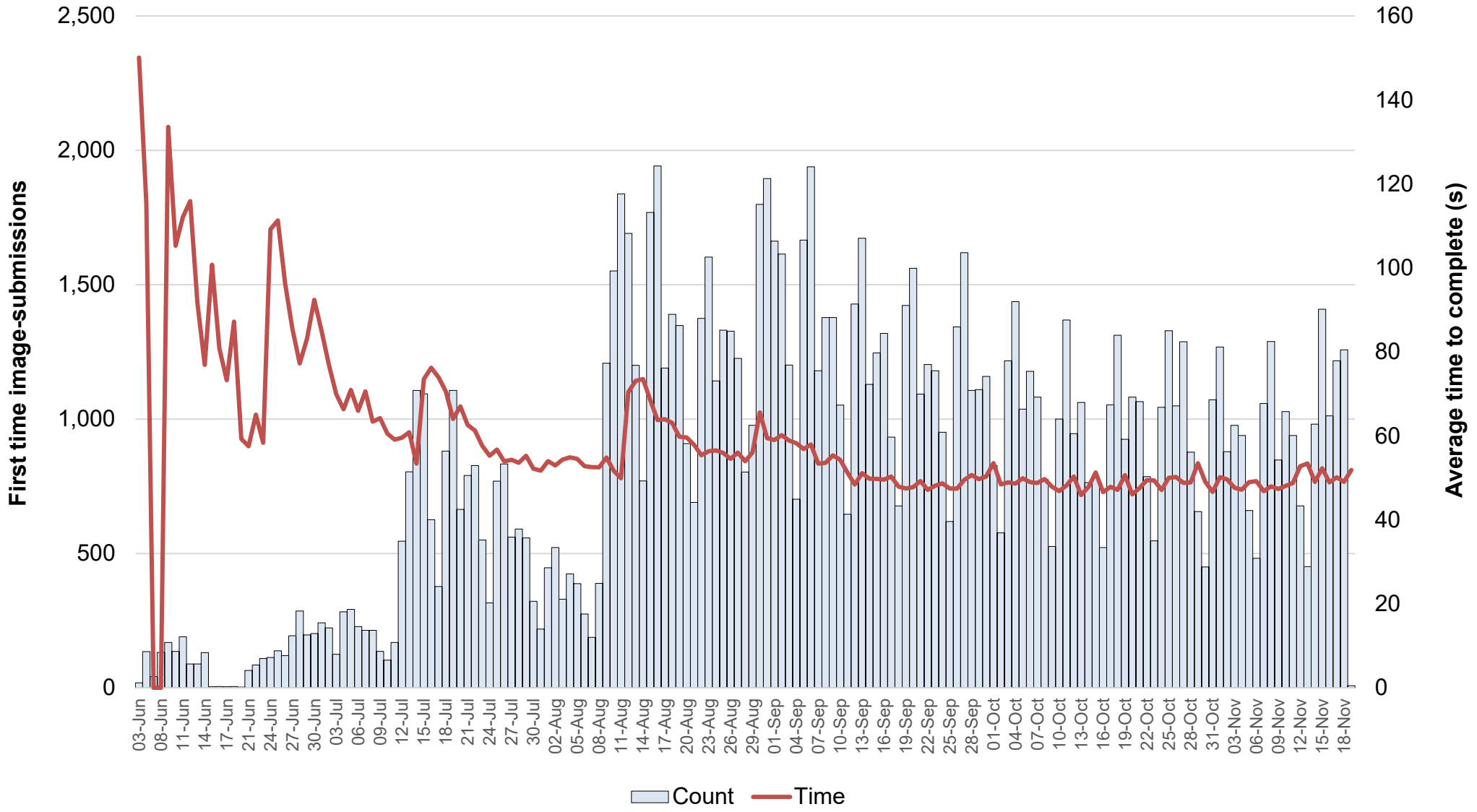
Table 13. Digital reader usage data June to December 2021

Cohort type	Date range	Total images	Images per day (1,000's)	Results (%)		Successful submission (%) at			Manual report	Journey (%) abandoned after	
				Valid	Invalid	1st try	2nd try	3rd try		1 invalid image	2 invalid images
n	03/06/2021 to 13/06/2021	2,104	0.19	84	16.0	88.0	5.3	1.2	1.9	1.2	0.4
n	14/06/2021 to 20/07/2021	66,181	1.79	85.8	14.2	90.3	5.3	1.5	2.1	0.5	0.3
i	21/07/2021 to 31/08/2021	313,054	7.45	85.3	14.7	89.9	5.3	1.5	2.2	0.6	0.4
i	01/09/2021 to 24/09/2021	295,436	12.31	85.9	14.1	90.2	4.9	1.5	2.1	1.0	0.3
e	01/10/2021 to 31/10/2021	371,045	11.97	86.7	13.3	-	-	-	-	-	-
e	01/11/2021 to 30/11/2021	327,570	10.92	86.8	13.2	89.7	5.5	1.6	2.6	0.7	0.2
e	01/12/2021 to 31/12/2021	465,506	15.02	84.4	15.6	89.2	5.6	1.6	2.6	0.8	0.2
	Total	**1,840,896**									

Analysis of the first 2 time periods was particularly important given that they reflect a naïve cohort (indicated with ‘n’) with no prior experience of the DR process. The second 2 time periods represent intermediate cohorts (indicated ‘i) with a mix of naïve and experienced users and the final 3 time periods represent experienced cohorts (indicated ‘e’) where most users have experience and addition of naïve users has little impact on overall metrics.

There was generally a very high number of successful image submissions at first tries and valid result rates, irrespective of cohort experience. This demonstrated that the system was successfully used from inception and while there is some improvement in these metrics over time, the fact that rates are high and generally consistent between naïve and experienced cohorts suggests that successfully completing the process is relatively insensitive to previous experience.

Figure 9. Average time taken for users to complete digital reader process plotted against new users submitting records (June to November 2021)



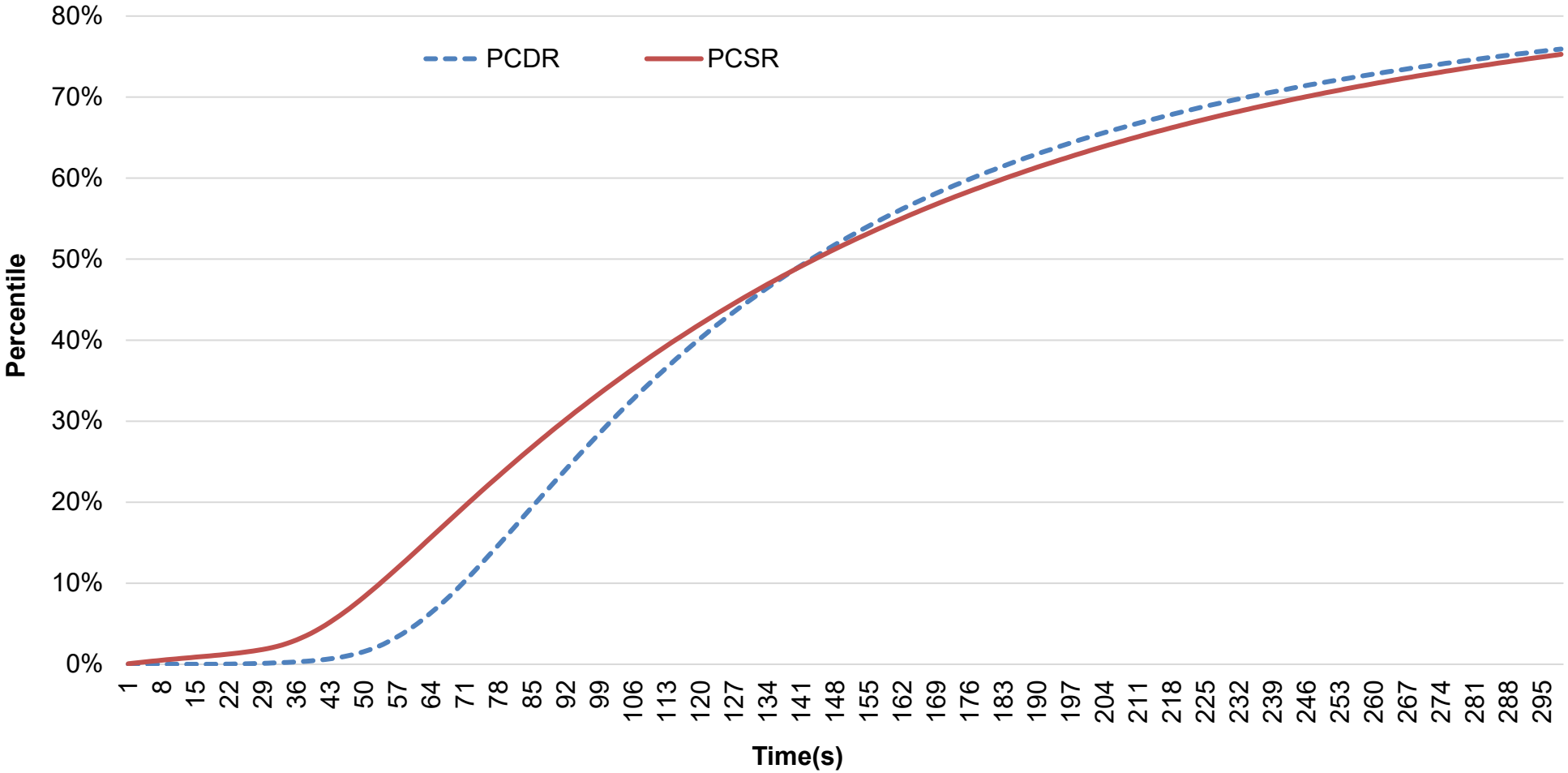
Where the impact of experience did appear to manifest itself was in the time taken to complete the DR process (Figure 9). The more experienced cohort (t=49 seconds, 95%CI:48.7 to 49.4s) was clearly able to complete the entire process successfully far more quickly than the naïve cohort (t=76.5s, 95%CI:68.9 to 84.1s).

The impact of an influx of naïve users on time taken to complete was considerable. Where new users were given access to the DR system there was a spike in the time required for completion (mid-June, mid-July, mid-August, late-August) and as users get accustomed to the process, the times generally improved back to a baseline. Additionally, as the proportion of new users in the overall pool decreased, their impact on the average time was also dampened, as can be seen by the gradual plateauing of average time taken despite relatively high numbers of newly onboarded individuals in the latter months.

Figure 9 time-series data was obtained from analytics provided by the DR web-portal which logged user interaction with the service. The main caveats of the data are that they only represent users who have active cookies (tracking code related to websites) and thus may exclude users or organisations who disable these.

We also compared time taken for DR versus self-read to ensure that operation of the service did not unduly burden users. Comparison of times taken (Figure 10) showed that the median time for both were comparable with DR users taking 143 seconds (IQR 93 to 286s) and self-read users taking 144s (IQR 81 to 296s) to complete the process.

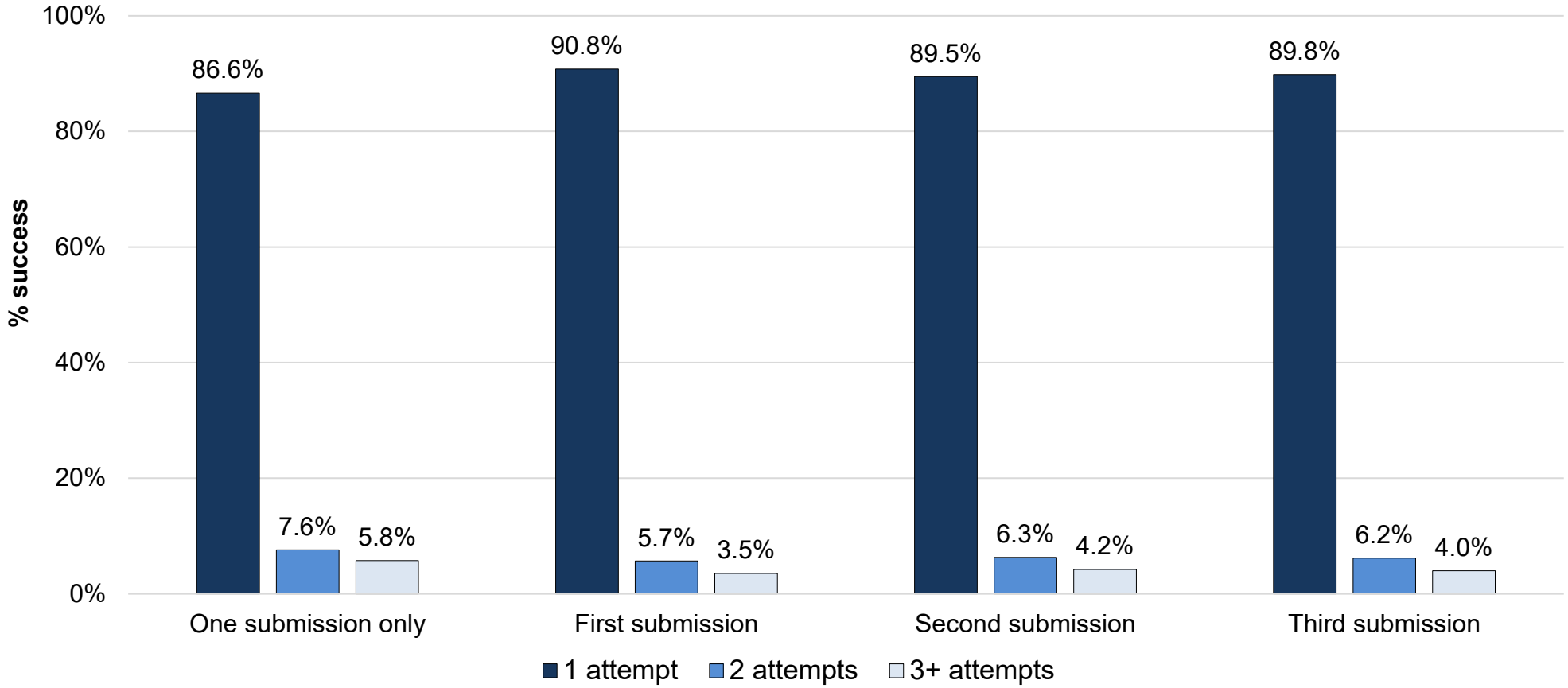
Figure 10. Percentile chart of time taken to complete the digital reader (PCDR) versus self-read (PCSR) processes



We also assessed whether there was an element of reporting bias within the cohorts by studying the performance metrics for users who only ever submitted a single DR test image versus those who submitted multiple LFD images over time. The hypothesis was that users who only used the system once may have done so because they experienced difficulty in doing so and thus were not willing to try again.

Success at nth try image submission data was recovered from the EDGE database (with the limitations identified earlier) for users who only ever submitted a single test (n=35,275) and this was compared to users with multiple submissions over the same period (n=100,837); the data collected were plotted in Figure 11. Overall, single use users have similar success at first try rates to repeat users, the latter also not seeing a significant improvement in success at first try rates despite continuous use of the service. From this we can infer that usability issues were likely not the reason some users chose not to continue using the DR process.

Figure 11. Number of attempts required for successful submission between single occasion and repeat users

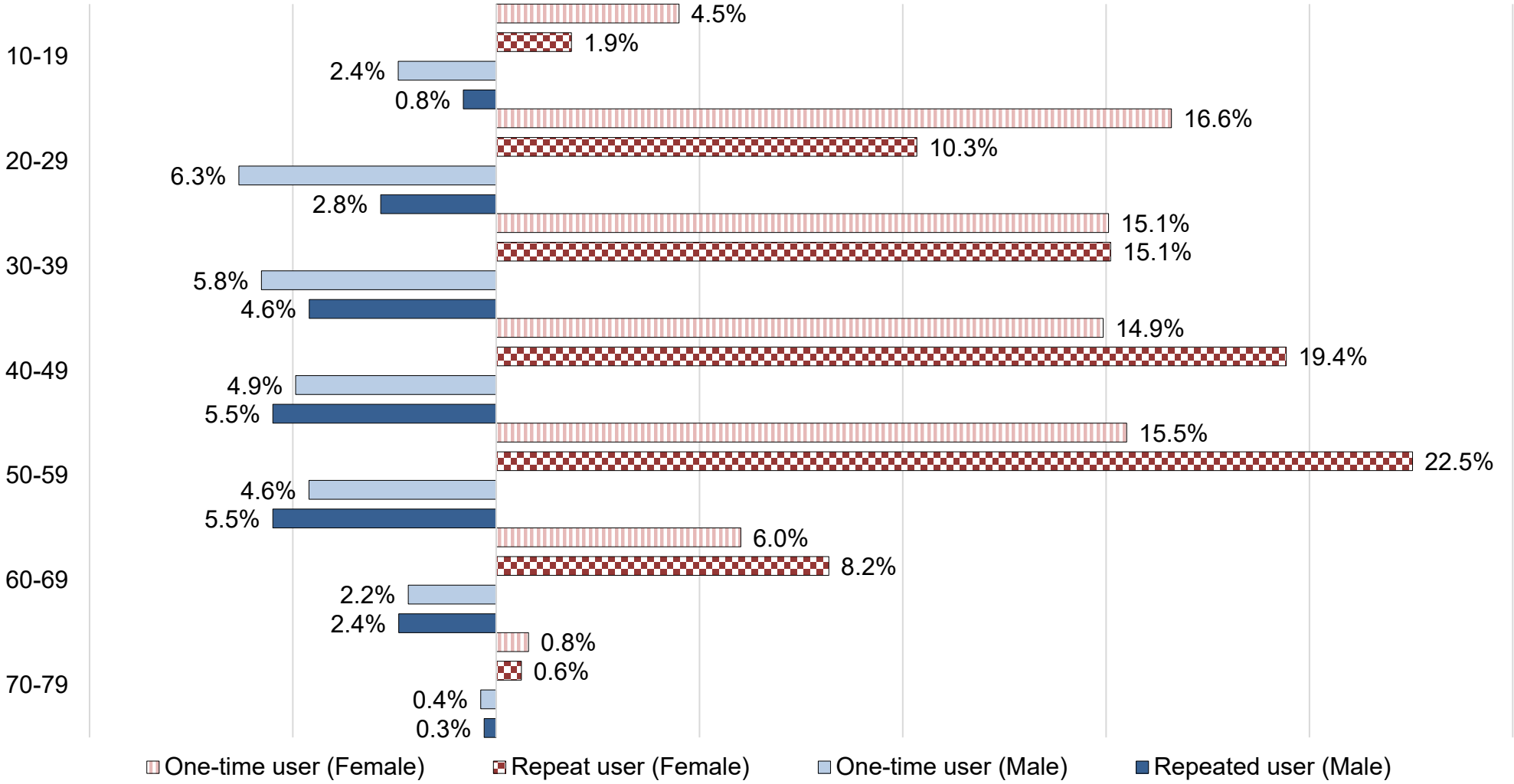


Comparison of the demographics between one-time users and repeat users demonstrated that there were some differences between age groups with one-time users having a greater representation of the younger age groups (10 to 29) whilst the repeat users had a greater representation in the higher age groups (40 to 59) (Table 14 and Figure 12).

Table 14. Distribution of age and gender of one-time users and repeat users

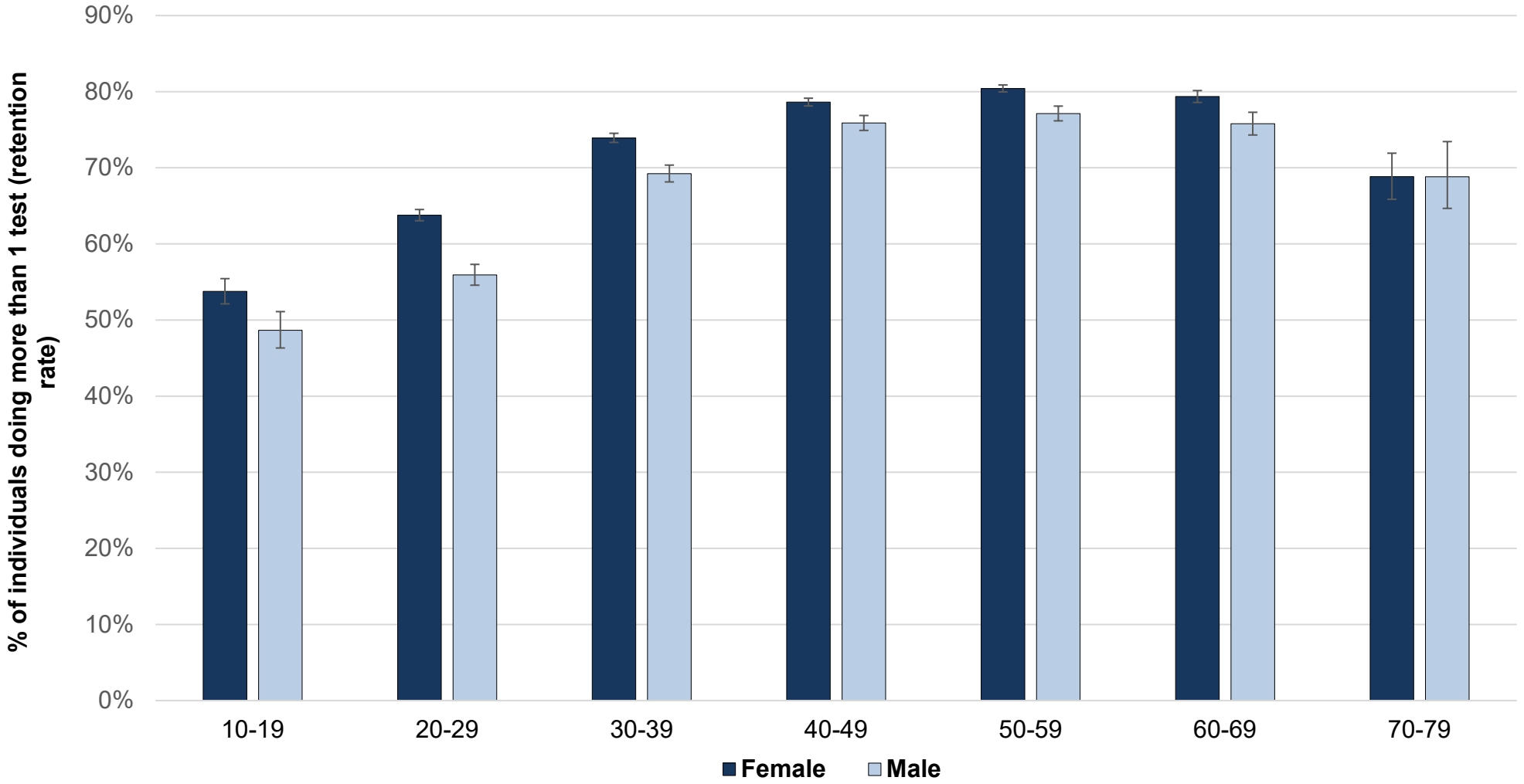
	One-time users				Repeat users			
	Female		Male		Female		Male	
Age	n	%	n	%	n	%	n	%
10 to 19	1,603	4.5%	859	2.4%	1,864	1.8%	814	0.8%
20 to 29	5,919	16.6%	2,255	6.3%	10,422	10.3%	2,861	2.8%
30 to 39	5,368	15.0%	2,058	5.8%	15,221	15.1%	4,631	4.6%
40 to 49	5,320	14.9%	1,758	4.9%	19,572	19.4%	5,532	5.5%
50 to 59	5,527	15.5%	1,641	4.6%	22,701	22.5%	5,534	5.5%
60 to 69	2,144	6.0%	771	2.2%	8,241	8.2%	2,413	2.4%
70 to 79	283	0.8%	135	0.4%	625	0.6%	298	0.3%
Total	26,164	73.3%	9,477	26.6%	78,646	78.0%	22,083	21.9%

Figure 12. Population pyramid comparing one-time and repeated DR process users



We also calculated the relative retention rates for age/sex (Figure 13) and age/ethnicity (Figure 14). The retention rates are defined as the proportion of repeat DR users in the cohort and provide a good proxy measure of ease of use, as we would expect a difficult and/or time-consuming process to have an adverse impact on repeat use.

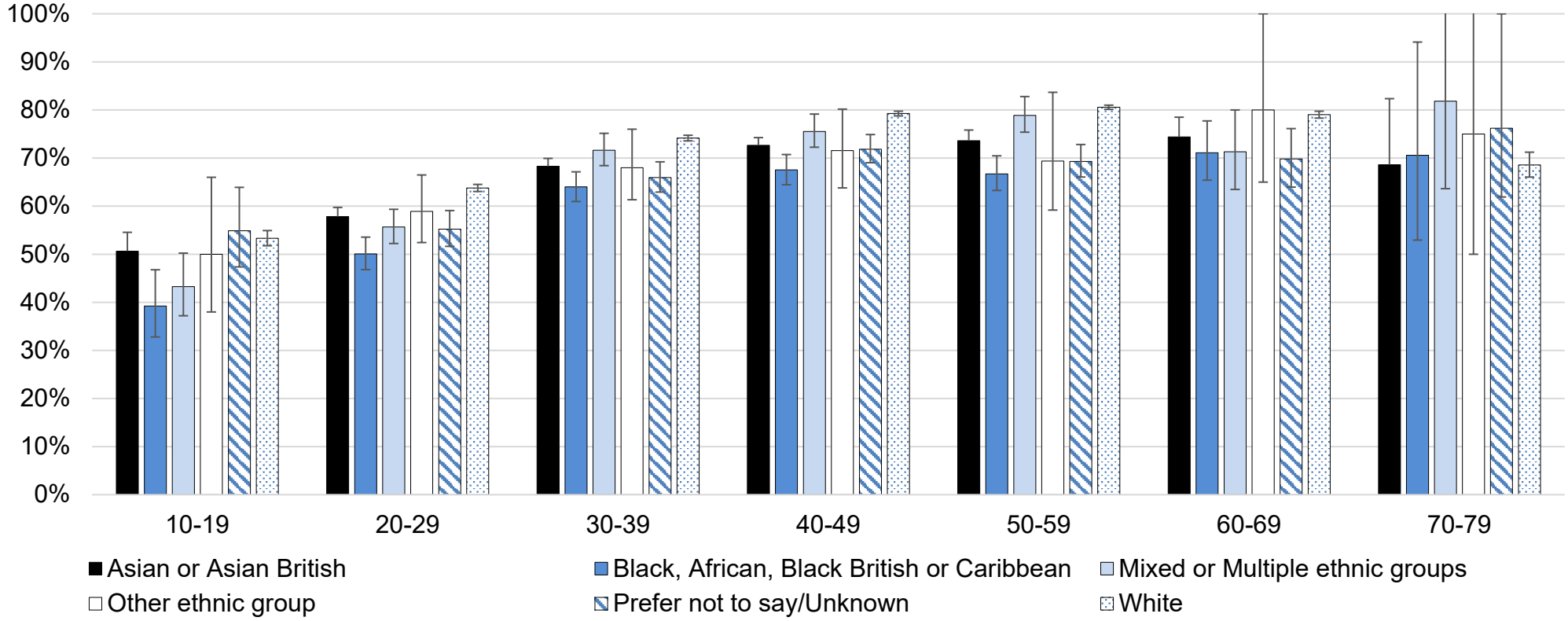
Figure 13. Percentage of repeat users by age and sex



With respect to age and sex, there is a lower retention rate in the 10 to 19 (52.1%) and 20 to 29 (61.9%) age groups compared to all other age groups with the highest retention rates seen in the 50 to 59 group (79.8%). In all age groups except 70 to 79, females were more likely to be repeat users than males ($p<0.05$).

The lower retention rate of the younger age groups may be concerning, especially given that they account for approximately 20% of the DR cohort. However, the real-world impact may not be substantial given that those in younger age groups are likely to be more resilient to scenarios which would impact interpretation of an LFD device (for example, age-related visual acuity problems). A similar pattern was also noted when looking at ethnicity across the age bands, with the younger (10 to 29) groups generally having lower retention rates than older ones. Overall, there does not appear to be any ethnic bias with respect to retention rates.

Figure 14. Percentage of repeat users by age and ethnicity



It is essential to note that we were unable to fully distinguish between users who simply stopped using the DR and those who stopped testing or reporting their test results. This analysis fell outside the remit of evaluation of the DR.

Objective 4. Value for money

1,376 images were included in the calculation of DR sensitivity and precision. Compared to self-read, baseline sensitivity of the DR during this period was 100% (CI:99.23% to 100%) and precision was 81.5% (CI: 78.44% to 84.72%). The 4x4 contingency table is presented in Table 15.

Table 15. 4x4 contingency table of ground truth versus digital reader

		Ground truth				
		Negative	Positive	Void	Invalid	Total
Digital reader	Negative	500	0	0	0	**500**
	Positive	107	480	2	0	**589**
	Void	229	0	44	14	**287**
	Invalid	0	0	0	0	**0**
	Total	**836**	**480**	**46**	**14**	**1,376**

A worst-case scenario model as previously described in the methods section was constructed for each context comparing the DR sample to all self-read results for staff testing in ASC and primary care (Table 16). Educational (schools, colleges, universities), private sector, public sector and secondary care settings are not included due to limited sample sizes.

The model assumed that both the DR and self-read data sets had the same underlying true prevalence and thus any difference in their reported positive rates were due to interpretative differences of the LFD tests between the cohorts.

Table 16. Context specific actual and WCS estimated modelling of digital reader versus self-read

Context	Read Mode	Total Tests	Negatives	Positives	Voids	% positives (95% CI)	% negatives	WCS estimated positives	% Positives updated	Gain (95% lower bound)
Adult social care	Digital	**7,640**	7,570	45	25	0.59% (0.43 to 0.79)	99.08%	37	0.48% (0.34 to 0.65)	89.42% (>34.29)
	Self	**722,145**	720,090	1,830	225	0.25% (0.24 to 0.27)	99.72%			
Primary care	Digital	**54,395**	53,860	360	175	0.66% (0.60 to 0.73)	99.02%	293	0.54% (0.48 to 0.60)	33.72% (>18.96)
	Self	**94,210**	93,785	380	45	0.40% (0.36 to 0.45)	99.55%			

Despite the relatively low numbers of DR images in ASC, under the WCS compared to self-reading, the DR demonstrated an improvement in sensitivity of 89.42% with the lower bound being 34.29%. Thus, even with the most pessimistic outlook, this still represents a significant gain over self-reading.

In primary care the gain in sensitivity was less marked but remained significant with an average gain of 33.72% and a lower bound at 18.96%.

Assuming that under optimum conditions, a trained and experienced user and the DR have equivalent performance in terms of interpreting LFDs, then across all devices and all viral concentrations the optimum effective sensitivity of an LFD device is approximately 80%. On a population level however, these conditions do not hold, and thus real-world sensitivity is around 40 to 60%.

We extended the analysis further to estimate the sensitivity in each of those 2 settings given the observed positivity in each journey, and number of positives that were potentially missed via self-report. Using the combined estimated sensitivity of the devices used and the digital reader from Objective 3 as a reasonable worst-case scenario (72.61%; Table 8), we estimated that the sensitivity of self-reports in ASC was around 40% while the sensitivity for self-reports by primary care staff was around 58%.

Although data for other contexts were not included due to limited sample size, there was moderate evidence that use of the DR in the private sector would result in an improvement in sensitivity comparable to that observed in primary care staff, and weak evidence that its use in educational settings would result in some benefit.

As of 10 March 2022, vaccination rates amongst the sub-groups of interest remain low with 69.7% of at-risk carers, and 78.1% of secondary care staff fully vaccinated and boosted. The overall rate was 73.1% for the adult population.[13] Prevalence in England was high at around 5% with the same figures reflected amongst social and healthcare settings. We included this scenario as well as a more optimistic 1% prevalence in our analysis. Given the reproductive rate of the current variants and the trends observed in the past 2 years, a prevalence of 1% was likely to be observed over the Summer. However, a prevalence of 5% is a plausible scenario from October 2022 onwards and throughout the winter. It is worth noting that the added benefit in terms of sensitivity improvement of the Digital Reader does not rely on the performance of the available LFDs, nor does it rely on the underlying prevalence since sensitivity is an intrinsic property of the device. Therefore, as long as LFD-type devices are available and a Digital Reader has been trained to read such devices, it is almost certain that there will always be an improvement in sensitivity when compared to human ability.

The underlying assumptions outlined here were used with the UKHSA VFM tool with the outputs in Tables 17 to 19.[14] We note that, with a conservative sensitivity estimate of 73% for the digital reader used in conjunction with LFDs currently available for use:

- a further 4,353 first-generation infections could be prevented per million tests at the current prevalence (March 2022) in adult social care, and 1,980 in primary care; if the prevalence were to drop to 1%, these values are estimated at 873 and 393 per million tests respectively
- a net benefit of at least £3.64 million for 1.5million tests would be observed in ASC under the scenario where 70% of the target population is fully vaccinated and boosted
- a net benefit of at least £2.24 million for 1.5million tests would be observed in primary care under the scenario where 70% of the target population is fully vaccinated and boosted

Table 17. Generation 1 excess infections prevented per million tests of digital reader implementation using ASC sensitivity (40%) and Primary care sensitivity (58%) as baseline, and 2 possible DR sensitivity scenarios (73% and 80%)

LFD plus DR Sensitivity	1% Prevalence (optimistic)		5% Prevalence (current)	
	ASC	Primary care	ASC	Primary care
73%	873	393	4,353	1,980
80%	1,060	580	5,280	2,907

[13] COVID-19 vaccinations
[14] The Canna model: assessing the impact of NHS Test and Trace

Table 18. Generation 1 net benefits (£millions) of digital reader implementation using ASC sensitivity (40%) as baseline, and 2 possible DR sensitivity scenarios (73% and 80%)

Cost scenario	Sensitivity	1% Prevalence (optimistic)			5% Prevalence (current)		
		Vaccination rate			Vaccination rate		
		60%	70%	80%	60%	70%	80%
0.2 per read, 1.5m tests	73%	0.73	0.59	0.43	4.49	3.74	2.99
	80%	0.93	0.75	0.57	5.49	4.57	3.66
0.3 per read, 1.5m tests	73%	0.63	0.49	0.33	4.39	3.64	2.89
	80%	0.83	0.65	0.47	5.39	4.47	3.56

Table 1. Generation 1 net benefits (£millions) of digital reader implementation using Primary care sensitivity (58%) as baseline, and 2 possible DR sensitivity scenarios (73% and 80%)

Cost scenario	Sensitivity	1% prevalence (optimistic)			5% prevalence (current)		
		Vaccination rate			Vaccination rate		
		60%	70%	80%	60%	70%	80%
0.2 per read, 1.5m tests	73%	0.34	0.23	0.13	2.90	2.39	1.87
	80%	0.64	0.48	0.34	4.40	3.64	2.88
0.3 per read, 1.5m tests	73%	0.19	0.08	-0.02	2.75	2.24	1.72
	80%	0.49	0.33	0.19	4.25	3.49	2.73

Conclusions

This report presents the evaluation of the real-world deployment of the LFD digital reader, one of the first widely used devices for aiding the interpretation of SARS-CoV-2 lateral flow devices.

We assessed the ability of the digital reader to correctly interpret LFD images by comparing DR outputs to those of an expert panel. The output accuracy analysis suggested that the LFD reader works well. The observed sensitivity and specificity reported were along similar lines to that reported by real world performance monitoring of LFD devices deployed by DHSC/UKHSA. A limitation of our study was the need to subset the number of tests studied in order to establish the ground truth, which as there was relatively low prevalence, meant that small changes in true positive cases would have a large effect on sensitivity. The upper bound of our sensitivity was > 90% demonstrating considerable benefits over self-reading which has reported sensitivities as low as 16% when compared to expert readers.

We then evaluated the concordance of the digital reader compared to RT-PCR and self-read LFD reports. Our analysis demonstrated that the DR has a good concordance with RT-PCR results and has higher accuracy at detecting positives and negatives compared to self-read LFD reports. It outperforms self-read LFD reports by detecting nearly twice the number of positive cases during early infection when the viral loads tend to be lower. DR has higher sensitivity than self-report results but has marginally lower specificity. The large sample size of negatives for self-reported LFD results is likely to have influenced the specificity and the accuracy of the DR and self-read concordance values.

Our evaluation of user journey performance for objective 3 suggested that the DR was easy to use, and no inherent bias was detected across the varying demographics that were analysed. While it is not clear why a large cohort of DR users only reported one LFT result, our analysis suggests that there was high retention of users in the older age groups while there was lower retention rate in the under 30s. A potential explanation is the novelty of the device encouraged first time users. Further analysis would be required to ascertain if the lower retention rate was due to users simply stopping to use the DR or if they stopped testing or reporting their results.

Finally, we performed a value for money analysis to assess the financial benefits of deploying the DR at a population level. VFM modelling suggested that compared to general population, at a prevalence rate of 5%, DR would prevent an additional 4,353 generation 1 infections per million tests in adult social care, and 1,980 in primary care.[15] In the current scenario[16] where 70% of the population is vaccinated and boosted, net benefit (over and above the cost of testing and the cost of the reader) of at least £3.64 million per 1.5 million tests would be observed in ASC, and a net benefit at least £2.34 million per 1.5 million tests would be observed through similar deployment of the DR for testing primary care staff.

[15] Generation 1 infections are not restricted to the stated staff groups as this represents onward transmission of any kind.
[16] Refers to time that analysis was conducted, namely late winter or early spring 2022.

In conclusion, we have demonstrated that the LFD digital reader outperforms human readers in the accurate interpretation of SARS-CoV-2 LFD tests, has high usability and provides a net cost benefit to the use of LFD in community testing, meaning it will save money over the short, medium and long term. We recommend it should be immediately implemented in all settings that use lateral flow devices.

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.

© Crown copyright 2022

Published: November 2022
Publishing reference: GOV-13102

You may re-use this information (excluding logos) free of charge in any format or medium, under the terms of the Open Government Licence v3.0. To view this licence, visit OGL. Where we have identified any third party copyright information you will need to obtain permission from the copyright holders concerned.

UKHSA supports the
Sustainable Development Goals