Real-world evaluation of AI-driven COVID-19 triage for emergency admissions: External validation & operational assessment of lab-free and high-throughput screening solutions

Dr. Andrew A. S. Soltan MB BChir[1,2]
Jenny Yang MSc[3]
Dr. Ravi Pattanshetty FRCEM[1]
Dr. Alex Novak FRCEM[1]
Dr. Yang Yang DPhil[3]
Dr. Omid Rohanian PhD[3]
Sally Beer[1]
Dr. Marina A. Soltan MRCP[5,6]
Prof. David R. Thickett FRCP[5,6]
Dr. Rory Fairhead BA[4]
CURIAL Translational Collaborative
Dr. Tingting Zhu DPhil[3]
Prof. David W. Eyre DPhil[1,7,8]
Prof. David A. Clifton DPhil[3]

Affiliations:

1. John Radcliffe Hospital, Oxford University Hospitals NHS Foundation Trust
2. RDM Division of Cardiovascular Medicine, University of Oxford
3. Institute of Biomedical Engineering, Dept. Engineering Science, University of Oxford
4. University of Oxford Medical School
5. The Queen Elizabeth Hospital, University Hospitals Birmingham NHS Foundation Trust
6. Institute of Inflammation and Ageing, University of Birmingham
7. Big Data Institute, Nuffield Department of Population Health, University of Oxford
8. NIHR Health Protection Research Unit in Healthcare Associated Infections and Antimicrobial Resistance at University of Oxford in partnership with Public Health England

Keywords:
SARS-CoV-2, COVID-19, Artificial Intelligence, Machine Learning, Screening test, Validation, Operational evaluation, Diagnosis, Electronic Health Records, Emergency Department

Corresponding:
Dr Andrew A. S. Soltan MB BChir MA
NIHR Academic Clinical Fellow (Cardiology), University of Oxford
John Radcliffe Hospital, Headley Way, Oxford, OX3 9DU
Andrew.Soltan@cardiov.ox.ac.uk

Abstract

Background

Uncertainty in patients' COVID-19 status contributes to treatment delays, nosocomial transmission, and operational pressures in hospitals. However, typical turnaround times for batch-processed laboratory PCR tests remain 12-24h. Although rapid antigen lateral flow testing (LFD) has been widely adopted in UK emergency care settings, sensitivity is limited. We recently demonstrated that AI-driven triage (CURIAL-1.0) allows high-throughput COVID-19 screening using clinical data routinely available within 1h of arrival to hospital. Here we aimed to determine operational and safety improvements over standard-care, performing external/prospective evaluation across four NHS trusts with updated algorithms optimised for generalisability and speed, and deploying a novel lab-free screening pathway in a UK emergency department.

Methods

We rationalised predictors in CURIAL-1.0 to optimise separately for generalisability and speed, developing CURIAL-Lab with vital signs and routine laboratory blood predictors (FBC, U&E, LFT, CRP) and CURIAL-Rapide with vital signs and FBC alone. Models were calibrated during training to 90% sensitivity and validated externally for unscheduled admissions to Portsmouth University Hospitals, University Hospitals Birmingham and Bedfordshire Hospitals NHS trusts, and prospectively during the second-wave of the UK COVID-19 epidemic at Oxford University Hospitals (OUH). Predictions were generated using first-performed blood tests and vital signs and compared against confirmatory viral nucleic acid testing. Next, we retrospectively evaluated a novel clinical pathway triaging patients to COVID-19-suspected clinical areas where either model prediction or LFD results were positive, comparing sensitivity and NPV with LFD results alone. Lastly, we deployed CURIAL-Rapide alongside an approved point-of-care FBC analyser (OLO; SightDiagnostics, Israel) to provide lab-free COVID-19 screening in the John Radcliffe Hospital's Emergency Department (Oxford, UK), as trust-approved service improvement. Our primary improvement outcome was time-to-result availability; secondary outcomes were sensitivity, specificity, PPV, and NPV assessed against a PCR reference standard. We compared CURIAL-Rapide's performance with clinician triage and LFD results within standard-care.

Results

72,223 patients met eligibility criteria across external and prospective validation sites. Model performance was consistent across trusts (CURIAL-Lab: AUROCs range 0.858-0.881; CURIAL-Rapide 0.836-0.854), with highest sensitivity achieved at Portsmouth University Hospitals (CURIAL-Lab:84.1% [95% Wilson's score CIs 82.5-85.7]; CURIAL-Rapide:83.5% [81.8 - 85.1]) at specificities of 71.3% (95% Wilson's score CIs: 70.9 - 71.8) and 63.6% (63.1 - 64.1). For 3,207 patients receiving LFD-triage within routine care for OUH admissions between December 23, 2021 and March 6, 2021, a combined clinical pathway increased sensitivity from 56.9% for LFDs alone (95% CI 51.7-62.0) to 88.2% with CURIAL-Rapide (84.4-91.1; AUROC 0.919) and 85.6% with CURIAL-Lab (81.6-88.9; AUROC 0.925). 520 patients were prospectively enrolled for point-of-care FBC analysis between February 18, 2021 and May 10, 2021, of whom 436 received confirmatory PCR testing within routine care and 10 (2.3%) tested positive. Median time from patient arrival to availability of CURIAL-Rapide result was 45:00 min (32-64), 16 minutes (26.3%) sooner than LFD

results (61:00 min, 37-99; log-rank p<0.0001), and 6:52 h (90.2%) sooner than PCR results (7:37 h, 6:05-15:39; p<0.0001). Sensitivity and specificity of CURIAL-Rapide were 87.5% (52.9-97.8) and 85.4% (81.3-88.7), therefore achieving high NPV (99.7%, 98.2-99.9). CURIAL-Rapide correctly excluded COVID-19 for 58.5% of negative patients who were triaged by a clinician to 'COVID-19-suspected' (amber) areas.

Impact
CURIAL-Lab & CURIAL-Rapide are generalisable, high-throughput screening tests for COVID-19, rapidly excluding the illness with higher NPV than LFDs. CURIAL-Rapide can be used in combination with near-patient FBC analysis for rapid, lab-free screening, and may reduce the number of COVID-19-negative patients triaged to enhanced precautions ('amber') clinical areas.

Funding:
Wellcome Trust/University of Oxford Medical & Life Sciences Translational Fund (Award: 0009350).

Abbreviations

AI	Artificial Intelligence
AUROC	Area under receiver operating characteristic curve
BH	Bedfordshire Hospitals NHS Foundation Trust
COVID-19	Coronavirus Disease 2019
CRP	C-Reactive Protein
ED	Emergency Department
EHR	Electronic Health Records
FBC	Full Blood Count
LFT	Liver Function Test
LFD	Rapid antigen lateral flow device
NHS	National Health Service
NPV	Negative Predictive Value
OUH	Oxford University Hospitals NHS Foundation Trust
PCR	Real Time Polymerase Chain Reaction
POCT	Point of Care Test
PPV	Positive Predictive Value
PUH	Portsmouth University Hospitals NHS Trust
SARS-CoV-2	Severe Acute Respiratory Syndrome Coronavirus 2
SD	Standard Deviation
U&Es	Urea, Creatinine and Electrolytes
UHB	University Hospitals Birmingham NHS Foundation Trust

Background

Reducing nosocomial transmission of SARS-CoV-2 has been identified as a priority in safeguarding patient and healthcare staff safety, particularly as individuals with existing medical conditions are at greatest risk of severe illness and death [1–5]. However, as the early clinical course of infection is often characterised by weakly-specific symptoms, and can be asymptomatic, viral testing is necessary to identify cases and is mandated for all UK hospital admissions.[4]

The mainstay of testing is batch-processed laboratory polymerase chain assay (PCR), which is imperfectly sensitive and requires specialist equipment[6–8]. Turn-around times have shortened throughout the pandemic, typically to within 12-24h in hospitals in high- and middle- income countries, but the interim uncertainty around patients' COVID-19 status may contribute to treatment delays and postpone transfer to wards, thereby contributing to nosocomial transmission and operational strain. Novel rapid testing solutions have been adopted, including point-of-care (POC) PCR, loop mediated isothermal amplification, and lateral flow antigen testing (LFD), despite limitations in throughput and sensitivity[9,10]. Where POC PCR is available, use is typically constrained to time-critical decisions, such as surrounding emergency or transplant surgery, due to supply[11,12]. Moreover, although LFD is lab-free and highly specific (>99.5%) [13,14], allowing for a role in community case-finding[15], multiple reports show more limited sensitivity (~40%- 70%)[16–18] leading up to the U.S. Food and Drug Administration issuing a Class 1 recall on June 10, 2021[19]. A recent study evaluating performance amongst unscheduled hospital admissions confirmed high specificity (99.6%), but relatively low sensitivity (62%)[10].

We recently demonstrated that an artificial-intelligence (AI) screening test (CURIAL-1.0) can rapidly detect COVID-19 amongst patients being admitted to hospital, by recognising SARS-CoV-2-induced abnormalities in routinely collected data[20]. A strength of our approach is the use of readily available blood test, blood gas & physiological measurements which are typically collected within 1h of presentation to hospitals in high- and middle-income countries, without requiring patient exposure to ionising radiation[21,22]. Explainability analyses revealed that features most informative to predictions were components of the Full Blood Count (FBC) and vital signs (Basophil count, Eosinophil count and Oxygen requirements), offering promise for clinically-guided optimisation to reduce prediction time.

Whereas many studies have investigated AI applications for diagnosis and prognosis during the pandemic[23–26], key reviews highlight sector-wide methodological and reporting concerns that threaten generalisability, questioning the suitability of many models to-date for clinical use[27–29]. Reviewing the contribution of AI to the COVID-19 response, a recent editorial highlighted the promise of CURIAL-1.0 amongst other solutions to support patients during the pandemic, discussing the importance of high-quality validation studies inclusive of diverse patient populations[30]. Moreover, additional work quantifying benefits in the real-world clinical setting would demonstrate the clinical added-value of such approaches.

Accordingly, in this study we investigate efficacy, generalisability and real-world operational benefits of AI-driven COVID-19 screening in emergency departments, using insight from explainability analyses to improve generalisability and exploit

advances in near-patient diagnostics for lab-free COVID-19 screening[31]. First, we externally validate two models with rationalised predictors, optimised separately for throughput (CURIAL-Lab; vital signs & routine blood tests) and speed (CURIAL-Rapide; vital signs & FBC only), across three independent UK NHS hospital trusts and prospectively for the second wave of the UK COVID-19 epidemic at Oxford University Hospitals. Next, we propose and investigate a novel clinical triage pathway using our AI models to enhance sensitivity of rapid antigen test (LFD)-based triage for unscheduled admissions. Lastly, we deploy CURIAL-Rapide alongside an approved point-of-care FBC analyser (OLO; SightDiagnostics, Israel) to provide rapid lab-free COVID-19 screening in the John Radcliffe Hospital's Emergency Department (Oxford, UK), evaluating real-world performance and operational characteristics at a time of falling COVID-19 community prevalence in the UK[32].

Methods

Diagnostic models to identify patients presenting with COVID-19:
We updated our previously described model, designed to identify patients presenting to hospital with COVID-19 using vital signs, blood gas and routine laboratory blood tests (CURIAL-1.0[20]), with additional training data. The model was trained previously using clinical data from patients presenting to emergency and acute medical services at Oxford University Hospitals (OUH) between Dec 1, 2017 and April 19, 2020; additional data on all COVID-19-positive patient presentations to June 30, 2020 were added (Appendix B)[20]. This was performed to encompass all COVID-19 cases presenting to OUH during the 'first wave' of the COVID-19 pandemic (Supplementary Figure S1). OUH consist of four teaching hospitals, serving a population of 600 000, and provides tertiary referral services to the surrounding region. Routine blood tests were full blood count (FBC), urea, creatinine and electrolytes (U&Es), liver function tests (LFTs), coagulation and C-reactive protein (CRP), due to their ubiquity within existing emergency care pathways and rapid results, typically within around 1 h.

Next, we eliminated predictors with lower relative feature importances to improve generalisability across hospitals. CURIAL-Lab uses a focussed subset of routinely performed blood tests (FBC, U&Es, liver function tests, CRP) and vital signs (Table 1), eliminating the coagulation panel and blood gas which are not universally performed and are less informative[20]. Separately, we optimised for result-time, developing a minimalist model (CURIAL-Rapide) considering only predictors that can be obtained by the patient bedside (FBC and vital signs). We selected the FBC due to recent approval of a point-of-care haematology analyser (OLO, SightDiagnostics, Israel) with a result-time of 10 minutes, and as explanability analyses showed FBC components were most informative[31]. Models were trained using the OUH first-wave dataset described above.

Table 1: Clinical predictors considered in CURIAL-1.0, CURIAL-Lab & CURIAL-Rapide, showing successive elimination of less informative predictors to optimise for generalisbility and result-time.
(ALT=alanine aminotransferase. APTT=activated partial thromboplastin time. CRP=C-reactive protein. eGFR=estimated glomerular filtration rate. INR=international normalised ratio. p50=pressure at which haemoglobin is 50% bound to oxygen. *c*=calculated)

		Model:		
Feature Set	Constituents	CURIAL-1.0 (Soltan et. al 2020)	CURIAL-Lab	CURIAL-Rapide
Vital Signs	Heart rate, respiratory rate, oxygen saturations, blood pressure, temperature, oxygen delivery device	✓	✓	✓
Full Blood Count (FBC)	Haemoglobin, haematocrit, mean cell volume, white cell count, neutrophil count, lymphocyte count, monocyte count, eosinophil count, basophil count, platelets	✓	✓	✓
Urea & Electrolytes (U&Es)	Sodium, potassium, creatinine, urea, eGFR	✓	✓	

Liver Function (LF) Tests & CRP	Albumin, alkaline phosphatase, ALT, bilirubin, C-Reactive Protein	✓	✓	
Coagulation	Prothrombin Time, INR, APTT	✓		
Blood Gas	Actual base excess, standard base excess, bicarbonate, calcium, chloride, estimated osmolality, fraction of carboxyhaemoglobin, glucose, haemoglobin, haematocrit, potassium, methaemoglobin, sodium, oxygen saturation, calculated lactate, calculated oxygen content, calculated p5O, partial pressure of carbon dioxide pH, partial pressure of oxygen	✓		

Evaluation for second wave of UK COVID-19 epidemic at OUH:
We evaluated performance of CURIAL-1.0, CURIAL-Lab and CURIAL-Rapide using an independent prospective set of all patients presenting to emergency departments and acute medical services at OUH between October 01, 2020 and March 6, 2021, representing the second wave of the UK COVID-19 epidemic (Supplementary Figure 1). To assess performance characteristics at different calibrations, we performed the analysis with thresholds set during training to achieve sensitivities of 80% and 90%. Further, we performed a sensitivity analysis to assess susceptibility to imputation strategy, training the models using three independent imputation methods (median, mean and age-based mean) and examining variance in performance during evaluation. Mean sensitivity, specificity, area under receiver operator characteristic curve (AUROC) and predictive values are reported, alongside standard deviations across the three imputation methods. Confidence intervals for sensitivity, specificity and predictive values were computed using Wilson's Method[33], and for AUROC with DeLong's method[34].

External validation at independent NHS trusts:
We externally validated CURIAL-Rapide and CURIAL-Lab, calibrated during training to a sensitivity of 90%, across three independent UK National Health Service (NHS) hospital Trusts by comparing model predictions to confirmatory molecular testing (SARS-CoV-2 laboratory PCR, SAMBA-II and Panther). Participating hospital trusts were University Hospitals Birmingham NHS Foundation Trust (UHB), Bedfordshire Hospitals NHS Foundation Trust (BH), and Portsmouth University Hospitals NHS Trust (PUH), serving a total population of ~3.5 million. We evaluated the models for all patients aged over 18 who had an unscheduled admission via emergency or acute medical pathways and received a blood draw on arrival during the specified date ranges. Screening against eligibility criteria, followed by anonymisation, were performed by the respective NHS trusts. Patients who dissented to EHR research, did not have confirmatory molecular testing for COVID-19, or had only an invalid confirmatory test result with no subsequent valid result, were excluded. For trusts where blood-gas results were available for electronic extraction, we also evaluated CURIAL-1.0. Evaluation periods and confirmatory testing method are listed in Appendix C.

Comparison with Lateral Flow Devices
To investigate suitability of CURIAL-Rapide, CURIAL-Lab and CURIAL-1.0 as rapid screening tests for unscheduled admissions, we compared sensitivities and negative

predictive values with LFD results from the Innova SARS-CoV-2 Antigen Rapid Qualitative Test which was used within the standard-of-care at OUH during the second wave of the UK COVID-19 epidemic (to March 6, 2021; Supplementary Figure S2). From December 23, 2020 patients admitted to OUH from acute and emergency care settings (Emergency Department, Ambulatory Medical Unit, Medical Assessment Unit) had LFDs performed routinely alongside PCR testing. Swabs of the nose and throat were collected for both tests by trained nursing or medical staff and LFDs were performed in the emergency/acute departments. Positive, negative and invalid LFD results were documented in the EHR. Swabs for PCR were transferred to the clinical laboratory in viral transport medium and tested by PCR (ThermoFisher TaqPath), forming the reference standard for evaluating model predictions and LFDs.

A combined algorithm to enhance the sensitivity of LFD testing

Next, we investigated whether CURIAL-Rapide, CURIAL-Lab and CURIAL-1.0 could enhance the sensitivity of LFDs for identifying COVID-19 amongst patients being admitted to hospital. We proposed and retrospectively evaluated a novel clinical triage pathway (Figure 1) labelling patients as COVID-19-suspected where they had either a positive CURIAL result or a positive LFD result. Due to high specificity, in our pathway patients with positive LFDs can be streamed directly to a COVID-19-positive clinical area, meanwhile patients with a negative LFD but a positive CURIAL result would be managed in an enhanced-precautions area pending PCR adjudication. The pathway aimed to enhanced negative predictive value for patients receiving both negative LFD and CURIAL results, reduce the false-negative rate and therefore supporting safe and rapid triage directly to a ‘green’ zone.

We assessed performance of this novel pathway retrospectively for all unscheduled admissions to OUH where patients received LFD testing, from introduction on December 23, 2020 to March 6, 2021. Comparison between the performances of CURIAL-Lab, CURIAL-Rapide and CURIAL-1.0 alone are performed with integrated clinical pathway for each model using McNemar’s Chi-Square test.

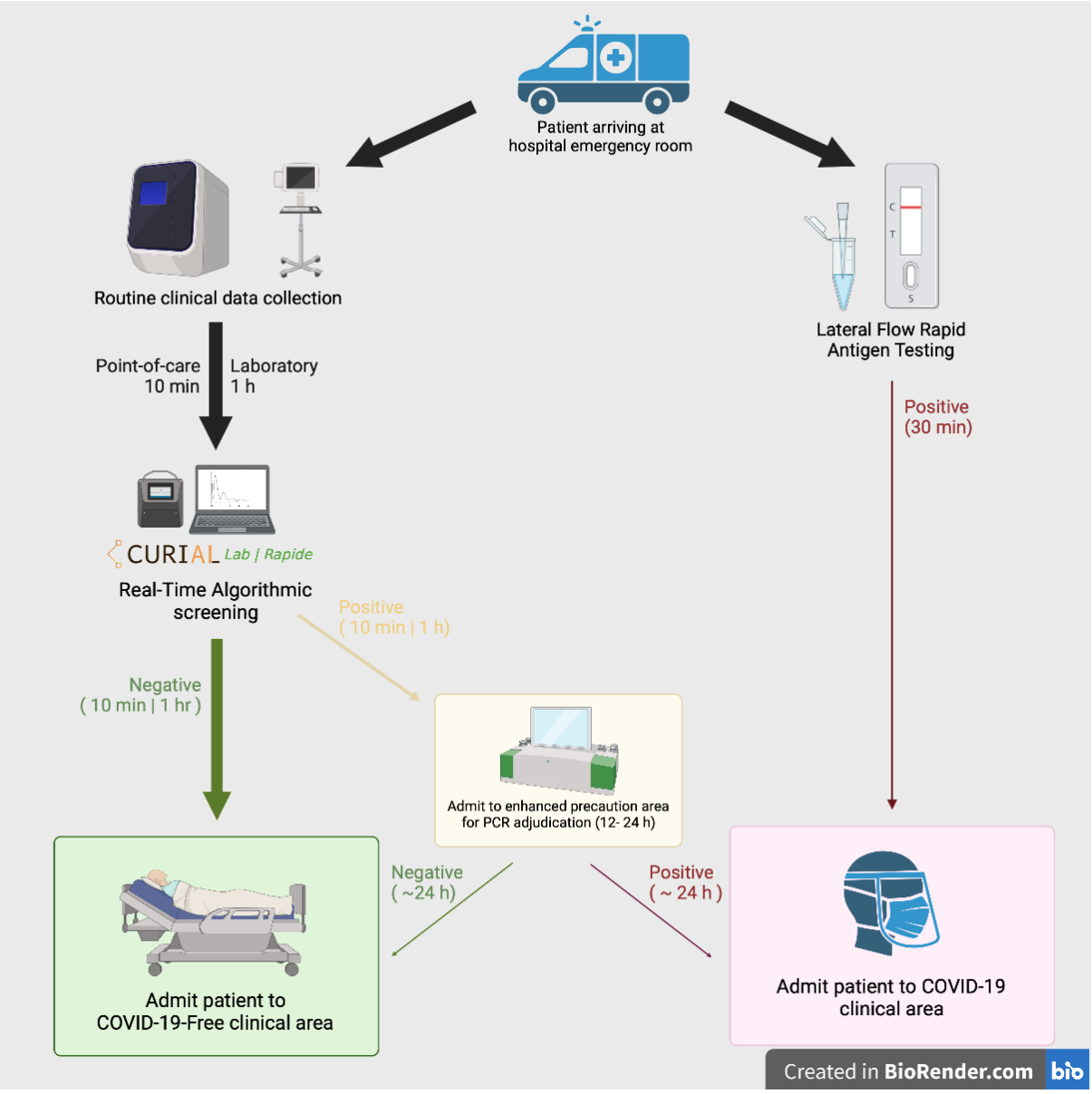


Figure 1: A rapid COVID-19 screening pathway for Emergency Department arrivals (a) routine blood tests and vital signs recordings are performed on patients' arrival to the emergency department, either using rapid point-of-care haematology analysers (~10 minutes; CURIAL-Rapide), or the existing laboratory analysis (~1h; CURIAL-Lab). (b) Real-time algorithmic analysis of blood tests & vital signs allows early, high-confidence identification of negative patients for safe streaming to COVID-19-Free clinical areas. (c) Patients with positive screening test results are admitted to enhanced precautions (amber) areas, pending confirmatory PCR result. (d) Patients testing positive via Lateral Flow Test can be streamed directly to COVID-19 ('red') clinical areas. Arrow thickness represents patient flow.

The study protocol for model development, external and prospective validation was approved by the National Health Service (NHS) Health Research Authority (IRAS ID 281832) and sponsored by the University of Oxford.

Prospective evaluation of CURIAL-Rapide in a lab-free clinical pathway

To prospectively assess operational and predictive performance of CURIAL-Rapide in a lab-free setting, we deployed two OLO rapid haematology analysers [SightDiagnostics, Tel Aviv] in the Emergency Department (ED) at the John Radcliffe Hospital, Oxford, as part of an OUH-approved service evaluation (Ulysses ID: 6907)[31]. We simultaneously aimed to improve routine clinical care by reducing the time for routine blood test results to become available in ED. The analysis plan and data requirements for the CURIAL-Rapide evaluation were determined prospectively and registered with the Trust service evaluation database.

We estimated a suitable review-point using Buderer's standard formulae[35]. Predicting a sensitivity of CURIAL-Rapide of 80% (matching model calibration), specificity of 75%, and prevalence of COVID-19 at 15% amongst patients in ED, we estimated a minimum sample size of 410 enrolled patients to determine sensitivity and 85 patients to determine specificity (95% confidence, precision 10%[36]). We therefore planned to review model performance once 500 patients had been enrolled, to allow for missing or invalid confirmatory tests.

The service evaluation operated from February 18, 2021 to May 10, 2021 between 8am and 8pm for patients meeting the eligibility criteria. Patients eligible were aged over 18, attending the ED with an acute illness and streamed to a bedded clinical area, and had consented to receive full blood count analysis and vital signs as part of their emergency care plan. We selected patients allocated to bedded clinical areas as non-ambulatory patients typically have higher acuity of illness, therefore being more likely to benefit from faster blood test results and having a higher probability of admission. Patients were identified on arrival using the FirstNet system (Cerner Millennium, Cerner, UK).

Eligible patients were enrolled for additional near-patient, lab-free full blood count analysis using the OLO, which in conjunction with vital signs were used to generate CURIAL-Rapide predictions. OLO FBC results were uploaded immediately to the EHR, making results available to clinicians and supporting routine care. We excluded patients with an invalid OLO result and no subsequent successful result, thereby ensuring data completeness. Routine COVID-19 testing was performed in line with trust policies, with LFDs (Innova SARS-CoV-2 Antigen Rapid Qualitative Test) performed in the department and paired multiplex PCR on-premises in a dedicated laboratory (ThermoFisher TaqPath). Patients who did not have confirmatory testing for COVID-19 within routine care were excluded from performance evaluation.

We recorded patients' arrival time to the hospital, measurement time of vital signs, and result times for LFD, PCR, OLO and laboratory FBC analysis. We also recorded the first-attending clinician's triage impression of COVID-19 status, using the locally adopted Green/Amber/Blue categorization system (Green representing a patient whose illness has no features of COVID-19, amber representing an illness with features potentially consistent with COVID-19, and blue representing laboratory confirmed COVID-19 infection)[37,38]. Where COVID-19 triage category had not been documented by the first-assessing physician, adjudication was performed through clinical review of notes using rules-based determination. Patients having documentation of a new continuous cough, temperature ≥37.8°C, or loss or change in sense of smell or taste were adjudicated as an 'amber' (COVID-19-suspected) stream, matching UK Government guidance on definition of a possible COVID-19

case[8]. Patients with PCR-confirmed COVID-19 in the 10 days preceding attendance were adjudicated to the 'Blue' (COVID-19-confirmed) stream. Patients with no features of COVID-19 and no documented clinical suspicion were adjudicated to the 'Green' stream.

We generated CURIAL-Rapide predictions using OLO results and vital signs, comparing predictions against results of confirmatory PCR testing. CURIAL-Rapide predictions were not made available to the attending clinician so as not to influence the clinical triage category or decisions to proceed to confirmatory testing for patients being discharged. Availability time for CURIAL-Rapide was the later of OLO result time and vital signs recording time as both are required to generate a prediction. The time-to-result for PCR, lateral flow, and CURIAL-Rapide tests were calculated as the time from a patient's first arrival in the ED to the time of a test result being available on the EHR.

We selected time-to-result as our primary outcome, recognising the role of rapid test results in reducing nosocomial transmission. Our secondary outcomes were sensitivity, specificity, PPV and NPV for CURIAL-Rapide & LFDs, and AUROC for CURIAL-Rapide, assessed against PCR results. Further detail is provided in Appendix D.

Role of the funding source: The funders of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the manuscript.

Results:

Patient Populations

Table 2: Summary population characteristics for (a) OUH pre-pandemic and COVID-19-cases training cohorts, (b) prospective validation cohort of patients attending OUH during the second wave of the UK COVID-19 epidemic, (c) independent validation cohorts of patients admitted to three independent NHS Trusts, (d) admissions to OUH during the second-wave receiving LFD testing, (e) patients enrolled to the CURIAL-Rapide lab-free service evaluation at OUH. The derivation of OUH cohorts is shown in Supplementary Figure S2. * indicates merging for statistical disclosure control.

	(a) Training		**(b) Prospective validation**	**(c) External Validation (Admissions)**			**(d) LFD Evaluation**	**(e) Lab-free deployment**
	Oxford University Hospitals (pre-pandemic & wave 1 cases, to 30 June 2020)		Oxford University Hospitals	Portsmouth University Hospital NHS Trust	University Hospitals Birmingham NHS Foundation Trust	Bedfordshire Hospitals NHS Foundation Trust	Oxford University Hospitals (wave 2 admissions)	Oxford University Hospitals ED
Cohort	Pre-pandemic cohort	COVID-19-cases cohort	October 1, 2020 – March 6, 2021	March 1, 2020 - February 28, 2021	December 01, 2019 - October 29, 2020	January 1, 2021 - March 31, 2021	December 23, 2020 – March 6, 2021	Feb 18, 2021 – May, 10, 2021
n, patients	114,957	701	22,857	37,896	10,293	1177	3,207	520
n, COVID-19 PCR / genome test positive	0	701	2,012	2,005 (5.29%)	439 (4.27%)	144 (12.2%)	355 (11.1%)	≤ 10
Sex: - Male (%)	53370 (46.43)	376 (53.64)	11409 (49.91)	20839 (54.99)	4831 (46.93)	627 (53.27)	1586 (49.45)	231 (44.42)
- Female (%)	61587 (53.57)	325 (46.36)	11448 (50.09)	17054 (45.0)	5462 (53.07)	549 (46.64)	1621 (50.55)	289 (55.58)
Age, yr (IQR)	60 (38-76)	72 (55-82)	67 (49-80)	69 (48-82)	63 (42-79)	68.0 (48-82)	70 (51-82)	76 (60-85)
LFD positive (%)							207 (6.45)	≤ 10
Ethnicity: -White (%)	93921 (81.7)	480 (68.47)	17387 (76.07)	28704 (75.74)	6848 (66.53)	1024 (87.0)	2491 (77.67)	419 (80.58)
-Not Stated (%)	13602 (11.83)	128 (18.26)	4127 (18.06)	8389 (22.14)	1061 (10.31)	≤ 10	513 (16.0)	80 (15.38)
-South Asian (%)	2754 (2.4)	22 (3.14)	441 (1.93)	170 (0.45)	1357 (13.18)	71 (6.03)	65 (2.03)	≤ 10
-Chinese (%)	284 (0.25)	*	51 (0.22)	42 (0.11)	41 (0.4)	≤ 10	*	≤ 10
-Black (%)	1418 (1.23)	25 (3.57)	279 (1.22)	187 (0.49)	484 (4.7)	36 (3.06)	45 (1.4)	≤ 10
-Other (%)	1840 (1.6)	34 (4.85)*	410 (1.79)	269 (0.71)	333 (3.24)	29 (2.46)	72 (2.25)*	≤ 10
-Mixed (%)	1138 (0.99)	12 (1.71)	162 (0.71)	135 (0.36)	169 (1.64)	13 (1.1)	21 (0.65)	≤ 10

Our updated training set (Supplementary Figure S2) comprised 114,957 patient presentations to OUH prior to the global COVID-19 outbreak (November 30, 2017 - December 01, 2019), considered as COVID-19-free presentations, and 701 patient presentations during the first wave of the UK COVID-19 epidemic (December 01, 2019 - June 30, 2021) who had a positive PCR test for COVID-19.

72,223 patients were included across four validation cohorts, of whom 4,600 had a positive confirmatory test for COVID-19 (Appendices A & B). In the external cohorts, patients admitted to PUH and BH trusts were similar ages (69 years (IQR 34) and 68 years (34); Kruskal-Walls p=0.9448), however patients admitted to UHB were significantly younger (63 years (37); p<0.0001 & <0.0001). A higher proportion of patients admitted to UHB were female (53.1%) than PUH and BH (45.0% and

46.7%; chi-square p<0.0001) and reported a South Asian ethnicity (13.2% versus 0.5% and 2.0%; chi-square p<0.0001).

Prevalence of COVID-19 was higher in the Bedfordshire cohort owing to the evaluation period matching the timeline of the second wave of the UK COVID-19 epidemic (11.1% versus 5.29% (PUH) and 4.27 (UHB); Fisher's exact test p<0.0001 & <0.0001). Summaries of vital signs, index routine blood tests and blood gases are presented in Supplementary Tables S2-S4.

Prospective evaluation of CURIAL-Lab & CURIAL-Rapide:

Table 3: Evaluation of the performance of (a) CURIAL-Rapide and (b) CURIAL-Lab, calibrated during training to achieve sensitivities of 80% and 90%, on an independent prospective set of all admissions to OUH during the second-wave of COVID-19, between October 1, 2020 and March 6, 2021. (c) Benchmark performance of CURIAL-1.0 (Soltan et al. 2020) on the prospective set. Mean values are reported alongside SD across three imputation methods.

Model	(a) CURIAL-Rapide: FBC + Vitals		(b) CURIAL-Lab: FBC, C&E, LFD, CRP + Vitals		(c) CURIAL-1.0: Blood Tests + Blood Gas + Vitals	
Calibration	80%	90%	80%	90%	80%	90%
Sensitivity	74.7 (0.4)	85.6 (0.6)	72.9 (0.8)	85.7 (0.9)	73.6 (0.3)	85.9 (1.4)
Specificity	78.6 (0.4)	59.1 (0.3)	87.3 (1.1)	68.6 (2.2)	86.6 (0.4)	67.1 (0.8)
PPV	14.9 (0.1)	9.46 (0.0)	22.4 (1.2)	12.0 (0.6)	21.5 (0.4)	11.6 (0.0)
NPV	98.4 (0.0)	98.8 (0.0)	98.5 (0.0)	99.0 (0.0)	98.5 (0.0)	99.0 (0.1)
F1	0.248 (0.003)	0.170 (0.001)	0.342 (0.014)	0.211 (0.009)	0.331 (0.005)	0.204 (0.001)
AUROC	0.843 (0.002)	0.843 (0.002)	0.878 (0.001)	0.878 (0.001)	0.875 (0.002)	0.875 (0.002)

Of 37,304 patients attending emergency departments and acute medical services across OUH during the second-wave of the COVID-19 pandemic (October 01, 2020 - March 06, 2021; Table 2 & Supplementary Figure S2), 14,447 were excluded as they did not receive confirmatory testing for COVID-19. We evaluated CURIAL-Lab and CURIAL-Rapide for all 22,857 patients receiving confirmatory testing (2056 testing positive), benchmarking performance against CURIAL-1.0 (Table 3). At the 80% sensitivity configuration, CURIAL-Lab performed similarly to CURIAL-1.0 (sensitives 72.9% & 73.6% respectively, and specificities 87.3% & 86.6%; McNemar chi-square test p=0.0823) but better than CURIAL-Rapide (sensitivity 74.7%, specificity 78.6%; p <0.0001) representing a trade-off between result-time and performance. CURIAL-Rapide and CURIAL-Lab achieved high NPVs (>98%) across both sensitivity configurations, achieving an AUROC of 0.843 and 0.878 respectively.

A sensitivity analysis to assess susceptibility of our models to imputation strategy demonstrated consistent performance across multiple imputations (Table 3). We therefore performed subsequent evaluation using a model trained solely using a single imputation strategy (population median).

External validation of CURIAL-Rapide, CURIAL-Lab and CURIAL-1.0
We externally validated CURIAL-Rapide & CURIAL-Lab at three independent hospital groups, across cohorts comprising 49,366 patient admissions (Figure 2, Supplementary Table S5). Performance was consistent across the trusts, with CURIAL-Lab achieving higher performance (AUROC range 0.858 - 0.881, 95% CIs range 0.838-0.912) than CURIAL-Rapide (AUROC range 0.836 - 0.854, 95% CIs range 0.814-0.889). Sensitivity of both models was higher when applied at Portsmouth University Hospitals (84.1%, 95% CI 82.5-85.7 & 83.5%, 81.8 - 85.1) compared to Bedfordshire Hospitals (74.3%, 66.6 - 80.7 & 74.3%, 66.6 - 80.7), at the expense of specificity (71.3%, 70.9 - 71.8 & 63.6%, 63.1 - 64.1 versus 84.8%, 82.5 - 86.9 and 81.8%, 79.3 - 84.0) possibly reflecting differences in confirmatory testing method at Bedfordshire Hospitals (SAMBA-II/Panther).

Both CURIAL-Rapide & CURIAL-Lab achieved high NPV across the three trusts in a prevalence-dependent fashion, with highest NPV at UHB where prevalence was lowest (4.27%, CURIAL-Rapide: 98.8% [98.5 - 99.0], CURIAL-Lab: 98.8% [98.5 - 99.0]). NPV was comparable at PUH where prevalence was similar (5.29%; CURIAL-Rapide: NPV 98.6% [98.4 - 98.7], CURIAL-Lab: 98.8% [98.6 - 98.9]). We additionally externally validated CURIAL-1.0 for the two trusts which electronically recorded blood-gas results, demonstrating comparable performance to CURIAL-Lab (Supplementary Table S5).

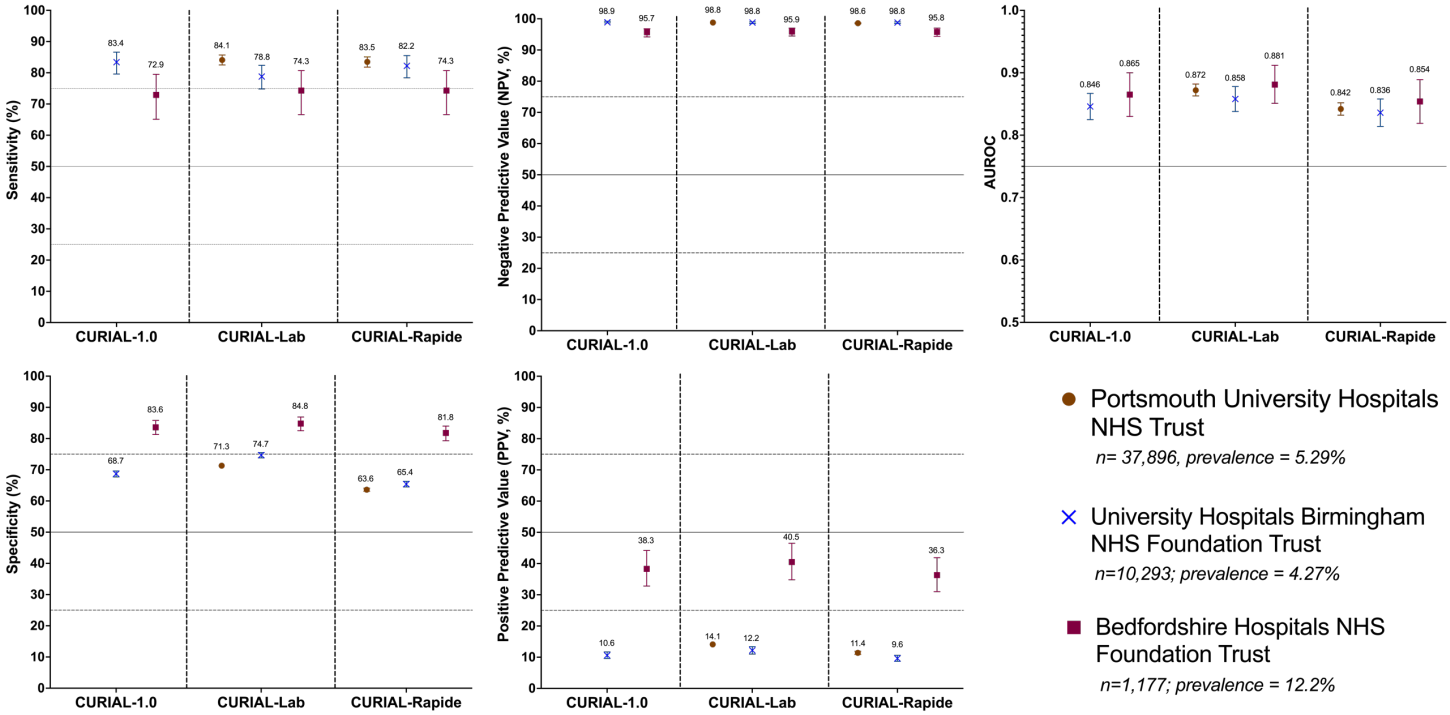


Figure 2: Performance of CURIAL-1.0, CURIAL-Lab & CURIAL-Rapide during external validation at three independent UK Hospitals trusts. All models were calibrated during training to achieve 90% sensitivity. Error bars show 95% confidence intervals. Numerical results are shown in Supplementary Table S5.

Comparing CURIAL triage performance with LFDs & evaluating a combined CURIAL-LFD clinical pathway to enhance LFD sensitivity

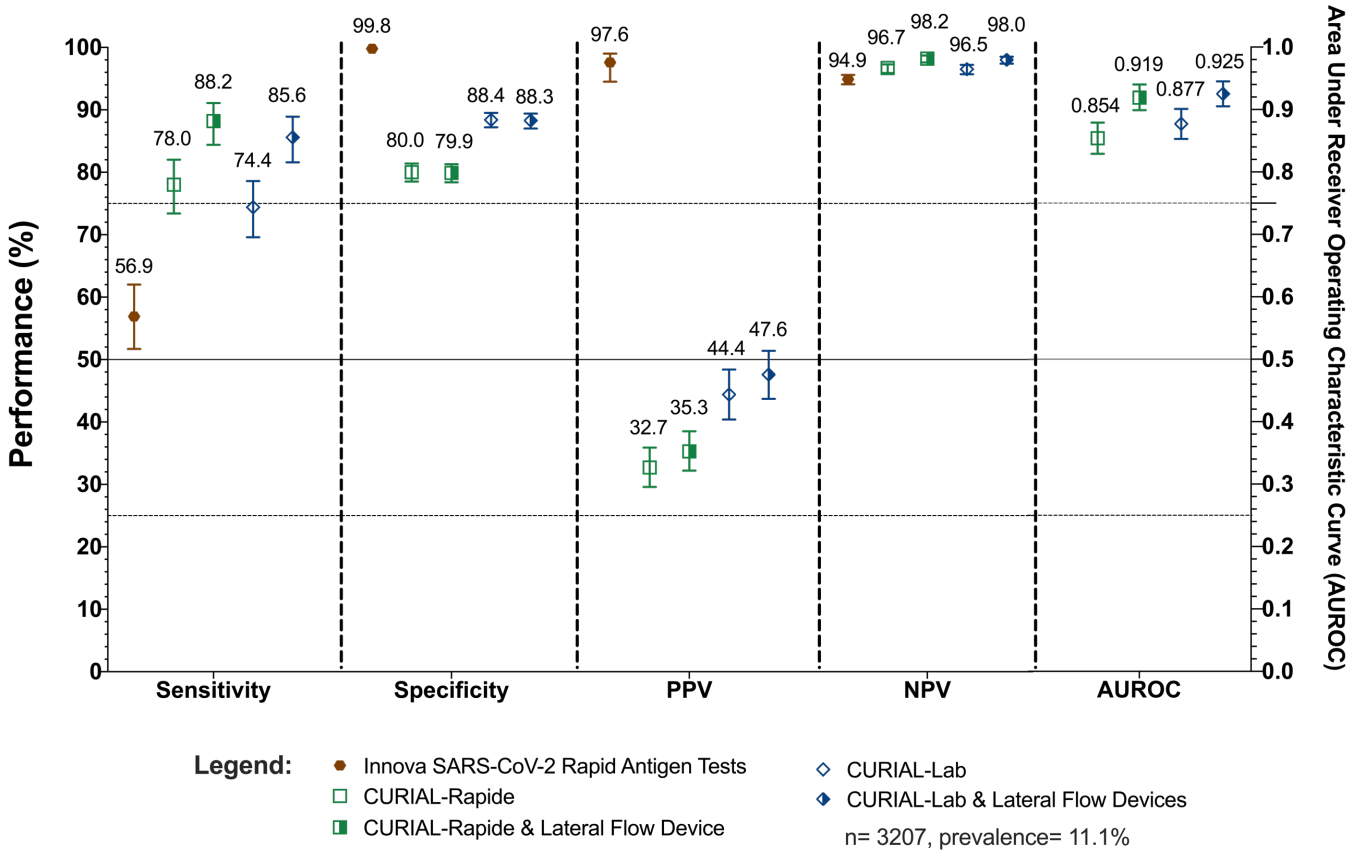


Figure 3: Performance characteristics of (a) INNOVA SARS-CoV2 Rapid Antigen Tests, (b) CURIAL-Rapide & CURIAL-Lab, calibrated during training to a sensitivity of 80%, and (c) combined clinical pathways considering either a positive CURIAL-Rapide/CURIAL-Lab result or a positive LFD test as a COVID-19 suspected case, at Oxford University Hospitals NHS Foundation Trust between December 23, 2020 & March 6, 2021. Error bars show 95% confidence intervals. Numerical results are shown in Supplementary Table S6.

We applied CURIAL-Rapide and CURIAL-Lab to the first-performed blood tests and vital signs for 3,207 patients admitted to OUH, receiving LFD testing from introduction on December 23, 2020 to March 6, 2021 (Figure 3, Supplementary Table S6). Patient eligibility is demonstrated in Supplementary Figure S2. One patient with an invalid LFD result was excluded from analysis. LFDs achieved a high specificity of 99.8% (99.6 - 99.9) during triage, but poor sensitivity of 56.9% (51.7 - 62.0).

CURIAL-Rapide & CURIAL-Lab were significantly more sensitive (78.0% [73.4 - 82.0] & 74.4% [69.9 - 78.6]) than LFDs, therefore achieving higher negative predictive values (CURIAL-Rapide: 96.7% [95.9-97.3], CURIAL-Lab: 96.5% [95.7 - 97.2], versus LFDs: 94.9% [94.1 - 95.6]). By contrast, the models were less specific (80.0% [78.5 - 81.4] & 88.4% [87.2 - 89.5]), thereby favouring higher reliability for safely excluding COVID-19 at the expense of the false positive rate.

Integrating positive LFD results with CURIAL-Rapide/CURIAL-Lab, as part of a combined clinical pathway (Figure 1), significantly improved overall sensitivity (to 88.2% [84.4 - 91.1] & 85.6% [81.6 - 88.9]) and NPV (to 98.2% [97.6 - 98.7] & 98.0% [97.4 - 98.5]), reducing COVID-19 status misclassification (McNemar's chi-square test, p=0.0003 & p=0.0004). AUROC was significantly improved to 0.919 (0.899 - 0.940) for a CURIAL-Rapide/LFD pathway, and 0.925 (0.905 – 0.945) for CURIAL-Lab/LFD pathway. Performance of CURIAL-Lab was similar to CURIAL-1.0 (p=0.860; Supplementary table S6).

Deployment & operational evaluation of CURIAL-Rapide at Oxford University Hospitals NHS Foundation Trust

520 patients were enrolled to the OLO/CURIAL-Rapide service evaluation between February 18, 2021 and May 10, 2021 (baseline characteristics, Table 2). 436 patients received confirmatory PCR testing within routine care, of whom 10 returned positive results (prevalence 2.3%). This reflected the falling prevalence of COVID-19 in the UK, coinciding with governmental restrictions and the national vaccination programme[32]. 348 patients received LFDs within routine care, with 4 positive results. Two patients with indeterminate PCR results were excluded from analysis, although both had a negative LFD result and were triaged by the assessing clinician to a COVID-19-free clinical pathway. A summary of OLO results and vital signs are shown in Supplementary Tables S2-S3.

Table 6: Operational and performance characteristics of (a) CURIAL-Rapide, (b) INNOVA SARS-CoV-2 Rapid Antigen Testing and (c) clinical triage by the first-attending clinician calculated against laboratory RT-PCR testing. Results are reported alongside interquartile range (time-to-result) or 95% confidence intervals.

	CURIAL-Rapide v1.0	**INNOVA SARS-CoV-2 Rapid Antigen Testing**	**First-attending Clinician Triage**	**Laboratory RT-PCR**
Time from patient arrival in ED to result availability, median (IQR)	45:00 min (32:00-64:00)	61:00 min (36:45-99:00)		07:37 hr (06:05-15:39)
Sensitivity (95% CIs)	87.5% (52.9 - 97.8)	50.0% (21.5 - 78.5)	75.0% (40.9 - 92.9)	
Specificity (95% CIs)	85.4% (81.3 - 88.7)	100.0% (98.9 - 100.0)	85.1% (81.0 - 88.4)	
Accuracy (95% CIs)	85.4% (81.4 - 88.7)	98.9% (97.2 - 99.6)	84.9% (80.8 - 88.2)	
Precision, PPV (95% CIs)	11.9% (5.9 - 22.5)	100.0% (51.0 - 100.0)	10.2% (4.7 - 20.5)	
NPV (95% CIs)	99.7% (98.2 - 99.9)	98.9% (97.2 - 99.6)	99.3% (97.6 - 99.8)	
AUROC (95% CIs)	0.907 (0.803 - 1.00)			

Median time from registration in the ED to availability of a CURIAL-Rapide result was 45:00 minutes (32:00-64:00), 16 minutes (26.3%) sooner than for LFDs (61:00 min, IQR 36:45-99:00; Wilcoxon Signed Rank test p<0.0001), and 6h 52 minutes (90.2%) sooner than RT-PCR results (07:37 h, IQR 06:05-15:39; p<0.0001). Kaplan Meier

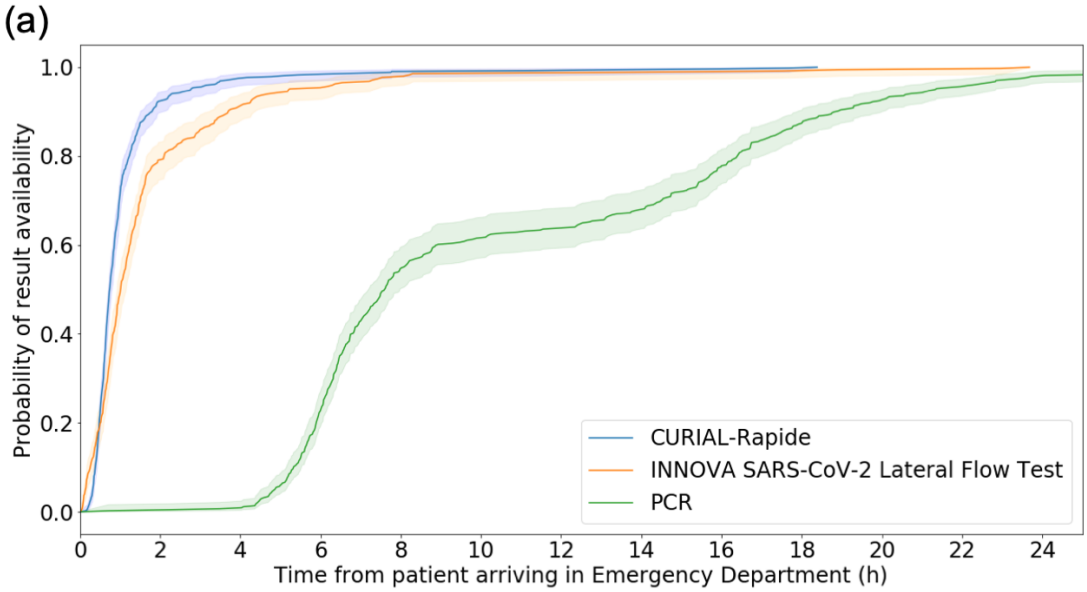
survival analyses (Figure 4a) showed CURIAL-Rapide results were available sooner than LFDs (log rank test, $p<0.0001$) and PCR results ($p<0.0001$). The median time-to-result for full blood count analysis was shorter with near-patient OLO analysis (44:00 min, 31:00-63:00 min) than laboratory analysis (76:00 min, 58:00-100:00 mins; $p <0.0001$), confirming an improvement to routine care.

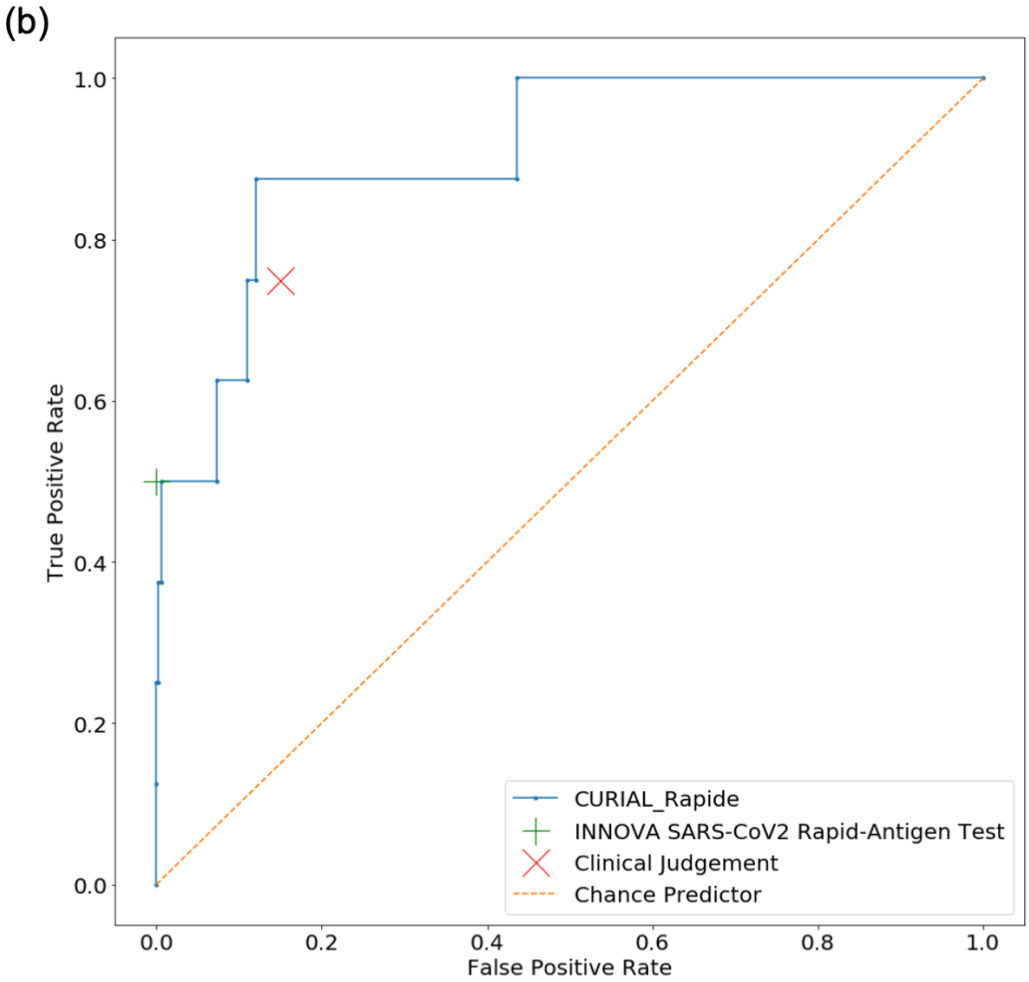
CURIAL-Rapide results had a negative predictive value of 99.7% (98.2-99.9), specificity of 85.4% (81.3-88.7), and AUROC of 0.907 (0.803-1.00). The point estimate of CURIAL-Rapide's sensitivity was 87.5%, however 95% CIs were wide owing to the lower-than-expected prevalence of COVID-19 (52.9-97.8)).

In one presentation, a patient given a 'negative' CURIAL-Rapide prediction went on to have a positive SARS-CoV-2 PCR test, although they had a negative LFD result and were triaged to a COVID-19-free ('green') clinical area. The patient did not have COVID-19 symptoms on this presentation. We noted that the patient had also been enrolled to the service evaluation 10 days prior; on that occasion having a positive CURIAL-Rapide prediction, positive Lateral Flow Test and a positive PCR test. This raises the possibility of a latent positive PCR result, detecting non-infectious residual viral fragments, on the date of the second presentation[39,40].

Rates of COVID-19 status misclassification were comparable between CURIAL-Rapide and clinician judgement (McNemar's Exact test; $p=0.91$). Moreover, of the 53 patients who were triaged to a 'COVID-19-suspected' (amber) pathway by the attending clinician but went on to test negative by PCR, 31 patients (58.5%) had a negative CURIAL-Rapide prediction demonstrating that the AI system could reduce operational strain by expediting clinical exclusion of infection. As all patients with positive LFD results also had positive CURIAL-Rapide predictions, a combined CURIAL-Rapide/LFD pathway did not impact classifier performance in this evaluation.

Figure 4: (a) Kaplan-Meier plots of time-to-result in hours from patient arrival in the Emergency Department for (i) CURIAL-Rapide, (ii) INNOVA SARS-CoV-2 Rapid-Antigen Tests and (iii) PCR swabs tests (Numbers at risk: 520, 348 and 436 respectively). CURIAL-Rapide results were available sooner than LFD testing (log rank test, p<0.0001) and PCR test results (p<0.001). (b) Receiver operating characteristic curve showing performance of (i) CURIAL-Rapide (ii) clinical judgement of the first-attending clinician, and (iii) INNOVA SARS-CoV-2 rapid antigen testing, against a PCR reference standard.





Discussion

National health policy recognises that effective in-hospital triage is necessary to safeguard patient and staff safety, mandating COVID-19 testing for all admissions[4]. However, despite significant innovation leading to new near-patient testing options, alongside reduced PCR result-times to typically within 12-24 h, there remain significant performance and logistical limitations that contribute to nosocomial transmission and operational strain. While many hospitals have adopted LFDs within acute admissions pathways[10], our study confirms a limited sensitivity (56.9%) indicating a clinically-meaningful false negative rate (Figure 3, Supplementary Table S6)[14–16,41].

In this study we demonstrate generalisability, efficacy, and real-world operational benefits of AI-driven COVID-19 screening in the acute care setting. Whereas rapid molecular testing options are frequently rationed[11,12], we show that a high-throughput AI-solution, CURIAL-Lab, rapidly excludes COVID-19 using routine data and generalises across three independent hospital groups (Figure 2). Moreover, we improve upon the speed of existing rapid testing solutions, demonstrating a median result-time of 45 minutes (32-64 min; CURIAL-Rapide) from patients' first arrival in an emergency department using near-patient haematology analysis. This decentralised approach may support time-critical decision making and assist triage in remote and primary care settings where laboratory facilities are less readily available.

In our external and prospective validation of CURIAL-Lab & CURIAL-Rapide, model performance was consistently high across four UK hospital groups (CURIAL-Lab: AUROC range 0.858-0.881; CURIAL-Rapide 0.836-0.854), with high negative predictive values confirming suitability as tests-of-exclusion for COVID-19. CURIAL-Lab expectedly achieved marginally superior performance to CURIAL-Rapide, representing a trade-off between result-time and specificity which would favour different clinical use-cases. Strengths of the validation include geographic breadth, including over 72,000 patients across three regions of the UK (Midlands, South East, and East of England), and temporal breadth across both waves of the UK COVID-19 epidemic therefore including vaccinated patients and patients with Coronavirus variants, across a range of prevalences (4.27%-12.2%). Moreover, the validation study considers the broad range of confirmatory COVID-19 tests (PCR, SAMBA-II & Panther) utilised across different centres, and address sector-wide concerns highlighted in key reviews of COVID-19 diagnostic and prognostic models by using external, representative cohorts of all unscheduled adult admissions[27,28]. Notable limitations include that the external validation is solely UK-based and the limitations of the confirmatory testing methods, with PCR testing having been shown to be imperfectly sensitive[6,7]. We were unable to quantify the number of vaccinated patients as we could not link our de-identified hospital datasets with vaccination records.

Our study finds that CURIAL-Rapide & CURIAL-Lab achieve significantly higher sensitivity and NPV than LFDs, improving upon standard-care by reducing risk of a COVID-19-positive patient being streamed to a COVID-19-free clinical area. Moreover, our study is the first to validate an application of an AI test to enhance sensitivity of LFDs in a real-world clinical setting, with our combined clinical pathway

(Figure 1; CURIAL-Lab/LFD) achieving both high sensitivity of 85.6% (81.6 - 88.9) and overall classification performance with AUROC of 0.925 (0.905 - 0.945). A significant beneficiary population includes patients streamed to COVID-19-free ('green') clinical areas on receiving negative LFD and CURIAL results, which are performed in parallel and available much sooner than PCR results. Moreover, our pathway identifies an enriched subpopulation at greater risk of testing positive for COVID-19 (negative LFD & positive CURIAL-Lab/Rapide compared to negative LFD alone), therefore enabling prioritisation for rapid confirmatory testing where availability is limited. Limitations of this analysis include that patients may have been inadvertently excluded if LFDs were recorded incorrectly on the EHR, and the analysis was performed only for a single trust as other participating sites did not electronically record LFD results.

We report the fastest result-time to date for AI-driven COVID-19 screening in a hospital emergency department, using lab-free haematology analysis to achieve a median reduction of 16 minutes (26.3%) over LFDs. Significantly, by demonstrating that CURIAL-Rapide correctly excluded COVID-19 for 58.5% of negative patients who were triaged by a clinician to a 'COVID-19-suspected' (amber) clinical area, we show a role for AI screening in reducing delays in transfers to wards. A strength of our service evaluation is its real-world context and operational focus, assessing time from first-arrival to result availability, and demonstrating added clinical value by comparison to LFD results and clinician impressions. In this study, we address the need for evidence of clinical utility that AI tools such as CURIAL offer to the pandemic response[30]; demonstrating both operational and safety improvements to standard-care.

A significant limitation of our service evaluation is that, although the a priori target sample size was achieved, the desired precision and power levels were not achieved for the metric of sensitivity due to sharply falling prevalence of COVID-19 in the UK, associated with the national vaccination programme and public health measures[32,42]. The evaluation was, however, adequate to determine specificity. As a service evaluation, we used a convenience series, with OLO operation limited to daytime and evening hours (8am-8pm) for logistical reasons. Moreover, although routine LFD testing was hospital policy, 33% of enrolled patients did not have a coded result in their EHR using the pre-specified form raising the possibility that these may have been recorded elsewhere or communicated verbally. As all patients who were LFD positive also had a positive CURIAL-Rapide result in our study, a larger evaluation is needed to assess whether integrating LFD results could further improve performance of CURIAL-Rapide in this context. Further evaluation would assess performance as a clinical decision support aid and for sensitivity across coronavirus variants.

A major strength of the CURIAL-Rapide and CURIAL-Lab solutions is the use of clinical data that is readily available and routinely collected for all patients admitted to hospital. Our approach optimises generalisability with CURIAL-Lab, applicable to virtually all unscheduled patient admissions to hospital, thereby facilitating COVID-19-screening without significant additional cost. Where faster exclusion of COVID-19 is helpful for operational or treatment reasons, CURIAL-Rapide can provide faster results by eliminating the need for blood sample transportation and laboratory processing, at an approximate cost of around ~£9 (~$12.50; inclusive of device rental and consumables). By contrast, alternative strategies for AI-assisted COVID-

19 diagnosis largely focus on chest imaging[23,28,43], which involve patient exposure to ionising radiation and have higher costs. Following successful COVID-19 vaccination programmes, falling community prevalence may reduce cost-effectiveness of universal PCR-testing for unscheduled admissions. Our results suggest that CURIAL-Lab could deliver significant cost savings by reducing the number of routine PCR tests by >85% (where prevalence <2%) while achieving high NPV, utilising data that would be collected within the routine course of a patients' care.

Our work demonstrates generalisability, efficacy, and real-world operational benefits of AI-driven COVID-19 screening for patients attending hospital. Future work would assess international generalisability, evaluate clinician-model interactions, and assess sensitivity of model performance across vaccination types and infection with variants of concern.

Acknowledgments: We express our sincere thanks to all patients and staff across the four participating NHS trusts; Oxford University Hospitals, University Hospitals Birmingham, Bedfordshire Hospitals, and Portsmouth University Hospitals NHS trusts. We particularly wish to thank all nursing, medical and allied healthcare professional colleagues in the John Radcliffe Hospital's Emergency Department for supporting evaluation of the OLO analyser and CURIAL-Rapide. We thank Emergency Medicine Research in OXford (EMROX), Dr Natasha Dole, Dr Jaspreet Bahra and Dr Ayisha Shaikh (Emergency Department Registrars), Ian Smith, Vikram Lyall and the OUH Point-of-care testing office. We additionally express our gratitude to Jingye Wang & Jolene Atia at University Hospitals Birmingham NHS Foundation trust, Phillip Dickson at Bedfordshire University Hospitals, and Paul Meredith at Portsmouth University Hospitals for assistance with data extraction.

Funding: This study was supported by the Wellcome Trust/University of Oxford Medical & Life Sciences Translational Fund (Award: 0009350) and the Oxford National Institute of Research (NIHR) Biomedical Research Campus (BRC). The funders of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the manuscript. AS is an NIHR Academic Clinical Fellow. DWE is a Robertson Foundation Fellow and an NIHR Oxford Biomedical Research Centre Senior Fellow. The views expressed are those of the authors and not necessarily those of the NHS, NIHR, or the Wellcome Trust.

Declarations: DWE reports personal fees from Gilead, outside the submitted work; DAC reports personal fees from Oxford University Innovation, personal fees from BioBeats, personal fees from Sensyne Health, outside the submitted work. No other authors report any conflicts of interest.

Contributions: AS, AN, RP, MS, DRT, TZ, DWE & DAC conceived of and designed the study. DWE, RF, AN, RP, MS, DRT, SB & CURIAL Translational Collaborative assisted with data collection/extraction and study operation. AS, JY, YY, OR pre-processed and verified the data, and performed the analyses. AS and DWE wrote the manuscript. All authors revised the manuscript.

Ethics: NHS Health Research Authority (HRA) approval was granted for the use of routine clinical and microbiology data from Electronic Health Records for

development and validation of artificial intelligence models to detect Covid-19 (CURIAL; NHS HRA IRAS ID 281832).

Data & Code Availability:
Data from OUH studied here are available from the Infections in Oxfordshire Research Database (https://oxfordbrc.nihr.ac.uk/research-themes-overview/antimicrobial-resistance-and-modernising-microbiology/infections-in-oxfordshire-research-database-iord/), subject to an application meeting the ethical and governance requirements of the Database. Data from UHB, PUH and BH are available on reasonable request to the respective trusts, subject to HRA requirements. Code and supplementary information for this paper are available online alongside publication.

Bibliography

1. The Guardian UK. 40,600 people likely caught Covid while hospital inpatients in England. *The Guardian* https://www.theguardian.com/world/2021/may/24/up-to-8700-patients-died-after-catching-covid-in-english-hospitals (2021).
2. Bhaskaran, K. *et al.* Factors associated with deaths due to COVID-19 versus other causes: population-based cohort analysis of UK primary care data and linked national death registrations within the OpenSAFELY platform. *The Lancet Regional Health - Europe* **6**, 100109 (2021).
3. Zhou, F. *et al.* Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *The Lancet* **395**, 1054–1062 (2020).
4. UK National Medical Director, Chief Nursing Officer for England, Chief People Officer & National Director for Emergency and Elective Care. Letter: Healthcare associated COVID-19 infections – further action. *NHS England and NHS Improvement*.
5. Clift, A. K. *et al.* Living risk prediction algorithm (QCOVID) for risk of hospital admission and mortality from coronavirus 19 in adults: national derivation and validation cohort study. *The BMJ* **371**, (2020).
6. Wikramaratna, P. S., Paton, R. S., Ghafari, M. & Lourenço, J. Estimating the false-negative test probability of SARSCoV- 2 by RT-PCR. *Eurosurveillance* **25**, (2020).
7. Kucirka, L. M., Lauer, S. A., Laeyendecker, O., Boon, D. & Lessler, J. Variation in False-Negative Rate of Reverse Transcriptase Polymerase Chain Reaction-Based SARS-CoV-2 Tests by Time Since Exposure. *Annals of internal medicine* vol. 173 262–267 (2020).
8. UK Government. COVID-19: investigation and initial clinical management of possible cases. *https://www.gov.uk/government/publications/wuhan-novel-coronavirus-initial-investigation-of-possible-cases/investigation-and-initial-clinical-management-of-possible-cases-of-wuhan-novel-coronavirus-wn-cov-infection#criteria* (2020).
9. Assennato, S. M. *et al.* Performance evaluation of the SAMBA II SARS-CoV-2 test for point-of-care detection of SARS-CoV-2. *Journal of Clinical Microbiology* **59**, (2021).
10. Young, B. C., Eyre, D. W. & Jeffery, K. Use of lateral flow devices allows rapid triage of patients with SARS-CoV-2 on admission to hospital. *Journal of Infection* (2021) doi:10.1016/j.jinf.2021.02.025.

11. University Hospitals Birmingham NHS Foundation Trust. Pathway for processing urgent COVID-19 samples. https://www.uhb.nhs.uk/coronavirus-staff/pathway-for-processing-urgent-covid-19-samples.htm (2020).
12. Portsmouth University Hospitals NHS Trust. *Trust Standard Operating Procedure COVID-19 Testing*. (2020).
13. Downs, L. O., Eyre, D. W., O'Donnell, D. & Jeffery, K. Home-based SARS-CoV-2 lateral flow antigen testing in hospital workers. *Journal of Infection* **82**, 282–327 (2021).
14. Wolf, A., Hulmes, J. & Hopkins, S. *Lateral flow device specificity in phase 4 (post marketing) surveillance*. https://www.ox.ac.uk/sites/files/oxford/media_wysiwyg/UK%20evaluation_PHE%20Porton%20Down%20%2 (2021).
15. Public Health England and 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*. https://www.gov.uk/government/publications/assessment-and-procurement-of-coronavirus-covid-19-.
16. Wise, J. Covid-19: Safety of lateral flow tests questioned after they are found to miss half of cases. *BMJ (Clinical research ed.)* **371**, m4744 (2020).
17. Iacobucci, G. Covid-19: Mass population testing is rolled out in Liverpool. *BMJ (Clinical research ed.)* **371**, m4268 (2020).
18. Iacobucci, G. Covid-19: MHRA is concerned over use of rapid lateral flow devices for mass testing. *BMJ* **373**, (2021).
19. Iacobucci, G. Covid-19: US regulator raises significant concerns over safety of rapid lateral flow tests. *BMJ* **373**, (2021).
20. Soltan, A. A. S. *et al.* Rapid triage for COVID-19 using routine clinical data for patients attending hospital: development and prospective validation of an artificial intelligence screening test. *The Lancet Digital Health* **3**, e78–e87 (2021).
21. The COvid-19 Multi-omics Blood ATlas (COMBAT) Consortium. A blood atlas of COVID-19 defines hallmarks of disease severity and specificity. *medRxiv* (2021) doi:10.1101/2021.05.11.21256877.
22. Richardson, S. *et al.* Presenting Characteristics, Comorbidities, and Outcomes among 5700 Patients Hospitalized with COVID-19 in the New York City Area. *JAMA - Journal of the American Medical Association* **323**, 2052–2059 (2020).
23. Mei, X. *et al.* Artificial intelligence–enabled rapid diagnosis of patients with COVID-19. *Nature Medicine* **26**, 1224–1228 (2020).
24. Soltan, A. A. S. *et al.* Artificial intelligence driven assessment of routinely collected healthcare data is an effective screening test for COVID-19 in patients presenting to hospital. *medRxiv* 2020.07.07.20148361 (2020) doi:10.1101/2020.07.07.20148361.
25. Jiao, Z. *et al.* Prognostication of patients with COVID-19 using artificial intelligence based on chest x-rays and clinical data: a retrospective study. *The Lancet Digital Health* **3**, e286–e294 (2021).
26. Laguarta, J., Hueto, F. & Subirana, B. COVID-19 Artificial Intelligence Diagnosis Using Only Cough Recordings. *IEEE Open Journal of Engineering in Medicine and Biology* **1**, 275–281 (2020).
27. Roberts, M. *et al.* Common pitfalls and recommendations for using machine learning to detect and prognosticate for COVID-19 using chest radiographs and CT scans. *Nature Machine Intelligence* **3**, 199–217 (2021).
28. Wynants, L. *et al.* Prediction models for diagnosis and prognosis of covid-19: Systematic review and critical appraisal. *The BMJ* **369**, (2020).

29. Laghi, A. Cautions about radiologic diagnosis of COVID-19 infection driven by artificial intelligence. *The Lancet Digital Health* **2**, e225 (2020).
30. The Lancet Digital Health. Artificial intelligence for COVID-19: saviour or saboteur? *The Lancet Digital Health* **3**, e1 (2021).
31. Bachar, N. *et al.* An Artificial Intelligence-Assisted Diagnostic Platform for Rapid Near-Patient Hematology. *medRxiv* (2021) doi:10.1101/2021.04.27.21255770.
32. UK Government. *Over half of UK adults vaccinated with second dose*. https://www.gov.uk/government/news/over-half-of-uk-adults-vaccinated-with-second-dose (2021).
33. Newcombe, R. G. & Altman, D. G. *Proportions and their differences, in Statisics with Confidence: Confidence intervals and statisctical guidelines, 2nd Ed.* (BMJ Books, 2000).
34. Sun, X. & Xu, W. Fast implementation of DeLong's algorithm for comparing the areas under correlated receiver operating characteristic curves. *IEEE Signal Processing Letters* **21**, 1389–1393 (2014).
35. Buderer, N. M. F. Statistical Methodology: I. Incorporating the Prevalence of Disease into the Sample Size Calculation for Sensitivity and Specificity. *Academic Emergency Medicine* **3**, 895–900 (1996).
36. Arifin, W. N. Sample size calculator. *https://wnarifin.github.io/ssc/sssnsp.html* https://wnarifin.github.io/ssc/sssnsp.html (2021).
37. Kung, C. te *et al.* Effective strategies to prevent in-hospital infection in the emergency department during the novel coronavirus disease 2019 pandemic. *Journal of Microbiology, Immunology and Infection* vol. 54 120–122 (2021).
38. University Hospitals Plymouth NHS Trust. Ward and Department -Coronavirus Infectious Disease-19 Zone Status. (2020).
39. Mina, M. J., Peto, T. E., García-Fiñana, M., Semple, M. G. & Buchan, I. E. Clarifying the evidence on SARS-CoV-2 antigen rapid tests in public health responses to COVID-19. *Lancet (London, England)* **397**, 1425–1427 (2021).
40. van Kampen, J. J. A. *et al.* Duration and key determinants of infectious virus shedding in hospitalized patients with coronavirus disease-2019 (COVID-19). *Nature Communications* **12**, (2021).
41. Dinnes, J. *et al.* Rapid, point-of-care antigen and molecular-based tests for diagnosis of SARS-CoV-2 infection. *Cochrane Database of Systematic Reviews* vol. 2021 (2021).
42. Office for National Statistics (UK). *Coronavirus (COVID-19) Infection Survey, UK: 21 May 2021*. https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/bulletins/coronaviruscovid19infectionsurveypilot/21may2021 (2021).
43. Islam, N. *et al.* Thoracic imaging tests for the diagnosis of COVID-19. *Cochrane Database of Systematic Reviews* vol. 2021 (2021).
44. Chen, T. & Guestrin, C. XGBoost: A scalable tree boosting system. in *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* vols. 13-17-August-2016 785–794 (Association for Computing Machinery, 2016).
45. Collier, D. A. *et al.* Point of Care Nucleic Acid Testing for SARS-CoV-2 in Hospitalized Patients: A Clinical Validation Trial and Implementation Study. *Cell Reports Medicine* **1**, (2020).
46. Bossuyt, P. M. *et al.* STARD 2015: An updated list of essential items for reporting diagnostic accuracy studies. *The BMJ* **351**, (2015).

47. Liu, X., Cruz Rivera, S., Moher, D., Calvert, M. J. & Denniston, A. K. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. *Nature medicine* vol. 26 1364–1374 (2020).

Supplementary Material

Real-world evaluation of AI-driven COVID-19 triage for emergency admissions: External validation & operational assessment of lab-free and high-throughput screening solutions

Andrew A.S. Soltan et al.

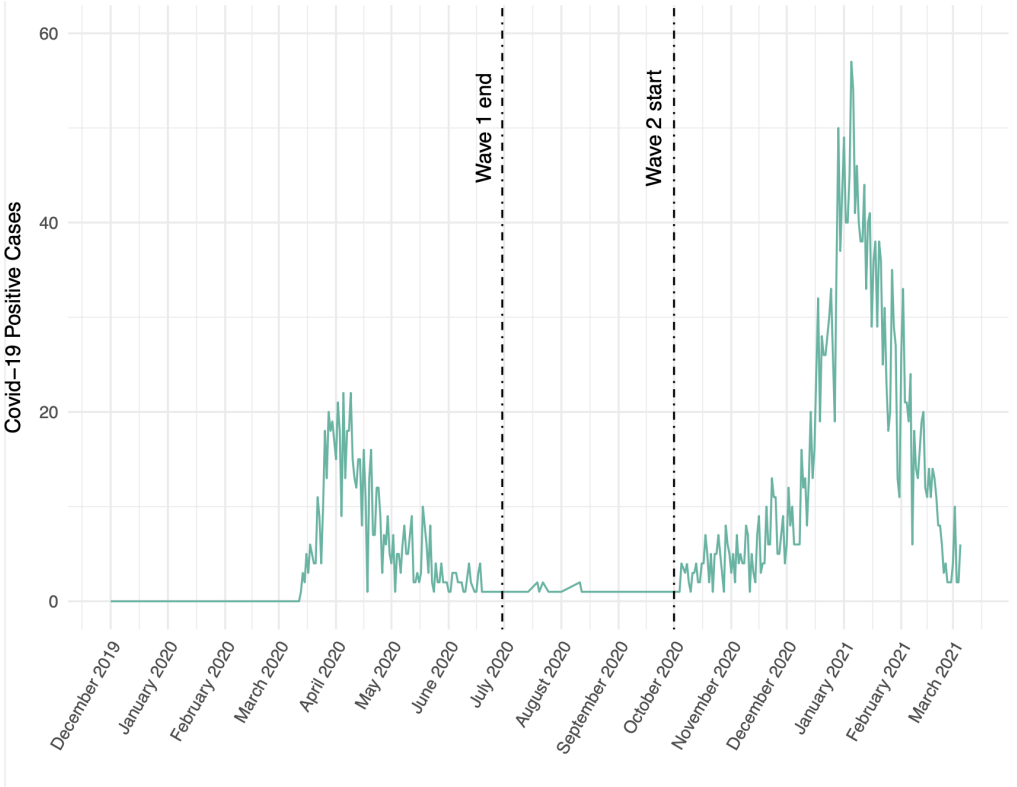
Appendix A

CURIAL-Rapide Translational Collaborative

We thank all healthcare professionals and students who have supported the CURIAL-Rapide Service evaluation. In particular, we wish to acknowledge:

Adam Watson
Akshay Bhargav
Alex Tough
Alice Rogers
Ayisha M.A. Shaikh
Carolina Valensise
Charlotte Lee
Claire Otasowie
David Metcalfe
Ekta Agarwal
Elham Zareh
Evelyn Thangaraj
Florence Pickles
Gabriella Kelly
Gayatri Tadikamalla
George Shaw
Heather Tong
Hettie Davies
Jaspreet Bahra
Jessica Morgan
Joe Wilson
Joseph Cutteridge
Katherine O'Byrne
Luiza Farache Trajano
Madeleine Oliver
Maria Pikoula
Maya Mendoza
Melissa Keevil
Muhammad Faisal
Natasha Dole
Oscar Deal
Rebecca Conway-Jones
Shajeel Sattar
Sneha Kundoor
Sumaiyah Shah
Vani Muthusami

Appendix B:



Supplementary Figure S1: Daily number of patients presenting to Oxford University Hospitals NHS Foundation Trust testing positive for COVID-19, between 1st December 2019 and 8th March 2021.

Model Development: Inclusion & Exclusion Criteria
As previously, we included all patients attending acute and emergency care settings at Oxford University Hospitals NHS foundation trust who received routine blood tests on arrival, considering presentations before December 1, 2019, and thus before the pandemic, as the COVID-19-negative (control) cohort. We considered presentations during the 'first wave' of the UK COVID-19 pandemic (December 1, 2019 to June 30, 2020) with PCR confirmed SARS-CoV-2 infection as the COVID-19-positive (cases) cohort. We excluded patients who opted out of electronic health record (EHR) research and those who did not receive laboratory blood tests or were younger than 18 years of age. Due to incomplete penetrance of testing during the first wave of the pandemic, and imperfect sensitivity of the PCR test, there is uncertainty in the viral status of patients presenting during the pandemic who were untested or tested negative. We therefore selected a pre-pandemic control cohort during training to ensure absence of disease in patients labelled as COVID-19-negative.
Clinical features extracted for each presentation included first-performed blood tests, blood gases, vital signs measurements and PCR testing for SARS-CoV-2 (Abbott Architect [Abbott, Maidenhead, UK], TaqPath [Thermo Fisher Scientific, Massachusetts, USA] and Public Health England-designed RNA-dependent RNA polymerase assays). A list of extracted clinical features is shown in Supplementary Table S1.

Normalisation
Data normalisation was implemented to mitigate overfitting and to avoid the reliance of the model on measurement units. Categorical data were handled by encoding as "1-hot" variables. Where a lab value was reported as being below the threshold of detection of the laboratory assay, the value was replaced with a numerical zero value. Where values were reported as being above the threshold of detection, clinically appropriate values were selected to maintain the significance of the high result. A summary of first-performed blood tests, vital signs, and blood gasses on arrival to hospital are shown Supplementary Tables S2-S4.

Missing Data
Multiple imputation strategies, population median, population mean, and age-based imputation, were separately used to impute missing data initially. As a sensitivity analysis to assess for effects of imputation strategy on model performance, we assessed performance of models trained using each imputation method prospectively for all patients attending emergency departments and acute medical services across OUH during the second-wave of the COVID-19 pandemic (October 01, 2020 and March 06, 2021; Table 2). Mean performance was reported alongside SD in Table 2, with narrow standard deviations in all performance metrics demonstrating resilience to imputation method. We therefore subsequently only used models trained with missing data imputed using population median, reporting results alongside 95% confidence intervals (CIs).

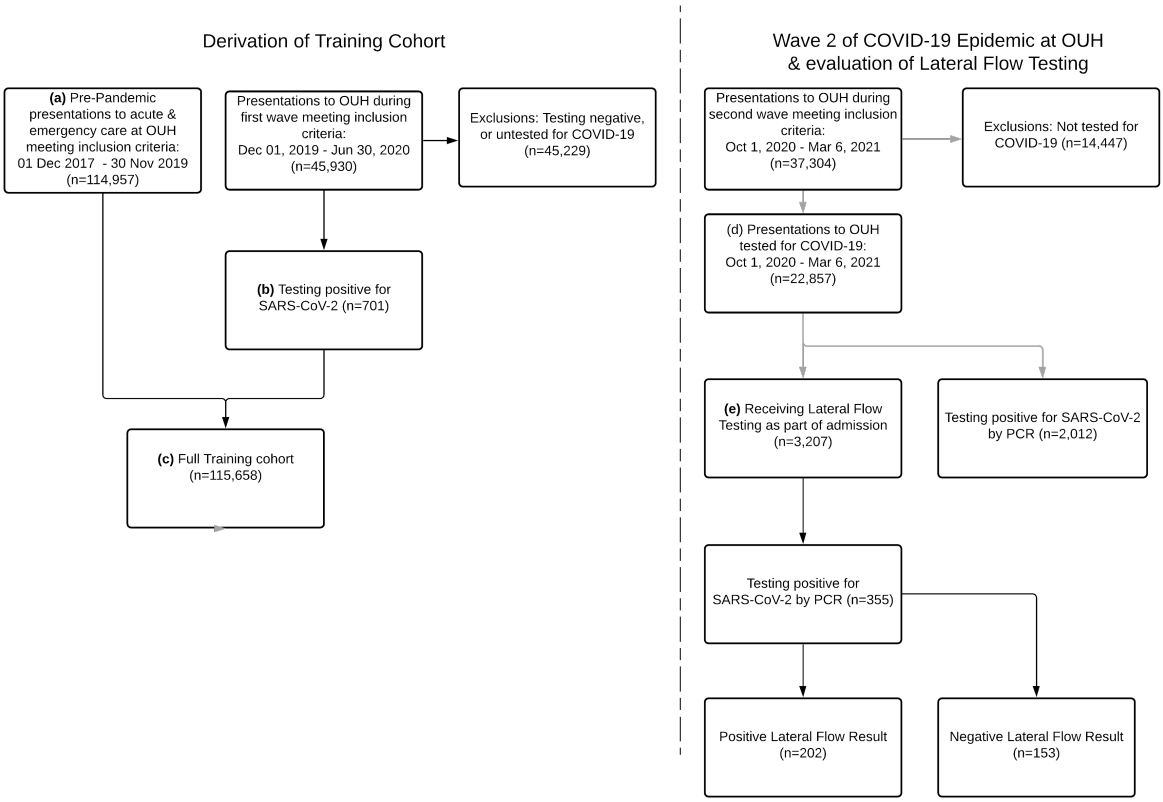
Model Training & Prospective Evaluation:
We repeated training and optimisation of our eXtreme Gradient BOOSTed tree model (XGBoost) to discriminate COVID-19-positive cases from pre-pandemic COVID-19-negative controls, for each of the three feature-sets (Table 1)[44]. During training using 'first wave' case, controls were matched for age, gender, and ethnicity at a ratio of 1:20. For missing data, we initially used three independent imputation

methods during training – median, mean and age-based mean – and assessed sensitivity of model performance to imputation strategy during testing. Thresholds were calibrated to achieve sensitivities of 80% and 90% during training, using stratified 10-fold cross validation.

XGBoost is a generalisation of boosting to an arbitrary differentiable loss function. XGBoost is more robust to outliers and has high predictive power. The scikit-learn (v0.23.2), LIBLINEAR (v2.41) and XGBoost (v1.2.0) modules for Python were used during model development and classifier evaluation.

Supplementary Table S1: Clinical data fields extracted from training and externally validating NHS sites, for all patients admitted to the trusts from 1st December 2019 onwards, and for up to 200,000 pre-pandemic admissions.

Clinical Descriptors:	**Presentation Blood Tests:**	**Presentation Blood Gas:**	**Premorbid Clinical Data**
Study ID	PresentationHAEMOGLOBIN	PresentationPOCT pC02	BaselineHAEMOGLOBIN
Presentation Date	PresentationWHITE CELLS	PresentationPOCT sO2	BaselineWHITE CELLS
Ethnicity	PresentationPLATELETS	PresentationPOCT pO2	BaselinePLATELETS
Age at presentation	PresentationMEAN CELL VOL.	PresentationPCT cBASE(Ecf)c	BaselineMEAN CELL VOL.
Gender (M/F)	PresentationRED CELL COUNT	PresentationPCT CO3(P,st)c	BaselineRED CELL COUNT
Comorbidities (ICD10)	PresentationNEUTROPHILS	PresentationPOCT Hctc	BaselineNEUTROPHILS
Outcome	PresentationHAEMATOCRIT	PresentationPOCT FO2Hb	BaselineHAEMATOCRIT
Vital Signs:	PresentationLYMPHOCYTES	PresentationPOCT ctO2c	BaselineLYMPHOCYTES
AdmissionRespRate	PresentationMEAN CELL HGB	PresentationPOCT cGLU	BaselineMEAN CELL HGB
AdmissionHeartRate	PresentationMONOCYTES	PresentationPOCT cK+	BaselineMONOCYTES
AdmissionBloodPressure	PresentationEOSINOPHILS	PresentationPOCT cNA+	BaselineEOSINOPHILS
AdmissionSpO2	PresentationBASOPHILS	PresentationPOCT cLAC	BaselineBASOPHILS
AdmissionOxygenDeliveryDevice	Presentation MCH	PresentationPOCT cCA++	BaselineMEAN CELL HGB CONC
AdmissionTemperature	PresentationMPV		BaselineSODIUM
Microbiology:	PresentationNRBC A		BaselineALBUMIN
SARS-CoV-2 PCR	PresentationNRBC %		BaselineALK.PHOSPHATASE
SARS-CoV-2 RESULT TYPE	PresentationSODIUM		BaselineALT
SARS-CoV-2 Antigen Test Result	PresentationALBUMIN		BaselineUREA
INFLUENZAPCR	PresentationALK.PHOSPHATASE		BaselineBILIRUBIN
RespiratoryPCR (Biofire)	PresentationALT		BaselineCREATININE
	PresentationUREA		BaselineeGFR
	PresentationBILIRUBIN		BaselinePOTASSIUM
	PresentationCREATININE		BaselineCALCIUM
	PresentationeGFR		BaselineADJUSTED CALC.
	PresentationPOTASSIUM		BaselineCRP
	PresentationCALCIUM		BaselineProthromb. Time
	PresentationADJUSTED CALC.		BaselineAPTT
	PresentationPHOSPHATE		BaselineINR
	PresentationCRP		BaselinePOCT pC02
	PresentationProthromb. Time		BaselinePOCT sO2
	PresentationPOCT ctHb		BaselinePOCT pO2
	PresentationGLUCOSE		BaselinePCT cBASE(Ecf)c
	PresentationAPTT		BaselinePCT CO3(P,st)c
	PresentationINR		BaselinePOCT Hctc
			BaselinePOCT FO2Hb
			BaselinePOCT ctO2c
			BaselinePOCT Cglu
			BaselinePOCT cK+
			BaselinePOCT cNA+
			BaselinePOCT cLAC
			BaselinePOCT cCA++



Supplementary Figure S2: Participant flow diagram showing patients attending OUH, who met inclusion and exclusion criteria, for (a) the pre-pandemic training cohort and (b) COVID-19-cases cohort, combining to form (c) a full training cohort for model development. Patients attending OUH during the second wave of the UK COVID-19 epidemic, between Oct 1, 2020 and Mar 6, 2020, meeting inclusion and exclusion criteria, formed (d) the second wave analysis cohort, of which a subset (e) received Lateral Flow Testing within routine care, as part of an admission.

Supplementary Table S2: Distribution of vital signs, reported as median and interquartile ranges, for each patient cohort.

	Training		**Prospective Test**	**External Validation (Admissions)**			**LFD Evaluation**	**Lab-free Evaluation**
	Oxford University Hospitals (pre-pandemic & wave 1 cases, to 30 June 2020)		Oxford University Hospitals	Portsmouth University Hospital NHS Trust	University Hospitals Birmingham NHS Foundation Trust	Bedfordshire Hospitals NHS Foundation Trust	Oxford University Hospitals (wave 2 receiving LFDs)	Oxford University Hospitals ED
	Prepandemic cohort	COVID-19-cases cohort	October 1, 2020 – March 6, 2021	March 1, 2020 - February 28, 2021	December 01, 2019 - October 29, 2020	January 1, 2021 - March 31, 2021	December 23, 2020 – March 6, 2021	Feb 18, 2021 – May, 10, 2021
Respiratory Rate (breath/min)	18.0 (16.0-19.0)	20.0 (18.0-24.0)	18.0 (16.6-19.0)	17.0 (16.0-19.0)	18.0 (17.0-20.0)	18.0 (16.0-20.0)	18.0 (17.0-20.0)	18.0 (17.0-20.0)
Heart Rate (beats/min)	82.0 (71.0-96.0)	88.0 (75.0-101.0)	84.0 (72.0-97.0)	82.0 (71.0-95.0)	86.0 (73.0-101.0)	84.0 (73.0-97.0)	87.0 (75.0-101.0)	82.0 (70.0-97.0)
Systolic Blood Pressure (mmHg)	132.0 (118.0-150.0)	131.0 (115.0-146.0)	134.0 (119.0-152.0)	128.0 (114.0-146.0)	136.0 (119.0-155.0)	131.0 (116.0-149.0)	136.0 (120.0-156.0)	146.5 (126.0-168.0)
Diastolic Blood Pressure (mmHg)	74.0 (65.0-84.0)	74.0 (64.0-84.0)	75.0 (65.0-85.0)	76.0 (67.0-84.0)	77.0 (68.0-87.0)	78.0 (68.0-88.0)	76.0 (65.0-87.0)	79.0 (68.0-90.8)

Oxygen Saturation (%)	97.0 (96.0-99.0)	96.0 (94.0-97.0)	97.0 (96.0-97.9)	97.0 (95.0-98.0)	97.0 (95.0-98.0)	97.0 (96.0-99.0)	97.0 (95.0-97.7)	97.0 (96.0-99.0)
Tympanic Temperature (C)	36.5 (36.1-36.9)	36.9 (36.3-37.6)	36.3 (36.0-36.7)	36.3 (36.0-36.8)	36.7 (36.4-37.2)	36.5 (36.4-36.9)	36.3 (35.9-36.8)	36.3 (35.8-36.8)

Supplementary Table S3: Distribution of blood test features, reported as median and interquartile ranges, for each patient cohort.

	Training		**Prospective Test**	**External Validation (Admissions)**			**LFD Evaluation**	**Lab-free Evaluation**
	Oxford University Hospitals (pre-pandemic & wave 1 cases, to 30 June 2020)		Oxford University Hospitals	Portsmouth University Hospital NHS Trust	University Hospitals Birmingham NHS Foundation Trust	Bedfordshire Hospitals NHS Foundation Trust	Oxford University Hospitals (wave 2 receiving LFDs)	Oxford University Hospitals ED (OLO FBC results)
	Prepandemic cohort	COVID-19-cases cohort	October 1, 2020 – March 6, 2021	March 1, 2020 - February 28, 2021	December 1, 2019 - October 29, 2020	January 1, 2021 - March 31, 2021	December 23, 2020 – March 6, 2021	Feb 18, 2021 – May, 10, 2021
HAEMOGLOBIN (g/L)	130.0 (116.0-142.0)	130.0 (114.0-144.0)	129.0 (114.0-142.0)	129.0 (114.0-143.0)	127.0 (113.0-140.0)	134.0 (119.0-146.0)	131.0 (116.0-144.0)	125.0 (112.0-137.5)
WHITE CELLS (10^9 l^{-1})	8.45 (6.46-11.18)	6.98 (5.14-9.72)	8.94 (6.7-12.06)	8.6 (6.7-11.3)	9.4 (7.1-12.6)	9.2 (6.9-12.5)	9.43 (7.05-12.73)	8.56 (6.68-11.37)
PLATELETS (10^9 l^{-1})	249.0 (199.0-307.0)	215.0 (163.0-283.5)	251.0 (198.0-314.0)	251.0 (199.0-312.0)	247.0 (196.0-311.0)	246.0 (196.0-310.0)	249.0 (195.0-313.0)	223.0 (183.5-270.5)
MEAN CELL VOL (fl)	89.6 (86.0-93.4)	90.2 (86.6-94.2)	90.2 (86.6-94.2)	89.0 (84.9-93.0)	89.9 (86.2-93.6)	88.0 (85.0-92.0)	90.0 (86.4-94.3)	90.2 (87.1-93.7)
NEUTROPHILS (10^9 l^{-1})	5.72 (3.99-8.36)	5.11 (3.48-7.49)	6.44 (4.4-9.55)	5.9 (4.2-8.6)	6.9 (4.7-10.0)	6.8 (4.7-9.73)	6.97 (4.68-10.19)	6.26 (4.36-9.05)
HAEMATOCRIT	0.39 (0.35-0.42)	0.4 (0.35-0.44)	0.39 (0.35-0.43)	0.39 (0.34-0.42)	0.38 (0.34-0.42)	0.39 (0.35-0.43)	0.4 (0.36-0.43)	0.37 (0.33-0.41)
LYMPHOCYTES (10^9 l^{-1})	1.51 (1.0-2.13)	0.96 (0.65-1.38)	1.31 (0.85-1.89)	1.5 (0.97-2.2)	1.3 (0.9-1.9)	1.27 (0.86-1.83)	1.26 (0.83-1.89)	1.25 (0.86-1.78)
MONOCYTES (10^9 l^{-1})	0.64 (0.48-0.85)	0.49 (0.35-0.74)	0.66 (0.48-0.89)	0.63 (0.48-0.85)	0.7 (0.5-0.9)	0.66 (0.48-0.92)	0.68 (0.49-0.93)	0.59 (0.43-0.78)
EOSINOPHILS (10^9 l^{-1})	0.1 (0.04-0.2)	0.01 (0.0-0.06)	0.07 (0.02-0.16)	0.1 (0.02-0.2)	0.1 (0.0-0.2)	0.06 (0.02-0.16)	0.06 (0.01-0.14)	0.09 (0.05-0.17)
BASOPHILS (10^9 l^{-1})	0.04 (0.03-0.06)	0.02 (0.01-0.03)	0.04 (0.02-0.06)	0.04 (0.02-0.06)	0.1 (0.0-0.1)	0.05 (0.03-0.07)	0.04 (0.02-0.06)	0.03 (0.01-0.04)
SODIUM (mM)	138.0 (136.0-140.0)	136.0 (134.0-139.0)	138.0 (135.0-140.0)	138.0 (136.0-140.0)	137.0 (134.0-139.0)	138.0 (136.0-140.0)	138.0 (135.0-140.0)	
ALBUMIN (g/L)	36.0 (32.0-39.0)	32.0 (28.0-35.0)	36.0 (31.0-39.0)	36.0 (31.0-40.0)	36.0 (32.0-40.0)	35.0 (31.0-39.0)	36.0 (31.0-39.0)	
ALKALINE PHOSPHATASE (IU/L)	80.0 (64.0-105.0)	82.0 (64.0-108.0)	84.0 (66.0-112.0)	84.0 (67.0-109.0)	90.0 (71.0-119.0)	94.0 (74.5-122.0)	86.0 (69.0-115.0)	
ALT (IU/L)	18.0 (13.0-28.0)	25.0 (17.0-41.0)	20.0 (13.0-33.0)	19.0 (13.0-30.0)	19.0 (13.0-30.0)	20.0 (13.0-31.0)	20.0 (13.0-33.0)	
UREA (mM)	5.3 (4.0-7.4)	5.9 (4.2-9.07)	5.7 (4.2-8.3)	5.2 (3.8-7.6)	6.2 (4.5-9.0)	5.8 (4.2-8.3)	5.9 (4.3-8.8)	
BILIRUBIN (umol/L)	9.0 (6.0-13.0)	9.0 (7.0-13.25)	9.0 (6.0-14.0)	10.0 (7.0-16.0)	10.0 (7.0-15.0)	10.0 (7.0-14.0)	10.0 (7.0-14.0)	
CREATININE (umol/L)	73.0 (60.0-93.0)	79.0 (65.0-106.0)	74.0 (60.0-97.0)	74.0 (60.0-96.0)	78.0 (62.0-105.0)	80.5 (65.75-104.0)	74.0 (60.0-98.0)	
eGFR (ml/min)	85.0 (63.0-150.0)	78.0 (53.0-150.0)	84.0 (58.0-150.0)	83.0 (60.0-90.0)	76.0 (52.0-90.0)	76.0 (54.0-90.0)	82.0 (56.0-150.0)	
POTASSIUM (mM)	4.0 (3.7-4.3)	4.0 (3.7-4.3)	4.0 (3.8-4.4)	4.2 (3.9-4.4)	4.1 (3.8-4.4)	4.3 (4.0-4.6)	4.1 (3.8-4.4)	
CRP (mg/L)	8.6 (2.3-39.0)	72.5 (23.8-143.6)	15.8 (3.5-67.4)	13.0 (3.0-71.0)	12.0 (3.0-61.0)	10.7 (2.8-48.78)	17.9 (3.6-77.5)	

Supplementary Table S4: Distribution of blood gas features, reported as median and interquartile ranges for each patient cohort.

	Training		**Prospective Test**	**External Validation (Admissions)**		**LFD Evaluation**
	Oxford University Hospitals (pre-pandemic & wave 1 cases, to 30 June 2020)		Oxford University Hospitals	University Hospitals Birmingham NHS Foundation Trust	Bedfordshire Hospitals NHS Foundation Trust	Oxford University Hospitals (wave 2 receiving LFDs)
	Prepandemic cohort	COVID-19-cases cohort	October 1, 2020 – March 6, 2021	December 01, 2019 - October 29, 2020	January 1, 2021 - March 31, 2021	December 23, 2020 – March 6, 2021
pCO2 (kPa)	5.57 (4.94-6.22)	5.34 (4.57-6.01)	5.61 (4.95-6.28)	5.7 (5.0-6.5)	5.72 (5.03-6.43)	5.68 (5.0-6.4)
O2 Sat (%)	64.5 (44.0-83.8)	65.15 (38.85-84.68)	65.3 (43.6-85.8)	69.8 (44.6-89.9)	68.0 (44.0-88.9)	60.35 (40.05-80.2)
pO2 (kPa)	4.68 (3.51-6.53)	4.62 (3.4-6.7)	4.79 (3.54-6.92)	5.0 (3.4-7.4)	4.86 (3.41-7.2)	4.52 (3.44-6.2)
BE Std (mM)	1.3 (-0.7-3.2)	1.3 (-0.8-3.3)	1.4 (-0.8-3.5)	-0.1 (-2.1-1.5)	2.2 (0.02-4.2)	1.6 (-0.6-3.8)
Bicarbonate (mM)	24.7 (23.2-26.0)	24.7 (23.2-26.3)	24.8 (23.2-26.3)	24.9 (22.6-27.1)	27.2 (24.72-29.6)	24.8 (23.1-26.3)
Haematocrit	41.6 (37.2-45.4)	41.3 (36.8-45.8)	41.3 (36.4-45.5)	43.0 (38.7-46.5)	41.8 (36.9-45.9)	42.0 (37.3-46.0)
Glucose (mM)	6.2 (5.4-7.5)	6.6 (5.7-8.2)	6.4 (5.5-8.0)	6.73 (5.81-8.48)	6.4 (5.5-8.0)	6.6 (5.6-8.3)
K+ (mM)	3.9 (3.7-4.3)	3.85 (3.6-4.2)	4.0 (3.7-4.3)	3.96 (3.66-4.3)	4.0 (3.7-4.3)	4.0 (3.7-4.3)
Na+ (mM)	138.0 (135.0-141.0)	137.0 (133.0-140.0)	138.0 (135.0-141.0)	140.0 (137.2-141.9)	139.0 (135.0-141.0)	138.0 (135.0-141.0)
cLAC (mM)	1.3 (0.9-1.9)	1.4 (1.24-2.0)	1.4 (1.24-1.9)	1.64 (1.25-2.27)	1.3 (1.0-1.9)	1.4 (1.24-2.1)
Ca2+ (mM)	1.18 (1.14-1.21)	1.12 (1.08-1.16)	1.17 (1.13-1.21)	1.21 (1.16-1.24)	1.17 (1.12-1.2)	1.17 (1.14-1.21)
Haemoglobin (g/dL)	136.0 (121.0-148.0)	134.0 (120.0-149.0)	135.0 (118.0-148.0)	133.1 (118.4-146.3)	136.0 (120.0-150.0)	137.0 (121.0-150.0)

Appendix C:
External validation at independent NHS Trusts:

We externally validated CURIAL-Rapide and CURIAL-Lab by applying the respective models to results of first-available blood test results and vital signs (Table 1), comparing model predictions to confirmatory SARS-CoV-2 viral genome test results. For trusts where blood-gas results were available for electronic extraction, we also evaluated CURIAL-1.0. Patients meeting inclusion criteria had an unscheduled acute or emergency care admission, during the specified periods, received a blood draw on arrival and were aged over 18. We excluded patients who did not have a valid confirmatory test result within a prespecified period, or who had opted out of EHR research. Screening against eligibility criteria, followed by anonymisation, was performed by the respective NHS Trusts.

Evaluation at Portsmouth Hospitals NHS Foundation Trust (PUH) considered all patients admitted to the Queen Alexandria Hospital, serving a population of 675,000 and offering tertiary referral services to the surrounding region, between March 1, 2020 and February 28, 2021. Confirmatory COVID-19 testing was by laboratory SARS-CoV-2 RT-PCR assay, considering any positive PCR result within 48hrs of admission as a true positive. As blood gas results were not available for electronic extraction, we evaluated only CURIAL-Rapide and CURIAL-Lab at Portsmouth.

Evaluation at University Hospitals Birmingham NHS Foundation (UHB) trust considered all patients admitted to The Queen Elizabeth Hospital, Birmingham, between December 01, 2019 and October 29, 2020. The Queen Elizabeth Hospital is a large tertiary referral unit within the UHB group which provides healthcare services for a population of 2.2 million across the West Midlands. Confirmatory COVID-19 testing was performed by laboratory SARS-CoV-2 RT-PCR assay.

Evaluation at Bedfordshire NHS Foundation Trust (BHT) considered all patients admitted to Bedford Hospital between January 1, 2021 and March 31, 2021. BHT provides healthcare services for a population of around 620,000 in Bedfordshire. Confirmatory COVID-19 testing was performed on the day of admission by point-of-

care PCR based nucleic acid testing [SAMBA-II & Panther Fusion System, Diagnostics in the Real World, UK, and Hologic, USA]. In an evaluation of the SAMBA-II against laboratory RT-PCR testing, the SAMBA-II achieved sensitivity of 96.9% and specificity of 99.1%[9,45].

We report sensitivity, specificity, positive and negative predictive values (PPV and NPV), AUROC and F1 alongside 95% CIs (Supplementary Table S5 & Figure 2), comparing model predictions to results of confirmatory viral testing (laboratory PCR and SAMBA-II). . Confidence intervals for sensitivity, specificity and predictive values were computed using Wilson's Method[33], and for AUROC with DeLong's method[34].

Supplementary Table S5: Performance of CURIAL-Rapide, CURIAL-Lab & CURIAL-1.0 (Soltan et al.) during external validation at three UK Hospitals trusts. All models were calibrated during training to achieve 90% sensitivity. Results are reported alongside 95% confidence intervals. (Acronyms – FBC: Complete Blood Count, U&E: Creatinine & Electrolytes, LFD: Liver Function Test, CRP: C-Reactive Protein)

	Portsmouth University Hospitals NHS Trust ***n= 37,896, prevalence = 5.29%***		**University Hospitals Birmingham NHS Foundation Trust** ***n=10,293; prevalence = 4.27%***			**Bedfordshire Hospitals NHS Foundation Trust** ***n=1,177; prevalence = 12.2%***		
	CURIAL-Rapide	CURIAL-Lab	CURIAL-Rapide	CURIAL-Lab	CURIAL-1.0	CURIAL-Rapide	CURIAL-Lab	CURIAL-1.0
Sensitivity (%)	83.5 (81.8 - 85.1)	84.1 (82.5 - 85.7)	82.2 (78.4 - 85.5)	78.8 (74.8 - 82.4)	83.4 (79.6 - 86.6)	74.3 (66.6 - 80.7)	74.3 (66.6 - 80.7)	72.9 (65.1 - 79.5)
Specificity (%)	63.6 (63.1 - 64.1)	71.3 (70.9 - 71.8)	65.4 (64.5 - 66.3)	74.7 (73.8 - 75.5)	68.7 (67.7 - 69.6)	81.8 (79.3 - 84.0)	84.8 (82.5 - 86.9)	83.6 (81.3 - 85.8)
PPV (%)	11.4 (10.9 - 11.9)	14.1 (13.5 - 14.7)	9.6 (8.7 - 10.6)	12.2 (11.0 - 13.4)	10.6 (9.6 - 11.7)	36.3 (31.0 - 41.9)	40.5 (34.8 - 46.5)	38.3 (32.8 - 44.2)
NPV (%)	98.6 (98.4 - 98.7)	98.8 (98.6 - 98.9)	98.8 (98.5 - 99.0)	98.8 (98.5 - 99.0)	98.9 (98.7 - 99.2)	95.8 (94.3 - 96.9)	95.9 (94.5 - 97.0)	95.7 (94.2 - 96.8)
F1	0.200	0.241	0.172	0.211	0.188	0.487	0.525	0.502
AUROC	0.842 (0.832 - 0.852)	0.872 (0.863 - 0.882)	0.836 (0.814 - 0.858)	0.858 (0.838 - 0.878)	0.846 (0.825 - 0.867)	0.854 (0.819 - 0.889)	0.881 (0.851 - 0.912)	0.865 (0.830 - 0.900)

Comparison with Lateral Flow Tests

We considered any positive lateral flow test which was followed by a positive PCR test within a +/- 48hr window of a patient being admitted to hospital to represent a true positive infection. As previously, model predictions were generated using blood tests performed from the first blood draw on arrival and first-recorded vital signs. In the integrated clinical pathway (Figure 1), patients were considered COVID-19-suspected if they had either a positive LFD result or CURIAL prediction. Results are show in Figure 3 and Supplementary Table S6.

Supplementary Table S6: Performance characteristics of (a) INNOVA SARS-CoV2 Rapid Antigen Tests, (b) CURIAL-Rapide & CURIAL-Lab, calibrated during training to a sensitivity of 80%, and (c) combined clinical pathways considering either a positive CURIAL-Rapide/CURIAL-Lab result or a positive LFD test as a COVID-19 suspected case, at Oxford University Hospitals NHS Foundation Trust between December 23, 2020 & March 6, 2021. Error bars show 95% confidence intervals.

n=3207 prevalence 11.1%		**CURIAL-Rapide: FBC & Vitals**		**CURIAL-Lab: FBC, C&E, LFD, CRP + Vitals**		**CURIAL-1.0: Blood Tests + Blood Gas + Vitals**	
Feature Sets	Innova SARS-CoV-2 Rapid Antigen Tests	CURIAL-Rapide	Innova Lateral Flow Tests + CURIAL-Rapide	CURIAL-Lab	Innova Lateral Flow Tests + CURIAL-Lab	CURIAL-1.0	Innova Lateral Flow Tests + CURIAL-1.0
Sensitivity	56.9% (51.7 - 62.0)	78.0% (73.4 - 82.0)	88.2% (84.4 - 91.1)	74.4% (69.6 - 78.6)	85.6% (81.6 - 88.9)	76.1% (71.4 - 80.2)	85.9% (81.9 - 89.2)
Specificity	99.8% (99.6 - 99.9)	80.0% (78.5 - 81.4)	79.9% (78.4 - 81.3)	88.4% (87.2 - 89.5)	88.3% (87.0 - 89.4)	88.5% (87.3 - 89.6)	88.4% (87.1 - 89.5)
PPV	97.6% (94.5 - 99.0)	32.7% (29.6 - 35.9)	35.3% (32.2 - 38.5)	44.4% (40.4 - 48.4)	47.6% (43.7 - 51.4)	45.2% (41.2 - 49.2)	47.9% (44.0 - 51.8)
NPV	94.9% (94.1 - 95.6)	96.7% (95.9 - 97.3)	98.2% (97.6 - 98.7)	96.5% (95.7 - 97.2)	98.0% (97.4 - 98.5)	96.7% (96.0 - 97.4)	98.1% (97.4 - 98.5)
F1	0.719	0.461	0.504	0.556	0.612	0.567	0.615
AUROC		0.854 (0.829 - 0.879)	0.919 (0.899 - 0.940)	0.877 (0.853 - 0.901)	0.925 (0.905 - 0.945)	0.887 (0.865 - 0.909)	0.926 (0.907 - 0.946)

Appendix D:
CURIAL-Rapide lab-free service evaluation

Service evaluation of the OLO haematology analyser/CURIAL-Rapide operated between February 18, 2021 and May 10, 2021 between 8am and 8pm.

Operator Training

We specified that clinical staff carrying out the service evaluation must ordinarily be employed by OUH, participate in the care of patients as part of their usual duties, have completed all statutory & mandatory training required by the trust for their role including for electronic health record systems, and be familiar and competent in using these systems as part of their usual role. We permitted student doctors meeting the above requirements to participate. Training to operate the OLO was provided by in-person device training, supported by demonstration and documentation from the device manufacturers, and a supporting online training video (made available at https://youtu.be/UofBAL7sAzc). Weekly quality-control checks were performed on the OLO analysers.

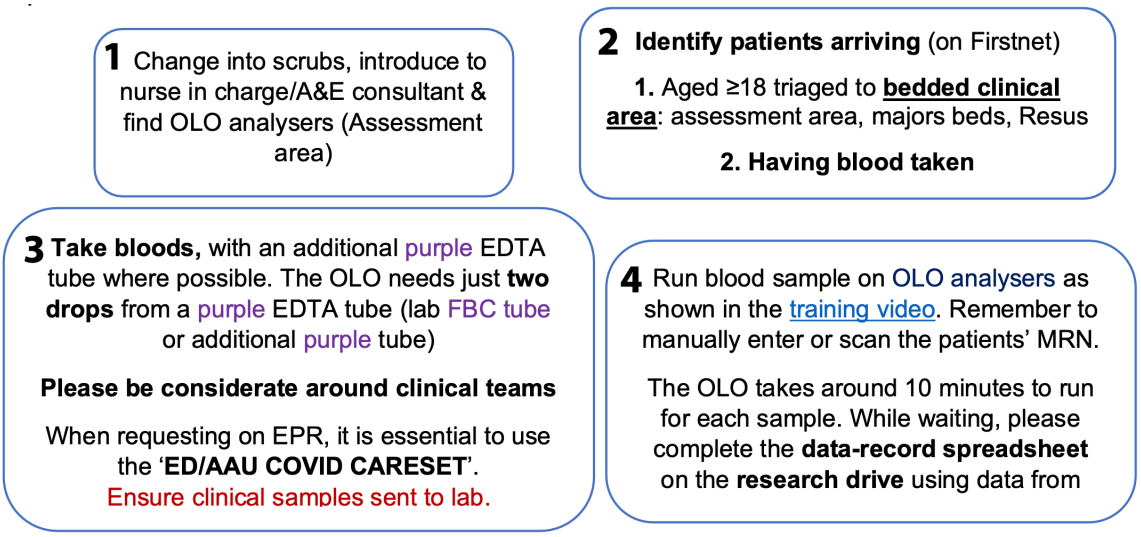
Enrolment:

OUH sites for eligibility: John Radcliffe Hospital
Inclusion: Adult patients (aged >18)
Clinical areas for sampling eligibility were ED Assessment area, ED Majors Beds and ED Resus. Patients who are not receiving blood tests on presentation to the emergency department as part of their care were not eligible.

Process:

Eligible patients were identified to take part in the service evaluation using the locally-adopted Cerner FirstNet system. Vital signs and blood draws were performed on arrival to the emergency department by healthcare professionals as part of routine care. Following trust procedures, vital signs were documented on the trust electronic health record [SEND; Sensyne Health], and blood bottles were labelled using printed labels from the electronic record. Two drops of venous blood (27uL) from a routinely-collected EDTA blood tube were extracted using a single-use sampling device, and prepared for OLO analysis by trained operators directed by on-screen instructions[31]. OLO results were uploaded immediately to the electronic medical record using the POCcelerator Data Management System [Siemens Healthineers GmbH, Erlangen, Germany], making results available to clinicians and

supporting routine patient care. Routine laboratory FBC analysis [Sysmex XN Automated, Sysmex UK] was used to confirm point of care results. Clinical care followed existing pathways and departmental procedures.



Supplementary Figure S3: Instructions to trained operators, specifying eligibility criteria for the service evaluation, sample handling and processing techniques.

Confirmatory COVID-19 Testing:

Confirmatory testing of patients enrolled in the OLO/CURIAL-Rapide service evaluation, and LFD comparison, followed OUH trust policies. Swabs of the nose and throat were routinely performed in the emergency department for all patients being admitted to OUH. Lateral Flow Testing (Innova SARS-CoV-2 Antigen Rapid Qualitative Test) was performed in the department, by trained nursing or medical staff, and results were documented on the electronic record. Swabs for PCR were transferred to the clinical laboratory in viral transport medium and tested by PCR (ThermoFisher TaqPath). Where patients were not tested for COVID-19 by confirmatory PCR, or did not receive blood tests or vital signs as part of routine care, we excluded the patients from the CURIAL-Rapide evaluation. We also excluded patients with an invalid OLO result and no subsequent successful result, thereby ensuring data completeness.

Analysis

Binary CURIAL-Rapide triage predictions (COVID-19-Suspected and COVID-19-Negative) were generated using a custom Python 3.0 application. Libraries used included scikit-learn, pandas, and NumPy. No other clinical data was made available to the algorithm. CURIAL-Rapide predictions were not made available to clinicians in this study, so as not to influence the clinical triage category or decisions to proceed to confirmatory testing.

We compared CURIAL-Rapide predictions, lateral flow results, and clinical triage category by first-assessing clinician against a PCR reference standard. We determined and report sensitivity, specificity, PPV, NPV and accuracy, alongside 95% confidence intervals. We calculated the time-to-result for each test, presenting mean with standard deviation for normally distributed data, and median with interquartile range for data with a skewed distributed (Table 6). Laboratory FBC

samples were not processed for 2 of the 520 patients, owing to sample or labelling errors. For paired samples, we compared time-to-result between each test using a one-tailed Wilcoxon Signed Rank test. We additionally performed a Kaplan-Meier survival analysis (Figure 4).

We report our study in compliance with the “Standards for Reporting Diagnostic accuracy studies” (STARD) standards[46,47].

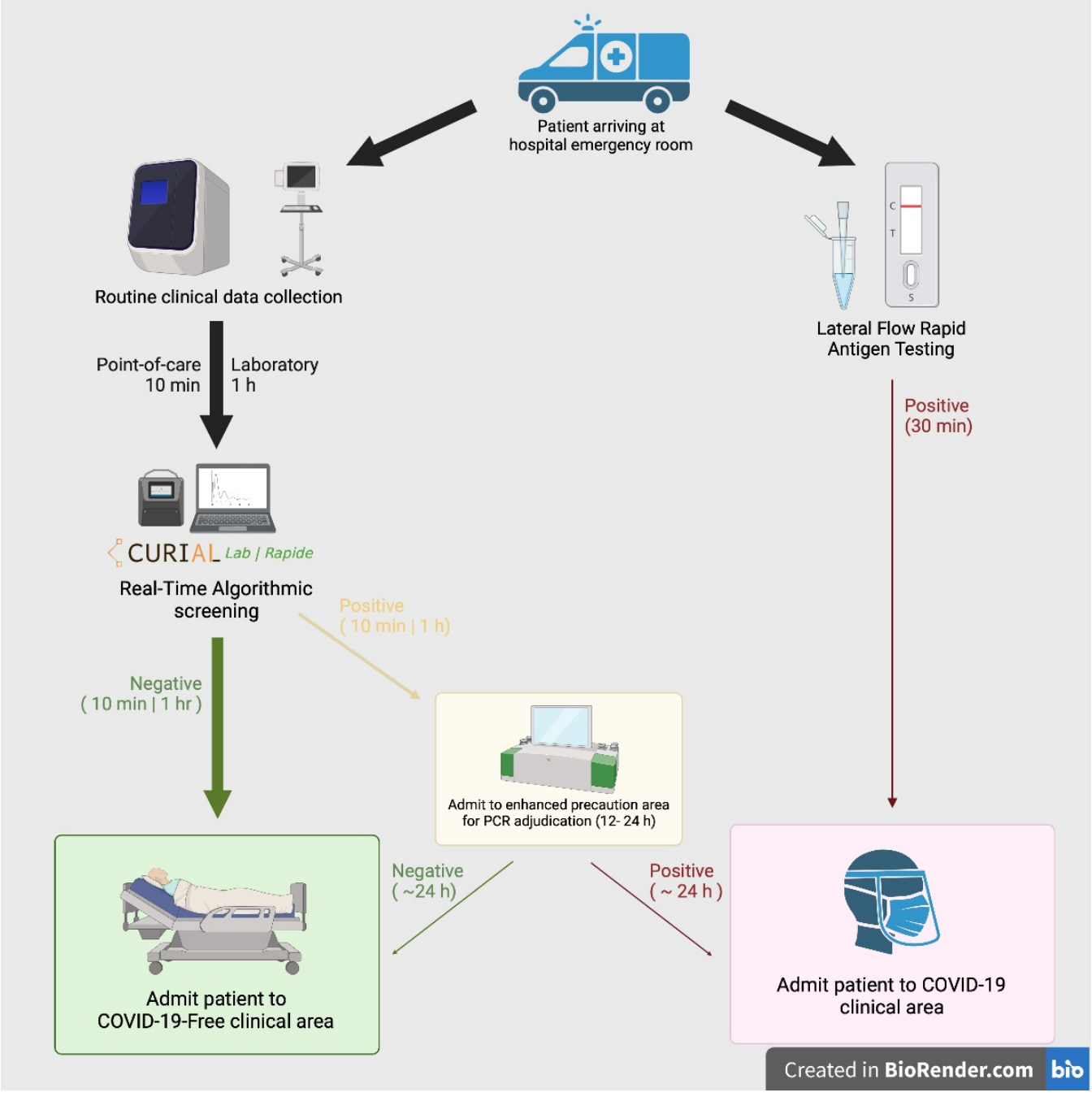


Figure 1: A rapid COVID-19 screening pathway for Emergency Department arrivals (a) routine blood tests and vital signs recordings are performed on patients' arrival to the emergency department, either using rapid point-of-care haematology analysers (~10 minutes; CURIAL-Rapide), or the existing laboratory analysis (~1h; CURIAL-Lab). (b) Real-time algorithmic analysis of blood tests & vital signs allows early, high-confidence identification of negative patients for safe streaming to COVID-19-Free clinical areas. (c) Patients with positive screening test results are admitted to enhanced precautions (amber) areas, pending confirmatory PCR result. (d) Patients testing positive via Lateral Flow Test can be streamed directly to COVID-19 ('red') clinical areas. Arrow thickness represents patient flow.

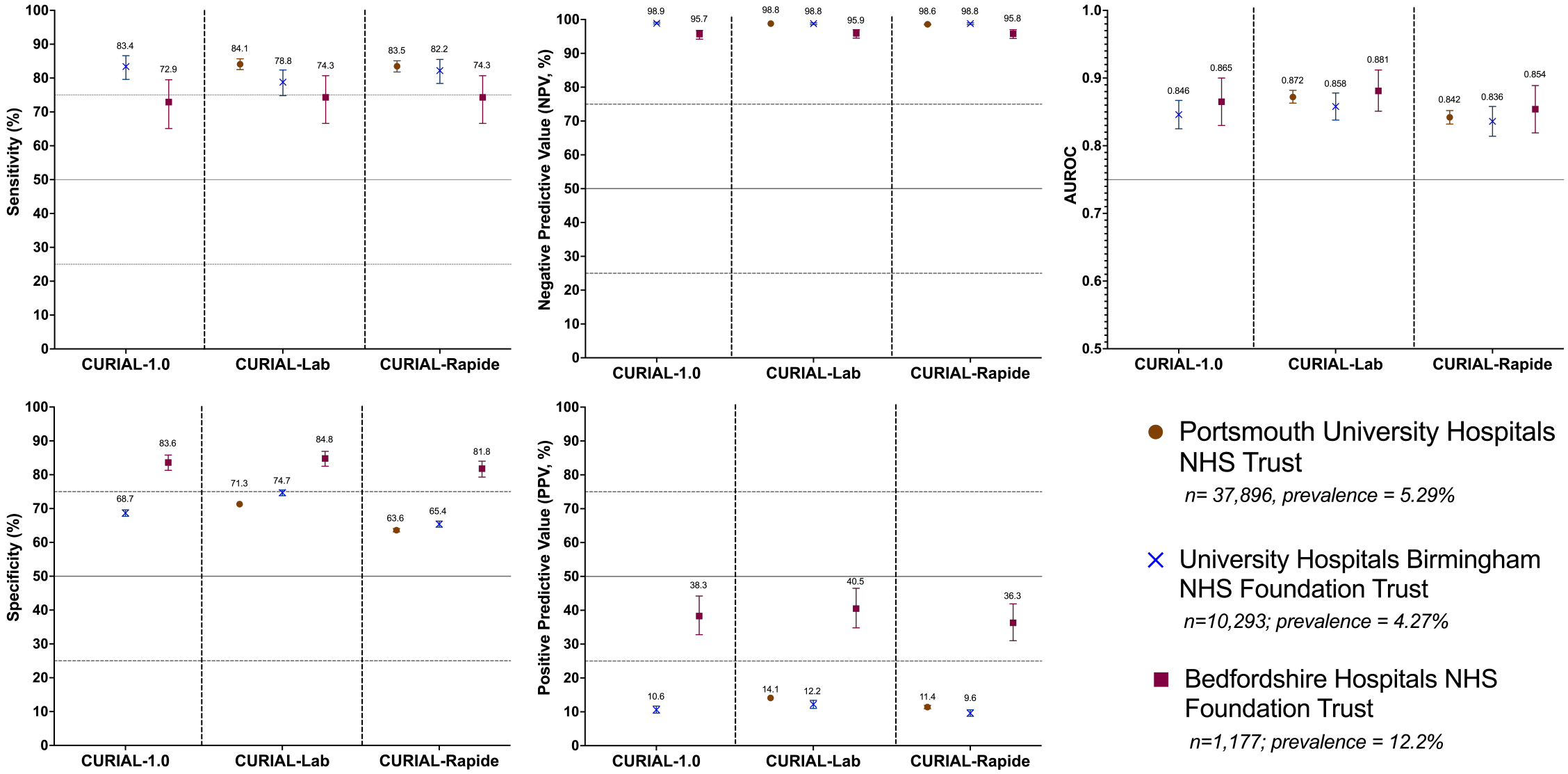


Figure 2: Performance of CURIAL-1.0, CURIAL-Lab & CURIAL-Rapide during external validation at three independent UK Hospitals trusts. All models were calibrated during training to achieve 90% sensitivity. Error bars show 95% confidence intervals. Numerical results are shown in Supplementary Table S5.

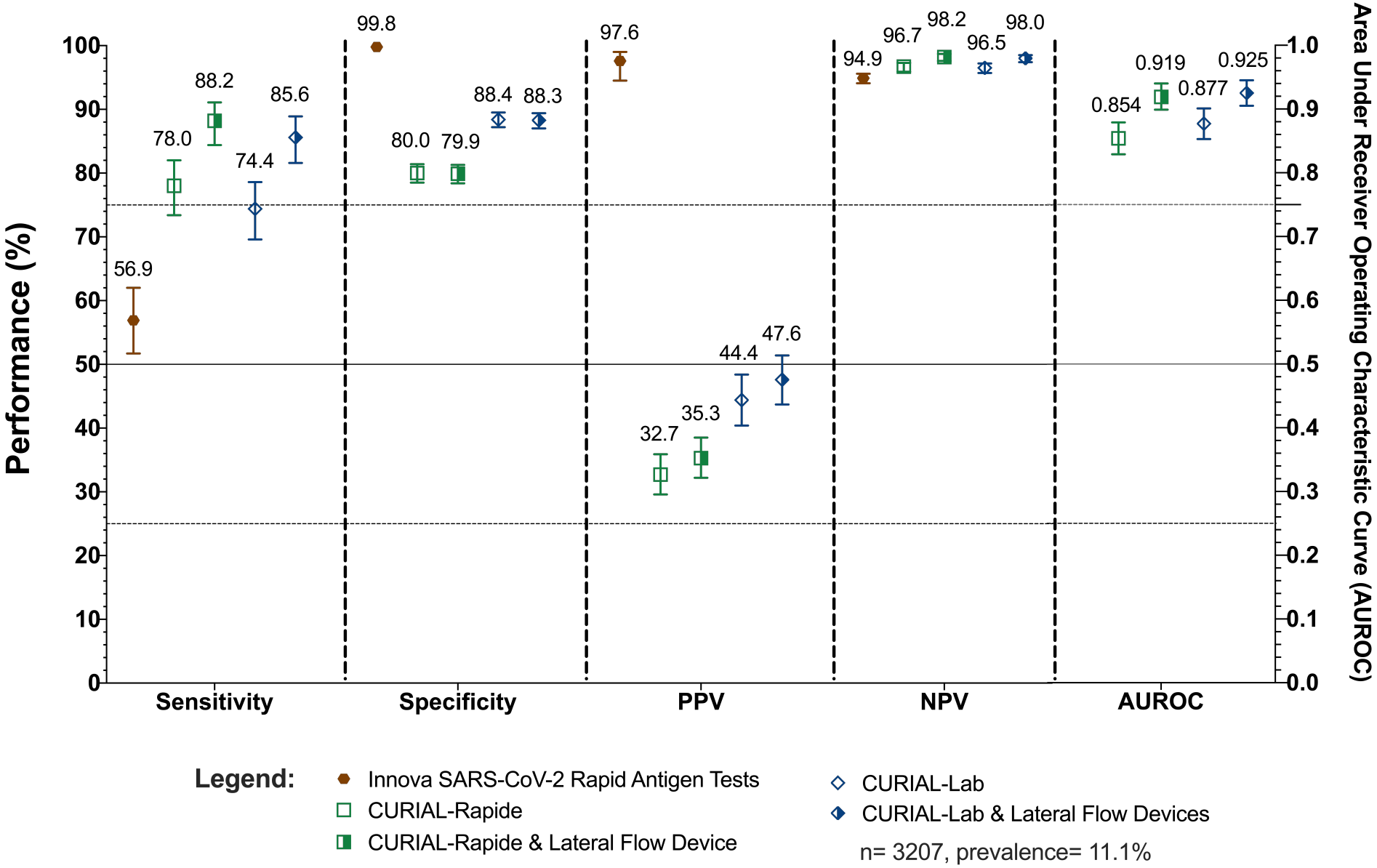


Figure 3: Performance characteristics of (a) INNOVA SARS-CoV2 Rapid Antigen Tests, (b) CURIAL-Rapide & CURIAL-Lab, calibrated during training to a sensitivity of 80%, and (c) combined clinical pathways considering either a positive CURIAL-Rapide/CURIAL-Lab result or a positive LFD test as a COVID-19 suspected case, at Oxford University Hospitals NHS Foundation Trust between December 23, 2020 & March 6, 2021. Error bars show 95% confidence intervals. Numerical results are shown in Supplementary Table S6.

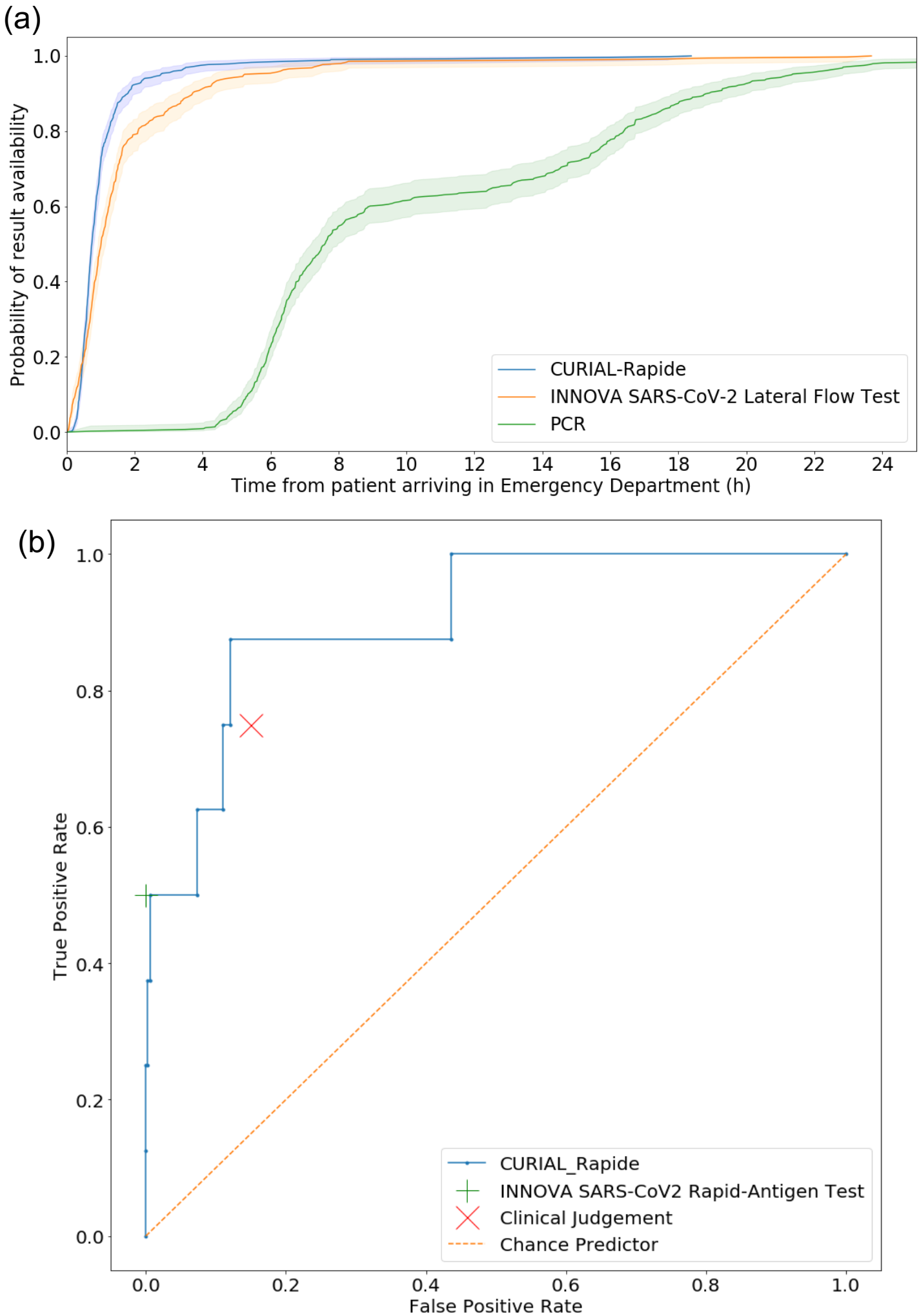


Figure 4: (a) Kaplan-Meier plots of time-to-result in hours from patient arrival in the Emergency Department for (i) CURIAL-Rapide, (ii) INNOVA SARS-CoV-2 Rapid-Antigen Tests and (iii) PCR swabs tests (Numbers at risk: 520, 348 and 436 respectively). CURIAL-Rapide results were available sooner than LFD testing (log rank test, $p<0.0001$) and PCR test results ($p<0.001$). (b) Receiver operating characteristic curve showing performance of (i) CURIAL-Rapide (ii) clinical judgement of the first-attending clinician, and (iii) INNOVA SARS-CoV-2 rapid antigen testing, against a PCR reference standard.