

1 Research note

2

3 **Performance of an artificial intelligence-based smartphone app for guided reading**
4 **of SARS-CoV-2 lateral-flow immunoassays**

5

6 David Bermejo-Peláez^{*1}

7 Daniel Marcos-Mencía^{*2}

8 Elisa Álamo¹

9 Nuria Pérez-Panizo³

10 Adriana Mousa¹

11 Elena Dacal¹

12 Lin Lin^{1,4}

13 Alexander Vladimirov¹

14 Daniel Cuadrado¹

15 Jesús Mateos-Nozal³

16 Juan Carlos Galán^{2,5}

17 Beatriz Romero-Hernández^{2,5}

18 Rafael Cantón^{2,6}

19 Miguel Luengo-Oroz^{#1}

20 Mario Rodríguez-Domínguez^{#2,5}

21

22 ^{*,#} Contributed equally

23

24 ¹Spotlab. Madrid, Spain

25 ²Servicio de Microbiología. Hospital Universitario Ramón y Cajal and Instituto Ramón
26 y Cajal de Investigación Sanitaria (IRYCIS). Madrid, Spain

27 ³Servicio de Geriátría. Hospital Universitario Ramón y Cajal and Instituto Ramón y
28 Cajal de Investigación Sanitaria (IRYCIS). Madrid, Spain

29 ⁴Biomedical Image Technologies, ETSI Telecomunicación, Universidad Politécnica de
30 Madrid & CIBER-BBN. Madrid, Spain

31 ⁵CIBER en Epidemiología y Salud Pública (CIBERESP). Instituto de Salud Carlos III.
32 Madrid, Spain

33 ⁶CIBER en Enfermedades Infecciosas (CIBERINFEC). Instituto de Salud Carlos III.
34 Madrid, Spain

35

36

37

38 Corresponding author: Miguel Luengo-Oroz

39 miguel@spotlab.ai

40 Spotlab

41 Paseo Juan XXIII, 36B

42 28040, Madrid, Spain

1 **Abstract**

2 **Objectives:** To evaluate an artificial intelligence-based smartphone application to
3 automatically and objectively read rapid diagnostic test (RDT) results and assess its
4 impact on COVID-19 pandemic management.

5 **Methods:** Overall, 252 human sera from individuals with PCR-positive SARS-CoV-2
6 infection were used to inoculate a total of 1165 RDTs for training and validation
7 purposes. We then conducted two field studies to assess the performance on real-world
8 scenarios by testing 172 antibody RDTs at two nursing homes and 92 antigen RDTs at
9 one hospital emergency department.

10 **Results:** Field studies demonstrated high levels of sensitivity (100%) and specificity
11 (94.4%, CI 92.8-96.1%) for reading IgG band of COVID-19 antibodies RDTs compared
12 to visual readings from health workers. Sensitivity of detecting IgM test bands was
13 100% and specificity was 95.8%, CI 94.3-97.3%. All COVID-19 antigen RDTs were
14 correctly read by the app.

15 **Conclusions:** The proposed reading system is automatic, reducing variability and
16 uncertainty associated with RDTs interpretation and can be used to read different RDTs
17 brands. The platform can serve as a real time epidemiological tracking tool and
18 facilitate reporting of positive RDTs to relevant health authorities.

1 **Introduction**

2 To control COVID-19 pandemic, timely and accurate early-detection strategies of
3 SARS-CoV-2 infections have been critical to slow down the spread of the virus. The
4 use of rapid diagnostic tests (RDTs), both for detection of antibodies and antigens, has
5 contributed to improve COVID-19 testing capacity reducing costs of diagnosis and
6 allowing for fastest results (1). First, COVID-19 RDTs were intended to be used just by
7 professional healthworkers, who have extensive experience in the use of this tool for
8 different infectious diseases (2, 3). Later, multiple Health Ministries have approved
9 home testing kits improving the accessibility to testing and taking pressure off from
10 health institutions. Nevertheless, self-testing strategies have some limitations, the
11 general population is not familiar with the use of RDTs and a minimum training is
12 needed for sampling, testing and result interpretation. Furthermore, as it has been seen
13 during the latest Omicron wave (4), many results go unreported, impairing post-testing
14 counseling and epidemiological surveillance.

15 Combining RDTs with digital tools, artificial intelligence (AI) and mobile health
16 approaches can help standardize result interpretation and facilitate immediate reporting
17 and monitoring of results (5). The use of AI to automatically interpret photographs of
18 RDTs has been also recently proposed (6, 7). Here, we describe the development and
19 field validation of a mobile-based tool for reading and reporting multiple types of
20 SARS-COV-2 RDTs which is connected to a real-time epidemiological monitoring web
21 platform.

22

23 **Methods**

24 This study was divided into two phases. First, the training and validation of an AI
25 algorithm for the automatic interpretation of RDTs. Second, two field studies to assess

1 the performance of the AI-based system for reading both COVID-19 antibodies and
2 antigen RDTs in real-world scenarios. Ethics approval for the study was obtained from
3 the Clinical Research Ethics Committee of the Ramón y Cajal University Hospital (Ref.
4 127/21).

5

6 *Algorithm Training and validation dataset*

7 For generating the training image dataset, twelve human sera from SARS-CoV-2
8 positive PCR patients with a positive ELISA test (Vircell Spain S.L.U., Granada, Spain)
9 were used. Each serum sample was serially diluted until it reached a negative result
10 when inoculated in a COVID-Ab test. Each dilution was tested in three replicates for
11 each of the three brands tested (2019-nCoV IgG/IgM Rapid Test Cassette (Hangzhou
12 AllTest Biotech Co., Ltd.); Panbio COVID-19 IgG/IgM Rapid Test Device (Abbott);
13 UNscience COVID-19 IgG/IgM Rapid Test (Wuhan UNscience Biotechnology Co.,
14 Ltd.), resulting in 433 RDTs inoculated (61 IgG+/IgM+; 166 IgG+/IgM-; 43 IgG-
15 /IgM+; 164 negative). Additionally, twelve COVID-Ag RDTs (Panbio COVID-19 Ag
16 Rapid Test Device (Abbott) (6+,6-) were also included to train the algorithm to read not
17 only three-band tests (such as the COVID-19 antibody tests used in this study) but also
18 two-band RDTs such as COVID-19 antigen tests. The entire training dataset consisted
19 of 3614 images.

20 For collecting the independent validation dataset, 240 human sera samples
21 independent from the ones used for training were used to inoculate 720 COVID-Ab
22 RDTs (each serum was tested in triplicate using the aforementioned brands). The
23 samples were selected ensuring all possible results are well represented along the
24 dataset (108 IgG+/IgM+; 321 IgG+/IgM-; 27 IgG-/IgM+; 264 negative).

1 Each RDT was visually read by multiple observers (from 3 to 5) and the ground truth
2 was established as the majority result from the total of analyzers. Each inoculated RDT
3 was digitized by using the TiraSpot mobile app (Spotlab, Madrid, Spain) for guided and
4 standardized acquisition ensuring correct positioning of RDTs in the image and using a
5 total of 9 smartphone models. Results data was uploaded to a web platform.

6

7 *Field validation studies*

8 The workflow for the field studies was as follows: a health professional digitized the
9 RDTs by using the app, was asked for recording the visual interpretation of the test
10 result, images were uploaded to the telemedicine platform and processed by the AI
11 algorithm, and discrepancies between the interpretation made by the health professional
12 and that obtained by the algorithm were subsequently reviewed by an external health
13 professional through the platform.

14 The first field study used the system as part of a seroprevalence study conducted in
15 two nursing homes in Madrid (Spain). A total of 172 vaccinated health care personnel
16 were included in this study from which a finger-prick blood sample was taken and
17 inoculated into SARS-CoV-2 Rapid Antibody Test (Roche). A trained nurse digitized
18 the RDTs and record their result using the application.

19 The second field validation field study tested the system to read also COVID-19
20 antigen tests composed of two bands (Panbio COVID-19 Ag Rapid Test Device,
21 Abbot). This study was carried out at the Emergency Department of the Ramón y Cajal
22 Hospital (Madrid, Spain), where 92 individuals' nasal swabs were inoculated in antigen
23 tests, and digitized by experienced health professionals using the app.

24 All images were acquired in very diverse real-world conditions (including different
25 environmental illuminations and shades).

1

2 **Results**

3 *AI algorithm training and app validation*

4 All images acquired with the app were uploaded to a cloud platform where the AI
5 algorithm processed the photographs to predict its result interpretation. As shown in
6 **Table 1A**, when comparing the visual interpretations (used as ground truth) against the
7 AI algorithm, the performance was high for all brands of RDTs tested, obtaining a mean
8 sensitivity and specificity of 98% and 100% respectively for detecting the IgG band and
9 a mean sensitivity and specificity of 80% and 89% for the detection of the IgM band.

10

11 *Validation in real-world scenarios*

12 From the 172 RDTs used in this study (5 IgG+/IgM+; 149 IgG+/IgM-; 1 IgG-/IgM+; 17
13 negative), we only found 9 discrepancies between test result interpretation made by the
14 health professional and the AI algorithm. From these 9 cases, two of them were
15 incorrectly classified by the algorithm due to an incorrect image acquisition with the
16 app. The remaining discrepant cases were further reviewed by a second professional,
17 and the AI-based system allowed to detect and modify the result with respect to the
18 initial healthworker interpretation in 4 cases by confirming the result predicted by the
19 algorithm.

20 The overall performance of the algorithm with respect to the ground truth is shown in
21 **Table 1B**. It should be noted that the performance of the system is high even when used
22 with a RDT different from those used for training the algorithm, suggesting its potential
23 use with any RDT on the market. The slight disparity in the performance of IgM band
24 identification in antibody RDTs between the validation set and this field study may be

1 explained by the presence of very faint signals which were almost invisible in the
2 photographs.

3 Regarding the second field study for reading COVID-19 antigen RDTs, we found
4 that all tests used and digitized using the TiraSpot app (58 negative, 30 positive) were
5 correctly interpreted by the proposed system (**Table 1C**), demonstrating that the system
6 can also be applied for reading two-band (control and test) as well as three-band tests
7 (IgG, IgM and control).

8

9 **Discussion**

10 We describe the usefulness of an app for reading and result interpretation of lateral-flow
11 RDTs for SARS-CoV-2 testing. Additionally, results are sent to a telemedicine platform
12 which allows for case identification and confirmation, quality control and real-time
13 monitoring.

14 Our AI algorithm demonstrates excellent performance, especially in prospective
15 validation in real-life scenarios and for both antibodies and antigen detection tests. The
16 algorithm performed as well in RDTs brands which were not used at all for training
17 purposes, making the solution suitable for other RDTs, including other diseases.
18 Compared with previous studies (6, 7), our system is able to identify individual bands of
19 the RDTs allowing complex results reading and sending them in real-time to a cloud
20 platform. A requirement and limitation of the proposed system is the correct acquisition
21 of the image (acquisition error in the field studies <0.8%).

22 In conclusion, the use of TiraSpot (**Figure 1**) is a useful tool for reporting, real-time
23 monitoring and quality control, as the results can be reviewed by specialists when
24 needed. This is especially important in contexts where massive testing is to be done and
25 the likelihood of subjectivity and errors in the interpretation of the result is higher. It is

- 1 also important in the validation of self-diagnostic tests performed by untrained users, as
- 2 it avoids the loss of information in case it is not notified by the user and provides an
- 3 efficient system to confirm and report data, which has been a key challenge during the
- 4 Omicron wave (4).

REFERENCES

1. European Centre for Disease Prevention and Control. Options for the use of rapid antigen detection tests for COVID-19 in the EU/EEA – first update, 26 October 2021. Stockholm: ECDC; 2021.
2. Weber NC, Klepser ME, Akers JM, Klepser DG, Adams AJ. Use of CLIA-waived point-of-care tests for infectious diseases in community pharmacies in the United States. *Expert Rev Mol Diagn.* 2016;16(2):253-64. doi: 10.1586/14737159.2015.1116388. Epub 2015 Nov 27. PMID: 26560318.
3. Bell SFE, Lemoire J, Debattista J, Redmond AM, Driver G, Durkin I, Coffey L, Warner M, Howard C, Williams OD, Gilks CF, Dean JA. Online HIV Self-Testing (HIVST) Dissemination by an Australian Community Peer HIV Organisation: A Scalable Way to Increase Access to Testing, Particularly for Suboptimal Testers. *International Journal of Environmental Research and Public Health.* 2021; 18(21):11252. <https://doi.org/10.3390/ijerph182111252>
4. Tegally H, Wilkinson E, Giovanetti M, Iranzadeh A, Fonseca V, Giandhari J, Doolabh D, Pillay S, San EJ, Msomi N, Mlisana K, von Gottberg A, Walaza S, Allam M, Ismail A, Mohale T, Glass AJ, Engelbrecht S, Van Zyl G, Preiser W, Petruccione F, Sigal A, Hardie D, Marais G, Hsiao NY, Korsman S, Davies MA, Tyers L, Mudau I, York D, Maslo C, Goedhals D, Abrahams S, Laguda-Akingba O, Alisoltani-Dehkordi A, Godzik A, Wibmer CK, Sewell BT, Lourenço J, Alcantara LCJ, Kosakovsky Pond SL, Weaver S, Martin D, Lessells RJ, Bhiman JN, Williamson C, de Oliveira T. Detection of a SARS-CoV-2 variant of concern in South Africa. *Nature.* 2021 Apr; 592(7854):438-443. doi: 10.1038/s41586-021-03402-9. Epub 2021 Mar 9. PMID: 33690265.
5. Kadam R, White W, Banks N, Katz Z, Dittrich S, Kelly-Cirino C. Target Product Profile for a mobile app to read rapid diagnostic tests to strengthen infectious disease surveillance. *PLoS One.* 2020 Jan 29; 15(1):e0228311. doi: 10.1371/journal.pone.0228311. PMID: 31995628; PMCID: PMC6988927.
6. Turbé, V., Herbst, C., Mngomezulu, T. *et al.* Deep learning of HIV field-based rapid tests. *Nat Med* **27**, 1165–1170 (2021). <https://doi.org/10.1038/s41591-021-01384-9>
7. David-A. Mendels, Laurent Dortet and Cécile Emeraud *et al.* Using artificial intelligence to improve COVID-19 rapid diagnostic test result interpretation. *Proceedings of the National Academy of Sciences of the United States of America.* Vol. 118(12). DOI: 10.1073/pnas.2019893118

1 **Table 1.** Performance of the AI algorithm for predicting RDTs results with respect to
 2 human visual reading in the validation set (**A**), in the field study for reading antibodies
 3 RDTs (Ab) (**B**) and in the field study when reading antigen RDTs (Ag) (**C**).
 4

| A | Model (Ab) | Band | AUC [95% CI] | SN [95% CI] | SP [95% CI] | N |
|----------|------------|------|------------------|------------------|------------------|------------|
| | Abbott | IgG | 99.5 [98.7,100] | 96.4 [94.1,98.8] | 100 [100,100] | 94-, 145+ |
| | | IgM | 92.5 [85.4,99.6] | 80.8 [75.8,85.8] | 90.7 [87.0,94.3] | 184-, 55+ |
| | UNScience | IgG | 100 [100,100] | 100 [100,100] | 100 [100,100] | 100-, 140+ |
| | | IgM | 89.5 [83.7,95.2] | 80.0 [74.9,85.1] | 88.6 [84.6,92.6] | 214-, 26+ |
| | AllTest | IgG | 99.8 [99.4,1] | 97.9 [96.1,99.7] | 100 [100,100] | 96-, 144+ |
| | | IgM | 90.6 [85.0,96.1] | 79.6 [74.5,84.7] | 86.0 [81.6,90.4] | 186-, 54+ |
| | Global | IgG | 99.8 [99.5,100] | 98.1 [97.1,99.1] | 100 [100,100] | 290-, 429+ |
| | | IgM | 90.8 [87.4,94.3] | 80.0 [77.1,82.9] | 89.0 [86.2,90.9] | 584-, 135+ |
| B | Model (Ab) | Band | AUC [95% CI] | SN [95% CI] | SP [95% CI] | N |
| | Roche | IgG | 100 [100,100] | 100 [100,100] | 94.4 [92.8,96.1] | 18-, 154+ |
| | | IgM | 99.6 [96.0,100] | 100 [100,100] | 95.8 [94.3,97.3] | 166-, 6+ |
| C | Model (Ag) | Band | AUC [95% CI] | SN [95% CI] | SP [95% CI] | N |
| | Abbott | NA | 100 [100,100] | 100 [100,100] | 100 [100,100] | 58-, 30+ |

Figure 1. Tiraspot system is composed of (1) a mobile app for test digitization and result recording, (2) an AI model for RDT result interpretation and (3) a web platform where all collected data can be visualized and allows for result corrections in the cases in which a discrepancy exists between AI and user interpretation.

