Letter to the Editor

Claudio Galli and Mario Plebani*

SARS-CoV-2 antibody performances: we need better criteria

https://doi.org/10.1515/cclm-2020-1358
Received May 9, 2018; accepted August 31, 2020; published online September 25, 2020

Keywords: sensitivity; specificity; motility; SARS-CoV-2 antibody.

To the Editor,

Unlike some reports suggesting that testing for severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) RNA (or antigen) alone will be sufficient to track and contain the ongoing pandemic, recently available data provide trustable evidence on the highly cost-effective usefulness of antibody assays [1]. Gudbjartsson and colleagues used two highly sensitive and specific immunoassays, and provided new insights on both the high rate of people infected with SARS-CoV-2 and not diagnosed by quantitative polymerase-chain-reaction (qPCR), that was found to be as high as 44% in Iceland, along with the duration of humoral immunity, as specific antibodies were found to remain stable for over 4 months after diagnosis [2].

These new available information highlights the essential need for accurate validation and evaluation of analytical and diagnostic performance specifications of SARS-CoV-2 immunoassays, especially concerning their specificity and sensitivity.

On June 25th, 2020 the Public Health England (PHE) released a report comparing the sensitivity and specificity of four automated assays for SARS-CoV-2 antibodies on 546 samples deemed as positive and 994 considered as negative for SARS-CoV-2 infection [3]. The conclusion was that all assays met the specificity criterion and just one met the sensitivity criterion previously established (98% for both). One month later, the World Health Organization (WHO) issued the target product profiles for COVID-19 diagnostics, indicating 95–97% sensitivity and 98–99% specificity as acceptable and desirable criteria for sensitivity and specificity, respectively, for tests aimed to fulfill moderate to high volume needs [4].

Those criteria deserve some comments. First, the 98% threshold for specificity seems conservative. The evaluation of assay specificity is conceptually simple relying on testing samples taken before the start of COVID-19 pandemic plus potentially interfering conditions such as other infectious diseases (including infections by other Coronaviruses). A 98% specificity is hardly conceivable with the use of any antibody assay for population screening, as will cause the positive predictive value (PPV) to deep dive and results to bring more harm than good, both for individuals and tested cohorts. In the first broad open population study available, carried out in Spain on over 61,000 individuals selected by the National Institute of Statistic to represent the whole country [5], the overall prevalence of disease was found to be 5% by a rapid assay and 4.6% using an automated method, respectively. Therefore, the estimated PPV for an assay yielding 98% specificity would have been slightly more than 71% (Table 1). The laboratory assay employed in that study (Abbott ARCHITECT SARS CoV-2 IgG) has a declared specificity of 99.6%, confirmed also in a prior external evaluation [6] and in the PHE study [3], and it is hence likely that the PPV in that cohort would have been very high.

On the other hand, concerns have been raised on assay sensitivity as in the PHE study all tests except one were quite distant from values declared in the respective inserts. Since this may hardly be attributable to assay performance, the problem shall be identified with the reference, highlighting at least three major biases. First, no international standard has been yet validated for SARS-CoV-2 antibodies [4], neither there is sufficient information on the natural history of this new infection for establishing widespread consensus on which target may be "measured". Then, relying on the period after symptoms onset or after a first positive result for SARS-CoV-2 RNA – as

***Corresponding author: Mario Plebani**, Professor, Department of Medicine-DIMED, University of Padova, Padova, Italy; and Department of Laboratory Medicine, University-Hospital of Padova, via Giustiniani 2, 35100 Padova, Italy, E-mail: mario.plebani@unipd.it. https://orcid.org/0000-0002-0270-1711
Claudio Galli, Medical & Scientific Affairs, Abbott Diagnostics, Rome, Italy

Table 1: Simulated performances of SARS-CoV-2 antibody assays on 50,000 subjects at two different prevalence estimates.

Prev.	Ref.	Sens.	Spec.	TP	FN	TN	FP	NPV	PPV
2%	WHO	95%	98%	950	50	48020	980	99.9%	49.2%
2%	WHO	95%	99%	950	50	48510	490	99.9%	66.0%
2%	WHO	97%	98%	970	30	48020	980	99.9%	49.2%
2%	WHO	97%	99%	970	30	48510	490	99.9%	66.0%
2%	PHE1	92.7%	99.9%	927	73	48951	49	99.8%	95.0%
2%	PHE2	90.2%	99.4%	902	98	48706	294	99.8%	75.4%
5%	WHO	95%	98%	2375	125	46550	950	99.7%	71.4%
5%	WHO	95%	99%	2375	125	47025	475	99.7%	83.3%
5%	WHO	97%	98%	2425	75	46550	950	99.8%	71.8%
5%	WHO	97%	99%	2425	75	47025	475	99.8%	83.6%
5%	PHE1	92.7%	99.9%	2317	183	47452	48	99.6%	98.0%
5%	PHE2	90.2%	99.4%	2255	245	47215	285	99.5%	88.8%

Reference sensitivity and specificity are indicated according to the recent World Health Organization (WHO) criteria for acceptable (95% sensitivity and 98% specificity) and ideal (97% sensitivity and 99% specificity) performance or based on the results for the Abbott ARCHITECT SARS-CoV-2 IgG assay obtained by Public Health England (PHE). On the latter, PHE1 indicates the estimated value and PHE2 the lowest 95% limit of confidence. Prev., prevalence; Ref., reference; Sens., sensitivity; Spec., specificity; TP, true positive; FN, false negative; TN, true negative; FP, false positive; NPV, negative predictive value; PPV, positive predictive value.

done in PHE and many other studies – would introduce further confounders. Symptoms are mostly self-declared and may be either under or overestimated by patients and physicians, whilst diagnostic testing may even be delayed for several days for medical reasons (e.g., few symptoms, late presentation) or basic organizational issues (assay availability, centralized testing, laboratory organization). Finally, there is clear evidence on the direct relationship between time to antibody detection (and titers) and clinical severity. Patients with severe/critical disease typically display earlier and stronger humoral response, whilst antibodies are usually detected at a later stage (or may even remain undetectable) in patients mild symptoms or in those who remain totally asymptomatic [7]. The options to assess antibody response to SARS-CoV-2 shall also be considered because is a choice between utilizing target antigens expressed by spike protein (S), the receptor binding domain (RBD) of the S1 subunit of the spike protein or nucleocapsid (N) genome regions. Either will bring some advantages, since N-specific antibodies usually appear earlier or at higher levels [8] whilst S-specific antibodies have been shown to last longer after a primary infection [9]. However, memory B cells shall produce high affinity antibodies able to confer protective immunity and exhibit a lower decay. This will be especially relevant in the so-called "phase 2" of the outbreak, when recurrent infections may occur, and considerable differences in antibody kinetics between phases 1 and 2 will influence calculations on antibody persistence [1].

SARS-CoV-2 antibody assays may be used along with molecular methods detecting SARS-CoV-2 RNA for diagnostic purposes [2] and specificity shall be regarded as the most relevant aspect, since an integration of anamnestic, epidemiological, clinical and imaging data will allow to reach an accurate diagnosis even when antibodies testing remains negative. When screening open populations or any group of people at low risk, pre-test probability will be low and assay sensitivity ranging between 90 and 98% will lead to very close values for negative predictive values (Table 1). At the same time, every 1% drop in specificity will lead to a huge decrease in positive predictive value. Prevalence estimates may then be offset. On an individual perspective a false positive test result may potentially place that person at risk of becoming infected with SARS-CoV-2, so that the risk of generating false test results positives shall be minimized [10, 11].

SARS-CoV-2 serology remains a challenging enterprise [12]. Sensitivity and specificity are intrinsic characteristics of any diagnostic method, so that their evaluation shall be based on clear evidences endorsed by accepted consensus. If those evidences are not consistent, as it happens for the sensitivity of SARS-CoV-2 serological assays so far, the acceptable values for each device shall be linked to the intended use and the means used for evaluating and eventually using test results shall be established beforehand to maximize their clinical utility.

Research funding: None declared.

Author contributions: All authors have accepted responsibility for the entire content of this manuscript and approved its submission.
Competing interests: Authors state no conflict of interest.

References

1. Alter G, Seder R. The power of antibody-based surveillance. N Engl J Med 2020. https://doi.org/10.1056/NEJMe2028079 [Epub ahead of print].
2. Gudbjartsson DF, Norddahl GL, Melsted P, Gunnarsdottir K, Holm H, Eythorsson E, et al. Humoral immune response to SARS-CoV-2 in Iceland. N Engl J Med 2020. https://doi.org/10.1056/NEJMoa2026116 [Epub ahead of print].
3. Public Health England. Evaluation of sensitivity and specificity of four commercially available SARS-CoV-2 antibody immunoassays; 2020. Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/898437/Evaluation__of_sensitivity_and_specificity_of_4_commercially_available_SARS-CoV-2_antibody_immunoassays.pdf [Accessed 14 Aug 2020].
4. World Health Organization. COVID-19 Target product profiles for priority diagnostics to support response to the COVID-19 pandemic v.0.1; 2020. Available from: https://www.who.int/publications/m/item/covid-19-target-product-profiles-for-priority-diagnostics-to-support-response-to-the-covid-19-pandemic-v.0.1 [Accessed 14 Aug 2020].
5. Pollán M, Pérez-Gómez B, Pastor-Barriuso R, Oteo J, Hernán MA, Pérez-Olmeda M, et al. Prevalence of SARS-CoV-2 in Spain (ENE-COVID): a nationwide, population-based seroepidemiological study. Lancet 2020. https://doi.org/10.1016/S0140-6736(20)31483-5 [Epub ahead of print].
6. Bryan A, Pepper G, Wener MH, Fink SL, Morishima C, Chaudhary A, et al. Performance characteristics of the Abbott Architect SARS-CoV-2 IgG assay and seroprevalence in Boise, Idaho. J Clin Microbiol 2020:e00941-20. https://doi.org/10.1128/JCM.00941-20.
7. Long Q-X, Tang X-J, Shi Q-L, Li Q, Deng H-J, Yuan J, et al. Clinical and immunological assessment of asymptomatic SARS-CoV-2 infections. Nature Med 2020. https://doi.org/10.1038/s41591-020-0965-6 [Epub ahead of print].
8. Van Elslande J, Decru B, Jonckheere S, Van Wijngaerden E, Houben E, Vandecandelaere P, et al. Antibody response against SARS-CoV-2 spike protein and nucleoprotein evaluated by 4 automated immunoassays and 3 ELISAs. Clin Microbiol Infect 2020. https://doi.org/10.1016/j.cmi.2020.07.038 [Epub ahead of print].
9. Seow J, Graham C, Merrick B, Acors S, Steel KJA, Hemmings O, et al. Longitudinal evaluation and decline of antibody responses in SARS-CoV-2 infection. medRxiv preprint 2020. https://doi.org/10.1101/2020.07.09.20148429 [Epub ahead of print].
10. Plebani M, Padoan A, Negrini D, Carpinteri B, Sciacovelli L. Diagnostic performances and thresholds: the key to harmonization in serological SARS-CoV-2 assays? Clin Chim Acta 2020;509:1–7.
11. Plebani M, Padoan A, Sciacovelli L, Basso D. Towards the rational utilization of SARS-CoV-2 serological tests in clinical practice. Clin Chem Lab Med 2020;58:e189–91.
12. Lippi G, Plebani M. SARS-CoV-2 antibodies titration. Ann Transl Med 2020;8:1032.