















BMJ Open Validation of a rapid SARS-CoV-2 antibody test in general practice

Julie Domen ¹, Jan Yvan Jos Verbakel ^{2,3}, Niels Adriaenssens ¹,
Beatrice Scholtes ⁴, Bart Peeters,⁵ Robin Bruyndonckx ^{1,6,7},
An De Sutter ⁸, Stefan Heytens ⁸, Ann Van den Bruel ³,
Isabelle Desombere ⁹, Pierre Van Damme ¹⁰, Herman Goossens ¹¹,
Laetitia Buret ⁴, Els Duysburgh ¹², Samuel Coenen ^{1,11}

To cite: Domen J, Verbakel JYJ, Adriaenssens N, *et al.* Validation of a rapid SARS-CoV-2 antibody test in general practice. *BMJ Open* 2023;**13**:e069997. doi:10.1136/bmjopen-2022-069997

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2022-069997>).

JD and JYJV are joint first authors.

Received 14 November 2022
Accepted 30 March 2023



© Author(s) (or their employer(s)) 2023. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

Correspondence to

Dr Julie Domen;
julie.domen@uantwerpen.be

ABSTRACT

Objectives To validate a rapid serological test (RST) for SARS-CoV-2 antibodies used in seroprevalence studies in healthcare providers, including primary healthcare providers (PHCPs) in Belgium.

Design A phase III validation study of the RST (OrientGene) within a prospective cohort study.

Setting Primary care in Belgium.

Participants Any general practitioner (GP) working in primary care in Belgium and any other PHCP from the same GP practice who physically manages patients were eligible in the seroprevalence study. For the validation study, all participants who tested positive (376) on the RST at the first testing timepoint (T1) and a random sample of those who tested negative (790) and unclear (24) were included.

Intervention At T2, 4 weeks later, PHCPs performed the RST with fingerprick blood (index test) immediately after providing a serum sample to be analysed for the presence of SARS-CoV-2 immunoglobulin G antibodies using a two-out-of-three assay (reference test).

Primary and secondary outcome measures The RST accuracy was estimated using inverse probability weighting to correct for missing reference test data, and considering unclear RST results as negative for the sensitivity and positive for the specificity. Using these conservative estimates, the true seroprevalence was estimated both for T2 and RST-based prevalence values found in a cohort study with PHCPs in Belgium.

Results 1073 paired tests (403 positive on the reference test) were included. A sensitivity of 73% (a specificity of 92%) was found considering unclear RST results as negative (positive). For an RST-based prevalence at T1 (13.9), T2 (24.9) and T7 (70.21), the true prevalence was estimated to be 9.1%, 25.9% and 95.7%, respectively.

Conclusion The RST sensitivity (73%) and specificity (92%) make an RST-based seroprevalence below (above) 23% overestimate (underestimate) the true seroprevalence.

Trial registration number NCT04779424.

INTRODUCTION

In 2020, the coronavirus SARS-CoV-2 emerged and spread throughout the world affecting the morbidity and mortality of millions of people. The COVID-19 has been causing a

STRENGTHS AND LIMITATIONS OF THIS STUDY

- ⇒ Phase III validation in a relevant target population, primary healthcare providers.
- ⇒ Simultaneous performance of both index and reference SARS-CoV-2 antibody test.
- ⇒ Careful interpretation of unclear index test results.
- ⇒ Inverse probability weighting to deal with missing reference test results by study design.
- ⇒ The index test's sensitivity and specificity depend on the study setting and, when used in settings with a higher test seroprevalence, underestimate the true seroprevalence.

pandemic for the past 3 years with several epidemic waves. In Belgium, the second wave, which started in autumn 2020, was responsible for almost the highest number of infectious cases per capita worldwide.¹ To monitor the pandemic typically PCR-confirmed cases are used. This method of surveillance is limited as mild and asymptomatic cases often do not reach the general practitioner (GP) or test centre. As a result, confirmed cases underestimate the true infection rate.²

Large-scale studies assessing the prevalence of antibodies against SARS-CoV-2 can be used to estimate exposure to the virus in a population as well as to monitor serological immunity to the virus after infection or vaccination. At the population level, seroprevalence studies give insights into the rate at which the virus has spread. They can guide policy making and timing of (booster) vaccination campaigns. Furthermore, they can be used to study the risk factors for SARS-CoV-2 infection.³

Estimating the seroprevalence among healthcare providers gives information on how the disease spreads in high-risk settings with many patient contacts, both symptomatic and asymptomatic.⁴⁻⁶ On the one hand, it identifies the burden of infections in healthcare providers, and on the other hand

it monitors the immunity of those at high risk for infection. The frequent interactions of primary healthcare providers (PHCP) with patients who might not have been diagnosed with COVID-19 because of their asymptomatic to mild infection make this setting of particular interest in this field.

Collecting samples from PHCPs, here GPs and other PHCPs within their practice, for seroprevalence studies is challenging since repeated samples from many PHCPs are needed, only few PHCPs work in the same practice, and the wide geographical spread of practices. As a result, the collection and analysis of venous samples on such a large scale is often not feasible.

Therefore, we used dried blood spots (DBS) and rapid serological tests (RST) in two consecutive prospective cohort studies to assess the prevalence of antibodies against SARS-CoV-2 among PHCPs in Belgium since the outbreak.^{4,7}

RSTs have been developed to identify the presence of antibodies against SARS-CoV-2 within 15 min. Compared with laboratory tests and DBS, a valid easy-to-use RST has the advantage of speeding up the availability of the test results and thus impacting clinical decision-making, lowering the burden on laboratories and eliminating the administrative barrier of returning samples. Cals and van Weert pointed out that a point-of-care test (POCT) in primary care should be valid, reliable, robust, easy to use and able to be interpreted correctly before its use is efficient.⁸ Although patient management will not rely on the results of a single RST, one can assume these criteria should also apply to RSTs in order for policy makers to confidently rely on up-to-date seroprevalence data.

Sciensano, the Belgian institute for public health, has validated five RSTs using samples from SARS-CoV-2 positive and negative cases, confirmed by combined reverse transcription-quantitative PCR and immunoassay positivity/negativity. Performance characteristics of these RSTs using fingerprick blood were also compared with the performance in serum. They identified one test, the OrientGene RST, with appropriate sensitivity (92.9%) and specificity (96.3%) for use in seroprevalence studies. The OrientGene had the highest overall percentage agreement with testing on serum (95.9%), and was therefore considered to be reliable outside optimal laboratory conditions.⁹ This lateral flow test contains colloidal gold conjugated to the SARS-CoV-2 Spike S1 protein, and targets immunoglobulin M (IgM) and/or IgG antibodies to SARS-CoV-2 when present.¹⁰ We used the OrientGene RST for our second seroprevalence study among PHCPs.⁴

A few laboratory validation studies with the OrientGene RST have been performed in the mean time, with a combined sensitivity for IgM and IgG ranging from 93.8% to 100% and a specificity ranging from 97.5% to 98.5%.^{11,12} Both studies treated PCR-positive participants as cases and prepandemic samples as controls. These laboratory studies are limited by highly selected samples, unrealistic nature of laboratory conditions and use of serum and plasma samples instead of fingerprick blood.

External validation is recommended for a better interpretation of the findings from seroprevalence studies using RST in real-world conditions.¹³ Therefore, we investigated the accuracy of this RST in an independent population, that is, when performed by GPs with fingerprick blood as part of a SARS-CoV-2 seroprevalence study among PHCPs in Belgium.⁴

METHODS

Study design

This is a phase III validation study,¹⁴ a large-scale prospective study which validates a test in the target population.^{14,15} Participants were enrolled on the basis of their result on the OrientGene RST at the first testing timepoint (T1) (24 December 2020; 8 January 2021) of a prospective cohort study assessing the seroprevalence of SARS-CoV-2 in healthcare providers in Belgium.⁴

Participants

Any GP working in primary care in Belgium (including those in training) and any PHCP from the same practice who physically manages (examines, tests, treats) patients were eligible for the prospective cohort study. They were invited to register online for the study and were asked to invite the other PHCPs in their practice to do the same. Online registration was available between 15 November 2020 and 15 January 2021. Information about the study was disseminated to GPs and PHCPs via professional organisations (Domus Medica and Collège de Médecine Générale), university networks and through professional media channels. This convenience sample was checked for geographical representativeness by comparing the distribution by region and by province of active GPs in Belgium in 2020 (source: www.ima-aim.be) with the distribution of GPs who participated at T1.⁴

All participants were asked at T1 if they wanted to participate in the validation study. A subsample of the participants who gave consent was asked to provide a serum sample for this validation study. This subsample is made up of all those participants that were seropositive for SARS-CoV-2 on the RST at T1 and a random sample of those who tested negative or had an unclear RST result at T1 (figure 1).

Patient and public involvement

Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Test methods

Index test

For the cohort study, each participant was sent RSTs by postal service and instructions by email before their first testing timepoint. Participants were instructed on how to use the RST both with written documentation and an instructional video. The participating GPs were responsible for the interpretation of the RST results of the other

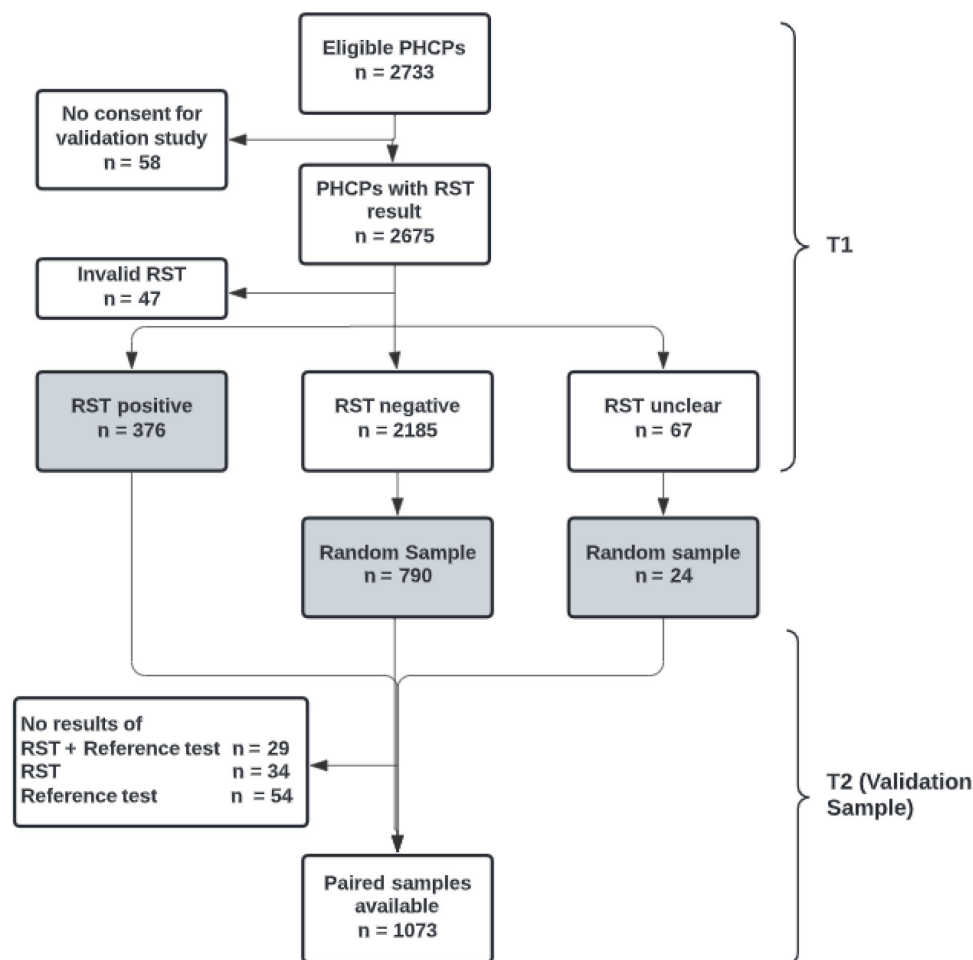


Figure 1 Participant flow. PHCP, primary healthcare provider; RST, rapid serological test.

PHCPs in their practice. Each PHCP entered their RST result into a secured online data capturing tool, LimeSurvey, hosted by Sciensano.¹⁶ The device includes a control line to confirm the validity of the test, along with two lines for the IgM and IgG antibodies, respectively (online supplemental figure S1). Possible RST results were invalid or valid, and, if valid, IgM positive or negative and IgG positive or negative (online supplemental figure S2). In our survey, we also included IgM unclear and IgG unclear to record unclear interpretation of valid RST results. The PHCPs included in this validation study were asked to perform the RST for their second testing timepoint (T2) (25–31 January 2021) immediately after collecting a venous blood sample for the reference test.

Reference test

For the reference test, participants were sent materials to collect a venous blood sample (Becton Dickinson Vacutainer SST II Advance; ref 368879) along with an envelope and stamp (in accordance with the UN3373 packaging norms) and instructions on how to send it to the central clinical laboratory of the University Hospital of Antwerp (UZA). Analysis at the laboratory was done within 24 hours of reception. The impact of this delay in processing of blood samples, time between sampling and

arrival at the laboratory, preprocessing storage temperature and vacutainer type on antibody seropositivity was investigated by Hodgkinson *et al.* They concluded that antibodies against epitopes of several infectious agents can be reliably stored in serum tubes at either room temperature or at 4°C for up to 6 days before analysis.¹⁷

Analysis of the venous blood samples was done with a reference standard using the following testing algorithm: serum samples were tested first on the Elecsys Anti-SARS-CoV-2 S assay (Roche, Basel, Switzerland); if the cut-off index was between 0.6 and 3.0, the sample was tested on the Atellica IM SARS-CoV-2 assay (Siemens, Munich, Germany), and in case of discordant results, further testing was done on the LIAISON SARS-CoV-2 S1/S2 IgG assay (DiaSorin, Saluggia, Italy) using a two-out-of-three ‘reference standard’. The analytical and clinical performance of these three commercially available, fully automated SARS-CoV-2 antibody assays was investigated at the UZA and the relevance of this testing algorithm was explained and illustrated (B Peeters, personal communication, 2020).¹⁸ Analytical performance of all three assays was acceptable (<7.6% analytical imprecision) and comparable with the results found in other studies.^{19–22}

Table 1 Characteristics of the participants in the validation sample and total study population at the second testing timepoint

	Validation sample n=1190	Study population at T2 n=2710
Age, median (IQR)	41 (32–56)	48 (32–56)
Female, n (%)	762 (64)	1800 (66)
Immunocompromising disease, n (%)	18 (2)	33 (1)
Profession, n (%)		
GP	1018 (86)	2339 (86)
GP in training	39 (3)	81 (3)
Other PHCP	125 (11)	274 (10)
Practice size, n (%)		
Solo	237 (20)	536 (20)
Duo	203 (17)	436 (16)
Group (<8 employees)	240 (20)	568 (21)
Large group (>7 employees)	493 (41)	1136 (42)
Vaccination, n (%)		
Not vaccinated	446 (37)	1101 (41)
Partially vaccinated	686 (58)	1584 (58)
Fully vaccinated	13 (1)	21 (1)
Vaccine type, n (%)		
Comirnaty (Pfizer/BioNTech)	696 (58)	1598 (59)
Spikevax (Moderna)	3 (0)	5 (0)
Jcovden (Johnson & Johnson)	0 (0)	1 (0)

If numbers do not add up to the column total, this is due to missing data.
IQR: interquartile range
GP, general practitioner; PHCP, primary healthcare provider.

Sample size

To be able to validate the RST accuracy in a primary care setting, that is, estimate the RST sensitivity (92.9%) with a lower limit of its 95% CI of 90% and its specificity (96.3%) with a lower limit of its 95% CI of 95%, a sample of 301 PHCPs seropositive on the reference standard (for sensitivity) and 810 PHCPs seronegative on the reference standard (for specificity) is required. This corresponds to, for

example, 6% seroprevalence in 5022 PHCPs. To reduce the burden on the participants and the costs of the study, all those with a positive RST and only a (random) sample of 900 PHCPs with a negative RST at T1 were assessed with the reference standard, and inverse probability weighting was applied to correct for missing reference standard data by design.^{23–25} Participants' characteristics of the validation sample and the total study population at T2 are shown in [table 1](#).

Statistical analysis

We calculated the sensitivity of the RST as the proportion of all subjects testing positive on our testing algorithm of all three assays in the venous blood sample (reference test) who tested positive with the RST (index test). We calculated the specificity of the RST as the proportion of all subjects testing negative on the reference test who tested negative on the index test. For both estimates, 95% CIs were calculated using the method described by Wilson.²⁶ We excluded subjects with missing results for the index test or reference test from the analysis.

Valid RST results were considered as positive if either IgM or IgG antibodies were positive; as unclear if IgM and IgG were unclear or if either IgM or IgG was unclear while the other was negative; and as negative if both IgM and IgG were negative. To determine the accuracy (sensitivity and specificity) of the RST based on the resulting three-by-two table including the missing variables ([table 2](#)), two separate analyses were performed, one categorising unclear results as positive and another categorising unclear results as negative.

For the determination of the overall sensitivity and specificity of the test, unclear results were considered as negative (positive) when the reference test on serum was positive (negative). This method seems clinically the most relevant to us.^{27–29} In order to determine the overall sensitivity of this test, unclear results were considered to be false negatives, as the presence of antibodies is questioned by the RST. For the determination of the overall specificity, we considered the unclear results to be false positives since the RST showed a possible positive result when in fact no antibodies were present.

Table 2 Reference test results of rapid serological test (RST) results in the validation sample and RST results in all participants at the second testing timepoint

		Reference test (on venous blood)				Total in validation sample	Total in all participants
		Positive	Negative	Invalid	Missing		
RST	Positive	337	77	0	14	428	660
	Unclear	25	17	0	1	43	97
	Negative	41	576	0	39	656	1897
	Invalid	2	13	1	3	19	56
	Missing	9	10	0	25	44	0
	Total	414	693	1	82	1190	2710

Reference test: ELISA tests on serum with a two-out-of-three testing algorithm.

A post hoc sensitivity analysis was performed to investigate if an interval of >6 days between collection and analysis of the serum sample affected our estimates of the sensitivity and specificity of the RST. Therefore, we assessed whether the accuracy in the subset with an interval of >6 days differed from the accuracy in the entire study sample. For this purpose, we divided the data into two independent subsamples: the subset of PHCPs with an interval of ≤6 days and the subset with an interval of >6 days.³⁰ Accuracy estimates of both subsamples were measured after inverse probability weighting. We performed a two-sample test for equality of proportions comparing the sensitivity and specificity of both groups and presented the differences in sensitivity and specificity with their 95% CI.

Subgroup analyses of the sensitivity and specificity were performed for vaccination status and age (<60 vs ≥60 years). Inverse probability weighting was done for those subgroups at T2 with extrapolation of the ratio negative and positive results on the RST for those subgroups.

We estimated the test prevalence at T2 as the proportion of positive tests out of the number of valid tests. The true prevalence at T2 was estimated using the overall sensitivity (73%) and specificity (92%) of the RST.

Finally, we estimated the true prevalence corresponding to RST-based prevalence values found during our cohort study in PHCPs in Belgium using the most conservative estimates for the RST sensitivity and specificity in R (V.4.2.1), with the package epiR which uses the Rogan-Gladen estimate for true prevalence.³¹ 95% CIs were calculated with the Wilson method.²⁶

RESULTS

Participants

The validation sample is similar to the study population at T2, with a majority being GPs. The median age is 41 in the

validation sample and 48 in the study population at T2. In both groups, there was a small majority of female participants, 59% had their COVID-19 vaccination (mainly messenger RNA vaccines) and only a few participants (18/1190 and 33/2710, respectively) reported having immunocompromising diseases (table 1).

Sensitivity and specificity

Out of 2733 PHCPs who participated at T1 of the prospective cohort study starting 24 December 2020, a total of 2675 gave their consent for the validation study. Up to 1190 participants were sampled for the validation of which 1073 paired samples were collected. All numbers are presented in figure 1 with corresponding missing data.

The median time between sample collection in the GP practice and analysis in the laboratory is 6 days (IQR 5–7). We obtained results for 1073 paired samples of which 403 were positive according to the reference testing algorithm. On the RST, 414 tested positive, 617 negative and 42 unclear. Online supplemental figure S3 shows the distribution of participants in the validation study.

The reference test results of the positive, negative, unclear and missing RST results are shown in table 2 together with the RST results in all participants who performed an RST at T2.

The two-by-two tables with the number of true/false positives and true/false negatives of the RST in comparison with the reference test are shown in table 3. Numbers are shown for both scenarios, unclear RST results considered as negative (left) and positive (right) using inverse probability weighting to extrapolate the reference test result in all participants at T2 based on the validation sample. Sensitivity and specificity with 95% CIs are also determined for the two scenarios in table 4. Sensitivity increases from 72.9% to 82.7%, and specificity decreases from 93.6% to 91.9% when changing the interpretation

Table 3 Reference test results of valid rapid serological test (RST) results in the validation sample with unclear RST results considered as negative (left) and positive (right) results and extrapolated to all participants at the second testing timepoint using inverse probability weighting

Reference test							
Validation sample		Positive	Negative	Total	Positive	Negative	Total
RST	Positive	337 (81.4%)	77 (18.6%)	414	362 (79.4%)	94 (20.6%)	456
	Negative	66 (10%)	593 (90%)	659	41 (6.7%)	576 (93.4%)	617
	Total	403 (37.6%)	670 (62.4%)	1073	403 (37.6%)	670 (62.4%)	1073
All participants at T2		Positive	Negative	Total	Positive	Negative	Total
	Positive	537	123	660	601	156	757
	Negative	200	1794	1994	126	1771	1897
	Total	737	1917	2654	727	1927	2654
Unclear results on RST considered as negative					Unclear results on RST considered as positive		
Reference test: ELISA tests on serum with a two-out-of-three testing algorithm.							

Table 4 Sensitivity and specificity of the rapid serological test (RST) with 95% CI for difference in interpretation of unclear RST, time intervals between sampling and analysis of the serum sample for the reference test, vaccination status and age groups

	Sensitivity (95% CI)	Specificity (95% CI)
Interpretation of unclear RST		
Negative	72.9 (69.5–76)	93.6 (92.4–94.6)
Positive	82.7 (79.7–85.3)	91.9 (90.6–93.1)
Time between sampling and analysis of serum		
≤6 days (n=655)	71.4 (68.0–74.5)	92.0 (90.7–93.0)
>6 days (n=435)	75.9 (72.6–78.9)	91.8 (90.5–93.0)
Vaccination status		
Yes (n=699)	79.4 (77.3–81.4)	90.4 (88.8–91.7)
No (n=446)	60.6 (57.6–63.5)	94.9 (93.3–96.0)
Age		
<60 years (n=977)	74.0 (72.1–75.8)	91.9 (90.7–93.0)
≥60 years (n=212)	76.4 (72.0–80.1)	90.8 (87.6–93.1)

of unclear test results on the RST from an absence to a presence of antibodies.

Sensitivity analysis and subgroup analyses

Both sensitivity and specificity did not differ significantly between the groups with different time intervals between sampling and analysis (≤6 days vs >6 days) (table 4).

The subgroup analysis for vaccination status showed a lower sensitivity and a higher specificity of the RST in non-vaccinated participants over vaccinated participants ($p<0.0001$). No difference was observed in the

performance of the RST between participants aged <60 and ≥60 years.

Test and true prevalence

The RST-based prevalence of antibodies against SARS-CoV-2 at T2 in this study population was 24.9%. The corresponding true prevalence using the most conservative estimates of RST sensitivity (72.9%) and specificity (91.9%) is 25.9% (95% CI 22.7%–29.1%).

Figure 2 shows the true prevalence of antibodies against SARS-CoV-2 corresponding to the other RST-based prevalence values estimated during our cohort study using the same values for the RST sensitivity and specificity.

DISCUSSION

Summary of findings

We evaluated the accuracy of the OrientGene RST for the detection of SARS-CoV-2 antibodies in PHCPs when performed by GPs and found a sensitivity of 72.9% (95% CI 69.5%–76.0%) and a specificity of 91.9% (95% CI 90.6%–93.1%) as most conservative estimates.

The subgroup analysis for vaccination status shows a difference in performance between the two groups with a lower sensitivity and a higher specificity in non-vaccinated participants. A possible explanation might be the lower number of antibodies in non-vaccinated participants as their infection might have been long before T2. On the other hand, the group of vaccinated people benefited from a recent immune response to a vaccination given in the month prior to the testing point. For the higher specificity in the non-vaccinated participants, we have no reliable explanation. However, since this is a post hoc

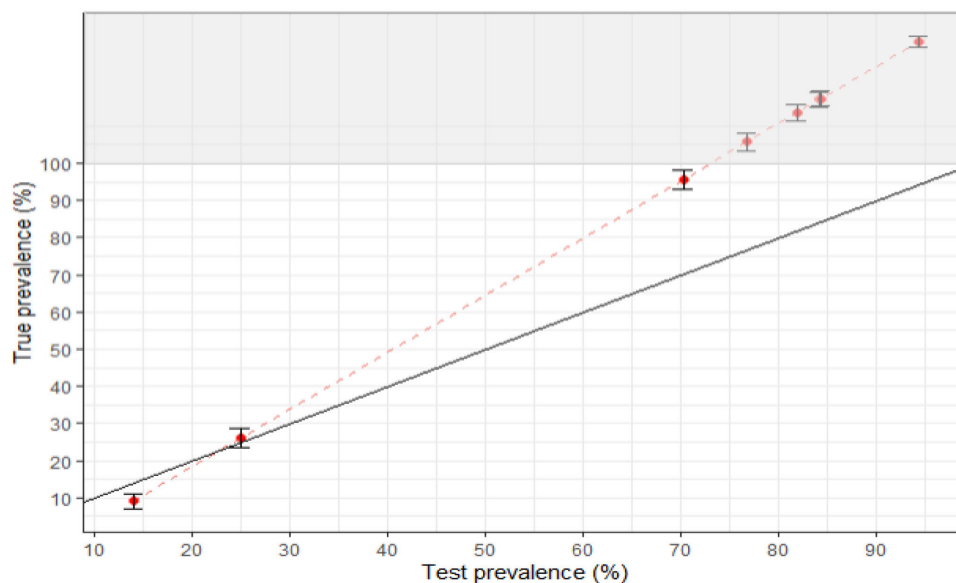


Figure 2 The estimated true prevalence* and 95% confidence intervals† for imperfect tests‡ based on prevalence values during our cohort study among primary healthcare providers (PHCPs) between 24 December 2020 and 26 December 2021.⁴ Since the true prevalence cannot exceed 100%, but the limited basic calculation of Rogan -Gladen, which uses a fixed sensitivity and specificity,³¹ can result in true prevalence values greater than 100%, these implausible values are clearly marked in grey. * Rogan -Gladen method.³¹ †Calculated with the Wilson method.²⁶ ‡Sensitivity of 72.9% and specificity 91.9%.

subgroup analysis, without formal hypothesis testing, we cannot exclude the possibility that it is an artefact of the data.

A test prevalence of 24.9% in this study population at T2 corresponds to a true prevalence of 25.9% (95% CI 23.5%–28.5%). A simulation of different test prevalence rates shows that RST-based values below 23% overestimate the true seroprevalence, while RST-based values above 23% underestimate the true seroprevalence. For example, an RST-based prevalence of 13.9% at T1 corresponds to a true prevalence of 9.1% (95% CI 7.2%–11.2%), while for an RST-based prevalence of 70.2% at T7, the true seroprevalence is expected to be 95.7% (95% CI 93%–98.3%).

Strengths and limitations

Strengths

We performed this phase III validation study in PHCPs, which is a relevant target population for the RST assessed. Choosing this population resulted in a high number of positive antibody tests early in the pandemic because of the early vaccination uptake of COVID-19 vaccination among GPs. As this validation study is part of a larger cohort study, it was possible to include individuals with negative results on the reference tests from the same population without using pre-pandemic samples. Therefore, calculation of the true prevalence was possible based on the RST prevalence value.

Index tests and reference tests were conducted at the same time by the GPs in their GP practices. First, this timing makes a direct comparison between the test results possible. Second, in a phase III study, it is desirable that results are interpreted by individuals who would do this as part of their usual clinical workload.¹⁴ This method provides realistic estimates when used in clinical practice.

Interpretation of the test result by a large group of individuals leads to a high likelihood of unclear results. These inconclusive results need to be incorporated in the analyses to be able to estimate the value of a test in clinical practice. However, no consensus exists for this incorporation. The ideal method takes into account how the test will be used in clinical practice.²⁸ In order to determine the overall sensitivity of this test, unclear results were considered to be false negatives, as the presence of antibodies is questioned by the RST. For the determination of the overall specificity, we considered the unclear results to be false positives since the RST showed a possible positive result when in fact no antibodies were present. This method, also explained by Garcia-Romero *et al*, seems clinically the most relevant to us.^{27–29} As shown in our results, adding the unclear results to the most favourable category, unclear as positive (negative) when the serum result is positive (negative), retrospectively leads to an overestimation of the test accuracy.³²

Limitations

Of the 1190 participants invited, there were 29 (2.4%) samples missing from both RST and serum samples, 34 (2.8%) RST results and 54 (4.5%) serum samples. As a

result, we are missing almost 10% of our intended sample size.

To avoid partial verification bias and to increase efficiency, we applied the reference standard to all those who were positive on the index test and a random sample of those who were negative. By using this test result-based sampling we adjusted the sensitivity and specificity for the sampling fraction in the negative index group.³³ We assume that the proportion of positive and negative reference test results for both positive and negative index test results is similar in all PHCPs, that is, whether they participated in the validation study or not.

In addition, participants were included based on their previous test result at T1. Many PHCPs had been vaccinated between T1 and T2. As a result, the number of reference test negatives required for the validation study was not achieved. In this validation study, 693 instead of 810 participants with a negative reference test, which we needed based on our sample size calculation, were included. This causes the 95% CI of the specificity to be wider than anticipated. However, since the parameter (specificity) estimate itself was below the lower limit of the 95% CI set for this parameter, it is unlikely that with a larger sample size the estimate for the specificity would have a lower limit of its 95% CI of 95% or above.

Up to now, no gold standard exists for the determination of antibodies against the SARS-CoV-2 virus. Therefore, a two-out-of-three testing algorithm was used in this validation study as reference standard.^{19–21} Nevertheless, these reference tests were extensively tested on PCR-positive and pre-pandemic samples or PCR-negative samples, and compared with other commercially available immunoassays. For the Elecsys Anti-SARS-CoV-2 S assay and the Atellica IM SARS-CoV-2 assay a superior agreement was observed with SARS-CoV-2-positive and negative samples.^{34 35}

The limited experience of those who read the index test, as this was only the second measuring point, could have affected the accuracy of the index test in this validation study due to the possibility of more unclear test results. To minimise this risk, participants were instructed with written information and an instructional video. However, as we also point out below, the REACT (REAL Time Assessment of Community Transmission) 2 study demonstrated a good agreement between the results read by participants and by trained observers.³⁶

Research shows that serum samples are stable for up to 6 days at room temperature to 4°C between collection and analysis.¹⁷ In our study, the median interval between collection and analysis of the serum sample was 6 (IQR 5–7) days. In a post hoc sensitivity analysis, we found no statistically or clinically significant impact of including participants with an interval of >6 days on our estimates for the sensitivity and specificity of the RST.

Subgroup analysis based on vaccination status showed a lower sensitivity and higher specificity in the non-vaccinated participants. Vaccine type may also impact the RST performance. The study population was primarily composed of GPs

who were prioritised in the vaccination campaigns. Given that Comirnaty (Pfizer) was the first vaccine to receive approval in Belgium, nearly all participating PHCPs received this vaccine. Therefore, subgroup analysis based on vaccine type was not possible. We acknowledge that the potential lack of diversity in the vaccine types in our participants may impact the generalisability of our findings.

Of note, in our cohort study we observed an RST-based prevalence of up to 93.9% in December 2021,^{4 37} which is higher than what could be expected with the sensitivity and specificity of the RST estimated in this validation study. As shown in figure 2, using imperfect sensitivity and specificity estimates overestimates/underestimates true seroprevalence. In the analysis of the true prevalence, we used fixed values for prevalence, sensitivity and specificity and did not account for uncertainty around them. Given a true prevalence of SARS-CoV-2 antibodies of 61%—more than 60% is to be expected in PHCPs in Belgium in December 2021—finding 939 positive RSTs out of 1000 corresponds to an RST sensitivity of at least 90% (higher if not all negative tests are considered to be false negative). Therefore, the sensitivity of the RST appears to be significantly higher than the estimated 72.9% when used in a high prevalence setting in those who received at least two doses of the vaccine and a booster (with a third dose of the vaccine and/or due to infection).

Comparison with literature

To date, a few studies have validated the OrientGene RST in laboratory settings. Jones *et al* found a sensitivity of 94.0% (95% CI 90.5%–96.3%) and a specificity of 95.8% (95% CI 94.8%–96.6%) based on individuals with self-reported PCR-confirmed infection as true positives and pre-pandemic samples as true negatives.¹² The validation by Sciensano showed a sensitivity of 92.9% and a specificity of 96.3%. The drop in test accuracy observed in our study can be explained by the variability of persons conducting and interpreting the index test in our study compared with only two observers interpreting the index test in the validation study by Sciensano.¹⁴

This loss in accuracy was also observed by the REACT 2 study. They investigated the difference in accuracy of several lateral flow immunoassays (LFIA) between (a) fingerprick (self-read), (b) fingerprick (trained observer read) and (c) serum in laboratory setting. Same as in our study, this study population consisted mainly of health-care providers. A good agreement was found between the results reported by the participants and those by the trained observers. The performance with fingerprick compared with serum in the laboratory showed only a moderate agreement at best (κ 0.56).³⁶ Specificity, calculated on pre-pandemic samples, was high for all LFIAs. Sensitivity was variable and moderate. Evaluation of each test in the intended setting is therefore necessary.

Contemporary role of the RST

Due to its performance characteristics, we do not recommend this RST for individual use to determine the immunity to SARS-CoV-2. However, in a high SARS-CoV-2

seroprevalence setting such as the current situation, this test may still have a role to play. By using this RST repeatedly in the same population, one gains insights into the duration of humoral immunity after vaccination. This information can be used to guide vaccine campaigns in specific populations. This is also suggested by Meyers *et al* who found a faster decline in SARS-CoV-2 antibodies' prevalence among nursing home residents.³⁸

Interpretation

At the start of our cohort study, in December 2020, we estimated a crude RST-based seroprevalence of 14% (366/2629). Accounting for its imperfect accuracy, we estimate a true prevalence of 9% (95% CI 6.31%–11.73%) at this time. In April 2021, this RST-based seroprevalence increased to 84% (2410/2859) which corresponds to a calculated, implausible true prevalence of 117%. In December 2021, that is, when 99% was fully vaccinated and 85% had received a booster vaccination, the RST-based seroprevalence even reached 94% (2356/2498) which corresponds to a true prevalence of 133%. Our findings suggest, on the one hand, that an imperfect RST with a sensitivity of only 73% and a specificity of 92% rather overestimates the true seroprevalence in the beginning of an epidemic when seroprevalence is low, but underestimates the true seroprevalence when seroprevalence is actually high. For both scenarios, therefore, some caution is required when interpreting the results of this RST as the SARS-CoV-2 seroprevalence. We emphasise prudence in using RST-based prevalence estimates, as well as estimates of sensitivity and specificity as fixed values to calculate the true prevalence, as there is uncertainty in these estimates.

CONCLUSION

The sensitivity and the specificity of the OrientGene RST were 72.9% and 91.9%, respectively, when performed by GPs with fingerprick blood as part of a large SARS-CoV-2 seroprevalence study among PHCPs in Belgium. As a result, RST-based estimates below 23% overestimate the true seroprevalence, and RST-based estimates above 23% underestimate the true seroprevalence.

Author affiliations

¹Department of Family Medicine and Population Health (FAMPOP), Centre for General Practice, University of Antwerp, Antwerpen, Belgium

²Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford, UK

³Department of Public Health and Primary Care, EPI-Centre, KU Leuven, Leuven, Belgium

⁴General Practice Department—Primary Care and Health Research Unit, Liege University, Liege, Belgium

⁵Department of Laboratory Medicine, University Hospital Antwerp, Edegem, Belgium

⁶Interuniversity Institute for Biostatistics and Statistical Bioinformatics (I-BioStat), Data Science Institute, Hasselt University, Hasselt, Belgium

⁷Epidemiology & Pharmavigilance, P95, Leuven, Belgium

⁸Department of Public Health and Primary Care, University of Ghent, Ghent, Belgium

⁹Department of Infectious Diseases in Humans, Sciensano, Brussels, Belgium

¹⁰Vaccine & Infectious Disease Institute, Centre for the Evaluation of Vaccination, University of Antwerp Faculty of Medicine and Health Sciences, Antwerpen (Wilrijk), Belgium

¹¹Laboratory of Medical Microbiology, Vaccine & Infectious Disease Institute (VAXINFECTIO), University of Antwerp, Antwerpen, Belgium
¹²Department of Epidemiology and Public Health, Sciensano, Brussels, Belgium

Twitter Jan Yvan Jos Verbakel @jan_verbakel

Acknowledgements JV was supported by the National Institute for Health and Care Research (NIHR) Community Healthcare MedTech and In Vitro Diagnostics Co-operative at Oxford Health NHS Foundation Trust.

Contributors All authors (JD, JYJV, NA, BS, BP, RB, ADS, SH, AVdB, ID, PVD, HG, LB, ED, SC) contributed to the study conception and design. Material preparation, data collection and analysis were performed by JD, SC, ED, NA, BS and RB. The first draft of the manuscript was written by JD and all coauthors commented on previous versions of the manuscript. All authors read and approved the final manuscript. SC is acting as guarantor.

Funding This work was supported by Sciensano (grant number CS-18390).

Disclaimer The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care.

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Ethics approval Ethical approval has been granted by the Ethics Committee of the University Hospital of Antwerp/University of Antwerp (Belgian registration number: 3002020000237). Alongside journal publications, dissemination activities include the publication of monthly reports to be shared with the participants and the general population through the publicly available website of the Belgian health authorities (Sciensano). Informed consent was obtained from all individual participants included in the study.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request. The relevant anonymised patient-level data that support the findings of this study are available from the corresponding author on reasonable request.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

Author note Els Duysburgh and Samuel Coenen contributed equally to this work as last author.

ORCID iDs

Julie Domen <http://orcid.org/0000-0003-3239-7834>
 Jan Yvan Jos Verbakel <http://orcid.org/0000-0002-7166-7211>
 Niels Adriaenssens <http://orcid.org/0000-0002-7957-1177>
 Beatrice Scholtes <http://orcid.org/0000-0001-5274-822X>
 Robin Bruyndonckx <http://orcid.org/0000-0002-4217-2869>
 An De Sutter <http://orcid.org/0000-0002-2540-8307>
 Stefan Heytens <http://orcid.org/0000-0003-1097-4987>
 Ann Van den Bruel <http://orcid.org/0000-0001-9012-2009>
 Isabelle Desombere <http://orcid.org/0000-0003-3521-7712>
 Pierre Van Damme <http://orcid.org/0000-0002-8642-1249>
 Herman Goossens <http://orcid.org/0000-0003-4891-4031>
 Laetitia Buret <http://orcid.org/0000-0002-5653-615X>
 Els Duysburgh <http://orcid.org/0000-0003-3345-7616>
 Samuel Coenen <http://orcid.org/0000-0002-1238-8052>

REFERENCES

- Roser M, Ritchie H, Ortiz-Ospina E, et al. Coronavirus pandemic (COVID-19). our world in data. 2020. Available: <https://ourworldindata.org/coronavirus>
- Meurisse M, Lajot A, Dupont Y, et al. One year of laboratory-based COVID-19 surveillance system in Belgium: main indicators and performance of the laboratories (March 2020-21). *Arch Public Health* 2021;79:188.
- McConnell D, Hickey C, Bargary N, et al. Understanding the challenges and uncertainties of seroprevalence studies for SARS-cov-2. *Int J Environ Res Public Health* 2021;18:4640.
- Adriaenssens N, Scholtes B, Bruyndonckx R, et al. Prevalence and incidence of antibodies against SARS-cov-2 among primary healthcare providers in Belgium during 1 year of the COVID-19 epidemic: prospective cohort study protocol. *BMJ Open* 2022;12:e054688.
- Duysburgh E, Mortgat L, Barbezange C, et al. Persistence of IgG response to SARS-cov-2. *Lancet Infect Dis* 2021;21:163-4.
- Mortgat L, Verdonck K, Hutse V, et al. Prevalence and incidence of anti-SARS-cov-2 antibodies among healthcare workers in Belgian hospitals before vaccination: a prospective cohort study. *BMJ Open* 2021;11:e050824.
- Mariën J, Ceulemans A, Bakokimi D, et al. Prospective SARS-cov-2 cohort study among primary health care providers during the second COVID-19 wave in Flanders, Belgium. *Fam Pract* 2022;39:92-8.
- Cals J, van Weert H. Point-Of-Care tests in general practice: hope or hype? *Eur J Gen Pract* 2013;19:251-6.
- Triest D, Geebelen L, De Pauw R, et al. Performance of five rapid serological tests in mild-diseased subjects using finger prick blood for exposure assessment to sars-cov-2. *J Clin Virol* 2021;142:104897.
- Covid-19 igg/igm rapid test cassette (whole blood/serum/plasma). 2023.
- Dimech J, Curley S, Bond K, et al. Post-Market validation of a further three serological assays for COVID-19: the University of Melbourne and the Royal Melbourne Hospital. 2021.
- Jones HE, Mulchandani R, Taylor-Phillips S, et al. Accuracy of four lateral flow immunoassays for anti sars-cov-2 antibodies: a head-to-head comparative study. *EBioMedicine* 2021;68:103414.
- Van den Bruel A, Aertgeerts B, Buntinx F. Results of diagnostic accuracy studies are not always validated. *J Clin Epidemiol* 2006;59:559-66.
- Zhou X-H, Obuchowski NA, McClish DK. Statistical methods in diagnostic medicine. Hoboken, NJ, 2011.
- Boelaert M, Bhattacharya S, Chappuis F, et al. Evaluation of rapid diagnostic tests: visceral leishmaniasis. *Nat Rev Microbiol* 2007;5:S31-9.
- Limesurvey GmbH. LimeSurvey: an open source survey tool/ LimeSurvey GmbH, Hamburg, Germany. Available: <http://www.limesurvey.org>
- Hodgkinson VS, Egger S, Betsou F, et al. Preanalytical stability of antibodies to pathogenic antigens. *Cancer Epidemiol Biomarkers Prev* 2017;26:1337-44.
- Huyghe E, Jansens H, Matheeußen V, et al. Performance of three automated SARS-cov-2 antibody assays and relevance of orthogonal testing algorithms. *Clin Chem Lab Med* 2020;59:411-9.
- Favresse J, Eucher C, Elsen M, et al. Clinical performance of the elecsys electrochemiluminescent immunoassay for the detection of sars-cov-2 total antibodies. *Clin Chem* 2020;66:1104-6.
- Egger M, Bundschuh C, Wiesinger K, et al. Comparison of the elecsys® anti-SARS-cov-2 immunoassay with the EDI. *Clin Chim Acta* 2020;509:18-21.
- Tré-Hardy M, Wilmet A, Beukinga I, et al. Validation of a chemiluminescent assay for specific SARS-cov-2 antibody. *Clin Chem Lab Med* 2020;58:1357-64.
- Kohmer N, Westhaus S, Rühl C, et al. Brief clinical evaluation of six high-throughput sars-cov-2 IgG antibody assays. *J Clin Virol* 2020;129:S1386-6532(20)30222-5.
- Naaktgeboren CA, de Groot JAH, Rutjes AWS, et al. Anticipating missing reference standard data when planning diagnostic accuracy studies. *BMJ* 2016;352:i402.
- Begg CB, Greenes RA. Assessment of diagnostic tests when disease verification is subject to selection bias. *Biometrics* 1983;39:207-15.
- Seaman SR, White IR. Review of inverse probability weighting for dealing with missing data. *Stat Methods Med Res* 2013;22:278-95.
- Wilson EB. Probable inference, the law of succession, and statistical inference. *J Am Stat Assoc* 1927;22:209-12.
- Landsheer JA. The clinical relevance of methods for handling inconclusive medical test results: quantification of uncertainty in medical decision-making and screening. *Diagnostics (Basel)* 2018;8:32.



- 28 Shinkins B, Thompson M, Mallett S, *et al.* Diagnostic accuracy studies: how to report and analyse inconclusive test results. *BMJ* 2013;346:bmj.f2778.
- 29 Garcia-Romero H, Garcia-Barrios C, Ramos-Gutierrez F. Effects of uncertain results on sensitivity and specificity of diagnostic tests. *Lancet* 1996;348:1745–6.
- 30 Hayes LJ, Berry G. Comparing the part with the whole: should overlap be ignored in public health measures? *J Public Health (Oxf)* 2006;28:278–82.
- 31 Rogan WJ, Gladen B. Estimating prevalence from the results of a screening test. *Am J Epidemiol* 1978;107:71–6.
- 32 Schuetz GM, Schlattmann P, Dewey M. Use of 3x2 tables with an intention to diagnose approach to assess clinical performance of diagnostic tests: meta-analytical evaluation of coronary CT angiography studies. *BMJ* 2012;345:e6717.
- 33 Kohn MA. Studies of diagnostic test accuracy: partial verification bias and test result-based sampling. *J Clin Epidemiol* 2022;145:179–82.
- 34 Riestler E, Findeisen P, Hegel JK, *et al.* Performance evaluation of the Roche elecsys anti-SARS-cov-2 S immunoassay. *J Virol Methods* 2021;297:S0166-0934(21)00210-X.
- 35 Ward MD, Mullins KE, Pickett E, *et al.* Performance of 4 automated sars-cov-2 serology assay platforms in a large cohort including susceptible covid-19-negative and covid-19-positive patients. *J Appl Lab Med* 2021;6:942–52.
- 36 Flower B, Brown JC, Simmons B, *et al.* Clinical and laboratory evaluation of SARS-cov-2 lateral flow assays for use in a national COVID-19 seroprevalence survey. *Thorax* 2020;75:1082–8.
- 37 Adriaenssens N, Scholtes B, Bruyndonckx R, *et al.* Prevalence, incidence and longevity of antibodies against SARS-cov-2 among primary healthcare providers in Belgium: a prospective cohort study with 12 months of follow-up. *BMJ Open* 2022;12:e065897.
- 38 Meyers E, Deschepper E, Duysburgh E, *et al.* Declining prevalence of SARS-cov-2 antibodies among vaccinated nursing home residents and staff six months after the primary bnt162b2 vaccination campaign in Belgium: a prospective cohort study. *Viruses* 2022;14:2361.