



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



ELSEVIER

Contents lists available at ScienceDirect

EClinicalMedicine

journal homepage: <https://www.journals.elsevier.com/eclinicalmedicine>

Research Paper

Multi-center nationwide comparison of seven serology assays reveals a SARS-CoV-2 non-responding seronegative subpopulation

Kfir Oved^{a,1}, Liraz Olmer^{b,1}, Yonat Shemer-Avni^c, Tamar Wolf^d, Lia Supino-Rosin^d, George Prajgrod^e, Yotam Shenhar^f, Irina Payorsky^f, Yuval Cohen^g, Yishai Kohn^g, Victoria Indenbaum^h, Rachel Lazar^d, Valeria Geylis^d, Michal Tepperberg Oikawa^h, Eilat Shinarⁱ, Evgeniy Stoyanovⁱ, Lital Keinan-Boker^{j,k}, Ravit Bassal^j, Shay Reicher^l, Ruti Yishai^{l,q}, Adina Bar-Chaim^m, Ram Doolmanⁿ, Yoram Reiter^o, Ella Mendelson^{h,k}, Zvi Livneh^p, Laurence S. Freedman^{b,2,*}, Yaniv Lustig^{h,2,*}

^a Canopy Immuno-Therapeutics and MeMed Diagnostics, Tirat Carmel, Israel

^b The Gertner Institute of Epidemiology and Health Policy Research, Sheba Medical Center, Tel Hashomer, Israel

^c Laboratory of Clinical Virology Clalit HMO and Soroka Medical Center, Beer-Sheva, Israel

^d Laboratory Division, Maccabi HMO, Rehovot, Israel

^e Laboratory Division, Meuhedet HMO, Rehovot, Israel

^f Laboratory Division, Leumit HMO, Or-Yehuda, Israel

^g Directorate of Defence Research and Development, Ministry of Defense, Tel Aviv, Israel

^h Central Virology Laboratory, Ministry of Health and Sheba Medical Center, Tel-Hashomer, Israel

ⁱ Magen David National Blood Services, Tel Hashomer, Israel

^j Israel Center for Disease Control, Ministry of Health, Chaim Sheba Medical center, Tel-Hashomer, Israel

^k Sackler Faculty of Medicine, Tel-Aviv University, Israel

^l Ministry of Health, Jerusalem, Israel

^m Asaf Harofeh Medical Center, Be'er Ya'akov, Israel

ⁿ Sheba Medical Center, Tel Hashomer, Israel

^o The Technion Institute of Technology, Haifa, Israel

^p Dept. of Biomolecular Sciences, Weizmann Institute of Science, Rehovot, Israel

^q Department of Laboratories, Public Health Services, Ministry of Health, Israel

ARTICLE INFO

Article History:

Received 30 September 2020

Revised 3 November 2020

Accepted 5 November 2020

Available online 19 November 2020

Keywords:

COVID-19

Serology

SARS-CoV-2

Antibodies

Validation study

Seronegative subpopulation

IgG

ABSTRACT

Background: An Israeli national taskforce performed a multi-center clinical and analytical validation of seven serology assays to determine their utility and limitations for SARS-CoV-2 diagnosis.

Methods: Serology assays from Roche, Abbott, Diasorin, BioMerieux, Beckman-Coulter, Siemens, and an in-house RBD ELISA were included. Negative samples from 2391 individuals representative of the Israeli population, and 698 SARS-CoV-2 PCR positive patients, collected between March and May 2020, were analyzed

Findings: Immunoassays sensitivities between 81.5%–89.4% and specificities between 97.7%–100% resulted in a profound impact on the expected Positive Predictive Value (PPV) in low (<15%) prevalence scenarios. No meaningful increase was detected in the false positive rate in children compared to adults. A positive correlation between disease severity and antibody titers, and no decrease in antibody titers in the first 8 weeks after PCR positivity was observed. We identified a subgroup of symptomatic SARS-CoV-2 positive patients (~5% of patients), who remained seronegative across a wide range of antigens, isotypes, and technologies.

Interpretation: The commercially available automated immunoassays exhibit significant differences in performance and expected PPV in low prevalence scenarios. The low false-positivity rate in under 20's suggests that cross-reactive immunity from previous CoV strains is unlikely to explain the milder disease course in children. Finding no decrease in antibody titers in the first 8 weeks is in contrast to some reports of short half-life for SARS-CoV-2 antibodies. The ~5% who were seronegative non-responders, using multiple

* Corresponding author.

E-mail address: Yaniv.lustig@sheba.health.gov.il (Y. Lustig).

¹ These authors contributed equally to this work.

² These authors contributed equally to this work.

assays in a population-wide manner, represents the proportion of patients that may be at risk for re-infection.

Funding: Israel Ministry of Health.

© 2020 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

1. Research in context

1.1. Evidence before this study

Serological assays are critical for identifying individuals who were previously infected with SARS-CoV-2. Most large diagnostic companies have developed automated high-throughput serology testing that allow monitoring of disease dynamics, identifying hot spots of infections, assessing herd immunity, and, in the future, will allow determining vaccination status. A literature search performed from March 1st up to Aug 20, 2020, showed that studies have been limited in the number of assays compared and the power and heterogeneity of the samples used, including small cohorts, lack of pediatric samples, gender biases, and disease severity biases.

1.2. Added value of this study

We found significant performance differences with major impact on PPV values in low prevalence scenarios. Additionally, we did not find evidence for prior cross-reactive immunity among children by past exposure to other corona strains. Our findings substantiate the correlation between antibody titers and disease severity and contradict evidence for decreased titers with time and short half-life. Finally, we found that roughly 1 out of 20 SARS-CoV-2 positive patients, did not seroconvert.

1.3. Implications of all the available evidence

The assay of choice for epidemiological assessments and diagnosis should be determined separately in each country and be based on the prevalence of SARS-CoV-2. Our results, together with previous evidence, suggest that antibody levels are maintained for at least 2 months, and are correlated with disease severity and hospitalization, suggesting that disease severity is not attributed to ineffective immune response. About 5% of the investigated population did not develop antibodies following SARS-CoV-2 infection, which puts a higher bound on the maximal serology sensitivity, and vaccine efficacy (~95%).

2. Introduction

Numerous serology assays have been developed for SARS-CoV-2 with considerable variation in underlying technology, ease of use, cost, and performance [1]. Clearly, accurate and scalable testing procedures are critical for achieving control over the current pandemic. Serological assays are the only measure to identify individuals who were previously infected with SARS-CoV-2. As such, they allow governments and large health organizations to continuously monitor disease dynamics, identify hot spots of infections, assess herd immunity, and better utilize resources. During the first and second quarter of 2020, most of the large diagnostics companies released their automated serology testing version to the market, mostly via the Emergency Use Authorization (EUA) FDA route [2]. A few comparisons between different assays have been published; however, most are limited in the number of assays compared and the power and heterogeneity of the samples used, including small cohorts, lack of pediatric samples, gender biases, disease severity biases etc. [3–5]. Therefore, a population-based validation was necessary to assess the actual performance of the different assays on the target population. The Israeli

Ministry of Health (MOH) established a national serology validation taskforce to validate and assess the performance of several assays in a nationwide, powered, and methodological manner [6]

Our goals for the study were to 1) validate and compare the performance of the different assays: despite declared performance of the different assays by the legal manufacturers, no external validation was performed due to the high urgency in launching these products. Of note, specificity in serological assays might differ between geographies, due to cross reactivity with background pathogens that might be more, or less prevalent in a specific territory. For example, Dengue was recently shown to have cross-reactivity with SARS-CoV-2 [7] and therefore endemic regions, might present higher false positive rates than other territories where this pathogen is less prevalent. 2) To assess the specificity in a large representative cohort of the target population and predict the Positive Predictive Value (PPV) under different prevalence scenarios: PPV is known to be highly sensitive to low prevalence and was shown in many cases, for example cancer screening, to diminish the overall utility of a diagnostic test due to the high ratio of false to true positive results [8]. The main parameters that determine the PPV are prevalence and specificity, and due to the relatively limited prevalence of SARS-CoV-2 exposure we wanted to quantify the PPV for each kit under low prevalence scenarios (<15%). The very slow increase in prevalence levels since the SARS-CoV-2 outbreak suggests that these low prevalence conditions are expected to remain until a wide and effective vaccination will be applied. 3) To assess the correlation between disease severity and antibody titers including titer decay: it was previously shown that antibody titers are correlated with disease severity [9,10]; however, some reports suggested that these titers rapidly decay and that the half-life time of SARS-CoV-2 antibodies is short [11]. If correct, such a behavior can significantly impact the creation of herd immunity and the probability for re-infection among previously exposed individuals. It might also have potential impact on the expected efficacy of vaccination and is therefore important to quantify in a population-wide manner. 4) To assess the proportion and characteristics of the patients who had no seroconversion ≥ 14 days from positive PCR (“seronegative non-responders”): recent reports on seroconversion were all based on limited sample size, limited number of assays in use (usually one) and thus created conflicting results, with some studies assessing no seronegative patients after 14–21 days [12], and others suggesting that the proportion of non-responding seronegative patients is over 20% [11]. Quantifying this sub-population is especially important due to its impact on herd immunity, potential vaccination efficacy, and psychological impact on the public.

3. Methods

3.1. Study design

The Israeli SARS-CoV-2 National Serology Validation Plan is based on a validation protocol (Supplementary Annx. 1). The diagnostic laboratories of all four Health Maintenance Organizations (HMOs) in Israel – Clalit, Maccabi, Leumit and Meuhedet and the Central Virology Laboratory (CVL) of the Ministry of Health (MOH), participated in the study which was designed and executed during March-July 2020. Each lab obtained and tested serum samples from different cohorts of non-SARS-CoV-2 and SARS-CoV-2 patients collected from all regions in Israel. The numbers of positive and negative samples tested in each laboratory are shown in Table 1. The Israeli MOH waived the

Table 1
Baseline cohort characteristics.

	Negative (n = 2391)		PCR Positive \geq 14d (n = 588)		PCR Positive <14d (n = 110)	
	N	(%)	N	(%)	N	(%)
CVL lab	1289	(53.9)	311	(52.9)	33	(30.0)
Clalit Lab	496	(20.7)	85	(14.5)	20	(18.2)
Leumit Lab	224	(9.4)	36	(6.1)	2	(1.8)
Maccabi Lab	177	(7.4)	125	(21.3)	32	(29.1)
Meuhedet Lab	205	(8.6)	31	(5.3)	23	(20.9)
Gender:						
Male	1037	(43.4)	327	(55.6)	53	(48.2)
Female	1354	(56.6)	226	(38.4)	39	(35.5)
Unknown			35	(6.0)	18	(16.4)
Age, y:						
0–6	222	(9.3)
6–10	149	(6.2)	1	(0.2)	.	.
11–20	317	(13.2)	55	(9.4)	17	(15.5)
21–30	497	(20.9)	137	(23.3)	38	(34.5)
31–40	325	(13.6)	91	(15.5)	6	(5.5)
41–50	281	(11.8)	86	(14.6)	14	(12.7)
51–60	192	(8.0)	76	(12.9)	13	(11.8)
61–70	234	(9.8)	37	(6.3)	10	(9.1)
>70	171	(7.2)	20	(3.4)	10	(9.1)
Unknown	1	(0.0)	85	(14.5)	2	(1.8)
Days after pos PCR:						
0–13	110	(100.0)
14–20	.	.	148	(25.2)	.	.
21–27	.	.	123	(20.9)	.	.
28–34	.	.	78	(13.3)	.	.
35–41	.	.	129	(21.9)	.	.
>42	.	.	110	(18.7)	.	.
Hospitalized:						
No	.	.	208	(35.4)	58	(52.7)
Yes	.	.	153	(26.0)	28	(25.5)
Unknown	.	.	227	(38.6)	24	(21.8)
Severity:						
Asymptomatic	.	.	20	(3.4)	10	(9.1)
Mild	.	.	406	(69.1)	44	(40)
Moderate/Severe	.	.	29	(4.9)	7	(6.4)
Unknown	.	.	133	(22.6)	49	(44.5)
Total	2391	(100.0)	588	(100)	110	(100)

requirement for an Institutional Review Board approval due to the urgent need for validation data, fully anonymized cohort, and the use of retrospective samples for the negative cohort.

3.2. Patients and samples

Overall 3089 samples, one per individual, were included; 2391 sera samples were collected prior to September 2019 and were considered SARS-CoV-2 negative samples, and 698 sera obtained between March 28th and May 24th, 2020 from individuals with a positive SARS-CoV-2 qRT-PCR sampled by nose/throat swab were considered positive samples. The overall distributions of age and other patient characteristics are shown in Table 1. Among the positives, 110 sera were collected less than 14 days, and 588, 14 or more days after the first positive PCR result.

3.3. Serology testing

Samples were tested using the following commercially available automated immunoassays according to manufacturer's instructions: ARCHITECT SARS-CoV-2 N IgG Immunoassay (Abbot, Illinois, U.S.A.), LIAISON® SARS-CoV-2 S1/S2 IgG (Diasorin, Saluggia, Italy), Elecsys® N Anti-SARS-CoV-2 (Roche, Mannheim, Germany), VIDAS® SARS-COV-2 RBD IgG (BioMérieux, Marcy-l'Etoile, France), Siemens SARS-CoV-2 RBD Total (COV2T) (Siemens, NY, USA) and Access SARS-CoV-2 RBD IgG assay (Beckman-Coulter, CA, U.S.A.). ELISA was performed based

on the Mount Sinai Hospital Clinical Laboratory SARS-CoV-2 IgG Antibody Test [13] which was modified accordingly: a 96 well microtiter Polysorb plate (Nunc, Thermo, Denmark) coated overnight with 1 μ g/ml of Receptor Binding Domain (RBD) antigen was blocked with 5% skimmed milk at 25 °C for 60 min and human serum samples (diluted 1:100 with 3% skimmed milk) were added to antigen coated wells. Following incubation at 25 °C for 120 min and incubation for 60 min after the addition of goat anti-human IgG horseradish peroxidase (HRP) conjugate (Jackson ImmunoResearch, PA, USA), Tetramethylbenzidine substrate (TMB) substrate was added followed by stop solution (1 M HCl) and the OD of each well was measured at 450 nm. ELISA index value below 0.9 was considered negative, between 0.9 and 1.1, equivocal and equal or above 1.1, positive. Supplementary Table 1 shows the antigen used in each assay for antibody detection.

3.4. Statistical methods

Specificity and sensitivity for each serological test were estimated as the proportion of correct results in the negative cohort and the PCR positive \geq 14 day cohort respectively. Equivocal results (occurring with Diasorin and ELISA tests) were excluded from the calculation. Confidence intervals were based on inverting the test based on the binomial distribution.

Positive predictive values (the probability that a person testing positive actually has the disease) and their 95% credible intervals were estimated using Bayesian methods, assuming independent

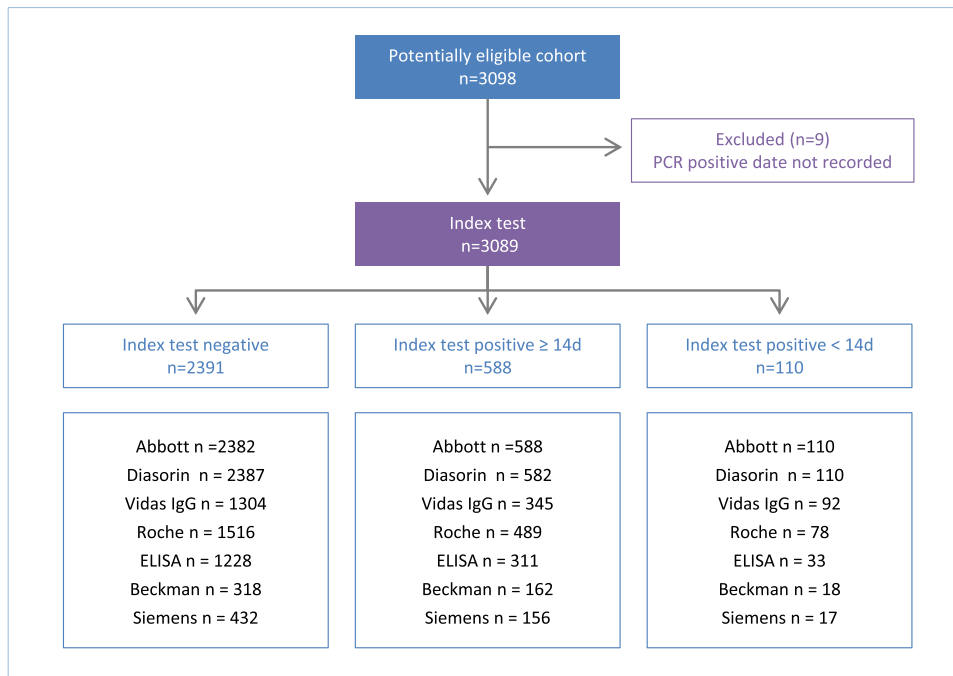


Fig. 1. Study selection process, total number of patients that were included and excluded, and details on the size of each tested group and kits based on the STARD guidelines. Out of 3098 available samples, 3089 samples were included in the study. Number of SARS-CoV-2 PCR positive (equal or above or below 14 days after first positive PCR) and negative samples tested with each of the serology assays are presented.

Uniform prior distributions for sensitivity and specificity. The motivation for Bayesian methods was to avoid over-interpretation of results showing 100% specificity (leading to 100% positive predictive value) based on finite samples. The Bayesian methods provide a more realistic estimate in such circumstances [14].

Optimal cut-off points were calculated based on minimizing the expected number of mistaken diagnoses for a given prevalence of disease. Confidence limits for the optimal cut-off points were calculated using the bootstrap percentile method [15], and the estimated number of mistakes was adjusted using 10-fold cross-validation.

The association between titer value and severity of disease, hospitalization and age was assessed using non-parametric methods (the Cuzick trend test [16] for age and severity of disease, and the Mann-Whitney test for hospitalization). The analysis of the association between titer value and the time from the positive PCR test included also those patients who were assessed less than 14d after the PCR test. The association was assessed using restricted cubic splines (5 knots placed at the 10th, 25th, 50th, 75th and 90th percentiles). Comparing characteristics of seronegative patients with those of seropositive patients was performed using Fisher's exact test, the choice for comparisons of proportions based on small numbers. Analyses were performed using the statistical software program SAS 9.4 (SAS Institute, Cary, NC, 2012).

3.5. 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 report. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

4. Results

4.1. Cohort characteristics and assays tested

We validated the performance of six commercially available automated IgG/total antibody serology products for SARS-CoV-2 and one

ELISA format, including the Roche Cobas Total Ig (Roche), Abbott SARS-CoV-2 IgG (Abbott), BioMerieux VIDAS IgG (VIDAS IgG), Beckman Coulter Access IgG (Beckman), Siemens ADVIA IgG (Siemens), Diasorin Liaison IgG (Diasorin) and an in-house RBD ELISA assay (ELISA, Supplementary Table 1). The study workflow based on the STARD guidelines is described in Fig. 1 and the cohort characteristics are detailed in Table 1. The varying numbers of samples per kit occurred from our focus on the Abbott and Diasorin kits that were due to be used in a national serological survey in Israel. Other kits were added according to the capacity of the individual laboratories and the volume of the blood samples. Assessing the sensitivity and specificity of the different assays revealed seemingly small gaps in specificity and more substantial gaps in sensitivity. The performance of each assay, sensitivity and specificity, with their corresponding confidence interval (CI), are shown in Table 2 and included 2391 negative and 588 positive samples collected ≥ 14 days after a positive PCR. The age distribution of the control (negative) cohort reflects the general population age distribution in the state of Israel and included a large proportion of pediatric and adolescent populations (aged <20 years), composing 28.8% of the cohort ($n = 688$) and the gender distribution was 1:1 (M:F 43.4:56.6%).

4.2. Assays performance

We calculated the sensitivity, specificity and the 95% confidence intervals for each one of the tested assays. Equivocal zone results, which were defined for the Diasorin ($n = 28$) and the ELISA format ($n = 20$), were not included in the calculation of sensitivity and specificity, although the FDA's approach is to use such results as false negative for sensitivity calculation and false positive for specificity calculation, which is expected to lower our reported results (<https://open.fda.gov/apis/device/covid19serology/>) (see Table 2). The best performing kits in terms of sensitivity were the ELISA with 89.4% (CI 85.4–92.3), the VIDAS IgG with 89.3% (CI 85.5–92.1), and the Roche, with 89.0% (CI 85.9–91.4). The best performing kits in terms of specificity were the Roche with 100% (CI 99.8–100), Beckman with 100% (CI 98.8–100), and the Siemens with 99.8% (CI 98.7–99.94), although

Table 2

Test performance of the different kits (equivocal results were excluded from the calculation of sensitivity and specificity and are not included in the sample size columns of the table).

Test Kit	Negative Patients (n)	% Specificity (95% CI ^e)	Positive PCR ≥14d Patients (n)	% Sensitivity (95% CI ^e)
Abbott	2382	99.5 (99.2–99.74)	588	84.7 (81.5–87.4)
Diasorin	2379 ^a	98.7 (98.2–99.1)	562 ^b	82.4 (79.0–85.3)
VIDAS IgG	1304	98.9 (98.2–99.4)	345	89.3 (85.5–92.1)
Roche	1516	100.0 (99.8–100.0)	489	89.0 (85.9–91.4)
ELISA	1209 ^c	97.7 (96.7–98.4)	310 ^d	89.4 (85.4–92.3)
Beckman	318	100.0 (98.8–100)	162	81.5 (74.6–86.6)
Siemens	432	99.8 (98.7–99.94)	156	85.9 (79.4–90.4)

a 8 equivocal results; b 20 equivocal results; c 19 equivocal results; d 1 equivocal result; e 95% confidence interval calculated using binomial distribution.

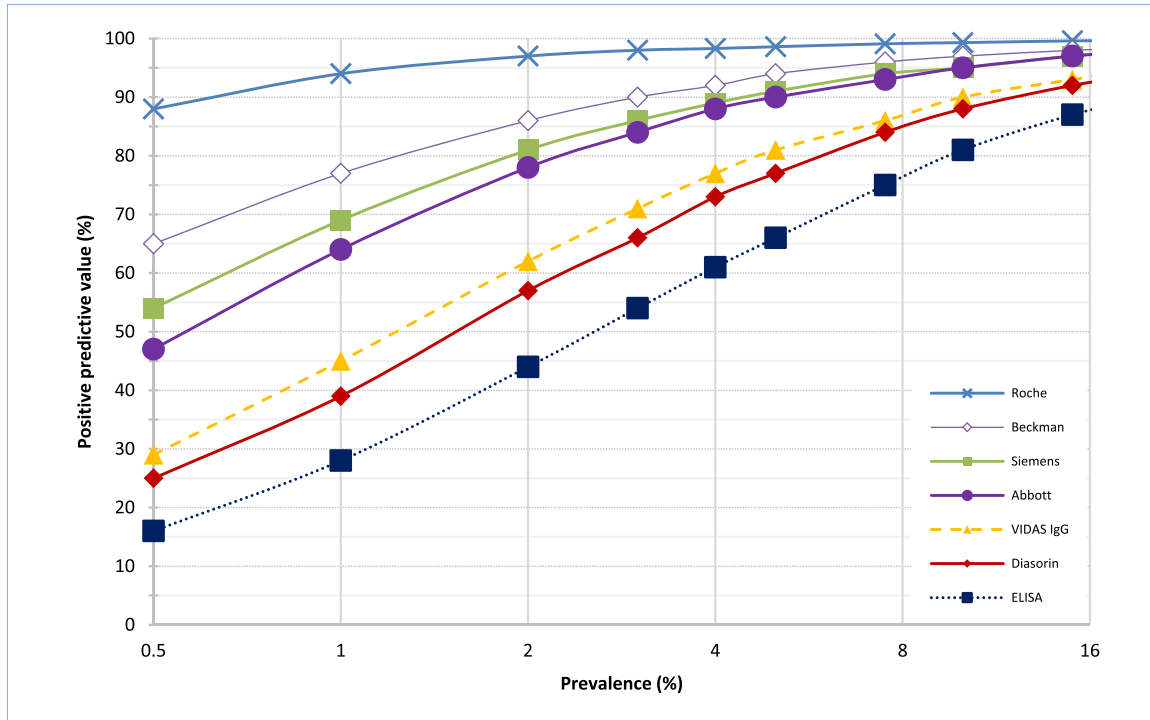


Fig. 2. Effect of SARS-CoV-2 prevalence on positive predictive values (PPV) of the different assays. The PPV of each assay was calculated and is depicted based on different SARS-CoV-2 prevalence in the population ranging between 0.5% and 15%.

the latter two results were based on substantially smaller numbers – note the wider confidence limits. The lowest performance was shown by the Diasorin assay that despite having 28 equivocal results in both groups (0.3% of the negative and 4% of the positive samples) showed 82.4% sensitivity (CI 79.0–85.3) and 98.7% specificity (CI 98.2–99.1). Another kit with low sensitivity despite high specificity was the Beckman (81.5% sensitivity (CI 74.6–86.6) and 100% specificity (CI 98.8–100)) (see Table 2).

4.3. Specificity in children

Since it was suggested that protective immunity in children might partially be due to cross-reactive immunity to past Coronavirus strains [17] we specifically investigated the specificity in children and adolescents under 20 years (688 samples which compose 28.7% of our SARS-CoV-2 negative cohort). Results show that specificity in children and adolescents were similar to those overall: Abbott 99.7% (642/644: CI 98.9–99.9), Diasorin 98.0% (633/646: CI 96.6–98.8), Vidas VIDAS IgG 97.4% (330/339: CI 95.0–98.6), Roche 100% (367/367: CI 99.0–100.0), ELISA 98.6% (351/356: CI 96.8–99.4), Beckman 100% (12/12: CI 73.5–100.0) and Siemens 100% (20/20: CI 83.2–100.0).

4.4. Seemingly narrow differences in specificity demonstrate profound effect on PPV in low prevalence scenarios

Despite the seemingly narrow differences in specificity, we calculated the effect of these differences on the Positive Predictive Values (PPV) of each assay, in prevalence scenarios lower than 15%. Such prevalence is expected in most territories, until an effective and wide vaccination program will be applied [18]. The assays' PPV were profoundly different and could reach up to 4-fold in PPV performance. Out of the commercial kits, Roche exhibited the best PPV curve with up to 88% [95% credible interval (CRI) 65–99.6] PPV in 0.5% prevalence while Diasorin demonstrated the poorest with 25% (CRI 19–32) PPV under the same conditions. The different performance in PPV was clear up until roughly 15% prevalence where the gaps in PPV diminished and all assays clustered into a narrow 15% ribbon (See Fig. 2). We then turned to determine for each kit the optimal cutoff value that optimizes total accuracy in different prevalence scenarios, using 10% cross validation. Results of this analysis suggests that total accuracy in most of the kits can benefit from higher cutoff values in very low prevalence scenarios <5% (see Supplementary Table 2).

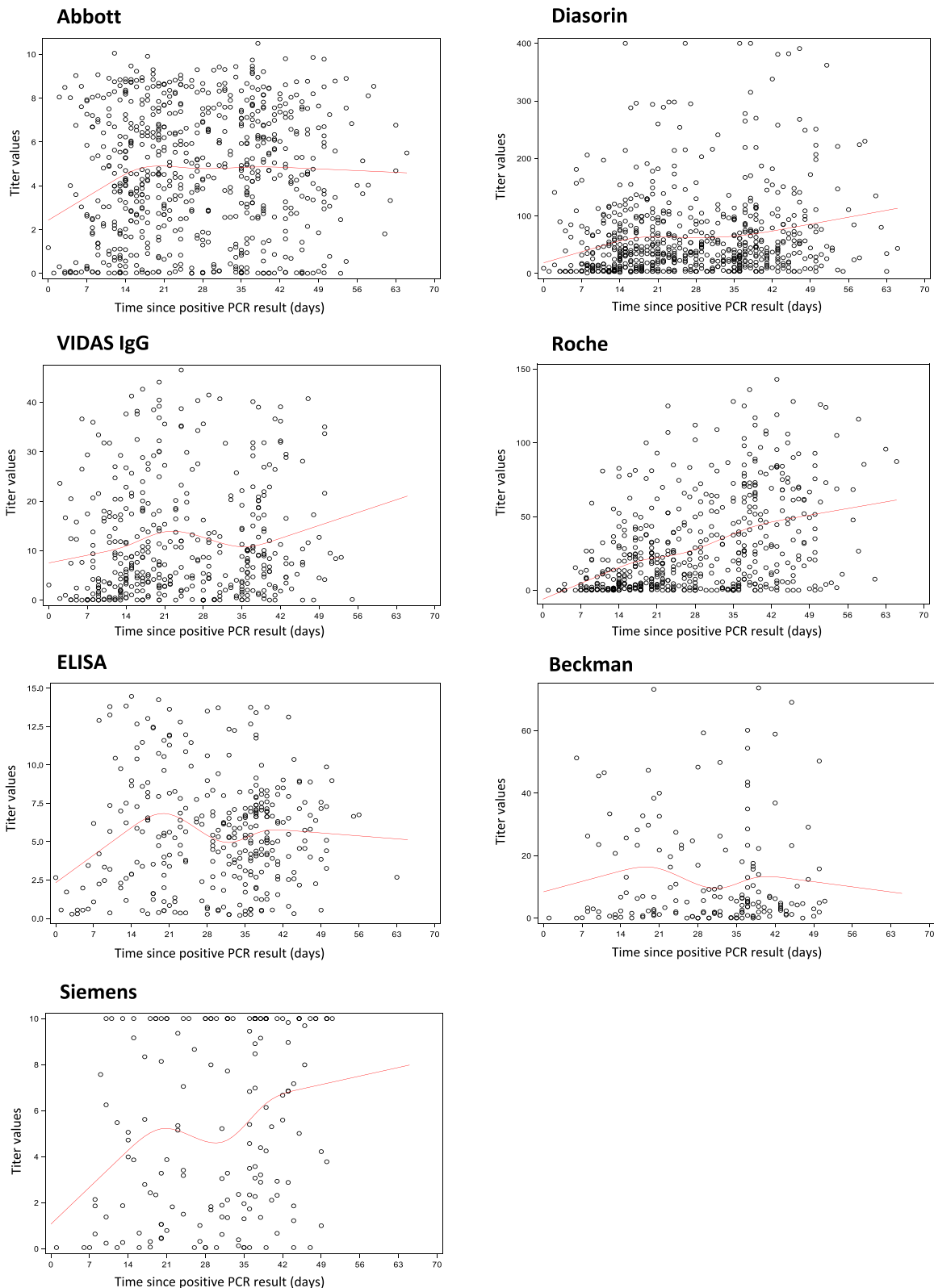


Fig. 3. SARS-CoV-2 antibody levels during the first 8 weeks after a positive PCR result. Antibody titer of each sample was recorded for each serological assay. Restricted cubic spline regression line with five knots located at the 10th, 25th, 50th, 75th and 90th percentiles are presented for each immunoassay.

4.5. Antibody titers do not decay during the first 8 weeks after exposure, and are correlated with disease severity

We recorded the antibody titer of each sample and examined the correlation between antibody titer and different cohort

characteristics. First, we compared the titers over time in the first 8 weeks after a positive PCR result for all samples with known dates post PCR test result. Such analysis of samples from multiple patients that were taken at different time points yielded an “averaged” temporal dynamics pattern. Despite some previous reports on antibody

Table 3

Test titer values versus hospitalization and severity of disease among PCR positive patients with blood samples taken ≥ 14 days after PCR result; Median titer levels (Interquartile Range).

Hospitalization	Abbott	Diasorin	VIDAS IgG	Roche	ELISA	Beckman	Siemens
Yes	5.9 (4.4)	70.0 (104.3)	24.8 (27.4)	24.4 (44.5)	6.7 (6.4)	8.4 (31.8)	6.8 (8.3)
No	4.4 (4.3)	39.3 (53.5)	6.9 (12.0)	24.8 (56.4)	4.6 (4.1)	5.1 (16.1)	5.2 (8.6)
Not known	4.5 (4.0)	39.4 (55.4)	7.9 (11.2)	22.3 (47.2)	5.7 (4.1)	4.9 (10.3)	5.6 (7.1)
Z ^a	5.10	3.89	4.37	0.65	3.87	1.07	0.44
P-value ^a	<0.0001	<0.0001	<0.0001	0.52	<0.0001	0.28	0.66
Severity							
Asymptomatic	5.0 (4.3)	39.2 (57.2)	6.6 (23.1)	11.9 (49.9)	^{-b}	^{-b}	^{-b}
Mild	5.3 (4.5)	42.3 (70.2)	8.6 (13.0)	29.8 (55.8)	5.3 (4.0)	4.9 (11.8)	5.3 (8.1)
Moderate/ Severe	6.5 (3.8)	112.0 (95.6)	25.7 (19.5)	21.9 (31.9)	8.1 (6.5)	19.7 (21.7)	9.7 (4.6)
Not known	4.2 (4.1)	39.2 (60.4)	7.3 (13.2)	6.6 (21.5)	6.3 (8.4)	6.5 (21.5)	3.6 (5.3)
Z ^c	2.06	4.16	4.19	0.24	2.48	1.73	1.96
P-value ^c	0.039	<0.0001	<0.0001	0.81	0.013	0.083	0.051

^a Mann–Whitney test statistic with P-value (excluding not knowns).

^b Less than 5 cases from which to estimate the median.

^c Cuzick trend test (excluding not knowns).

Table 4

Number and percent of negative samples from SARS-CoV-2 positive individuals in 3 or more immunoassays.

Tests Performed	Numbers negative on all tests performed	Percent
All 7 tests	4/95	4.2
6 tests	3/65	4.6
5 tests	1/70	1.4
4 tests	10/155	6.5
All combinations ≥ 4	18/385	4.7
3 tests	20/201	10.0
All combinations ≥ 3	38/586	6.5

titer decay [11], in all the assays we tested, the average antibody titer was either stable or increased in levels over a period from 2 weeks until 8 weeks after a positive PCR. The kits that presented stable levels of antibodies between 2 and 8 weeks were the Abbott, VIDAS IgG, ELISA and Beckman while increased antibody titers were observed in

the Roche, Siemens, and Diasorin assays (See Fig. 3). Importantly, the antibody titer did not vary with the type of antigen used. Subsequently, we analyzed the correlation between the assay titers and disease severity. As previously reported [12], we found the titers in Abbott, Diasorin, VIDAS IgG and ELISA to be significantly positively correlated while titers in Beckman and Siemens showed a trend for positive correlation with disease severity between asymptomatic to mild to moderate/severe patients. Similarly, we also observed significantly higher titers in hospitalized versus non-hospitalized patients (see Table 3) in Abbott, Diasorin, VIDAS IgG and ELISA immunoassays.

4.6. Roughly 1 out of 20 PCR positive patients remains seronegative despite testing with multiple assays and antigens

Some of the PCR positive SARS-CoV-2 patients that we tested were consistently seronegative despite applying various commercially available assays that utilized different antigens, and different isotypes (IgG or total antibody). We defined patients as seronegative

Table 5

Characteristics of SARS-CoV-2 seropositive and seronegative in 4 or more immunoassays populations.

	Seronegatives (n = 18)		Seropositives ≥ 14 d with ≥ 4 tests (n = 367)		P*
	N	(%)	N	(%)	
Sex:					
Male	11	(61.1%)	190	(51.8%)	0.44
Female	5	(27.8%)	145	(39.5%)	
Unknown	2	(11.1%)	32	(8.7%)	
Age, y:					
0–17	0	(0%)	8	(2.2%)	0.71
18–29	6	(33.3%)	98	(26.7%)	
30–49	4	(22.2%)	115	(31.3%)	
50–69	5	(27.8%)	59	(16.1%)	
70+	0	(0%)	8	(2.2%)	
Unknown	3	(16.7%)	79	(21.5%)	
Time since positive PCR result, d:					
14–20	5	(27.8%)	88	(24.0%)	0.46
21–27	4	(22.2%)	80	(21.8%)	
28–41	9	(50.0%)	153	(41.7%)	
42+	0	(0%)	46	(12.5%)	
Hospitalized:					
No	9	(50.0%)	130	(35.4%)	0.76
Yes	3	(16.7%)	60	(16.4%)	
Unknown	6	(33.3%)	177	(48.2%)	
Severity:					
Asymptomatic	0	(0%)	9	(2.4%)	0.76
Mild	12	(66.7%)	248	(67.6%)	
Moderate/Severe	2	(11.1%)	26	(7.1%)	
Unknown	4	(22.2%)	84	(22.9%)	

* P-value from Fisher's Exact Test (excluding the unknown category).

non-responders if they were tested with 4 or more independent assays and remained negative across all tests. We found the proportion of this subpopulation to be 4.7% of the symptomatic PCR positive patients with blood collection 14 or more days after the positive PCR result. If we loosen the criteria and take into account patients that were negative in 3 or more different assays, the proportion increases to 6.5% (See Table 4). We then compared the seronegative subpopulation against the seropositive patients to search for potential unique characteristics, but found no clear differences in age, hospitalization, disease severity, time from positive PCR and gender (see Table 5), although numbers are too small to rule out such differences. This measurement represents the first population-wide estimate using multiple commercially available tests and different antigens to quantify the proportion of seronegative non-responding patients at risk for re-infection.

5. Discussion

This study was designed and executed during March–July 2020 by the national serology validation taskforce, to test the performance and limitations of multiple commercially available automated immunoassay kits. During normal days, a validation process of such magnitude is not required due to the rigorous methodologies and approval process that are being applied by the FDA in approving in-vitro diagnostics (IVD). These regulatory processes, preceded by meticulous work of highly experienced teams within the large IVD companies, allow the end users in labs across the world to be confident in the performance declared by each vendor. During the SARS-CoV-2 pandemic, an urgent need for testing emerged, which had to be satisfied according to an extremely challenging timeline never previously required. This was exacerbated by the lack of sufficient samples to support verification and validation. The FDA approved an emergency use authorization process (EUA), which was necessary, yet, lacked the rigor of the classical procedures and thus decreased confidence in assays' performance [2]. Indeed, many kits with EUA were later retracted by the vendors or banned by the FDA due to diminished performance [19]. The state of Israel, like numerous other countries, signed supply contracts with multiple IVD vendors to get access to diagnostic testing once launched [20].

To fill the performance gap, a national team of experts with a wide range of expertise was established, to methodologically evaluate the performance of the different kits (see protocol as Supplementary Annex. 1). The team included experts in virology, clinical chemistry, immunology, IVD development and validation, and biostatisticians, from various sectors including industry, government, HMOs, and public and private research institutes. The work was split between five labs and included testing the serology assays manufactured by Roche, Abbott, BioMerieux, Beckman Coulter, Siemens, and an in-house RBD ELISA assay. We tested 2391 negative samples that were collected prior to September 2019, aged 0–103y that closely represented the distribution of ages within the general population and resembles a unique cohort due to the large number of pediatric samples composed of 222 (9.3%), 149 (6.2%) and 317 (13.2%) samples from children and adolescent age 0–6, 6–10 and 11–20 years old, respectively. Since seasonal Human Coronavirus (HCoVs) infections is prevalent in Israel with more than 10% of Influenza-like illness molecularly diagnosed annually as HCoVs [21] and antibodies against HCoVs are widespread in human sera [22] cross-reactivity with HCoVs antibodies should have been recognized by evaluating antibodies against SARS-CoV-2 in the Israeli non-COVID-19 population tested in this study. In light of the mild disease course in children, where less than 1% of infected cases were <10 years and ~3.5% of SARS-CoV-2 infected children had lymphocytopenia [23], one hypothesis suggested that protective immunity in children might partially be due to cross-reactive immunity to past Coronavirus strains. Our measurements of the negative cohort did not yield

significantly more false positive results in the pediatric versus adult population, thus largely excluding this hypothesis. The positive cohort that we tested included 698 SARS-CoV-2 PCR positive patients with different degrees of severity, time from positive PCR result, age etc. The 588 samples obtained ≥ 14 days after first positive PCR allowed evaluation of the true sensitivity of the different kits in comparison to other studies examining specific assays and smaller cohorts [3, 4]. All samples were collected between March 28th and May 24th, 2020. A comparison of the seven assays revealed a range of sensitivities and specificities, where some of these, deviated from the declared performance in the product inserts published in the FDA EUA (<https://www.fda.gov/medical-devices/emergency-situations-medical-devices/emergency-use-authorizations-medical-devices>). The most sensitive commercially available kits were Roche and VIDAS IgG, while the least sensitive were Diasorin and Beckman. Despite the seemingly narrow differences in specificity, the impact of such differences on PPV in low prevalence settings can be dramatic. The global prevalence of previous exposure to SARS-CoV-2 mostly vary between 0 and 5% [24, 25]. These relatively low numbers are not expected to significantly increase, considering the social distancing measures that are being taken, until a successful immunization will be widely applied [18]. We therefore estimated the PPV of each assay under the different prevalence scenarios ranging between 0.5–15% and were able to show profound differences as high as 4-fold between different diagnostic products. Such differences might impact the assay of choice for epidemiological assessments, or alternatively, must be factored in, while performing such studies. Importantly, we are aware of assays that were evaluated by our colleagues around the globe reporting real world data of 80% specificity or less, especially for lateral flow rapid assays, like the ones that were purchased by the UK government [26]. Such performance represents 5% PPV in a 1% prevalence scenario, which makes these tests a poor alternative for epidemiological studies and might cause unnecessary psychological distress for patients.

We then analyzed the factors that correlated with antibody titers among different patients. A few groups suggested that the half-life time of antibodies to certain SARS-CoV-2 proteins might be as low as 8 weeks, and might reflect a risk to lose protective immunity [27] and to potentially underestimate the number of infected individuals. We therefore analyzed the average antibody titer levels among individuals that were sampled at different time points after positive PCR results. Such analysis, although lacking the antibody titer dynamics of a certain individual, is more representative of the overall antibody response behavior within a wide population. Since we used 7 different assays and different patients for each time point, we could factor out elements such as antigen differences, assay performance, and inter-patient variability. The patterns we observed showed a stable antibody titer after 14 days and up to 8 weeks after positive PCR, in 4 out of 7 kits (Abbott, VIDAS IgG, ELISA and Beckman) while in 3 out of 7 kits, the Roche, Siemens, and Diasorin we observed an increase in titers up to 8 weeks post infection. These results contradict previous reports that suggested a short half-life of SARS-CoV-2 antibodies [27] and support long-term immunity after infection as was shown in other studies [28]. In addition, we observed strong correlation between antibody titers and disease severity. In 4 kits (Abbott, Diasorin, VIDAS IgG and the ELISA), the titers showed a significant increase from asymptomatic to mild to moderate/severe cases, while in 2 other kits (Beckman and Siemens) this increase did not reach statistical significance. Accordingly, a similar pattern of higher titers in hospitalized compared with non-hospitalized patients was detected, further supporting this observation. Interestingly, only Roche, which detects total Ig against the N antigen, demonstrated no difference in titer levels between severe and non-severe cases but a titer increase over time. The differences in antibody titers, kinetics and correlation with disease severity observed among the immunoassays tested here may reflect

heterogeneity of antibodies against N, S and RBD antigens [29], cumulative or synergistic binding of the different IgA, IgM and IgG antibody isotopes [30], or assay specific determinants. Future studies should investigate the long-term kinetics of antibodies against each antigen, the correlation with neutralization capacity and the overall prognostic value of each assay.

Overall, these data suggest that an ineffective humoral response is not a potential reason for severe disease. More investigation is needed to understand whether increased antibody production is the cause or the effect of more severe disease.

After measuring the serological response among 588/698 PCR positive patients, who were tested ≥ 14 days post infection, we observed one specific subpopulation of patients who remained seronegative despite repeated testing with different assays utilizing different antigens, isotopes, and methods. Previous anecdotal reports about this phenomenon were conflicting, ranging between no seronegative patients [12], to over 20% [11]. Our findings suggest that the number is about 1 out of 20 patients, with 4.7% being negative in 4 or more tests, and 6.5% in 3 or more tests. We did not find any clear correlation between seronegative non-responders and age, gender, time from positive PCR, disease severity or hospitalization, although the number of seronegative persons is small. To the best of our knowledge, this is the first multi-test, multi-antigen, population-wide quantification of seronegative patients that can support the real magnitude of this problem. No seroconversion, and the potential for re-infection among individuals is attracting significant media and public concerns and also potentially impacts herd immunity and the burden on the healthcare system. Our findings suggest that approximately 5% of patients fall into this category and their impact on herd immunity or the healthcare system is expected to be limited; yet, these individuals are at increased risk for re-infection. Indeed, in recent study of the first characterized definite SARS-CoV-2 re-infection case, no antibody response was found after the first infection episode [31]. As T cells also play an important role in immunity to SARS-CoV-2 [32] and may be responsible for protection against re-infection despite lack of antibodies against SARS-CoV-2, further studies are required to better identify this subpopulation in order to effectively protect it. Seronegative non-responders are also expected at a certain proportion after immunization and the proportion of vaccine non-responders was traditionally higher than the natural disease itself [33]. According to the data presented here within the Israeli population, we therefore estimate that the proportion of vaccination failure may be higher than 5% regardless of the vaccine identity.

Despite the unprecedented high numbers of samples and assays that were tested in this study of both SARS-CoV-2 positive and negative individuals, a limitation of this study is that not all assays were tested with all the samples, due to insufficient sample volume and that no longitudinal sera were analyzed for the antibody titer profiles over time. Nevertheless, our dataset was large enough to draw statistically solid conclusions regarding performance and antibody kinetics. In addition, since most SARS-CoV-2 cases in the general population are mild, our “all comers” cohort is mostly composed of mild infections and has a lower proportion of moderate and severe cases, despite their higher clinical importance. Finally, all the samples tested here were obtained within the Israeli population. It is possible that different background immunity status across geographical regions may impact the specificity and sensitivity of immunoassays and therefore other large scale studies from different regions of the world are urgently needed.

In summary, we performed a nationwide validation process for seven commercially available automated serology testing and found significant performance differences with major impact on PPV values in low prevalence scenarios. In addition, we were able to largely exclude prior cross-reactive immunity among children by past exposure to other corona strains. Our findings support correlation between antibody titers and disease severity and contradict evidence

for decreased titers with time and short half-life. We quantify in a population-wide, multi-assay fashion, the number of seronegative, non-responding patients at risk for re-infection and estimate that to be roughly 5%.

6. Funding

Israel Ministry of Health provided the Abbot and Diasorin kits.

7. Authors' contributions

Kfir Oved performed literature search, designed the study, analyzed the results, contributed to the analysis and wrote the article. Liraz Olmer and Laurence S. Freedman designed the study, analyzed the results interpreted the data and contributed to writing the article. Yonat Shemer-Avni, Tamar Wolf, Lia Supino-Rosin, George Prajgrod, Yotam Shenhar, Irina Payorsky, Yuval Cohen, Yishai Kohn, Victoria Indenbaum, Rachel Lazar, Valeria Geylis, Michal Tepperberg Oikawa, Eilat Shinar, Evgeniy Stoyanov, Lital Keinan-Boker, Ravit Bassal, Shay Reicher, Ruti Yishai, Adina Bar-Chaim, Ram Doolman and Yoram Reiter performed the data collection and some data interpretation. Ella Mendelson and Zvi Livneh reviewed the study design and the article. Yaniv Lustig performed literature search, designed the study, collected the data, contributed to the analysis and to writing the paper.

Data sharing statement

De-identified participant data will be available upon request to the corresponding author following publication.

Declaration of Competing Interest

YL is supported by the Nehemia Rubin Excellence in Biomedical Research – The TELEM Program of Chaim Sheba Medical Center. All other authors have nothing to declare.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi: [10.1016/j.eclinm.2020.100651](https://doi.org/10.1016/j.eclinm.2020.100651).

References

- [1] Lisboa Bastos M, Tavaziva G, Abidi SK, Campbell JR, Haraoui LP, Johnston JC, et al. Diagnostic accuracy of serological tests for covid-19: systematic review and meta-analysis. *BMJ* 2020;370:m2516.
- [2] Theel ES, Slev P, Wheeler S, Couturier MR, Wong SJ, Kadkhoda K. The role of antibody testing for SARS-CoV-2: is there one? *J Clin Microbiol* 2020;58(8).
- [3] Charlton CL, Kanji JN, Johal K, Bailey A, Plitt SS, MacDonald C, et al. Evaluation of six commercial mid to high volume antibody and six point of care lateral flow assays for detection of SARS-CoV-2 antibodies. *J Clin Microbiol* 2020;58(10): e01361–20. doi: [10.1128/JCM.01361-20](https://doi.org/10.1128/JCM.01361-20).
- [4] Kohmer N, Westhaus S, Ruhl C, Ciesek S, Rabenau HF. Brief clinical evaluation of six high-throughput SARS-CoV-2 IgG antibody assays. *J Clin Virol* 2020;129:104480.
- [5] Montesinos I, Gruson D, Kabamba B, Dahma H, Van den Wijngaert S, Reza S, et al. Evaluation of two automated and three rapid lateral flow immunoassays for the detection of anti-SARS-CoV-2 antibodies. *J Clin Virol* 2020;128:104413.
- [6] Ishai R. New updates on serological testing for determining the presence of antibodies against SARS-CoV-2 (Hebrew). Laboratory Division I MoH.; 2020. URL: <https://oemmnrcbldboiebnladdacbfmadadm/> <https://govextra.gov.il/media/27425/bz-276288420.pdf>.
- [7] Lustig Y, Keler S, Kolodny R, Ben-Tal N, Atias-Varon D, Shlush E, et al. Infectious antigenic cross-reactivity between SARS-CoV-2 and dengue viruses. *Clin Infect Dis* 2020.
- [8] Grimes DA, Schulz KF. Uses and abuses of screening tests. *Lancet* 2002;359(9309):881–4.
- [9] Klein SL, Pekosz A, Park HS, Ursin RL, Shapiro JR, Benner SE, et al. Sex, age, and hospitalization drive antibody responses in a COVID-19 convalescent plasma donor population. *J Clin Invest* 2020;130(11):6141–50. doi: [10.1172/JCI142004](https://doi.org/10.1172/JCI142004).

- [10] Lynch KL, Whitman JD, Lacanienta NP, Beckerdite EW, Kastner SA, Shy BR, et al. Magnitude and kinetics of anti-SARS-CoV-2 antibody responses and their relationship to disease severity. *Clin Infect Dis* 2020 Online ahead of print. doi: 10.1093/cid/ciaa979.
- [11] Long QX, Tang XJ, Shi QL, Li Q, Deng HJ, Yuan J, et al. Clinical and immunological assessment of asymptomatic SARS-CoV-2 infections. *Nat Med* 2020;26(8):1200–4.
- [12] Long QX, Liu BZ, Deng HJ, Wu GC, Deng K, Chen YK, et al. Antibody responses to SARS-CoV-2 in patients with COVID-19. *Nat Med* 2020;26(6):845–8. doi: 10.1038/s41591-020-0897-1.
- [13] Amanat F, Stadlbauer D, Strohmeier S, Nguyen THO, Chromikova V, McMahon M, et al. A serological assay to detect SARS-CoV-2 seroconversion in humans. *Nat Med* 2020;26(7):1033–6. doi: 10.1038/s41591-020-0913-5.
- [14] van de Schoot R, Broere JJ, Perryck KH, Zondervan-Zwijnenburg M, van Loey NE. Analyzing small data sets using Bayesian estimation: the case of posttraumatic stress symptoms following mechanical ventilation in burn survivors. *Eur J Psychotraumatol* 2015;6:25216.
- [15] Efron B. The Jackknife, the bootstrap, and other resampling plans. CBMS-NSF regional conference series in applied mathematics; Philadelphia, PA.
- [16] Cuzick J. A Wilcoxon-type test for trend. *Stat Med* 1985;4(1):87–90.
- [17] Fierz W, Walz B. Antibody dependent enhancement due to original antigenic sin and the development of SARS. *Front Immunol* 2020;11:1120.
- [18] Scudellari M. How the pandemic might play out in 2021 and beyond. *Nature* 2020;584(7819):22–5.
- [19] FDA. Coronavirus (COVID-19) update: fda revokes emergency use authorization for chembio antibody test. U.S. Food and Drug Administration; 2020. URL <https://www.fda.gov/news-events/press-announcements/coronavirus-covid-19-update-fda-revokes-emergency-use-authorization-chembio-antibody-test>.
- [20] Magid J, Bachner M. Health Ministry announces purchase of 2.4 million antibody tests. *Times Israel* 2020 17 April 2020 URL https://www.timesofisrael.com/liveblog_entry/health-ministry-announces-purchase-of-2-4-million-antibody-tests/.
- [21] Friedman N, Alter H, Hindiyeh M, Mendelson E, Shemer Avni Y, Mandelboim M. Human coronavirus infections in israel: epidemiology, clinical symptoms and summer seasonality of HCoV-HKU1. *Viruses* 2018;10(10):515. doi: 10.3390/v10100515.
- [22] Premkumar L, Segovia-Chumbez B, Jadi R, Martinez DR, Raut R, Markmann A, et al. The receptor binding domain of the viral spike protein is an immunodominant and highly specific target of antibodies in SARS-CoV-2 patients. *Sci Immunol* 2020;5(48):eabc8413. doi: 10.1126/sciimmunol.abc8413.
- [23] Lu X, Zhang L, Du H, Zhang J, Li YY, Qu J, et al. SARS-CoV-2 Infection in Children. *N Engl J Med* 2020;382(17):1663–5.
- [24] Bogogiannidou Z, Vontas A, Dadouli K, Kyritsi MA, Soteriades S, Nikoulis DJ, et al. Repeated leftover serosurvey of SARS-CoV-2 IgG antibodies, Greece, March and April 2020. *Euro Surveill* 2020;25(31).
- [25] Pollan M, Perez-Gomez B, Pastor-Barriuso R, Oteo J, Hernan MA, Perez-Olmeda M, et al. Prevalence of SARS-CoV-2 in Spain (ENE-COVID): a nationwide, population-based seroepidemiological study. *Lancet* 2020;396(10250):535–44.
- [26] Kirkpatrick DD, Bradley JUK. Paid \$20 million for new coronavirus tests. They didn't work.. *New York Times*. 2020 May 7, 2020 URL <https://www.nytimes.com/2020/04/16/world/europe/coronavirus-antibody-test-uk.html>.
- [27] Bruni M, Cecatiello V, Diaz-Basabe A, Lattanzi G, Mileti E, Monzani S, et al. Persistence of Anti-SARS-CoV-2 antibodies in non-hospitalized COVID-19 convalescent health care workers. *J Clin Med* 2020;9(10):3188. doi: 10.3390/jcm9103188.
- [28] Wajnberg A, Amanat F, Firpo A, Altman DR, Bailey MJ, Mansour M, et al. Robust neutralizing antibodies to SARS-CoV-2 infection persist for months. *Science* 2020.
- [29] McAndrews KM, Dowlatshahi DP, Dai J, Becker LM, Hensel J, Snowden LM, et al. Heterogeneous antibodies against SARS-CoV-2 spike receptor binding domain and nucleocapsid with implications for COVID-19 immunity. *JCI Insight* 2020;5(18):e142386. doi: 10.1172/jci.insight.142386.
- [30] Chen Y, Tong X, Li Y, Gu B, Yan J, Liu Y, et al. A comprehensive, longitudinal analysis of humoral responses specific to four recombinant antigens of SARS-CoV-2 in severe and non-severe COVID-19 patients. *PLoS Pathog* 2020;16(9):e1008796.
- [31] To KK, Hung IF, Ip JD, Chu AW, Chan WM, Tam AR, et al. COVID-19 re-infection by a phylogenetically distinct SARS-coronavirus-2 strain confirmed by whole genome sequencing. *Clin Infect Dis* 2020 Online ahead of print. doi: 10.1093/cid/ciaa1275.
- [32] Braun J, Loyal L, Frentsch M, Wendisch D, Georg P, Kurth F, et al. SARS-CoV-2-reactive T cells in healthy donors and patients with COVID-19. *Nature* 2020 Online ahead of print. doi: 10.1038/s41586-020-2598-9.
- [33] Wiedermann U, Garner-Spitzer E, Wagner A. Primary vaccine failure to routine vaccines: why and what to do? *Hum Vaccin Immunother* 2016;12(1):239–43.