COVID 19 Serology National Validation Plan

1. SCOPE

1.1. In-Scope

The purpose of this protocol is to partially verify and validate newly developed COVID-19 serology tests that did not pass through a full authorization cycle by the FDA and were not tested on a wide population in the State of Israel (which may have its own characteristic benign coronavirus epidemiology). Importantly, this protocol is primarily focused on evaluating the performance of the different kits for the purpose of quantifying the level of past exposure among the Israeli population and allow healthcare organizations, policy makers, and others, to get reliable information about prevalence in different territories and populations, herd immunity, as well as epidemics dynamics. It is important to note that the primary goal is not to validate and verify the use of serology for the diagnosis and management of an individual patient (see 1.2 Out-of-scope), yet, we believe that some of the data collected in this protocol will support the management of individual patients and allow a better understanding of the different serological characteristics of COVID-19.

1.2. Out-of-Scope

The following topics are not aimed to be fully covered in this protocol:

- This protocol is not aimed to evaluate the utility of serology as a diagnostic tool at the individual patient level, neither as a patient management tool. For example, dynamics of titers, isotype switch, IgM serology, and neutralization are out of scope. For the same reason, we will prioritize IgG over IgA. However, we believe that the data that will be gathered, can be used by different stakeholders to better learn the utility of serology in patient diagnosis and management in the future.
- Reagent stability (tests, controls, and calibrators) will not be tested.
- Calibrator traceability between the different kits will not be tested.

2. MATERIALS FOR VALIDATION STUDIES

2.1. General

The validation plan will be executed using the Test system configuration and based on the manufacturer instructions. An updated SW version will be utilized, unless instructed differently.

2.2. Analyzers

Serial numbers of the different analyzers will be documented in the analytical validation report. All maintenance procedures will be performed according to the operation manual. Each lab will be responsible for documenting the device type, serial number, exact model, and proper calibration. The Validation will include the following devices:

- Diasorin Liaison XL
- Biomerieux VIDAS (all models, will not be tested in Maccabi Lab)
- Abbott Architect 2000
- Roche Elecsys
- Siemens ADVIA
- Beckman Coulter access

2.3. Cartridges

The validation will be conducted using cartridges from at least one production lot. However, the lot to lot variability is outside of the scope of this study. Lot number and expiration date will be documented as stated on the cartridge label.

2.4. Calibrators and Controls

The validation will be conducted using calibrators that were verified and supplied by the vendors and used based on their instructions. Each lab will run a control for each kit at the beginning of every working day and will record the reading obtained.
Due to the lack of external controls, the Central Virology Lab will send 4 representative samples that will serve as clinical controls with set values that can be compared between labs.

2.5. Laboratories

The following five laboratories will participate in the validation study:

- Meuheded Central Lab
- Maccabi Mega Lab
- Clalit Soroka Virology Lab
- Leumit Central Lab
- The Central Virology laboratory, ministry of health, Sheba Medical Center**.

** The samples will be tested on instruments located at the Mega-Lab and Endocrinology institute in the laboratories wing of Sheba Medical Center.

2.6. Definition of Positive, Negative and Equivocal serological results

The definition of positive, negative, and equivocal results will follow that recommended by the manufacturer in the insert that they provide.

SAMPLES USED AND POPULATION CHARACTERISTICS

2.7. Not more than one sample from the same individual will be included.

2.8. Negative samples

In each laboratory, a minimum number of samples will be tested.

- The Central Virology laboratory – Sheba will test >1000 negative samples
- Maccabi Mega Lab will test >200 negative samples (No children samples)
- Clalit Soroka Virology Lab will test >400 negative samples
- Meuhedet central lab will test >200 negative samples from Jerusalem/north.
- Leumit Central Lab will test >200 negative samples

Special controls:

- Acute CMV (n=60) (Soroka 30 and Maccabi 30).
- West Nile Virus (IgG) (n=30) (Central Virology Laboratory).
- Dengue (IgG) (n=30) (Central Virology Laboratory).

All the negative samples had been collected between January 2014 to September 2019. We assume no exposure to the virus was possible before autumn 2019.
The following information on each sample will be recorded: sample source (hospital, kupat cholim, MDA) month and year of sample, age of patient, gender, area of residence. The result of the test will be recorded as *IgG level and result (positive, equivocal or negative)*. Stratification of the sample by age, and by area of residence (where possible), will be performed as detailed below:

Age distribution (based on the MOH statistics): the final two columns are for the numbers to be entered by the Central Viral Laboratory

Age group	%	Central - Sheba		Maccabi,Meuhedet & Leumit*		Clalit	
		n_{min}	n_{max}	n_{min}	n_{max}	n_{min}	n_{max}
0-1	6%	60	72	11	13	24	27
2-5	6%	60	72	11	13	24	27
6-10	6%	60	72	11	13	24	27
11-20	15%	150	180	27	33	60	67
21-30	15%	150	180	27	33	60	67
31-40	13%	130	156	23	29	52	59
41-50	12%	120	144	22	26	48	54
51-60	11%	110	132	20	24	44	50
61-70	8%	80	96	14	18	32	36
> 71	8%	80	96	14	18	32	36
Total	100%	1000	1200	180	220	400	450

* The numbers and age distribution of that sample provided by Maccabi, Leumit and Meuhedet will serve as an ideal goal, yet, might be distributed differently.

Counties (based on the MOH statistics): the final two columns are for the numbers to be entered by the Central Viral Laboratory

Counties	%	Central - Sheba		Maccabi,Meuhedet & Leumit*		Clalit	
		n_{min}	n_{max}	n_{min}	n_{max}	n_{min}	n_{max}
South	15%	150	180	27	33	60	67
Central	26%	260	312	47	57	104	117
TLV Rabati	17%	170	204	31	37	68	77
Jerusalem	13%	130	156	23	29	52	58
North	17%	170	204	31	37	68	77
Haifa	12%	120	144	21	27	48	54
Total	100	1000	1200	180	220	400	450

* The numbers and age distribution of that sample provided by Maccabi, Leumit and Meuhedet will serve as an ideal goal, yet, might be distributed differently.

- The gender distribution will be 1:1 between males and females with up to 10% allowed deviation, i.e. between 2:3 and 3:2.
- In case allowed by the manufacturer, we will use both serum and plasma samples, however, most samples are expected to be serum.
- Pediatric samples will not necessarily be distributed uniformly due to lack of samples, and, if possible, the three subgroups 0-1, 2-5, and 6-10 will be selected in equal numbers.

Repeat testing: For samples that test positive on negative cohort samples, the test will be repeated, and the second test result also recorded.

2.9. Positive samples

In each laboratory, a minimum number of patients who have been tested positive for SARS-CoV-2 RNA will be measured.

- Meuhedet Central Lab will test 100 samples
- Maccabi Mega Lab will test ~80 patients (Dan Hotel – Asymptomatic/Mild)
- Soroka Central Virology Lab will test 50 samples
- The Central Virology laboratory – Sheba will test between 200-350 samples
- Leumit central lab will test 50 samples

All patients considered positive for the purpose of this study will be tested for IgG. Therefore, we will include patients with samples collected **at least 14 days** after the day of a positive PCR test result.

We will add a separate additional cohort of patients with samples collected **less than 14 days** after a positive PCR test result for the purpose of assessing seroconversion dynamics and assay sensitivity.

- Meuhedet Central Lab will test between 30-50 samples
- Maccabi Mega Lab will test 20-40 patients (Dan Hotel – Asymptomatic/Mild)
- Soroka Central Virology Lab will test a few samples
- The Central Virology laboratory – Sheba will test ~ 50 samples

2.9.1. Baseline variables (parameters that must be available for >90% of samples):

The following information on each sample will be recorded:

- Sample source (Hospital, Kupat Cholim, MDA)
- Date of sample collection
- Age of patient
- Gender
- County
- Severity of symptoms (none, mild, moderate, severe)
- Time from symptoms onset
- Hospitalization (y/n), if yes – length of stay (days)
- Test results positive, equivocal, or negative.

Stratification of the sample by severity of symptoms will be performed as below.

- **Symptomatic – mild**: PCR positive patients with signs symptoms typical for COVID-19 infection that did not require medical care or hospitalization. "Melonit" cases are included in the "mild" unless they required further hospitalization.
- **Symptomatic – moderate**: PCR positive patients with signs symptoms typical for COVID-19 infection that required hospitalization without respiratory or hemodynamic support (Not ventilated or ECMO assisted)
- **Symptomatic – severe**: PCR positive patients with signs symptoms typical for COVID-19 infection that required hospitalization AND at least 24 hours of respiratory (ventilation) and/or hemodynamic support.
- **Asymptomatic**: PCR positive individuals without any reported signs or symptoms of an acute infectious disease.

Meuhedet Central Lab:

Type	Percentage	n_{min}	n_{max}
Symptomatic – mild	30%		NA
Symptomatic – moderate	30%		NA
Symptomatic – severe	30%		NA
Asymptomatic	10%		NA

Soroka Central Virology Lab:

Type	Percentage	n_{min}	n_{max}
Symptomatic – mild	30%		NA
Symptomatic – moderate	30%		NA
Symptomatic – severe	30%		NA
Asymptomatic	10%		NA

Maccabi Mega Lab:

Type	Percentage	n_{min}	n_{max}
Symptomatic – mild	~50%*		NA
Symptomatic – moderate	~50%*		NA
Symptomatic – severe			NA
Asymptomatic			NA

* The percentage of moderate symptomatic patients provided by Maccabi will serve as an ideal goal, yet might be distributed differently

The Central Virology laboratory

Type	Percentage	n_{min}	n_{max}
Symptomatic – mild	30%	75	NA
Symptomatic – moderate	30%	75	NA
Symptomatic – severe	30%	75	NA
Asymptomatic	10%	25	NA

Leumit Central laboratory

Type	Percentage	n_{min}	n_{max}
Symptomatic – mild	35%		NA
Symptomatic – moderate	35%		NA
Symptomatic – severe			NA
Asymptomatic	35%		NA

Repeat testing: For samples that test negative on positive cohort samples, the test will be repeated, and the second test result also recorded. After repeat testing has been performed on 30 such false negatives, repeat testing may be stopped.

2.10. Samples storage, handling, and labeling

- HMOs and Central virology laboratory are treating all samples as positive, take the relevant safety measures and do not perform neither RT-PCR nor heat Inactivation.
- The samples collected for this study, will be defined as a **National COVID-19 Biobank** and will be stored in the National Virology Laboratory in Tel-Hashomer hospital and the different labs participating in the validation. The samples will be collected, labeled with an internal sample number, aliquoted if necessary and store in -80C for long term preservation.

3. LIMIT OF DETECTION AND QUANTITATION

Serology testing is a challenging assay to determine limited of detection and qualification due to multiple reasons. In addition, even LOD and LOQ values should be carefully compared between kits due to the lack of harmonization. For the purpose of this study, we will use the manufacturer declarations as stated in the product "Instructions For Use" (IFU), if available. Actual determination of LOD and LOQ and outside of the scope of this protocol.

As for the tested kits, the stated LOD and LOQ levels are as follows:

Vendor	Product	Stated LOD	Stated LOQ
Diasorin	LIAISON® SARS-CoV-2 S1/S2 IgG	3.8 AU/ml	NA
Abbott	Architect 2000 SARS-CoV-2 IgG	ND	ND
Biomerieux	VIDAS® SARS-COV-2 IgG	ND	ND
Roche	Elecsys® Anti-SARS-CoV-2	ND	ND
Siemens	SARS-CoV-2 Total (COV2T)	ND	ND
Beckman Coulter	Access SARS-CoV-2 IgG assay	ND	ND

4. PRECISION/REPRODUCIBILITY

Reproducibility and precision should be assessed between day, user, device and lots. For the purpose of this study, we will rely partly on the manufacturer declarations as stated in the product IFU, if available. Determination of intra-day, intra-device and other reproducibility assessments are outside of the scope of this protocol. However, we will assess inter-day and inter-laboratory reproducibility through repeat assessments of manufacturer-provided controls at each laboratory at the beginning of each day and by testing the same 4 COVID 19 positive samples. Between-day within-laboratory variation and between-laboratory variation hwill be reported as standard deviations.

As for the tested kits, the stated Precision results:

Vendor	Product	Intra-Run	Inter-device	Intra-day	Overall
Diasorin	LIAISON® SARS-CoV-2 S1/S2 IgG	3.17%	1.53%	2.78%	4.98%
Abbott	Architect 2000 SARS-CoV-2 IgG	3.5%	ND	ND	3.55%
Biomerieux	VIDAS® SARS-COV-2 IgG	5.5%	ND	ND	8.9%
Roche	Elecsys® Anti-SARS-CoV-2	ND	ND	ND	ND
Siemens	SARS-CoV-2 Total (COV2T)	5.9%	8.8%	ND	ND
Beckman Coulter	Access SARS-CoV-2 IgG assay	ND	ND	ND	ND

5. INTERFERENCE / CROSS REACTIVITY

Some of the tested kits were evaluated for interference and cross reactivity. The kits that were not tested, might suffer from inferior performance, especially since the most affected population are elderly patients with multiple co-morbidities and multi-pharmacy. However, verification of this aspect is outside of the scope of this study. The following substances and infectious agents were tested by the vendors:

Vendor	Product	Interferant + concentration
Diasorin	LIAISON® SARS-CoV-2 S1/S2 IgG	Biotin 3500 ng/mL Triglycerides 3000 mg/dL Hemoglobin 1000 mg/dL Unconjugated bilirubin 40 mg/dL Conjugated bilirubin 40 mg/dL Cholesterol total 400 mg/dL Paracetamol 500 mg/mL Ibuprofen 500 mg/mL ACE inhibitors – Not tested Beta blockers – Not tested AT2 blockers – Not tested
Abbott	Architect 2000 SARS-CoV-2 IgG	ND
Biomerieux	VIDAS® SARS-COV-2 IgG	ND
Roche	Elecsys® Anti-SARS-CoV-2	Biotin ≤ 4912 nmol/L or ≤ 1200 ng/m Hemolysis, bilirubin, RF, and pharmaceutical compounds other than biotin have not been tested and an interference cannot be excluded
Siemens	SARS-CoV-2 Total (COV2T)	Hemoglobin 1000 mg/dL Bilirubin, conjugated 40 mg/dL Bilirubin, unconjugated 40 mg/dL Triglycerides (Intralipid) 2000 mg/dL Biotin 3500 ng/mL
Beckman Coulter	Access SARS-CoV-2 IgG assay	Bilirubin (conjugated) 43 mg/dL Bilirubin (unconjugated) 43 mg/dL Hemoglobin 300 mg/dL Triglycerides (Triolein) 1,500 mg/dL

Vendor	Product	Cross reactant (Positive/Total)
Diasorin	LIAISON® SARS-CoV-2 S1/S2 IgG	Anti-nuclear autoantibodies (ANA) 0/10 Anti-HBV 1/10 Anti-HCV 0/10 Anti-Influenza A 1/10 Anti-Influenza B 0/10 Anti-respiratory syncytial virus 0/10 Anti-borrelia burgdorferi 0/10 Anti-Mycoplasma pneumoniae 0/10 Anti-EBV antibodies 0/10 Anti-CMV 0/10 Anti-HSV1/HSV2 0/10 HAMA 0/10 Anti-Parvovirus B19 0/10 Rheumatoid factor 1/10 Anti-Rubella 0/10 Anti-VZV 0/10 Anti-Human CoV OC43 0/3 Anti-Human CoV HKU1 0/1 Anti-Human CoV unknown strain 0/4
Abbott	Architect 2000 SARS-CoV-2 IgG	Adenovirus 0/5 Antinuclear Antibody 0/5 Autoimmune Hepatitis 0/5 Cytomegalovirus (CMV) IgG 1/5 CMV Immunoglobulin Class M (IgM) 0/5 dsDNA Antibody 0/5 Epstein-Barr Virus (EBV) IgG 0/5 EBV IgM 0/5 Escherichia coli (E. coli) Antibody 0/5 HAMA 0/5 Hemodialysis Patients 0/5 Hepatitis A Virus (HAV) 0/5 Hepatitis B Core (HBc) IgM 0/4 Hepatitis B Virus (HBV) 0/5 Hepatitis C Virus (HCV) 0/5 Hepatitis D Virus (HDV) 0/5 Herpes Simplex Virus (HSV) 5 0 5 Heterophilic Antibody Positive 0/5 HIV Ab 0/5 Human T-Lymphotropic Virus (HTLV) 1 0/5 HTLV Type 2 0/5 Influenza A 0/7 Influenza B 0/5 Influenza (Type Unknown) 0/8 Influenza Vaccine 0/5 Lupus 0/5 Monoclonal Hyper IgG 0/5 Picornavirus 0/5 Polyclonal Hyper IgG 0/3 Pregnant Females 0/5 Pregnant Females, Multiparous 0/5 Respiratory Syncytial Virus (RSV) 0/5 RF 0/5 Rubella IgG 0/5 Toxoplasmosis IgG 0/5

		Varicella Zoster Virus 0/5 Anti-Human CoV – No testing performed
Biomerieux	VIDAS® SARS-COV-2 IgG	Pregnant Females 0/5 Anti-Nuclear Antibody 0/5 Rheumatoid factor 0/5 Human Anti-Mouse Antibody 0/5 Borrelia burgdorferi 0/6 Plasmodium Falciparum 0/3 Toxoplasma gondii 0/6 Treponema pallidum 0/3 Trypanosoma cruzi 0/5 Hepatitis A Virus 0/3 Hepatitis B Virus 0/5 Hepatitis C Virus 0/5 Hepatitis E Virus 0/6 Herpes Simplex Virus 0/6 Human Immuno-deficiency Virus 1/5 Cytomegalovirus 0/3 Measles Virus 0/3 Mumps Virus 0/3 Rubella Virus 0/6 Dengue Virus 0/3 West Nile Virus 0/3 Yellow Fever Virus 0/3 Zika Virus 0/5 Influenza A and B Virus 0/10 Respiratory Syncytial Virus 0/10
Roche	Elecsys® Anti-SARS-CoV-2	Common cold panel – 0/40 Common Corona panel (HKU1, NL63, 229E or OC43) – 0/40
Siemens	SARS-CoV-2 Total (COV2T)	Anti nuclear antibody (ANA) 0/5 Chlamydia IgG 0/5 Cytomegalovirus (CMV) IgG 0/5 Epstein Barr virus (EBV) IgG 0/5 Epstein Barr virus (EBV) IgM 0/5 Graves' disease 0/5 Hepatitis A infection (HAV) IgM 0/5 Hepatitis B core antigen (anti-HBc) IgM 0/5 Hepatitis C infection (HCV) antibody 0/5 Human anti-mouse antibody (HAMA) 0/5 Human herpes virus (HHV) 0/3 Human immunodeficiency virus (HIV) antibody 0/10 Influenza antibody 0/10 Measles antibody 0/5 Parvovirus B19 antibody 0/5 Rheumatoid factor (RF) 0/5 Varicella zoster virus (VZV) antibody 0/5
Beckman Coulter	Access SARS-CoV-2 IgG assay	Anti-Influenza A 0/5 Anti-Influenza B 0/5 Anit-Hepatitis C Virus (HCV) 0/5 Anti-Hepatitis B Virus (HBV) 0/5 Anti-HIV 0/10 Anti-Nuclear Antibodies (ANA) 0/5 Anti-Adenovirus Positive IgG 0/2 Cytomegalovirus (CMV) IgG 0/7 Rheumatoid Factor (RF) 0/5

As a separate control with the aim of testing important and potentially cross-reactive serums, we will test groups of samples as described in section 3.1.

6. LINEARITY AND HOOK EFFECT

Linearity and hook effects as determined according to CLSI EP6-A guidelines are outside of the scope of this study.

7. FREEZE-THAW STABILITY

Freeze Thaw stability will be assessed. For the purpose of re-using thawed samples. This study will examine the stability of the test result (AU/ml), for 10 paired specimens, at the central virology laboratory on a single device. The serology tests will be executed on each fresh sample type and repeated three times (three runs) for each paired sample type, which was frozen for at least 24 hours and thawed in the day of use. In total, we will collect 10 specimens across the titer range (0-200 AU/ml based on the Diasorin assay). The titers will be recorded after every run.

7.1. Freeze Thaw and heat inactivation Data Analysis

Table XXX: Data from 10 fresh frozen and heat inactivated paired specimens

Sample	N	Fresh Mean	Thaw#1 Mean	Thaw#2 Mean	Thaw#3 Mean	Heat inactivation
Sample #n/10	3					

7.1.1. We will plot a graph of the mean titers of the three runs per sample with the fresh mean on the x-axis and the frozen mean on the y-axis for 1, 2, and 3 freeze thaw cycles.

7.1.2. Ordinary linear regression will be used to compare the results from the fresh specimens and the frozen. The intercept and slope will be estimated and the intercept will be tested against zero (is there a shift bias introduced by thawing?) and slope against 1 (is there a proportional bias introduced by thawing?)

7.1.3. A 95% CI for the regression will be calculated using standard methods based on normal distribution assumptions.

7.1.4. The acceptance criteria for equivalency for the titers will be (i) that the tests of intercept and slope in 8.1.2 are not rejected, and (ii) that the standard deviation of the differences between fresh and frozen results will be less than one twentieth (1/20) of the width of the equivocal interval (e.g. if the equivocal

interval is from 12 to 15 AU/ml, then the standard deviation of the differences will be less than 0.15 AU/ml).

8. MATRIX EQUIVALENCY STUDY

The following sample types in the different kits were tested by the vendor as equivalent and will be approached as similar for the purpose of this study:

Vendor	Product	Allowed and tested	Not allowed/Not tested
Diasorin	LIAISON® SARS-CoV-2 S1/S2 IgG	Serum Plasma Li. heparin Plasma Na. heparin Plasma K. EDTA	
Abbott	Architect 2000 SARS-CoV-2 N IgG	Serum Plasma EDTA	- Heat-inactivated - Pooled samples - Hemolysis - Microbial contaminant - Fungal growth
Biomerieux	VIDAS® SARS-COV-2 IgG	Serum Plasma Li. heparin	- Up to three freeze thaw cycles
Roche	Elecsys® Anti-SARS-CoV-2	Serum Plasma Li-heparin Plasma K2-EDTA Plasma K3-EDTA	- The samples may be frozen twice
Siemens	SARS-CoV-2 Total (COV2T)	Serum Plasma Li-heparin Plasma EDTA	
Beckman Coulter	Access SARS-CoV-2 IgG assay	Serum Plasma Heparin Plasma EDTA Plasma Citrate	- Thaw samples only once

9. CARRY OVER

The likelihood of 'carry over' in the automated immunoassay machines is considered low. This is because the nozzle is being washed between every patient. However, we do not have enough data on the non-specific binding of the COVID-19 antibodies and will perform a carryover study. The study will be conducted only at the central virology laboratory on a single device using low titer negative samples ("N"), and high titer positive samples ("P"). In this study, 3 sample sets will be run in the following sequence: P, P, P, P, P, N, P, N, P, N, N, N, N, N. The titers of the individually measured samples will be recorded. The impact of high positive samples on negative samples will be tested. The analysis will be based on testing for a downward trend in the antibody level measured in the 7 "N" samples. This is expected under carryover, because successive "N" samples have a decreasing number of

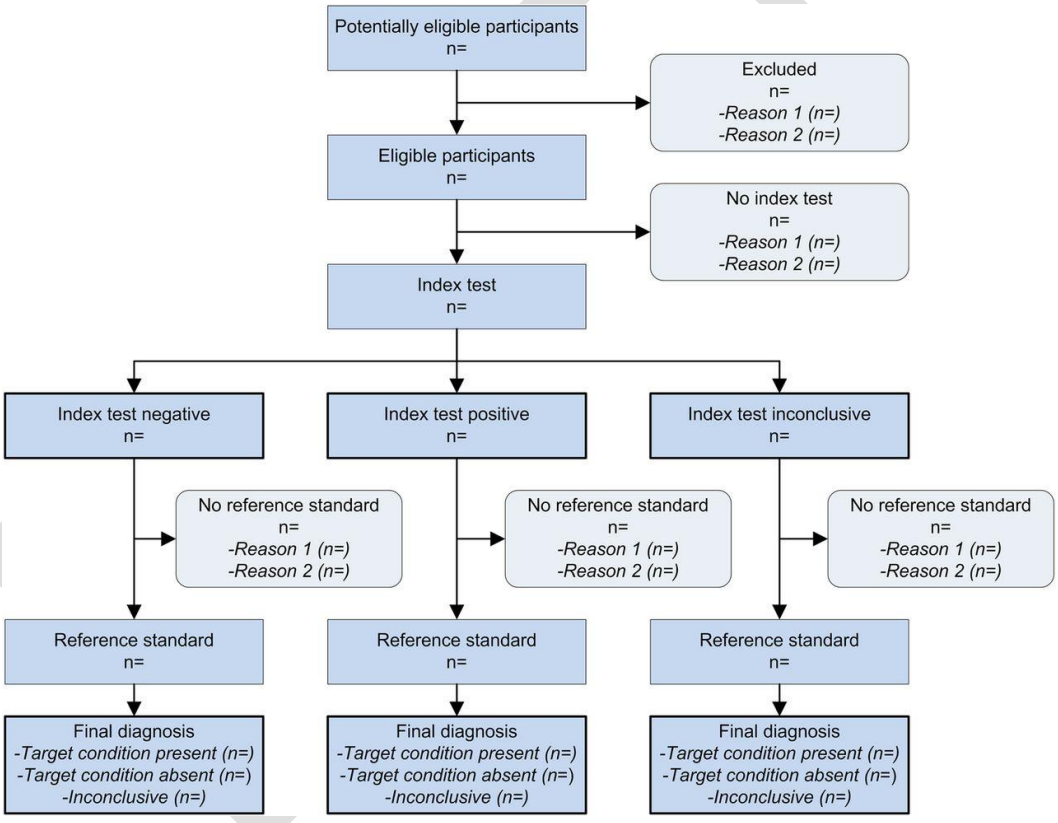
"P" samples run before them – in the previous 5 runs before each "N" sample, the number of "P" samples is 5, 4, 3, 2, 2, 1, and 0, respectively.

10. STATISTICAL CONSIDERATIONS OF THE CLINICAL ANALYSIS

Analyses will primarily be conducted and reported separately on the data of each laboratory. Following this, if results appear compatible across laboratories, they will be combined by meta-analysis methods (see Section 11.8).

10.1. Descriptive analysis of study population

The study workflow will be depicted in a graphical chart as recommended in the STARD statement (Bossuyt et al. 2003) as described below:



All continuous variables will be summarized using the following descriptive statistics: n, mean, standard deviation and median. The frequency and percentages will be reported for categorical features. Table XXX (in the Appendix) depicts some examples of the variables that will be included in the analysis using mock data.

10.2. Primary outcome

The primary outcome of this study will be to assess the performance of the kits to differentiate between previously infected seropositive (≥14 days from a positive PCR test result) and non-infected seronegative samples.

10.3. Primary Outcome Measures

The assays AUC, Sensitivity and Specificity will be determined, excluding patients with an equivocal result.

Specificity as a primary outcome measure, will be calculated as a crude value (no. positive tests in negative cohort/no. in negative cohort). In addition, as a secondary analysis, we will also calculate specificity as a population value (after adjustment for age and geography to the population distributions shown in Sections 3.1 and 3.2) The 95% CI, based on binomial distribution methods, will be reported in all cases.

In an additional secondary analysis, sensitivity will be re-estimated among the subgroup of PCR positive patients who are measured as seronegative on all kits.

10.4. Secondary outcomes

- To assess the performance of the kits to differentiate between early seropositive patients (<14 days from a positive PCR test result) and late seropositive patients (≥14 days from infection positive PCR test result) using a logistic regression on the number of days from symptoms onset.
- To assess the performance of the kits to differentiate between pediatric seropositive patients (<20 years of age) and non-pediatric population (>20 years of age) using a logistic regression on patient's age.
- Additional Subgroup analysis will be performed based on the following variables: disease severity, sample source, etc.
- The proportion of false positive results that are replicated on a repeat test will be estimated; also for false negative results.

- In order to explore possible ways of increasing sensitivity or specificity by combining test results, the following will be examined:

  (a) sensitivity and specificity under the rule that those found positive on either one of the kits is called positive

  (b) specificity and sensitivity under the rule that only those found positive on all the kits will be called positive.

It is understood that for rule (a) sensitivity will increase, but specificity will decrease; whereas for rule (b) specificity will increase, but sensitivity will decrease.

10.5. PPV and NPV dependency on bacterial prevalence

The prevalence of COVID-19 patients is time and setting dependent. For example, early during the epidemics the observed prevalence was as low as 0.1% while in endemic areas (e.g. Bnei Barak) we might find a much higher prevalence. Measures such as sensitivity, and specificity are invariant to the background prevalence, while measures such as PPV and NPV are prevalence dependent. We will simulate the expected PPV and NPV as a function of prevalence starting from 1% up to 90% in a continuous manner based on the estimated sensitivity (population value) and specificity.

10.6. Examining assay performance using different [AU/ml] cutoffs

The assays in focus have been developed with limited time and sample availability for optimal cutoff determination. Other factors such as background population and past epidemics might impact the optimal cutoff set for the Israeli population. Since we will perform a well powered study on a representative sampling of the Israeli population, we might be able to identify cutoff that further optimized compared to the manufacturer's instructions.

To test the different cutoff performance as a function of sensitivity and specificity, we will calculate performance when applying thresholds increasing in +/-5% of the suggested cutoff up to 300% of the original cutoff value. We will than plot the total accuracy, as a function of the different cutoffs and expect to see a non-monotonic function that will have a peak around the recommended cutoff.

10.7. Samples with equivocal results

Equivocal results will be handled in the following manner. First, the proportion of equivocal results will be reported for each test. Second, sensitivity and specificity will be estimated by three methods: (i) excluding equivocal results from the numerator and denominator (the primary analysis); (ii) treating all equivocal results as positive (secondary analysis); and (iii) treating all equivocal results as negative (secondary analysis).

10.8. Comparison of laboratories and combination of results across laboratories

(i) The characteristics of the samples from the four laboratories and their results will be compared using statistical methods. Four main comparisons will be made: (a) distributions of age, gender, geographical location, and disease severity; (b) proportion of equivocal results; (c) specificity; and (d) sensitivity. Since the populations from the four laboratories are likely to differ somewhat, the comparisons (b), (c) and (d) will be adjusted for age and gender, and for the positive cohorts also disease severity, using logistic regression. Adjustment for geographical location may not be possible, due to several of the laboratories drawing their samples from the local population.

(ii) If the above estimated test properties are found to differ substantially (i.e. with statistical significance at the 5% level, or with specificity differing by more than 1% from that of the largest sample, or sensitivity by more than 5% from that of the largest sample), then results will be presented side by side, and no attempt to combine them will be made. If they are not found to differ then results for proportion of equivocal results, specificity and sensitivity will be combined by using weighted averages of the laboratory-specific estimates, where the weights (inverse of the variances) reflect the precision of each estimate.

Appendix:

	Total (n=)	Positive (n=)	Negative (n=)	Equivocal (n=)
Age (y) (Mean, SD)				
Gender, Male (n, %)				
Time from symptoms (d) (Median, IQR)				
Hospitalization duration (d) (Median, IQR)				
Mechanical ventilation (n, %)				
Sample source Community (n, %)				
Sample source MDA (n, %)				
Sample source Hospital (n, %)				

The Validation protocol was written by the following team members:

Dr. Kfir Oved, Canopy Immuno-Therapeutics, MeMed Dx

Dr. Yaniv Lustig, Central Virology Laboratory, The Ministry of Health, Sheba Medical Center

Prof. Laurence S. Freedman, Gertner Institute for Epidemiology and Health Policy Research

Prof. Yonat Shemer Avni, Laboratory of Clinical Virology Clalit HMO and Soroka Medical Center

Dr. George Prajgrod, Laboratory Division, Meuhedet HMO

Dr. Lia Supino, Laboratory Division, Maccabi HMO

Irina Payorsky, Laboratory Division, Leumit HMO

The serology committee responsible for overseeing serological validation included in addition to the writers of the Validation protocol also the following persons:

Prof. Zvi Livneh, Weizmann Institute of Science

Major Yishai Cohen, MAFAT

Dr. Yuval Cohen, MAFAT

Dr. Tamar Wolf Laboratory Division, Maccabi HMO

Prof. Ella Mendelson, Central Virology Laboratory, The Ministry of Health, Sheba Medical Center