



OPEN

## Early and long term antibody kinetics of asymptomatic and mild disease COVID-19 patients

Shai Efrati<sup>1,2</sup>✉, Merav Catalogna<sup>1</sup>, Ramzia Abu Hamed<sup>2,3</sup>, Amir Hadanny<sup>1</sup>, Adina Bar-Chaim<sup>3</sup>, Patricia Benveniste-Levkovitz<sup>3</sup>, Refael Strugo<sup>4</sup> & Osnat Levtzion-korach<sup>2,5</sup>

Most patients infected with SARS-CoV-2 are asymptomatic or mildly symptomatic. However, the early and late antibody kinetics, and the association between antibody levels, clinical symptoms, and disease phase in these patients have not yet been fully defined. Confirmed SARS-CoV-2 patients and their household contacts were evaluated over a period four months. The evaluation procedure included symptom monitoring, viral load and serology analysis every ten days. A total of 1334 serum samples were collected from 135 patients and analyzed using three assays for IgG-N, IgG-S and IgM antibodies. Of the study participants, 97% were seropositive during the study, and two distinct clusters were identified. These clusters were significantly different in their inflammatory related symptoms. Peak IgG-S was 40.0 AU/ml for the non-inflammatory cluster and 71.5 AU/ml for the inflammatory cluster ( $P = 0.006$ ), whereas IgG-N peaks were 4.3 and 5.87 ( $P = 0.023$ ) respectively. Finally, a decision tree model was designed to predict the disease phase based on the serological titer levels, and had an overall accuracy of 80.7%. The specific profile of seroconversion and decay of serum antibodies can be used to predict the time-course from the acute infection.

Following a SARS-CoV-2 infection, most patients develop detectable serum antibodies to the receptor-binding domain of the viral spike protein along with associated neutralizing activities<sup>1–8</sup>. The magnitude of the antibody response may be associated with disease severity, and it has been reported that patients with mild infections may not develop detectable amounts of neutralizing antibodies<sup>3,4,9</sup>. However, the exact nature of seroconversion with respect to patient risk factors and disease severity is still controversial<sup>4,6,7,10,11</sup>.

Since data in the early stages of the disease were gathered mostly from hospitalized patients or relatively late, after the onset of the infection<sup>2,12–14</sup>, there is a lack of knowledge about the early antibody kinetics in non-hospitalized patients with asymptomatic and mildly symptomatic disease. Moreover, the data regarding IgG levels against the spike (S) and nucleocapsid (N) antigens are based on blood samples collected at non-consequent timepoints<sup>2–8</sup>.

Recently, artificial intelligence (AI) was used in various aspects of the disease: public health and clinical decision making, fast detection, and rapid diagnosis<sup>15–17</sup>. Specifically, AI models were designed to predict the prevalence of asymptomatic COVID-19 carriers<sup>18</sup>. However, only limited results are available regarding classification of asymptomatic carriers, and predicting the course of the disease based on antibody kinetics<sup>19</sup>. The aim of the current study is therefore to evaluate early and late antibody kinetics in asymptomatic and mildly symptomatic cases, and to provide further insights into the association between antibody levels and disease phase in a longitudinal household study design.

### Results

**Cohort characteristics.** Between May 2020, and January 2021, a total of 458 individuals 18 years old or older with a positive COVID-19 RT-PCR, were screened for eligibility. Of them, a total of 137 primary confirmed SARS-CoV-2 infection patients were enrolled in the study and signed an informed consent. An additional 77 household members consented to participate. Of the household cohort, 59/77 (76.6%) individuals were found negative for SARS-CoV-2 infection. Among the positive patients, 20 participants withdrew their consent to participate before study visit 6, and were excluded from the analysis. Seven patients were vaccinated after visit 9, and their results after vaccination were excluded from the analysis. One hospitalized patient died after visit 8

<sup>1</sup>Research and Development Unit, Shamir Medical Center, Zerifin, Israel. <sup>2</sup>Sackler School of Medicine, Tel Aviv University, Tel Aviv, Israel. <sup>3</sup>Clinical Chemistry Laboratory, Shamir Medical Center, Zerifin, Israel. <sup>4</sup>Magen David Adom (MDA), Tel Aviv, Israel. <sup>5</sup>Medical Management, Shamir Medical Center, Zerifin, Israel. ✉email: efratishai@outlook.com

due to COVID 19 related complications. Accordingly, 123/135 patients (91.1%) completed the study procedures (100 days). The study's workflow is described in Supplementary Fig. 1.

**Symptoms.** Three groups were analyzed according to their symptom severity: 22 patients (16.3%) were asymptomatic (had no symptoms consistent with COVID-19), 96 patients (71.1%) were mildly symptomatic, and 17 patients (12.4%) were hospitalized during the study period with severe respiratory symptoms. Fifty-eight patients (43.9%) had no known medical risk factors for COVID-19. The median time between the onset of symptoms to the first blood sample was nine days (IQR, 7–12). Cohort baseline characteristics, demographics, high risk comorbidities, and COVID-19 symptoms data are provided in Table 1.

**Clustering within the mild-symptoms cohort.** To explore possible associations between pre-conditions, symptom combination and severity, and antibody kinetics among the outpatient mildly symptomatic COVID-19 patients, two sub-group clusters were identified using the unsupervised k-medoids clustering algorithm. Figure 1 shows a t-SNE plot of the two clusters (n = 40 and 56 respectively). The resulting clusters and their medical parameters are listed in Supplementary Table 1. These clusters were significantly different in their inflammatory related symptoms: Cluster 2 (Inflammatory) was characterized by higher frequencies of fever, chills, fatigue, muscle and joint ache, and cough, while Cluster 1 was characterized by non-inflammatory related symptoms.

**Kinetics of SARS-CoV-2 antibody responses.** We analyzed the longitudinal antibody response to the SARS-CoV-2 infection in 1334 samples taken from the 1485 planned blood samples (adherence of 88.7% to study procedures). All samples were tested for IgG-S, IgG-N and IgM antibodies.

A dynamic trend of PCR positive and seropositivity in the study's COVID-19 patients is shown in Fig. 2, and Supplementary Table 2 regarding the study groups. The full kinetics along the study's follow-up periods is illustrated in fitted curves for IgM, IgG-S and IgG-N, plotted against the study visits in Fig. 3. Detailed result figures are presented in Supplementary Figs. 2–5.

Peak IgM levels were 3.1 [IQR, 1.1–6.5] and 4.5 [IQR, 1.9–12.4] ( $P = 0.498$ ) in the non-inflammatory and inflammatory groups respectively. Peak IgG-S levels were 40.0 AU/ml [IQR, 26.0–66.5] and 71.5 AU/ml [IQR, 31.5–111.0] ( $P = 0.006$ ) and for IgG-N were 4.3 [IQR, 2.6–5.5] and 5.7 [IQR, 4.0–7.6] ( $P = 0.023$ ) in the non-inflammatory and inflammatory groups respectively. In the asymptomatic group, peak antibody concentrations were 2.6 [IQR, 1.0–7.9], 58.9 AU/ml [IQR, 39.6–98.5], and 4.8 [IQR, 2.8–6.5] for IgM, IgG-S and IgG-N respectively. Interestingly, there was no significant difference between the asymptomatic and the mild symptomatic groups (Table 2).

As expected, in the mild-symptoms cohort, IgM was associated with earlier seroconversion, with 49.6% having positive detectable levels at day 9 [IQR, 6–12]. Maximal peak concentrations were reached at day 18 [IQR, 13–22], where 76% of the mild cohort had positive detectable IgM levels. Regarding IgG, more patients had positive detectable levels of IgG-N as compared to IgG-S at day 9 [IQR, 6.5–12], 38.5% vs. 23.4% respectively ( $P = 0.004$ ). The peak antibody concentration of IgG-N was on day 30 [24–38] while the peak level of IgG-S was on day 47 [IQR, 34–76] (Table 2).

Antibody decay at visit 11 was calculated with respect to the peak seroconversion rate, as shown in Table 2. In the mild-symptoms cohort, IgM levels were reduced by 81.2% [IQR, 62.7–89.6%]. The decay of the IgG-N levels was more significant compared to the decay of the IgG-S levels 50.3% [IQR, 29.9–65.1%] vs 16.2% [IQR, 4.5–32.4%] respectively ( $P < 0.0001$ ). Regarding the mild-symptoms clusters, IgG-S was reduced by 16.2% [IQR, 3.9–32.5%], and 16.2% [IQR, 7.3–32.3%] ( $P = 0.828$ ), and IgG-N levels were reduced by 59.9% [IQR, 40.1–72.3%], and 44.8% [IQR, 27.9–61.9%] ( $P = 0.032$ ) in the non-inflammatory and inflammatory group clusters respectively. Antibody decay levels at visit 11 in the asymptomatic group were 72.6% [53.6–83.2], 54.5% [29.2–68.8] and 23.0% [2.4–34.9] for IgM, IgG-N and IgG-S respectively. The decay was not significantly different from the mild-symptoms group.

We also followed 17 severely ill hospitalized patients as a reference. Among these patients, the peak antibody concentrations were significantly higher ( $P < 0.001$ ) than the mildly ill patients. The peak IgM level was 13.20 [IQR, 3.55–23.67], for IgG-S it was 179.0 (AU/ml) [IQR, 105.77–211.0], and for IgG-N it was 8.02 [IQR, 6.96–8.93] (Table 2).

Four (3%) asymptomatic participants were PCR positive at screening, but seronegative during the entire surveillance period. Six (4.4%) symptomatic patients had IgG-S and IgG-N titers below the cutoff level. During the follow up period (median visit 10, [IQR, 5–10]), IgG-S titers decayed below the cutoff level in seven (5.2%) patients (of them, four patients from the non-inflammatory cluster).

**Decision tree model's performance.** A total of 998 out of 1334 records were used for the DT dataset training. Records with a missing value, or outlier records, as detected according to the boxplot analysis, were eliminated. Figure 4A shows a scatterplot of the antibody results distribution, representing the relation between IgG-N, IgG-S and IgM antibodies detected in three disease phases: infection phase, inflammation phase and recovery phase. Supplementary Table 3 and Supplementary Fig. 6 (confusion matrix) present the statistical results of the model performance for validation of the DT model. The overall accuracy of the model was 80.7%, with 73.6% cases of the infection phase, 70.3% of the inflammatory phase, and 86.0% of the recovery phase being correctly classified. Figure 4B shows the model's ROC curves. The ROC curve shows good to excellent performance: AUC, 0.96, 0.88, 0.91 for the infection, inflammation and recovery phases respectively.

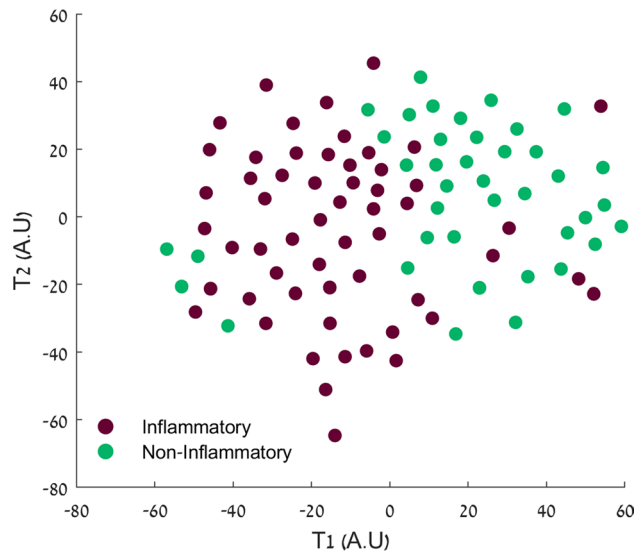
|  | Asymptomatic  | Mild symptoms | Severe symptoms |
|--|---------------|---------------|-----------------|
| N  | 22            | 96            | 17              |
| Age (Y)                                  | 41.86 ± 14.96 | 41.72 ± 13.89 | 54.00 ± 12.95   |
| Males                                    | 16 (72.7)     | 43 (44.8)     | 8 (47.1)        |
| Female                                   | 6 (27.3)      | 53 (55.2)     | 9 (52.9)        |
| BMI (Kg/m <sup>2</sup> )                 | 26.37 ± 4.31  | 26.71 ± 5.80  | 30.04 ± 5.03    |
| Primary confirmed patients               | 19 (86.4)     | 86 (89.6)     | 17 (100.0)      |
| Infected households                      | 3 (13.6)      | 10 (10.4)     | 0 (0.0)         |
| <b>High risk conditions</b>              |               |               |                 |
| BMI <sup>†</sup> > 30                    | 2 (9.1)       | 13 (13.5)     | 6 (35.3)        |
| Age > 60 Y                               | 4 (18.2)      | 28 (29.2)     | 9 (52.9)        |
| Cancer                                   | 1 (4.5)       | 4 (4.2)       | 1 (5.9)         |
| Diabetes mellitus                        | 2 (9.1)       | 7 (7.3)       | 3 (17.6)        |
| Hypertension                             | 1 (4.5)       | 5 (5.2)       | 3 (17.6)        |
| Heart disease                            | 2 (9.1)       | 3 (3.1)       | 2 (11.8)        |
| Immune deficiency                        | 0 (0.0)       | 5 (5.2)       | 0 (0.0)         |
| Asthma                                   | 2 (9.1)       | 5 (5.2)       | 1 (5.9)         |
| Other chronic lung diseases              | 0 (0.0)       | 0 (0.0)       | 2 (11.8)        |
| Chronic liver disease                    | 0 (0.0)       | 0 (0.0)       | 0 (0.0)         |
| Chronic kidney disease                   | 0 (0.0)       | 2 (2.1)       | 1 (5.9)         |
| Hematologic disease\disorder             | 0 (0.0)       | 0 (0.0)       | 0 (0.0)         |
| Chronic neurological impairment\ disease | 1 (4.5)       | 3 (3.1)       | 0 (0.0)         |
| Organ or bone marrow recipient           | 2 (9.1)       | 4 (4.2)       | 2 (11.8)        |
| Smoking                                  | 4 (18.2)      | 11 (11.5)     | 0 (0.0)         |
| <b>Symptoms</b>                          |               |               |                 |
| Fever ≥ 38 °C                            | 0 (0.0)       | 46 (47.9)     | 12 (70.6)       |
| Dry cough                                | 0 (0.0)       | 57 (59.4)     | 11 (64.7)       |
| Sore throat                              | 0 (0.0)       | 33 (34.4)     | 3 (17.6)        |
| Runny nose                               | 0 (0.0)       | 33 (34.4)     | 4 (23.5)        |
| Shortness of breath                      | 0 (0.0)       | 20 (20.8)     | 10 (58.8)       |
| Abdominal pain                           | 0 (0.0)       | 20 (20.8)     | 4 (23.5)        |
| Headache                                 | 0 (0.0)       | 58 (60.4)     | 8 (47.1)        |
| Problem in smell sensation               | 0 (0.0)       | 41 (42.7)     | 10 (58.8)       |
| Problem in taste sensation               | 0 (0.0)       | 22 (22.9)     | 7 (41.2)        |
| Chills                                   | 0 (0.0)       | 26 (27.1)     | 3 (17.6)        |
| Vomiting                                 | 0 (0.0)       | 16 (16.7)     | 2 (11.8)        |
| Nausea                                   | 0 (0.0)       | 27 (28.1)     | 4 (23.5)        |
| Diarrhea                                 | 0 (0.0)       | 18 (18.8)     | 4 (23.5)        |
| Rash                                     | 0 (0.0)       | 9 (9.4)       | 1 (5.9)         |
| Conjunctivitis                           | 0 (0.0)       | 16 (16.7)     | 3 (17.6)        |
| Muscle aches                             | 0 (0.0)       | 47 (48.9)     | 6 (35.3)        |
| Joint ache                               | 0 (0.0)       | 27 (28.1)     | 4 (23.5)        |
| Loss of appetite                         | 0 (0.0)       | 30 (31.2)     | 6 (35.3)        |
| Nose bleed                               | 0 (0.0)       | 25 (26.0)     | 3 (17.6)        |
| Fatigue                                  | 0 (0.0)       | 56 (58.3)     | 9 (52.9)        |
| Days from onset of symptoms              | 0 (0.0)       | 8.99 ± 1.55   | 11.00 ± 1.88    |

**Table 1.** Baseline characteristics. Data presented as n (%); continuous data, mean ± SD. <sup>†</sup>The body-mass index is the weight in kilograms divided by the square of the height in meters.

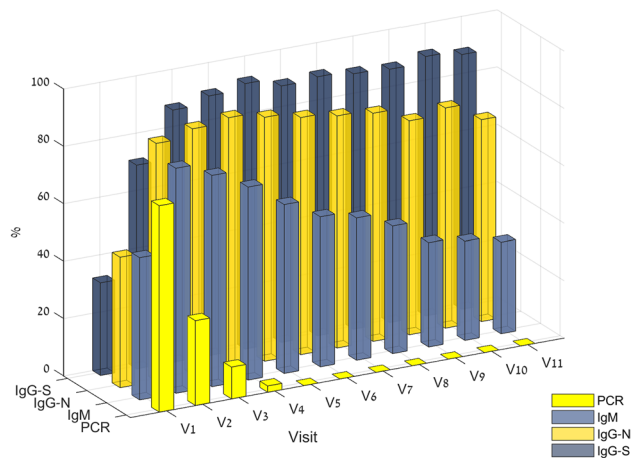
## Discussion

The clinical spectrum of patients infected with COVID 19 ranges from asymptomatic to critically ill, with the majority having mild symptoms that do not require hospitalization<sup>20,21</sup>. This study highlights the importance of the dynamic course of SARS-CoV-2 in non-hospitalized patients. A rigorous testing process and a replicative laboratory method enabled us to examine the association between serological tests and clinical symptoms in asymptomatic and mildly symptomatic patients.

It was previously reported that some cases of asymptomatic and mildly symptomatic patients failed to mount neutralizing antibodies<sup>1–4,9,10</sup>. However, this study and other studies<sup>5,7,12,22</sup> indicate that by using current



**Figure 1.** Mild symptoms cohort ( $n=96$ ) classification results: t-stochastic neighbor embedding (t-SNE) plot based on HAMMIG distance metrics in data transformation binary results. T1 and T2 are t-SNE results after dimension reduction of the 38 input parameters into 2D visualization. Cluster 1 (green,  $N=40$ ) represents patients suffered from non-inflammatory related symptoms, and Cluster 2 (red,  $N=56$ ) represents patients suffered from inflammatory related symptoms.

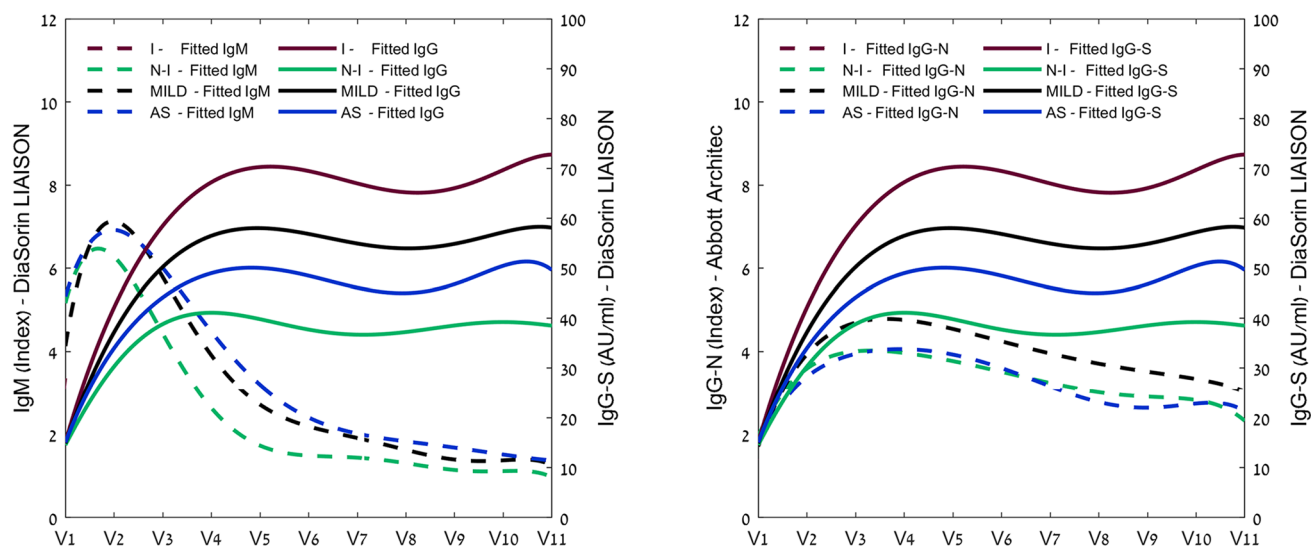


**Figure 2.** Dynamic trend of PCR positive, and seropositivity in study COVID19 patients ( $N=135$ ). Of note, the median time between the onset of symptoms to the first visit was 9 days (IQR, 7–12).

laboratory methods, the vast majority (>95%) do developed detectable levels of IgM, IgG-S and IgG-N. Compared to the asymptomatic and the mild symptomatic cohorts, severe respiratory symptomatic patients generated two–fivefold higher antibody titers.

We also identified subclasses within the outpatient mild cohort. Using an unsupervised clustering technique, two distinct clusters were identified based on their symptoms. Interestingly, there was a significant difference in IgG-S and IgG-N antibody concentrations between the clusters throughout the study period. These results suggest that systemic symptoms (i.e., fever, fatigue, pain, dry cough) are associated with higher antibody titers in mild symptomatic patients. Interestingly, the decay rate in antibody concentrations during the study period was similar in both clusters.

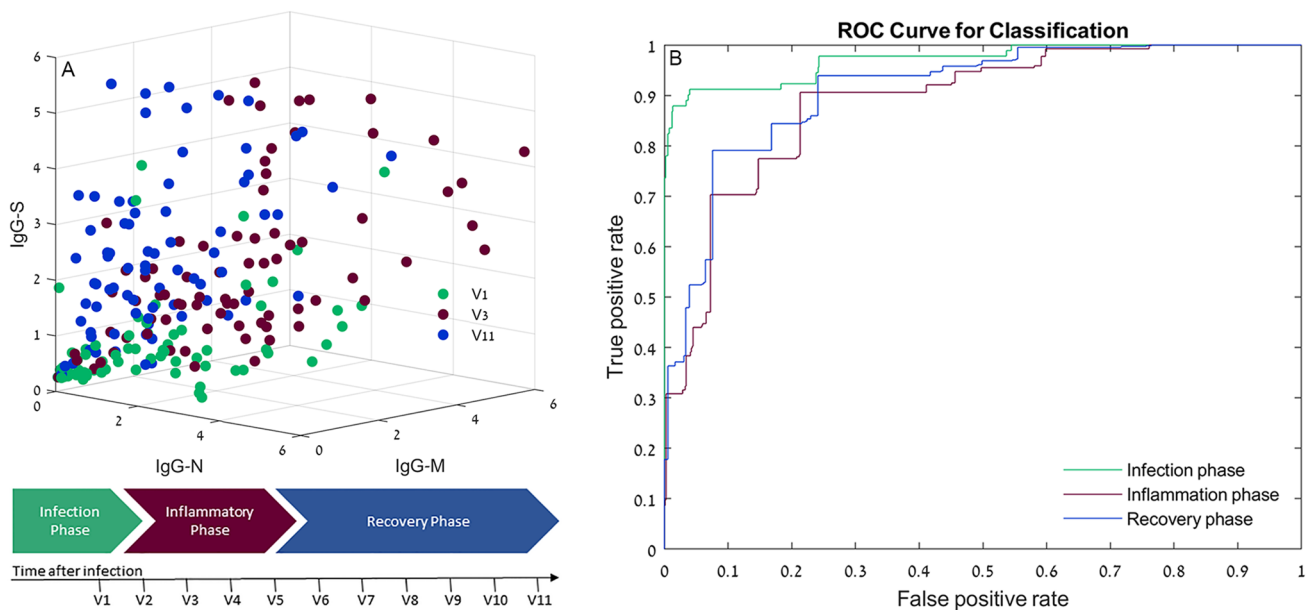
The minimal level of antibodies required for infection immunity has yet to be determined and will require additional long-term studies including re-infected patients. It should be noted that in addition to protective antibodies, immunity for recurrent infections includes SARS-CoV-2-specific memory lymphocytes with potent antiviral functions. Memory T-cells proliferate and secrete antiviral cytokines upon antigen reencounter, whereas memory B-cells participate in the adaptive immune response, followed by differentiation into plasma cells and the production of virus neutralizing antibodies<sup>8,23</sup>. In a study on 15 mildly symptomatic patients, it was demonstrated that in addition to neutralizing antibodies, neutralizing plasma, memory B and memory T cells also persist and may even increase over a period of 3 months<sup>8</sup>. In another study, which included 21 randomly selected



**Figure 3.** Longitudinal profile of SARS-CoV-2 antibodies in asymptomatic and mild COVID19 patients. (A) IgM vs. IgG-S, (B) IgG-S vs. IgG-N antibodies. Data are fitted by a quadratic polynomial regression model, and presented for asymptomatic patients (N=22, blue), the mild cohort (N=96, black) and for non-inflammatory cluster (N=40), and the inflammatory cluster (N=56). For full details, see also Supplementary Figs. 2–4. AS asymptomatic, I mild inflammatory related symptoms, N-I mild non-inflammatory related symptoms, MILD all mild symptoms.

|                              | Asymptomatic                    | Mild symptoms                   | P-Value1 | Severe symptoms                     | P-Value2 | Non-inflammatory                | Inflammatory                     | P-Value3 |
|------------------------------|---------------------------------|---------------------------------|----------|-------------------------------------|----------|---------------------------------|----------------------------------|----------|
| N                            | 22                              | 96                              |          | 17                                  |          | 40                              | 56                               |          |
| <b>IGG-S</b>                 |                                 |                                 |          |                                     |          |                                 |                                  |          |
| Peak concentration (AU/ml)   | 58.9 [39.6–98.5]<br>65.3 ± 44.1 | 50.8 [28.4–93.8]<br>68.1 ± 55.8 | 0.746    | 185.0 [105.8–220.8]<br>174.1 ± 98.3 | < 0.001  | 40.0 [26.0–66.5]<br>49.8 ± 35.1 | 71.5 [31.5–111.0]<br>81.3 ± 63.9 | 0.006    |
| Concentration at V11 (AU/ml) | 49.3 [36.1–58.5]<br>50.1 ± 29.7 | 39.3 [23.1–82.0]<br>57.9 ± 54.3 | 0.621    | 148.0 [54.6–199.0]<br>136 ± 86.4    | < 0.001  | 28.5 [18.0–52.1]<br>38.4 ± 27.7 | 47.9 [30.0–106.0]<br>72.6 ± 64.3 | 0.004    |
| Time to peak (days)          | 48.0 [30–81]<br>50.8 ± 36.8     | 47.0 [34.5–76.5]<br>55.8 ± 28.8 | 0.599    | 56.0 [32–71]<br>55.2 ± 25.2         | 0.878    | 40.0 [30.0–86.0]<br>55.4 ± 32.5 | 49.0 [37.5–74.0]<br>56.1 ± 26.2  | 0.907    |
| <b>IGG-N</b>                 |                                 |                                 |          |                                     |          |                                 |                                  |          |
| Peak concentration (AU)      | 4.8 [2.8–6.5]<br>4.4 ± 2.8      | 5.0 [3.4–7.2]<br>5.1 ± 2.6      | 0.292    | 8.0 [7.0–8.9]<br>7.9 ± 1.5          | < 0.001  | 4.3 [2.6–5.5]<br>4.4 ± 2.7      | 5.7 [4.0–7.6]<br>5.6 ± 2.5       | 0.023    |
| Concentration at V11 (AU)    | 1.9 [0.6–4.0]<br>2.6 ± 2.3      | 2.4 [1.1–4.0]<br>3.0 ± 2.5      | 0.548    | 4.9 [3.6–7.5]<br>5.3 ± 2.4          | 0.003    | 1.6 [0.7–2.7]<br>2.3 ± 2.3      | 3.0 [1.4–5.2]<br>3.5 ± 2.0       | 0.028    |
| Time to peak (days)          | 28.5 [0–35]<br>24 ± 16.9        | 30.0 [24.5–38.5]<br>34.8 ± 19.3 | 0.046    | 25.0 [20–36]<br>34.6 ± 26.7         | 0.497    | 30.0 [21.0–37.0]<br>33.6 ± 19.2 | 30.0 [26.0–39.0]<br>35.7 ± 19.5  | 0.603    |
| <b>IGG-M</b>                 |                                 |                                 |          |                                     |          |                                 |                                  |          |
| Peak concentration (AU/ml)   | 2.6 [1.0–7.9]<br>4.6 ± 10.6     | 3.6 [1.7–8.2]<br>5.0 ± 13.8     | 0.430    | 13.2 [3.5–23.7]<br>22.6 ± 31        | < 0.001  | 3.1 [1.1–6.5]<br>5.3 ± 16.7     | 4.5 [1.9–12.4]<br>7.5 ± 9.7      | 0.498    |
| Concentration at V11 (AU/ml) | 1.0 [0.4–2.0]<br>3.4 ± 7.2      | 0.6 [0.4–1.3]<br>1.3 ± 1.7      | 0.104    | 1.1 [0.4–5.1]<br>3 ± 4              | 0.009    | 0.6 [0.3–0.9]<br>1.0 ± 1.2      | 0.7 [0.4–1.8]<br>1.5 ± 1.2       | 0.179    |
| Time to peak (days)          | 16.0 [0–22]<br>20.8 ± 25.8      | 18.5 [13.0–22.0]<br>22.1 ± 20.7 | 0.833    | 19.0 [12–26]<br>19.9 ± 10.6         | 0.801    | 16.0 [11.0–20.0]<br>19.4 ± 20.2 | 19.5 [14.0–23.5]<br>19.8 ± 10.4  | 0.270    |

**Table 2.** Clinical characteristics, and titer levels of COVID-19 patients with different immune response phenotypes. \*P-Value1- Asymptomatic vs. Mild symptomatic; P-Value2- Mild symptomatic vs. Severe symptoms; P-Value3- Non-inflammatory (CLUSTER 1) vs. Inflammatory (CLUSTER 2). Data are presented as median [IQR], and mean ± SD.



**Figure 4.** Distribution of antibody results. (A) 3D scatterplots representing the relation between IgG-N, IgG-S and IgM antibodies detected in three phases of the disease: Infection phase, inflammation phase and recovery phase. Values are expressed in antibody level/cutoff. (B) ROC curves for the classification tree model: Applied on the infection phase, inflammation phase and recovery phase data. The x-axis represents the fraction of negative examples classified as positives. The y-axis shows the fraction of positive examples classified as positives. The probabilities for class prediction were estimated by leave-one-out cross validations.

patients, memory B cells persisted for 6.2 months after the acute infection<sup>23</sup>. Unfortunately, neither memory T nor memory B-cells can be evaluated at standard laboratories. Therefore, daily clinical practices will have to rely on IgG antibody measurements as evaluated in this study.

Previous studies conducted in moderate and severe hospitalized patients have examined the relationship between IgM, IgG-S and IgG-N responses at the early stage of the disease. Seroconversion for anti-N occurs significantly faster than for anti-S in COVID-19 patients, which was also demonstrated in our study. The combination of the two may improve the early serological detection rate<sup>24,25</sup>. In this study, we examined the dynamics between these antibodies for a longer period. The additional information we provide regarding the different kinetics over time may help predict the time of disease onset and identify post-infected subjects based on a serological blood test. The ability to predict the phase of the disease mainly in asymptomatic patients and patients that suffer from symptoms that are not clearly associated with COVID-19 could be highly important for epidemiological studies and decision making during the SARS-CoV-19 pandemic.

The study has several limitations. Even though the study has a relatively medium size dataset, rigorous testing has enabled us to develop accurate predictors, and to demonstrate a full antibody kinetics evaluation. Approaching mildly symptomatic outpatients raised a technical challenge related to contacting quarantined patients by protected medical staff at the patients' houses. Nevertheless, the adherence rate in our study was high (88.7%). Another limitation of this study is the limited number of asymptomatic patients. Even though households of patients were sampled, it is very challenging to identify a significant cohort of fully asymptomatic infected individuals. Larger asymptomatic cohorts may delineate significant changes in antibody kinetics compared to mild symptomatic patients. In addition, longer term changes of over one year in antibody kinetics are needed and will be evaluated in a follow-up study of this cohort.

In conclusion, this study, for the first time, marks the early and long-term antibody kinetics of asymptomatic and mildly symptomatic cases, representing the majority of patients infected with SARS-CoV-19. This study supports the relationship between disease severity and antibody titer levels also when mild symptoms are presented. The specific profile of seroconversion and the decay of IgG-N, IgG-S and IgM antibodies enables us to predict the time course from the acute infection.

## Methods

**Patients and recruitment procedure.** Patients 18 years or older with a positive COVID-19 infection result which was performed in the Shamir Medical Center laboratory, were offered study enrollment, irrespective of clinical signs and symptoms. Patients were excluded if pregnant, or unable to sign an informed consent. To evaluate patients for eligibility and consent, patients were contacted by telephone. Upon consent, a medical staff representative arrived at the patient's house for a full explanation and obtained their informed consent in addition to anyone else in the household over 18 years old.

**Study design.** This study was a prospective clinical trial performed on laboratory confirmed SARS-CoV-2 infected patients and their household contacts. The evaluation procedure included symptom monitoring, viral

respiratory load and serological analysis. Sequential symptom information, and specimens were collected from primary cases and from their household contacts every 10 days ( $\pm 2$  days) for a period of four months. The study was approved by Shamir Medical Center's institutional review board (IRB) (No. 105-20) and all participants signed an informed consent prior to their inclusion. All research was performed according to the relevant guidelines and regulations. This study was registered with ClinicalTrials.gov, number NCT04348422.

**Patient data and symptom monitoring.** Epidemiologic, demographic, pre-existing conditions, contact and exposure history data were collected by interview. In addition, during the study, each participant was asked to complete a symptom questionnaire which covers the symptoms every 10 days, and any changes in quarantine and exposure status (questionnaires were provided in the study protocol).

**SARS-CoV-2 serology.** Whole blood samples were collected into EDTA and gel tubes using a standard technique at the patient's house or in the hospital every 10 days by protected medical staff. Blood samples were kept at 2–8 °C degrees and transferred to Shamir Medical Center laboratory within two hours. COVID-19 serological tests were performed using the following commercially available, FDA approved, automated immunoassays:

*Abbott architect instrument SARS-CoV-2 IgG (H07891R03, Abbott, Illinois, USA).* A chemiluminescent microparticle immunoassay (CMIA), for quantitative detection of IgG in human serum or plasma, against the SARS-CoV-2 nucleoprotein. The assay uses a double-antigen sandwich immunoassay design employing microparticle-bound antigen and acridinium-labeled human anti-IgG. The recommended manufacturer's index cutoff value is 1.40.

*Liaison SARS-CoV-2 S1/S2 IgG (311450, DiaSorin, Saluggia, Italy).* A chemiluminescent immunoassay (CLIA) for quantitative determination of anti-S1 and anti-S2 specific IgG antibodies using magnetic beads coated with S1 and S2 antigens. The analyzer automatically calculates SARS-CoV-2 S1/S2 IgG antibody concentrations expressed as arbitrary units (AU/ml), with a positive cutoff level of 15.0 AU/ml.

*Liaison SARS-CoV-2 IgM (311470, Diasorin, Saluggia, Italy).* A chemiluminescence immunoassay (CLIA) for the qualitative determination of specific IgM antibodies to SARS-CoV-2 in human serum samples with a positive cutoff level of 1.1 AU/ml.

Compared with ELISA, the Abbott IgG assay was reported to have a sensitivity and specificity of 92.7% and 99.9% respectively. The Liaison IgG assay sensitivity was reported as 96.2% and the specificity as 98.9%<sup>26,27</sup>.

**PCR test for viral detection.** Real-time quantitative reverse transcriptase polymerase chain reactions (qRT-PCRs) were performed using the TaqPath Covid-19 RT-PCR Kit (A48067; ThermoFisher Scientific, Massachusetts, USA), a fast, highly sensitive, multiplex and robust RT-qPCR assay for the detection of SARS-CoV-2<sup>28</sup>. Nucleic acids were isolated according to manufacturer's instructions. Briefly, 200  $\mu$ l viral transport medium (VTM) was taken from the patient swab sample inside a class 2 safety cabinet and mixed with 150  $\mu$ l lysis buffer, 1  $\mu$ l carrier RNA, and extraction controls (MS2, was provided as part of the kit). After incubation at room temperature for at least 15 min, samples were processed using the liquid handler Biomek i7 automated workstation (Beckman, Coulter) for RNA isolation. Primers and probes to target the SARS-CoV-2 E, N (N1 and N2 targets), and S genes, were included in the kit. A positive result for SARS-CoV-2 detection was determined by amplification of at least two of the three genes targeted, using a cutoff threshold cycle (CT) value of 37.

**Statistical methods.** Descriptive statistics: Continuous data are expressed as means  $\pm$  standard-deviations (SD), and as median and interquartile range (IQR). Independent t-tests with two-tail distribution were performed to compare variables between groups, when a normality assumption held according to a Kolmogorov–Smirnov test. Categorical data are expressed in numbers and percentages. A value of  $p < 0.05$  is considered significant.

Longitudinal serological data were fitted by a quadratic polynomial regression model and analyzed at a 95% confidence level ( $p < 0.05$ ). The model parameter estimates were iteratively determined using the Levenberg–Marquadt optimization method. Boxplot analysis was used to present data distribution, and to detect outliers.

Clustering model: To explore possible associations between preconditions, symptom combination and severity, and the response of the mild cohort's immune system, an unsupervised k-medoids clustering algorithm was used. The input data binary matrix  $S(i,j)$  represents the patient's  $j$  vector of parameters  $i = 1: 38$  (training parameters are listed in Supplementary Table 1). Distances between the two vectors were computed using the Hamming distance function, and the data was partitioned into two clusters (medoids), due to the relatively small sample size<sup>29</sup>. The clusters were then visualized using the t-stochastic neighbor embedding (t-SNE) plot based on Hamming distance metrics in data transformation binary results<sup>30</sup>.

Prediction model: A decision tree (DT) model was used to predict the phase of the disease based on the relationship between antibody titers. A decision tree is a nonparametric supervised learning method used for classification and prediction. In this study, a classification tree type was used, which employs the CART algorithm for binary classification and the Gini diversity index split criterion for optimization<sup>31</sup>. The input training vectors  $V(i,j)$  are the [antibody(i) concentration/ cut-off level(i)] of the sample  $i$ , where  $i = 1: 3$  assays, and  $j = 1: 1$  number of samples. Each vector was labeled with one of the three phases of the disease: infection phase (visit 1), inflammation phase (visits 2–4) and recovery phase (visits 5–11)<sup>32</sup>, as shown in Fig. 4A. To train and estimate

the predictive performance of the model, a leave-one-out cross-validation (LOOCV) approach was used. In this strategy, the training and testing procedure is done  $N$  times, where  $N$  is the number of observations. Each observation is considered as the validation set, and the other  $(N-1)$  observations are used for training the model. This method was chosen due to its robustness and low bias in small sample size models<sup>33</sup>. A confusion matrix and area under the ROC curve were used to assess the model's performance.

Data were statistically analyzed using the Matlab Statistics, Machine Learning, and Curve Fitting Toolbox, R2020b (Mathworks, Natick, MA).

Received: 11 March 2021; Accepted: 14 June 2021

Published online: 02 July 2021

## References

- Seow, J. *et al.* Longitudinal evaluation and decline of antibody responses in SARS-CoV-2 infection. *medRxiv* (2020).
- Escribano, P. *et al.* Detection of SARS-CoV-2 antibodies is insufficient for the diagnosis of active or cured COVID-19. *Sci. Rep.* **10**, 19893. <https://doi.org/10.1038/s41598-020-76914-5> (2020).
- Chen, X. *et al.* Disease severity dictates SARS-CoV-2-specific neutralizing antibody responses in COVID-19. *Signal Transduct. Target Ther.* **5**, 180. <https://doi.org/10.1038/s41392-020-00301-9> (2020).
- Long, Q. X. *et al.* Clinical and immunological assessment of asymptomatic SARS-CoV-2 infections. *Nat. Med.* **26**, 1200–1204. <https://doi.org/10.1038/s41591-020-0965-6> (2020).
- Gudbjartsson, D. F. *et al.* Humoral immune response to SARS-CoV-2 in Iceland. *N. Engl. J. Med.* **383**, 1724–1734. <https://doi.org/10.1056/NEJMoa2026116> (2020).
- Wajnberg, A. *et al.* Robust neutralizing antibodies to SARS-CoV-2 infection persist for months. *Science* **370**, 1227–1230. <https://doi.org/10.1126/science.abd7728> (2020).
- Dan, J. M. *et al.* Immunological memory to SARS-CoV-2 assessed for up to 8 months after infection. *Science* <https://doi.org/10.1126/science.abf4063> (2021).
- Rodda, L. B. *et al.* Functional SARS-CoV-2-specific immune memory persists after mild COVID-19. *Cell* **184**, 169–183. <https://doi.org/10.1016/j.cell.2020.11.029> (2021).
- Rijkers, G. *et al.* Differences in antibody kinetics and functionality between severe and mild severe acute respiratory syndrome coronavirus 2 infections. *J. Infect. Dis.* **222**, 1265–1269. <https://doi.org/10.1093/infdis/jiaa463> (2020).
- Robbiani, D. F. *et al.* Convergent antibody responses to SARS-CoV-2 in convalescent individuals. *Nature* **584**, 437–442. <https://doi.org/10.1038/s41586-020-2456-9> (2020).
- Wang, K. *et al.* Longitudinal dynamics of the neutralizing antibody response to SARS-CoV-2 infection. *Clin. Infect. Dis.* <https://doi.org/10.1093/cid/ciaa1143> (2020).
- Iyer, A. S. *et al.* Dynamics and significance of the antibody response to SARS-CoV-2 infection. *medRxiv* <https://doi.org/10.1101/2020.07.18.20155374> (2020).
- Lee, Y. L. *et al.* Dynamics of anti-SARS-Cov-2 IgM and IgG antibodies among COVID-19 patients. *J. Infect.* **81**, e55–e58. <https://doi.org/10.1016/j.jinf.2020.04.019> (2020).
- Lynch, K. L. *et al.* Magnitude and kinetics of anti-severe acute respiratory syndrome coronavirus 2 antibody responses and their relationship to disease severity. *Clin. Infect. Dis.* **72**, 301–308. <https://doi.org/10.1093/cid/ciaa979> (2021).
- Huang, S., Yang, J., Fong, S. & Zhao, Q. Artificial intelligence in the diagnosis of COVID-19: Challenges and perspectives. *Int. J. Biol. Sci.* **17**, 1581 (2021).
- Vaishya, R., Javaid, M., Khan, I. H. & Haleem, A. Artificial intelligence (AI) applications for COVID-19 pandemic. *Diabetes Metab. Syndr.* **14**, 337–339 (2020).
- Mei, X. *et al.* Artificial intelligence-enabled rapid diagnosis of patients with COVID-19. *Nat. Med.* **26**, 1224–1228 (2020).
- Lee, C., Apio, C. & Park, T. Estimation of undetected asymptomatic COVID-19 cases in South Korea using a probabilistic model. *Int. J. Environ. Res. Public Health* **18**, 4946 (2021).
- Sun, C. *et al.* Accurate classification of COVID-19 patients with different severity via machine learning. *Clin. Transl. Med.* **11**, 2 (2021).
- Wu, Z. & McGoogan, J. M. Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China: Summary of a report of 72314 cases from the Chinese center for disease control and prevention. *JAMA* **323**, 1239–1242. <https://doi.org/10.1001/jama.2020.2648> (2020).
- Oran, D. P. & Topol, E. J. Prevalence of asymptomatic SARS-CoV-2 infection: A narrative review. *Ann. Intern. Med.* **173**, 362–367. <https://doi.org/10.7326/M20-3012> (2020).
- L'Huillier, A. G. *et al.* Antibody persistence in the first 6 months following SARS-CoV-2 infection among hospital workers: A prospective longitudinal study. *Clin. Microbiol. Infect.* <https://doi.org/10.1016/j.cmi.2021.01.005> (2021).
- Gaebler, C. *et al.* Evolution of antibody immunity to SARS-CoV-2. *Nature* <https://doi.org/10.1038/s41586-021-03207-w> (2021).
- Van Elslande, J. *et al.* Antibody response against SARS-CoV-2 spike protein and nucleoprotein evaluated by four automated immunoassays and three ELISAs. *Clin. Microbiol. Infect.* **26**, 1557 (2020).
- Sun, B. *et al.* Kinetics of SARS-CoV-2 specific IgM and IgG responses in COVID-19 patients. *Emerg. Microb. Infect.* **9**, 940–948 (2020).
- Meschi, S. *et al.* Performance evaluation of Abbott ARCHITECT SARS-CoV-2 IgG immunoassay in comparison with indirect immunofluorescence and virus microneutralization test. *J. Clin. Virol.* **129**, 104539 (2020).
- Ainsworth, M. *et al.* Performance characteristics of five immunoassays for SARS-CoV-2: A head-to-head benchmark comparison. *Lancet. Infect. Dis.* **20**, 1390–1400 (2020).
- Reijns, M. A. *et al.* A sensitive and affordable multiplex RT-qPCR assay for SARS-CoV-2 detection. *PLoS Biol.* **18**, e3001030 (2020).
- Park, H.-S. & Jun, C.-H. A simple and fast algorithm for K-medoids clustering. *Expert Syst. Appl.* **36**, 3336–3341 (2009).
- Van der Maaten, L. & Hinton, G. Visualizing data using t-SNE. *J. Mach. Learn. Res.* **9**, 2 (2008).
- Breiman, L., Friedman, J., Stone, C. J. & Olshen, R. A. *Classification and Regression Trees* (CRC Press, 1984).
- Xia, L. *et al.* The course of mild and moderate COVID-19 infections—The unexpected long-lasting challenge. *Open Forum Infect. Dis.* <https://doi.org/10.1093/ofid/ofaa286> (2020).
- Wong, T.-T. Performance evaluation of classification algorithms by k-fold and leave-one-out cross validation. *Pattern Recogn.* **48**, 2839–2846 (2015).

## Acknowledgements

The study was funded by a research grant from the Sagol Network for Neuroscience established by Mr. Sami Sagol. This study was performed in collaboration with the Israeli Ministry of Health. We would like to thank Dr.



Mechael Kanovsky for his editing of this manuscript. Special thanks to Roy Sagi for managing the trial, and for Hila Goldner-Yerushalmi, Moran Adler, Ron-El Goldman, Hofit Muzan and Dana Shimonov from the Research and Development Unit, and for Marina Sova, Shima Esawi, Shahaf Cohen, and Basima Alokbi from the Clinical Chemistry Laboratory, Shamir Medical Center for their dedicated work. We also would like to thank Racheli Ikar, Nati Frid, Igor Artel, Rafi Sheva, Ofir Shaish, Lior Marmelstein, Re'em Naimy, Linoy Yerushalmy, Rami Avraham Shapira, Eli Ben Simchon, Maya Kaplan Oved, and Eti Ventura Daushin from Magen David Adom (MDA) for their assistance in collecting the blood samples.

### Author contributions

S.E., A.H., O.L.K., conceived and designed the study. R.A.H., A.B.C., P.B.L. and R.S. designed and conducted the experiments. M.C. performed the data analysis. S.E., M.C., and A.H. wrote the manuscript that was reviewed and accepted by all authors before submission.

### Competing interests

The authors declare no competing interests.

### Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-93175-y>.

**Correspondence** and requests for materials should be addressed to S.E.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021