The antibody response to SARS-CoV-2 increases over 5 months in patients with anosmia/dysgeusia

Riccardo Levi[1,2]*, Leonardo Ubaldi[1,2]*, Chiara Pozzi[2], Giovanni Angelotti[2], Maria Teresa Sandri[1,2], Elena Azzolini[1,2], Michela Salvatici[2], Victor Savevski[2], Alberto Mantovani[1,2,3] and Maria Rescigno[1,2,#]

[1]Humanitas University, Department of Biomedical Sciences, Via Rita Levi Montalcini, 20090 Pieve Emanuele, MI, Italy.
[2]Humanitas Clinical and Research Center - IRCCS, Via Manzoni 56, 20089 Rozzano, MI, Italy.
[3]The William Harvey Research Institute, Queen Mary University of London, London, UK

*These authors contributed equally to the work

[#]Corresponding author
E-mail: maria.rescigno@hunimed.eu
Humanitas University
Via Rita Levi Montalcini, 4
20090 Pieve Emanuele (MI) Italy

Abstract

The factors involved in the persistence of antibodies to SARS-CoV-2 are unknown. We evaluated the antibody response to SARS-CoV-2 in personnel from 10 healthcare facilities and its association with individuals' characteristics and COVID-19 symptoms in an observational study. We enrolled 4735 subjects (corresponding to 80% of all personnel) over a period of 5 months when the spreading of the virus was drastically reduced. For each participant, we determined the rate of antibody increase or decrease over time in relation to 93 features analyzed in univariate and multivariate analyses through a machine learning approach. In individuals positive for IgG (≥ 12 AU/mL) at the beginning of the study, we found an increase [$p= 0.0002$] in antibody response in symptomatic subjects, particularly with anosmia/dysgeusia (OR 2.75, 95% CI 1.753 – 4.301), in a multivariate logistic regression analysis. This may be linked to the persistence of SARS-CoV-2 in the olfactory bulb.

Introduction

It is becoming clear that the antibody response to SARS-CoV-2 can last at least 6 months in symptomatic patients [1], but it seems to decline in asymptomatics [2]. Similarly, a reduction of antibody response in asymptomatic individuals was shown in a study with a fewer number of individuals (n = 37) [3]. The antibody response in COVID-19 patients is associated with the establishment of a memory B cell response which is higher at 6 months [1], however, it is not clear whether there are features that correlate with this sustained B cell response. We previously showed that an anti-SARS-CoV-2 serological analysis allowed us to follow the diffusion of the virus within healthcare facilities in areas differently hit by the virus [4]. At 5 months of distance, we analyzed the duration of this antibody response and evaluated whether there were features correlating with maintenance, reduction or increase of the antibody response.

Results

We assessed the correlation of the rate of antibody increase or decrease with the different analyzed features. In Tables 1 and 2 are reported the rates for individual classes of features with relative statistical analysis. As shown, in the 5 months of observation, females sustained the antibody response better than males ($p = 0.01$); similarly non-medical healthcare professionals (specifically, healthcare partner operators) had higher antibody rates ($p = 0.0009$). The levels of antibodies increased in hospitals located in the Bergamo area (Castelli and Gavazzeni $p < 0.0001$) (Table 1) which was more hit by COVID-19 (37 – 43% of individuals with IgG ≥ 12)[7]. More important, the IgG rate in individuals which were positive for IgG (IgG ≥ 12 AU/mL; n = 613) at the beginning of the study was increased ($p<0.000001$) over time, and this increase was either minor in asymptomatics (n = 91, $p = 0.00003$) and paucisymptomatics (n = 203) or strong in symptomatics (n = 319, $p = 0.0006$) (Table 2). On the contrary, those that had an intermediate IgG titer (3.8 < IgG < 12 AU/mL considered as negative) displayed all a significant reduction in IgG rate ($p < 0.000001$) (Table 1). This may be due to a noise in test analysis as these subjects are considered as negative for SARS-CoV-2 IgG according to manufacturer. Many symptoms, including fever, cough, muscle pain, asthenia, tachycardia and anosmia/dysgeusia, correlated with an increase of antibodies in the 5-month observation period (Table 2).

As we noticed that the distribution of the rate feature presented a high value of kurtosis (see methods) we restricted the data set to subjects with IgG rates either below the 10th percentile [< -0.033 (n = 454)] or above the 90th percentile [> 0.005 (n = 445)] to prevent a bias-variance problem in machine learning models. The accuracy of these rates was confirmed by a linear regression analysis. In Figure 1a and 1b are shown the regression diagnostic plots of predicted values against residuals of training

and test data according to the threshold (< -0.033 AU/ml*day and > 0.005 AU/ml*day). In Table 3 is shown the Chi-squared analysis for the populations below or above the set threshold rates. We found that as for the previous analysis, males reduced the level of antibodies more than females, even though this difference was no longer statistically significant ($p = 0.06$). The levels of antibodies increased in hospitals located in the Bergamo area (Castelli and Gavazzeni: $p = 0.0032$ and $p = 0.0005$, respectively) which was more hit by COVID-19 (37 – 43% of individuals with IgG ≥ 12) [7] while it decreased in Humanitas Rozzano ($p=0.0806$) which was less heavily hit (10% of individuals with IgG ≥ 12) [7] (Table 3). The rate decreased in asymptomatic (65% of subjects fell in the group < -0.033; $p < 0.000001$), remained constant in paucisymptomatic and increased in symptomatic individuals (62% of subjects were in the group > 0.005; $p < 0.000001$) (Table 4). Interestingly, among the different symptoms, fever, cough, muscle pain, asthenia, dyspnea, tachycardia, chest pain and anosmia/dysgeusia all correlated with a higher number of individuals falling into the group with rate > 0.005, indicating that these symptoms were strongly associated with sustained/increased antibody response ($0.000001 < p < 0.05$, Table 4). Among these, anosmia/dysgeusia was associated with the highest percentage of subjects presenting with increased IgG rate (69%; $p < 0.000001$, Table 4). Having observed differences according to sex, role and site, and since many symptoms are linked, we performed a multivariate logistic regression analysis based in a machine learning approach (see methods). We performed a Bagging classifier of 7 logistic analysis on a training (Accuracy = 76.26; ROC AUC = 76.30; Recall = 81.14) and a test dataset (Accuracy = 72.00; ROC AUC = 72.12; Recall = 81.08). In Figure 1c is shown the multivariate logistic regression analysis. We found that the increased rate was associated primarily with anosmia/dysgeusia (regression coefficient=1.0, 95% CI 0.56 – 1.46) and with chest pain (regression coefficient=0.84, 95% CI 0.24 – 1.44), while the decreased rate was associated to subjects with intermediate IgG (3.8 < IgG < 12) (regression coefficient = -1.61, 95% CI -2.03 – -1.0), which may be related to a noise in the instrument testing, and with past neoplasia (regression coefficient = -1.38, 95% CI -2.4 – -0.37). Interestingly, 54% of subjects with chest pain also presented loss of smell/taste while only 22% of subjects with smell/taste dysfunction also had chest pain, suggesting that IgG increase in the symptomatic population is primarily linked to anosmia and dysgeusia (not shown). In figure 1d are shown the odds ratio relative to figure 1c, which for chest pain is 2.32 (95% CI 1.27 – 4.24), for anosmia/dysgeusia is 2.75 (95% CI 1.75 – 4.30), for subjects with intermediate IgG (3.8 < IgG < 12) is 0.2 (95% CI 0.13 – 0.30) and for subjects with past neoplasia is 0.25 (95% CI 0.09 – 0.69). Overall, these results indicate that although many symptoms are associated with an increase of IgG abundance in the observation time, only anosmia/dysgeusia and chest pain are associated to a higher IgG rate in the multivariate logistic

110 regression analysis. By contrast the population with past neoplasia or intermediate levels of IgG (3.8
111 < IgG < 12 AU/mL) are the one that display a reduction in IgG.
112
113 **Discussion**
114 We analyzed the 5-month duration of an antibody response to SARS-CoV-2 in personnel from 9
115 healthcare facilities and an international medical school (Humanitas University) in Northern Italy in
116 areas differently hit by the virus [4]. We show that the antibody response is stable both in symptomatic
117 and asymptomatic/paucisymptomatic individuals and is increased in females and in non-medical
118 healthcare professionals. Previously, it has been shown in a study conducted in the British population
119 that the antibody response declines of nearly 22% in symptomatic individuals and of 64% in
120 asymptomatic individuals [2]. However, this study was based on a prick qualitative test and thus the
121 decline may be related to the sensitivity of the test. We also observed that the antibody response
122 declined when we analyzed the group (3.8 < IgG < 12 AU/mL) with IgG between the limit of
123 detection (3.8 AU/mL) and the threshold of positivity (IgG ≥ 12 AU/mL), as set by manufacturer.
124 Whether this is linked to a noise of the instrument that may change according to the test or to a real
125 reduction in an antibody response that may or may not be specific to SARS-CoV-2, remains to be
126 established. When we analyzed the extremes, i.e. the individuals with higher rates of antibody
127 increase or decrease (< -0.033 and > 0.005 AU/mL*day) we observed that asymptomatics had higher
128 negative rates while symptomatics tended to continue increasing the antibody levels suggesting that
129 extreme changes in rate separate the symptomatics from the asymptomatics. As during the
130 observation time there was very limited viral diffusion in Northern Italy, as confirmed also by the
131 finding that only 2 individuals became IgG positive and 2981 remained IgG negative throughout the
132 study (all excluded from the analysis), we can conclude that the sustained or augmented antibody
133 response may not be linked to a re-exposure to the virus. In an attempt to address what improved the
134 antibody response, we found that several symptoms were associated to increased rates of antibodies,
135 however, in a multivariate logistic analysis only anosmia/dysgeusia and chest pain were linked with
136 the highest regression coefficients. Chest pain and anosmia are long-lasting symptoms in COVID-19
137 patients [8]. In addition, anosmia and/or dysgeusia are very common as they are found in around 50-
138 70% of subjects affected by COVID-19 [9, 10]. In our cohort (Table 2), 49% of IgG positive subjects
139 had anosmia/dysgeusia, 28% chest pain and 13.7% both anosmia/dysgeusia and chest pain,
140 suggesting that indeed these two symptoms may, either alone or in combination, associate with IgG
141 increase. We and others previously found that anosmia/dysgeusia together with fever were the
142 symptoms that mostly characterized SARS-CoV-2 exposure [4, 11]. In agreement, anosmia and
143 dysgeusia have been proposed to be used to track SARS-CoV-2 diffusion [12]. Interestingly, SARS-

144 CoV-2 can infect the olfactory epithelium [13, 14], including olfactory sensory neurons, support cells
145 and immune cells, that express the viral entry receptors ACE2 and TMPRSS2 [13, 15, 16]. Here, the virus
146 can persist long and induce local inflammation [14] and olfactory bulb abnormalities [17, 18, 19]. In
147 agreement, the loss of smell and taste can persist in individuals even with RT-PCR SARS-CoV-2
148 negativity in the nasopharyngeal swab [14, 20].
149 Overall, these data suggest that increased antibody response in patients with anosmia/dysgeusia may
150 be linked to persistence of the virus in the olfactory bulb which through local inflammation and
151 release of antigens, maintains and boosts the antibody response. This study opens new perspectives
152 on the immunity to SARS-CoV-2 and warrants further investigation on the role of anosmia/dysgeusia
153 on antibody response through the design of prospective observational studies coupling the testing of
154 SARS-CoV-2 persistence in the olfactory bulb, loss of smell or taste and antibody titers.
155
156 **Methods**
157 Study population
158 This observational study has been approved by the international review board of Istituto Clinico
159 Humanitas for all participating institutes (clinicaltrial.gov NCT04387929). Accrual was on a
160 voluntary basis: it started on April 28th and more than 80% of personnel participated (n = 4735). The
161 study foresees 4 blood collections every 3/4 months. 10 different centers participate: Istituto Clinico
162 Humanitas (ICH), Rozzano (MI); Humanitas Gavazzeni, Bergamo; Humanitas Castelli, Bergamo;
163 Humanitas Mater Domini (HMD), Castellanza (VA); Humanitas Medical Center, HMC, Varese;
164 Humanitas University, Pieve Emanuele (MI); Humanitas San Pio X, Milano; Humanitas Cellini,
165 Torino; Humanitas Gradenigo, Torino; Clinica Fornaca, Torino. All participants signed an informed
166 consent and filled a questionnaire before blood collection. We analyzed 93 features (72 categorical
167 and 17 numerical and 4 temporal) including, age, sex, location, professional role, time between
168 sample collections, COVID-19 symptoms (fever, sore throat, cough, muscle pain, asthenia,
169 anosmia/dysgeusia (loss of smell and taste), gastrointestinal disorders, conjunctivitis, dyspnea, chest
170 pain, tachycardia, pneumonia), home exits and smart-working, comorbidities (diabetes, asthma,
171 neoplasia, autoimmunity, cardiovascular disorders, hepatic disorders). After excluding for employees
172 that became positive for SARS-CoV-2 IgG (n = 2) during the observation period and those that
173 dropped from phase I or for which we were missing at least two features, we analyzed 4534
174 participants (4.25% drop out). Here we show the results of the end of phase II (second blood
175 sampling).
176
177 IgG measure

For the determination of IgG anti SARS-CoV-2, the Liaison SARS-CoV-2 S1/S2 IgG assay (DiaSorin, Saluggia (VC), Italy) was used [5]. The method is an indirect chemiluminescence immunoassay for the determination of anti-S1 and anti-S2 specific antibodies. According to kit manufacturer, the test discriminates among negative (< 12AU/mL; with 3.8 as the limit of IgG detection) and positive (≥ 12.0 AU/mL) subjects. However, we considered also individuals with IgG comprised between 3.8 and 12 AU/mL (which we called IgG med: 3.8 < IgG < 12.0 AU/mL). Consistency and reproducibility of the antibody test in samples collected in the two time points was confirmed for a limited number of individuals (n = 50) displaying different degrees of IgG positivity.

Statistical analysis and model

We first cleared the dataset by eliminating data from all of those subjects that did not develop an IgG response over time (IgG ≤ 3.8 at the beginning and at the end of the examination) (n = 2981). We then analyzed the rate of antibody response defined as:

$$RATE = \frac{IgG\ phase\ II - IgG\ phase\ I}{\Delta days} = [AU\ /mL * day]$$

Positive rates mean increased antibody response, while negative rates indicate reduction of antibody response between the two analyzed time points.

For statistical analysis, we performed both a univariate and a multivariate analysis. We applied Wilcoxon-Mann-Whitney statistical non-parametric test to compare the antibody rate distribution between classes of subjects (Table 1 and Table 2).

We analyzed the distribution of the rate feature and found a high value of kurtosis (461) around the median value of 0.016, hence to perform a multivariate analysis we restricted the data set to subjects with IgG rates either below the 10th percentile or above the 90th percentile to prevent a bias-variance problem in machine learning models and subjected the data to a linear regression analysis between the training and test data sets, where the target variable (rate of antibodies) was standardized using the Yeo-Johnson method [6]. We then applied Chi-squared statistical test to evaluate differences between classes and the rate thresholds described above (Tables 3 and 4). In order to evaluate the possible interactions between features and the rate of antibody response, we developed a multivariate approach to perform a binary classification between subjects who increased or decreased the level of antibodies. A set of 7 logistic regressions has been applied on data using a bootstrap procedure (samples are drawn with replacement) and the output of each classifier has been averaged by a Bagging classifier to obtain the final output. The selection of hyperparameters of the machine learning

model and the feature selection has been performed with a Bayesian optimization approach based on cross validation (4 folds, stratified by outcome).

Data availability

Data supporting the findings of this study are available within the paper. All other data are available from the corresponding author upon reasonable requests.

References

1. Dan JM, *et al.* Immunological memory to SARS-CoV-2 assessed for greater than six months after infection. *bioRxiv*, 2020.2011.2015.383323 (2020).
2. Ward H, *et al.* Declining prevalence of antibody positivity to SARS-CoV-2: a community study of 365,000 adults. *medRxiv*, 2020.2010.2026.20219725 (2020).
3. Long QX, *et al.* Clinical and immunological assessment of asymptomatic SARS-CoV-2 infections. *Nat Med* **26**, 1200-1204 (2020).
4. Sandri MT, *et al.* SARS-CoV-2 serology in 4000 health care and administrative staff across seven sites in Lombardy, Italy. *medRxiv*, 2020.2005.2024.20111245 (2020).
5. Bonelli F, *et al.* Clinical And Analytical Performance Of An Automated Serological Test That Identifies S1/S2 Neutralizing IgG In COVID-19 Patients Semiquantitatively. *J Clin Microbiol*, (2020).
6. Yeo IK, Johnson RA. A new family of power transformations to improve normality or symmetry. *Biometrika* **87**, 954-959 (2000).
7. Sandri M, Elena Azzolini, Valter Torri, Sara Carloni, Michele Tedeschi, Massimo Castoldi, Alberto Mantovani, Maria Rescigno. IgG serology in health care and administrative staff populations from 7 hospital representative of different exposures to SARS-CoV-2 in Lombardy, Italy. *MedrXiv*, (2020).
8. Carfì A, Bernabei R, Landi F, Group ftGAC-P-ACS. Persistent Symptoms in Patients After Acute COVID-19. *JAMA* **324**, 603-605 (2020).
9. Butowt R, von Bartheld CS. Anosmia in COVID-19: Underlying Mechanisms and Assessment of an Olfactory Route to Brain Infection. *Neuroscientist*, 1073858420956905 (2020).
10. Lechien JR, *et al.* Clinical and epidemiological characteristics of 1420 European patients with mild-to-moderate coronavirus disease 2019. *J Intern Med* **288**, 335-344 (2020).

11. Kumar N, Bhartiya S, Desai S, Mutha A, Beldar A, Singh T. Seroprevalence of Antibodies Against SARS-CoV-2 Among Health Care Workers in Mumbai, India. *Asia Pac J Public Health*, 1010539520977307 (2020).
12. Pierron D*, et al.* Smell and taste changes are early indicators of the COVID-19 pandemic and political decision effectiveness. *Nat Commun* **11**, 5152 (2020).
13. Brann DH*, et al.* Non-neuronal expression of SARS-CoV-2 entry genes in the olfactory system suggests mechanisms underlying COVID-19-associated anosmia. *Sci Adv* **6**, (2020).
14. De Melo GD*, et al.* COVID-19-associated olfactory dysfunction reveals SARS-CoV-2 neuroinvasion and persistence in the olfactory system. *bioRxiv*, 2020.2011.2018.388819 (2020).
15. Bilinska K, Jakubowska P, Von Bartheld CS, Butowt R. Expression of the SARS-CoV-2 Entry Proteins, ACE2 and TMPRSS2, in Cells of the Olfactory Epithelium: Identification of Cell Types and Trends with Age. *ACS Chem Neurosci* **11**, 1555-1562 (2020).
16. Chen M*, et al.* Elevated ACE-2 expression in the olfactory neuroepithelium: implications for anosmia and upper respiratory SARS-CoV-2 entry and replication. *Eur Respir J* **56**, (2020).
17. Politi LS, Salsano E, Grimaldi M. Magnetic Resonance Imaging Alteration of the Brain in a Patient With Coronavirus Disease 2019 (COVID-19) and Anosmia. *JAMA Neurol* **77**, 1028-1029 (2020).
18. Tsivgoulis G*, et al.* Olfactory bulb and mucosa abnormalities in persistent COVID-19-induced anosmia: a magnetic resonance imaging study. *Eur J Neurol*, (2020).
19. Laurendon T*, et al.* Bilateral transient olfactory bulb edema during COVID-19-related anosmia. *Neurology* **95**, 224-225 (2020).
20. Yan CH, Prajapati DP, Ritter ML, DeConde AS. Persistent Smell Loss Following Undetectable SARS-CoV-2. *Otolaryngol Head Neck Surg* **163**, 923-925 (2020).

Acknowledgments

This work was partially supported by a philantropic donation by Dolce & Gabbana, by the Italian Ministry of Health (Ricerca corrente) and by Fondazione Humanitas per la Ricerca.

We would like to thank all the employees that volunteered to participate to this study, all the nurses and personnel that collected the samples and the laboratory technicians that run the serological and rinopharyngeal tests. We would also like to thank the Humanitas management and staff, Drs Patrizia Meroni, Michele Lagioia and Michele Tedeschi, who warmly supported this study for the safety of the employees.

Author contribution

R.L. and L.U.: performed data analysis; C.P.: contributed to data analysis and manuscript writing; G.A. and V.S.: contributed to data analysis; M.T.S.: coordinated and supervised the laboratory analyses; E.A.: coordinated the recruitment and sampling of subjects (project administration) and participated in clinical study design; M.S.: carried out the laboratory analyses; A.M.: conceptualization and funding acquisition; M.R.: conceived the study, analyzed the data and wrote the manuscript.

Competing interests

The authors declare no competing interests.

290 **Tables**

291 Table 1. Demographic distribution of antibody rates.

		counts	min	25 perc	50 perc	75 perc	max	mean	St.Dev.	p_value [a]
Sex	F	1105	-3.39	-0.04	-0.02	0.02	16.32	0.07	0.61	0.0139
	M	448	-0.88	-0.04	-0.02	0	10.37	0.03	0.53	
Age	21-30	300	-3.39	-0.04	-0.02	0.01	4.51	0.03	0.41	0.2644
	31-40	365	-0.92	-0.04	-0.02	0.01	2.26	0.02	0.21	0.2302
	41-50	455	-2.49	-0.04	-0.02	0.03	3.11	0.06	0.37	0.1083
	51-60	309	-0.88	-0.04	-0.01	0.02	16.32	0.1	0.96	0.1442
	60+	124	-0.68	-0.04	-0.02	-0.01	10.37	0.12	0.99	0.0586
BMI [b]	18.5≤ BMI <25	940	-0.92	-0.04	-0.02	0.01	10.37	0.03	0.41	0.3821
	BMI ≥ 30	106	-3.39	-0.05	-0.02	0.14	16.32	0.23	1.69	0.2182
	25≤ BMI <30	347	-0.57	-0.04	-0.02	0.01	3.11	0.05	0.29	0.3933
	BMI < 18.5	73	-0.22	-0.04	-0.02	0.02	1.27	0.05	0.26	0.3959
IgG class phase 1	IgG ≥ 12	613	-3.39	-0.08	0.02	0.23	16.32	0.18	0.92	2.1 E-10
	IgG ≤ 3.8	74	0.01	0.01	0.02	0.03	0.08	0.02	0.01	1.5 E-15
	3.8 < IgG < 12	866	-0.13	-0.03	-0.02	-0.01	0.83	-0.02	0.04	8.0 E-22
Role	Other [c]	200	-0.57	-0.04	-0.02	-0.01	16.32	0.09	1.17	0.0116
	Anesthesiologist	19	-0.83	-0.04	-0.02	0.01	0.09	-0.05	0.2	0.4305 [b]
	Biologist	18	-0.18	-0.05	-0.02	-0.02	0.03	-0.04	0.05	0.0978 [b]
	Surgeon	67	-0.88	-0.05	-0.02	0.07	0.61	0.02	0.21	0.2653
	Physiotherapist	21	-0.12	-0.04	-0.02	0.01	0.35	0.01	0.1	0.4204
	Nurse	398	-3.39	-0.04	-0.02	0.04	3.11	0.08	0.41	0.0581
	Physician	210	-0.92	-0.04	-0.01	0.02	10.37	0.08	0.75	0.0804
	Healthcare Partner Operator	149	-2.49	-0.03	-0.01	0.17	1.55	0.1	0.38	0.0009
	Front office (PARC)	108	-0.32	-0.04	-0.02	0.01	2.55	0.03	0.27	0.2892
	Researcher	50	-1.5	-0.03	-0.02	-0.01	4.51	0.06	0.68	0.1514
	Cleaning service	29	-0.65	-0.03	-0.02	0.03	1.72	0.09	0.41	0.2414
	Transport service	14	-0.39	-0.04	-0.01	0.02	2.22	0.13	0.62	0.4359 [b]
	Staff	188	-0.4	-0.04	-0.02	-0.01	0.98	0	0.15	0.0026
	Student	20	-0.19	-0.03	-0.02	-0.01	0.21	-0.02	0.08	0.3705 [b]
	Laboratory technician	31	-0.17	-0.03	-0.01	-0.01	0.51	0.02	0.15	0.3243
	Radiology technician	31	-0.18	-0.03	-0.02	0.01	1.01	0.04	0.22	0.4002
Site	Other [c]	21	-0.19	-0.04	-0.02	-0.01	2.22	0.08	0.49	0.2080
	Casa di Cura Cellini	51	-0.17	-0.05	-0.03	-0.01	1.27	0.01	0.22	0.0111
	Clinica Fornaca di Sessant	47	-0.4	-0.05	-0.03	0	0.53	0	0.15	0.0833
	Humanitas Castelli	87	-0.28	-0.04	0.02	0.21	3.11	0.24	0.62	4.7 E-05
	Humanitas Gavazzeni	313	-0.88	-0.04	-0.01	0.15	10.37	0.13	0.68	0.0001
	Humanitas Gradenigo	109	-2.49	-0.04	-0.02	-0.01	0.67	-0.01	0.27	0.0586
	Humanitas Mater Domini	105	-0.17	-0.03	-0.02	-0.01	0.63	0	0.1	0.3412
	Humanitas Medical Care	23	-0.19	-0.04	-0.02	-0.01	0.02	-0.03	0.04	0.0969
	Humanitas Rozzano	667	-3.39	-0.04	-0.02	0	16.32	0.04	0.7	0.0338
	Humanitas San Pio X	98	-0.68	-0.03	-0.02	0	2.39	0.04	0.29	0.2968
	Humanitas University	32	-0.2	-0.04	-0.02	-0.01	0.19	-0.03	0.08	0.0590

[a] = Wilcoxon-Mann-Whitney test
[b] = Some subjects did not indicate their BMI
[c] = Refers to volunteers and other professionals that operate in several structures
[d] = Minority class is less or equal to 20 (Wilcoxon-Mann-Whitney test is not reliable)

292
293
294
295

Table 2. Antibody rates according to symptoms.

			counts	min	25 perc	50 perc	75 perc	max	mean	St. Dev.	p_value [a]
Class symptoms phase 1 (subjects with IgG ≥ 12)	Asymptomatic		91	-1.5	-0.13	-0.05	0.09	2.39	0.04	0.5	0.00003
	Paucisymptomatic		203	-3.39	-0.08	0.02	0.21	4.51	0.14	0.56	0.32865
	Symptomatic		319	-2.49	-0.06	0.07	0.28	16.32	0.24	1.16	0.00057
Symptoms phase 1 (subjects with IgG ≥ 12)	Fever	No	350	-3.39	-0.09	0.01	0.21	4.51	0.13	0.02725	0.02725
		Yes	263	-2.49	-0.06	0.05	0.25	16.32	0.24		
	Low-grade fever	No	481	-3.39	-0.08	0.02	0.23	16.32	0.19	0.17265	0.17265
		Yes	132	-0.39	-0.07	0.05	0.26	2.55	0.14		
	Cough	No	372	-3.39	-0.08	0.01	0.19	16.32	0.14	0.01120	0.01120
		Yes	241	-2.49	-0.07	0.07	0.31	10.37	0.24		
	Sore throath	No	353	-3.39	-0.09	0.02	0.23	16.32	0.19	0.08309	0.08309
		Yes	260	-2.49	-0.07	0.04	0.25	4.51	0.16		
	Muscle pain	No	299	-1.5	-0.1	0	0.2	4.51	0.13	0.00763	0.00763
		Yes	314	-3.39	-0.06	0.06	0.27	16.32	0.22		
	Asthenia	No	341	-3.39	-0.1	0	0.21	16.32	0.17	0.00574	0.00574
		Yes	272	-2.49	-0.06	0.07	0.25	10.37	0.19		
	Anosmia / dysgeusia	No	313	-3.39	-0.12	-0.01	0.2	16.32	0.14	0.00006	0.00006
		Yes	300	-0.86	-0.05	0.06	0.28	10.37	0.22		
	Gastrointestinal disorders	No	403	-3.39	-0.08	0.03	0.21	10.37	0.17	0.46477	0.46477
		Yes	210	-2.49	-0.08	0.02	0.25	16.32	0.19		
	Conjunctivitis	No	517	-3.39	-0.08	0.02	0.21	16.32	0.18	0.16050	0.16050
		Yes	96	-2.49	-0.07	0.04	0.32	1.72	0.14		
	Dyspnea	No	493	-3.39	-0.08	0.02	0.23	16.32	0.16	0.34700	0.34700
		Yes	120	-2.49	-0.07	0.05	0.26	10.37	0.24		
	Chest pain	No	502	-3.39	-0.08	0.01	0.24	10.37	0.15	0.08088	0.08088
		Yes	111	-0.39	-0.04	0.07	0.22	16.32	0.3		
	Tachycardia	No	512	-3.39	-0.08	0.01	0.21	4.51	0.13	0.02353	0.02353
		Yes	101	-2.49	-0.06	0.08	0.32	16.32	0.4		
	Pneumonia	No	568	-3.39	-0.08	0.02	0.22	16.32	0.15	0.18692	0.18692
		Yes	45	-0.88	-0.06	0.06	0.38	10.37	0.51	1.69	

[a] = Wilcoxon-Mann-Whitney test

312 Table 3. Chi-squared analysis of groups < 10th percentile and > 90th percentile.

		< 10 perc	> 90 perc	p_value
Sex	F	321	340	0.0627
	M	133	105	
Age	21-30	93	83	0.5429
	31-40	108	94	0.3804
	41-50	121	142	0.0970
	51-60	88	98	0.3711
	60+	44	28	0.0793
BMI [a]	18.5≤ BMI <25	267	255	0.6963
	BMI ≥ 30	58	72	0.1750
	25≤ BMI <30	106	93	0.4214
	BMI < 18.5	23	25	0.8261
Role	Other [b]	63	37	0.0109
	Anesthesiologist	6	7	0.9701
	Biologist	5	2	0.4640
	Surgeon	25	21	0.7007
	Physiotherapist	6	7	0.9710
	Nurse	113	132	0.1255
	Physician	56	61	0.6082
	Healthcare Partner Operator	38	64	0.0062
	Front office (PARC)	31	28	0.8494
	Researcher	13	9	0.5485
	Cleaning service	7	9	0.7698
	Transport Service	5	5	0.7747 [c]
	Staff	69	44	0.0214
	Student	5	4	0.9759 [c]
	Laboratory Technician	6	7	0.9701
	Radiology Technician	6	8	0.7587
Site	Other [b]	6	3	0.5223 [c]
	Casa di Cura Cellini	19	8	0.0573
	Clinica Fornaca di Sessant	18	12	0.3828
	Humanitas Castelli	23	47	0.0032
	Humanitas Gavazzeni	97	142	0.0005
	Humanitas Gradenigo	36	22	0.0918
	Humanitas Mater Domini	20	23	0.7041
	Humanitas Medical Care	8	3	0.2379 [c]
	Humanitas Rozzano	190	160	0.0806
	Humanitas San Pio X	27	21	0.5025
	Humanitas University	10	4	0.1905 [c]

[a] = Some subjects did not indicate their BMI
[b] = Refers to volunteers and other professionals that operate in several structures
[c] = Minority class is less or equal to 5 (chi-square test is not reliable)

313
314

Table 4. Chi-squared analysis of groups < 10th percentile and > 90th percentile per symptoms.

			<10perc	>90perc	p_value	%Yes<10perc [a]	%Yes>90perc [a]
Symptoms	Fever	No	360	284	3.9E-07		
		Yes	94	161		37	63
	Low-grade fever	No	388	359	0.0678		
		Yes	66	86		43	57
	Cough	No	335	284	0.0016		
		Yes	119	161		43	58
	Sore throath	No	302	271	0.0923		
		Yes	152	174		47	53
	Muscle pain	No	310	246	0.0001		
		Yes	144	199		42	58
	Asthenia	No	340	268	3.7E-06		
		Yes	114	177		39	61
	Anosmia/dysgeusia	No	364	247	4.0E-15		
		Yes	90	198		31	69
	Gastrointestinal disorders	No	336	313	0.2485		
		Yes	118	132		47	53
	Conjunctivitis	No	400	382	0.3633		
		Yes	54	63		46	54
	Dyspnea	No	405	375	0.0370		
		Yes	49	70		41	59
	Chest pain	No	415	367	0.0001		
		Yes	39	78		33	67
	Tachycardia	No	406	371	0.0107		
		Yes	48	74		39	61
	Pneumonia	No	440	420	0.0889		
		Yes	14	25		36	64
	Total of Symptoms in phase I	Asymptomatic	150	80	3.4E-07	65	35
		Paucisymptomatic	179	158	0.2520	53	47
		Symptomatic	125	207	5.6E-09	38	62
Comorbidities	Chronic Obstructive Bronchopneumopathy	No	451	444	0.6305 [b]		
		Yes	3	1			
	Asthma	No	430	412	0.2408		
		Yes	24	33			
	Dyslipidemia	No	413	404	0.9834		
		Yes	41	41			
	Past Neoplasia	No	431	438	0.0063		
		Yes	23	7			
	Hypertension	No	401	409	0.0915		
		Yes	53	36			
	Past Coronaropathies	No	454	443	0.4702 [b]		
		Yes	0	2			
	Atrial Fibrillation	No	450	441	0.7439 [b]		
		Yes	4	4			
	Past Stroke/ TIA	No	452	445	0.4878 [b]		
		Yes	2	0			
	Steatosis/Cyrrosis	No	448	444	0.1359 [b]		
		Yes	6	1			
	Chronic kidney failure	No	453	445	0.9920 [b]		
		Yes	1	0			
	Other liver diseases	No	452	441	0.6641 [b]		
		Yes	2	4			
	Rheumatological diseases	No	445	431	0.3715		
		Yes	9	14			
	Other diseases of the immune system	No	418	409	0.9726		
		Yes	36	36			
	Diabetes mellitus	No	452	443	0.6305 [b]		
		Yes	2	2			
	Gotta	No	452	445	0.4878 [b]		
		Yes	2	0			

[a] = Percentage of subjects with symptoms (Yes) per rate class
[b] = Minority class is less or equal to 5 (chi-square test is not reliable)

Figure

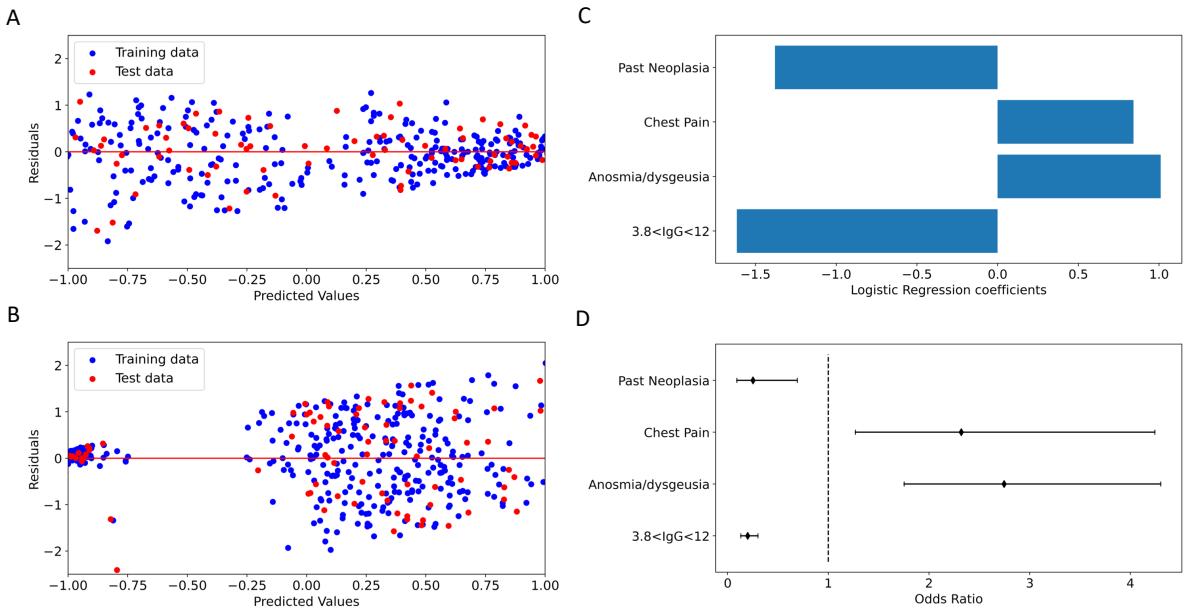

Figure 1: a, Dataset < 10% percentile Regression Diagnostic plot of predicted values against residuals of training and test data; **b**, Dataset >90% percentile Regression Diagnostic plot of predicted values against residuals of training and test data; **c**, Barplot with Logistic Regression coefficients for most important features; **d**, Odds ratio of Logistic Regression with confidence intervals (95%) for the most important features.