

Clinical diagnosis of COVID-19. A multivariate logistic regression analysis of symptoms of COVID-19 at presentation.

Fleitas, Pedro E., Paz, Jorge A., Simoy, Mario I., Vargas, Carlos, Cimino, Rubén O., Krolewiecki, Alejandro J. y Aparicio, Juan P.

Cita:

Fleitas, Pedro E., Paz, Jorge A., Simoy, Mario I., Vargas, Carlos, Cimino, Rubén O., Krolewiecki, Alejandro J. y Aparicio, Juan P. (2021). *Clinical diagnosis of COVID-19. A multivariate logistic regression analysis of symptoms of COVID-19 at presentation. e-ISSN: 2248-2997 - GERMS, 11, 221-237.*

Dirección estable: <https://www.aacademica.org/jorge.paz/141>

ARK: <https://n2t.net/ark:/13683/prpd/g0F>



Esta obra está bajo una licencia de Creative Commons.
Para ver una copia de esta licencia, visite
<https://creativecommons.org/licenses/by-nc-nd/4.0/deed.es>.

Acta Académica es un proyecto académico sin fines de lucro enmarcado en la iniciativa de acceso abierto. Acta Académica fue creado para facilitar a investigadores de todo el mundo el compartir su producción académica. Para crear un perfil gratuitamente o acceder a otros trabajos visite: <https://www.aacademica.org>.

Clinical diagnosis of COVID-19. A multivariate logistic regression analysis of symptoms of COVID-19 at presentation

Pedro E. Fleitas^{1*}, Jorge A. Paz², Mario I. Simoy³, Carlos Vargas⁴, Rubén O. Cimino⁵, Alejandro J. Krolewiecki⁶, Juan P. Aparicio⁷

Abstract

Introduction The objective of this cross-sectional study was to describe the main symptoms associated with COVID-19, and their diagnostic characteristics, to aid in the clinical diagnosis.

Methods An analysis of all patients diagnosed by RT-PCR for SARS-CoV-2 between April and May 2020 in Argentina was conducted. The data includes clinical and demographic information from all subjects at the time of presentation (n=67318, where 12% were positive for SARS-CoV-2). The study population was divided into four age groups: pediatric (0-17 years), young adults (18-44 years), adults (45-64 years), and elderly (65-103 years). Multivariate logistic regression was used to measure the association of all symptoms and to create a diagnostic model based on symptoms.

Results Symptoms associated with COVID-19 were anosmia, dysgeusia, headache, low-grade fever, odynophagia, and malaise. However, the presentation of these symptoms was different between the different age groups. In turn, at the time of presentation, the symptoms associated with respiratory problems (chest pain, abdominal pain, and dyspnea) had a negative association with COVID-19 or did not present statistical relevance. On the other hand, the model based on 16 symptoms, age and sex, presented a sensitivity of 80% and a specificity of 46%.

Conclusions There were significant differences between the different age groups. Additionally, there were interactions between different symptoms that were highly associated with COVID-19. Finally, our findings showed that a regression model based on multiple factors (age, sex, interaction between symptoms) can be used as an accessory diagnostic method or a rapid screening of suspected COVID-19 cases.

Keywords COVID-19, symptoms, clinical diagnosis.

Introduction

In December 2019, cases of pneumonia of unknown cause were reported in Wuhan, China. These cases were epidemiologically associated with a wild animal food market.¹ A new coronavirus was identified as the causative agent

and provisionally named 2019 novel coronavirus (2019-nCoV).¹⁻³ In January 2020, China reported 5900 cases of 2019-nCoV in 33 provinces or municipalities.² In February 2020, the World Health Organization (WHO) officially named the coronavirus that causes the pandemic as SARS-

Received: 10 February 2021; revised: 28 May 2021; accepted: 31 May 2021.

¹MSc, Instituto de Investigaciones de Enfermedades Tropicales (IIET-CONICET), Universidad Nacional de Salta, Alvarado 751, San Ramon de la Nueva Oran, zip code A4530, Salta, Argentina, Cátedra de Química Biología, Facultad de Ciencias Naturales, Universidad Nacional de Salta, Av. Bolivia 5150, zip code A4400, Salta, Argentina; ²PhD, Instituto de Estudios Laborales y del Desarrollo Económico (IELDE), Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Universidad Nacional de Salta, Av. Bolivia 5150, zip code A4400, Salta, Argentina; ³PhD, Instituto de Investigaciones en Energía no Convencional (INENCO), Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET),

Universidad Nacional de Salta, Av. Bolivia 5150, zip code A4400, Salta, Argentina, Instituto Multidisciplinario sobre Ecosistemas y Desarrollo Sustentable, Universidad Nacional del Centro de la Provincia de Buenos Aires, Pinto 399, zip code B7000, Tandil, Buenos Aires, Argentina; ⁴MSc, Instituto de Investigaciones de Enfermedades Tropicales (IIET-CONICET), Universidad Nacional de Salta, Alvarado 751, San Ramon de la Nueva Oran, zip code A4530, Salta, Argentina; ⁵PhD, Instituto de Investigaciones de Enfermedades Tropicales (IIET-CONICET), Universidad Nacional de Salta, Alvarado 751, San Ramon de la Nueva Oran, zip code A4530, Salta, Argentina, Cátedra de Química Biología, Facultad de Ciencias Naturales, Universidad Nacional de Salta, Av. Bolivia 5150, zip code A4400, Salta, Argentina; ⁶MD, PhD,

CoV-2, and COVID-19 the disease that this virus causes.⁴ At the beginning of May 2021, the number of people infected with SARS-CoV-2 is over 155 million worldwide with more than 3 million deaths.⁵ The rapid spread of the virus has demonstrated its ability to overwhelm the health systems of developed countries; this implies a greater danger in countries with vulnerable health systems.⁶ These particular characteristics underline the importance of adopting social-distancing measures including quarantines, lockdowns and limitations in international travel implemented by most governments.

The identification and isolation of infected people through contact tracing or passive detection is of vital importance for controlling disease spread.⁷ Currently, SARS-CoV-2 real-time polymerase chain reaction (RT-PCR) from nasopharyngeal and throat swabs is the diagnostic procedure of choice.⁸ RT-PCR has a sensitivity of 56% and a specificity of 86%;⁹ however, the collection of quality samples in different stages of the infection is a problem that has caused a high rate of false negatives.^{8,10} Additional diagnostic approaches include serological tests based on the detection of IgM or IgG antibodies. These serological tests have a sensitivity range of 77% to 98% with specificities greater than 91%;¹¹ but they have a low sensitivity during the first week of infection (30%), therefore being impractical for management of early acute cases.¹² It should be noted that the sensitivity of serological tests increases with time, while the sensitivity of RT-

PCR decreases.¹¹ On the other hand, chest CT has proven to be a valuable tool in the detection of COVID-19 cases. Chest CT has a sensitivity of 97% with a specificity of 25%.⁷ Another way to detect COVID-19 cases is by identifying the symptoms that characterize the clinical presentation of the disease. The main clinical manifestations of COVID-19 include fever, cough, fatigue, headache, myalgia among other less frequent symptoms.¹³⁻¹⁵ However, the presentation of symptoms varies among children, young adult and the elderly.^{14,16-18} In addition, most of the studies on COVID-19 restrict the analysis to confirmed cases,^{13,15,18,19} which limits its capacity to calculate the sensitivity and specificity of each sign and symptom.

The aim of this study was to analyze the symptoms included in a large database of suspected cases of COVID-19 that were evaluated by RT-PCR in Argentina, in order to identify the symptoms associated with COVID-19. The objectives of this study were: (i) to identify symptoms associated with COVID-19 in different age groups, (ii) to identify the sensitivity and specificity of the main symptoms associated with COVID-19 and (iii) to create a model that helps to discriminate between cases of COVID-19 and other infections with similar clinical presentations.

Methods

Data source

The data was provided by the National Ministry of Health of Argentina (Ministerio de Salud de la Nación Argentina. Dirección Nacional de Epidemiología y Análisis de la Situación de Salud. Área de Vigilancia). Data on age, sex and RT-PCR results are available online.²⁰ The data on symptoms related to these individuals was requested to the Ministry of Health of the Nation of Argentina for academic reasons according to the Access to Public Information Law No. 27275.²¹ In addition, according to article 5 of Law 25,326 of Argentina,²² the consent or approval of an ethics committee was not necessary because the data comes from publicly accessible sources.

Data was obtained from epidemiological files filled out by physicians at all health facilities in

Instituto de Investigaciones de Enfermedades Tropicales (IIET-CONICET), Universidad Nacional de Salta, Alvarado 751, San Ramon de la Nueva Oran, zip code A4530, Salta, Argentina; ⁷PhD, Instituto de Investigaciones en Energía no Convencional (INENCO), Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Universidad Nacional de Salta, Av. Bolivia 5150, zip code A4400, Salta, Argentina, Simon A. Levin Mathematical, Computational and Modeling Sciences Center, Arizona State University, P.O. Box 873901, Tempe, AZ 85287-3901, USA.

*Corresponding author: Pedro Emanuel Fleitas, pedro.fleittas@gmail.com

Article downloaded from www.germs.ro
 Published June 2021
 © GERMS 2021
 ISSN 2248 - 2997
 ISSN - L = 2248 - 2997

Argentina evaluating suspected cases of COVID-19. This epidemiological file registers a group of 23 signs and symptoms identified at presentation.

This cross-sectional study included 67318 people evaluated through RT-PCR for SARS-CoV-2 throughout the country from 1 April 2020 to 24 May 2020 and included in the registry of the National Ministry of Health. The inclusion criterion was through the definition of suspected cases of COVID-19 by the Ministry of Health (eFile 1, supplementary material). Several variables including age, sex, RT-PCR result (positive or negative), and symptoms at presentation were recorded in the epidemiological file for each individual. Symptoms recorded (presence or absence) included anosmia, dysgeusia, arthralgia, headache, confusion, seizures, diarrhea, dyspnea (defined as the subjective perception of shortness of breath by the patient), abdominal pain, chest pain, low-grade fever (37.5-37.9°C), high-grade fever ($\geq 38^\circ\text{C}$), respiratory failure (need for supplementary oxygen through face mask or mechanical ventilation to maintain an adequate oxygen supply in blood and tissues), conjunctival injection, irritability, malaise, myalgia, food refusal, tachypnea (respiratory rate >20 /minute in adults), use of accessory muscles for breathing (UAMB), cough and vomiting. Before 18 May 2020 the form included “odynophagia” and “sore throat”. After 18 May 2020 only odynophagia remained in the form. Therefore, to avoid eliminating data, these two variables were combined as odynophagia. We did not impose any further exclusion criteria to limit selection bias and all data was included in the analysis. This study was carried out according to the STROBE statement (eFile 2, supplementary material).

Diagnosis by RT-PCR

The diagnosis by RT-PCR was used as a reference standard for the design of the regression model. All kits for use in Argentina were approved by the local regulatory agency (Administración Nacional de Medicamentos, Alimentos y Tecnología Médica – ANMAT). A sample was taken with a swab from each nostril and pharynx. From these samples, an RNA

extraction was performed for the subsequent RT-PCR.

Data analysis

Symptoms of COVID-19 have been reported to differ in different age groups,^{16,18,23} therefore, four age categories were used for the analysis, pediatric (0-17 years), young adult (18-44 years), adult (45-64 years) and the elderly (65-103 years).

The frequencies of comorbidities and symptoms were compared between the different age groups using Chi-square Pearson test. A p value <0.05 was considered significant.

The association of symptoms (presence versus absence) and sex (male versus female) with an RT-PCR positive result for SARS-CoV-2 was studied using multivariate logistic regression analysis considering pairwise interactions in each age group. The effect of the different covariates was quantified by the odds ratio values and a p value <0.05 was considered statistically significant. In addition, the link between associated symptoms and sex in the infected population was analyzed with the same methodology for each of the age groups.

Predictive models

Modeling was performed using logistic regression,

$$p = \frac{1}{1 + e^{-A}}$$

where A depends on the set of symptoms, sex, and age. This was done on the premise that such models should incorporate the fewest significant variables and symptoms possible, and that these variables be collected without the need for a physical exam or laboratory testing (e.g., respiratory failure). Symptoms and sex were modeled as dichotomous qualitative variables. Age was modeled as a continuous quantitative variable. In addition, it was evaluated if the age squared presented statistical significance. The age squared was incorporated because the frequency of the population (total and infected) varies as a function of age in a parabolic way.

To develop and test the models, the database (n=67318) was randomly divided into two data

sets (Training and Testing datasets) with the same percentage of RT-PCR positive cases in each one. The Training set (n=53854) was used for the design and internal validation of the models, while the Testing set (n=13464) was used for external validation and evaluation of the predictive capacity of the model. Models were compared using the Akaike Information Criterion, the area under the receiver-operator characteristic curve (AUC) and the predictive value.

Variables without statistical significance were removed by stepwise regression with bidirectional elimination based on Akaike information criterion (AIC). In addition, we performed a 100-fold cross-validation method in which the training base was partitioned in 100 randomly selected subsets and the model was fitted for each one. All statistical analyzes were performed with R software, packages base, caret, caTools, epiDisplays and ggplot2,^{24,26} and GraphPad Prism version 5.00 for Windows (La Jolla California USA, www.graphpad.com). R scripts can be seen in supplementary material (eFile 3, supplementary material).

Results

Data description

A total of 67318 individuals were included in this study; among them, 12% had a positive RT-PCR for SARS-CoV-2 with a median age of 37 years-old (IQR: 26-51 years), while the group with a negative RT-PCR had a median age of 40 years-old (IQR: 25-59 years).

Significant differences were identified in the positivity rate of RT-PCR according to the interval between the onset of symptoms and the time of sampling. It was observed that for the day, ranges of 0-3 days, 4-15 days and 16-30 days, the proportion of positives was 10.8%, 13.7% and 8.6%, respectively. A significant difference was observed between the three intervals of days ($p < 0.001$).

Infection rates varied among different age groups, with the highest rates being in the 15 to 60 year-old group (Figure 1). The pediatric group (n=10653) presented a rate of 7.9%, and a median age of 4 years (IQR 1-10 years); the young adult group (n=28909) presented a rate of 15.0%,

and a median age of 32 years (IQR 26-37 years); the adult group (n=14712) presented a rate of 12.8%, and a median age of 53 years (IQR 48-58 years); and the elderly group (n=13044) presented a rate of 7.0%, and a median age of 77 years (IQR 71-85 years).

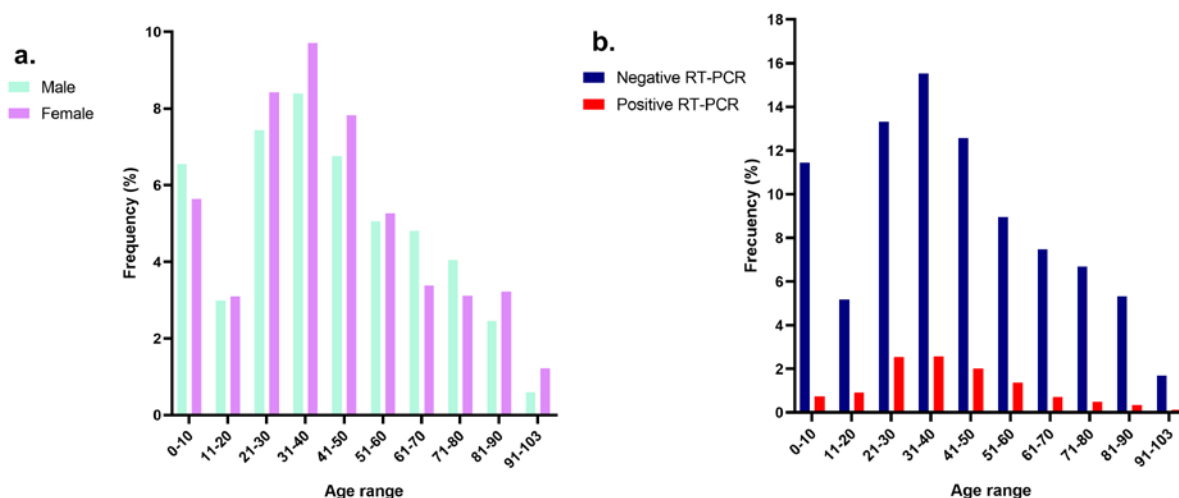
Regarding comorbidities, only the frequency of asthma and tuberculosis did not show statistical differences between age groups (Table 1). The most frequent comorbidities were hypertension, heart failure and diabetes in all groups over 18 years of age and asthma in the pediatric group.

The frequency of most symptoms was different between the different age groups (Table 1). Only the frequency of diarrhea, vomiting, abdominal pain, conjunctival injection and sex did not present differences between the different age groups in the infected population. On the other hand, the most frequent symptoms in all age groups were cough, high fever, odynophagia, headache, malaise and myalgia. It should be noted that these symptoms not only presented the highest frequencies in the infected population but also in the total population (Table 1). However, the symptoms with the highest ratio between the infected population and the total population were different between age groups. In the pediatric group, the symptoms with the highest ratio were anosmia, dysgeusia, low fever, headache and myalgia; in the group of adults and young adults the highest ratios were seen for low fever, headache and cough; in the elderly group the symptoms with the highest ratios were anosmia dysgeusia, odynophagia, low fever and headache.

Clinical features associated with COVID-19

The association between sex, symptoms and infection with SARS-CoV-2 was studied using multivariate logistic regression analysis considering pairwise interactions. The odds ratios (OR) for each of the symptoms evaluated in the four age groups are shown in Figure 2. The complete analysis is included in supplementary material (eFile 4, supplementary material). Sex was significant only for the young adult group, with males displaying higher odds than females of being infected (OR=1.36, $p=0.001$).

Figure 1. Study population and age distribution



a. Frequency of males and females in the total population. b. Frequency of SARS-CoV-2 RT-PCR positive and negative cases in relation to age (years).

Table 1. Frequency of comorbidities and symptoms among the total and infected populations

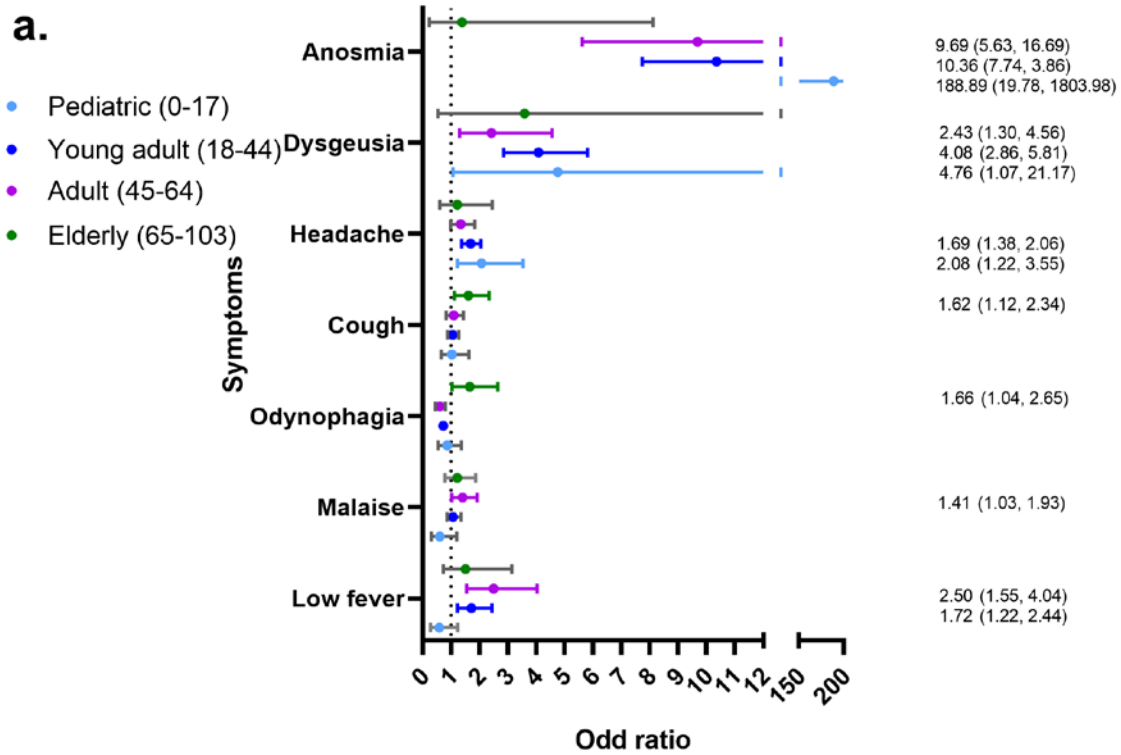
	Total population					Infected population					Ratio between infected population and total population			
	0-17 years (n=10653)	18-44 years (n=28909)	45-64 years (n=14712)	65-103 years (n=13044)	P value	0-17 years (n=842)	18-44 years (n=4336)	45-64 years (n=1880)	65-103 years (n=910)	P value	0-17 years	18-44 years	45-64 years	65-103 years
Sex (male)	5630 (52.8%)	13480 (46.7%)	7268 (49.4%)	6670 (51.3%)	<0.001	412 (48.9%)	2118 (48.4%)	966 (51.4%)	466 (51.2%)	0.220	0.07	0.16	0.13	0.07
Comorbidities														
Hypertension	23 (0.2%)	865 (2.9%)	3120 (21.2%)	6071 (46.5%)	<0.001	0 (0%)	95 (2.1%)	341 (18.1%)	420 (46.1%)	<0.001	0.00	0.11	0.11	0.07
Heart failure	64 (0.6%)	121 (0.4%)	586 (3.9%)	2457 (18.8%)	<0.001	4 (0.4%)	10 (0.2%)	34 (1.8%)	120 (13.1%)	<0.001	0.06	0.08	0.06	0.05
Diabetes	38 (0.3%)	620 (2.1%)	1673 (11.3%)	2441 (18.7%)	<0.001	0 (0%)	100 (2.3%)	216 (11.4%)	152 (16.7%)	<0.001	0.00	0.16	0.13	0.06

	Total population					Infected population					Ratio between infected population and total population			
Neurological disease	296 (2.7%)	400 (1.3%)	460 (3.1%)	1972 (15.1%)	<0.001	11 (1.3%)	27 (0.6%)	35 (1.8%)	125 (13.7%)	<0.001	0.04	0.07	0.08	0.06
Chronic obstructive pulmonary disease	46 (0.4%)	167 (0.5%)	864 (5.8%)	1847 (14.1%)	<0.001	1 (0.1%)	8 (0.1%)	32 (1.7%)	90 (9.8%)	<0.001	0.02	0.05	0.04	0.05
Cancer	179 (1.6%)	328 (1.1%)	791 (5.3%)	1272 (9.7%)	<0.001	5 (0.5%)	13 (0.2%)	49 (2.6%)	50 (5.4%)	<0.001	0.03	0.04	0.06	0.04
Obesity	71 (0.6%)	923 (3.1%)	1274 (8.6%)	1190 (9.1%)	<0.001	5 (0.5%)	137 (3.1%)	13 (7.3%)	72 (7.9%)	<0.001	0.07	0.15	0.11	0.06
Renal insufficiency	39 (0.3%)	254 (0.8%)	470 (3.1%)	966 (7.4%)	<0.001	1 (0.1%)	11 (0.2%)	29 (1.5%)	54 (5.9%)	<0.001	0.03	0.04	0.06	0.06
Asthma	529 (4.9%)	1329 (4.5%)	616 (4.1%)	411 (3.1%)	<0.001	29 (3.4%)	141 (3.2%)	52 (2.7%)	29 (3.1%)	0.729	0.05	0.11	0.08	0.07
Immune disease	164 (1.5%)	669 (2.3%)	606 (4.1%)	401 (3%)	<0.001	2 (0.2%)	38 (0.8%)	32 (1.7%)	16 (1.7%)	0.001	0.01	0.06	0.05	0.04
Chronic liver disease	9 (0.1%)	87 (0.3%)	197 (1.3%)	190 (1.4%)	<0.001	0 (0%)	6 (0.1%)	17 (0.9%)	7 (0.7%)	<0.001	0.00	0.07	0.09	0.04
Previous bronchitis	643 (6%)	76 (0.2%)	83 (0.5%)	104 (0.7%)	<0.001	17 (2%)	1 (0.1%)	8 (0.4%)	6 (0.6%)	<0.001	0.03	0.01	0.10	0.06
Tuberculosis	27 (0.2%)	227 (0.7%)	152 (1%)	92 (0.7%)	<0.001	4 (0.4%)	50 (1.1%)	23 (1.2%)	7 (0.7%)	0.227	0.15	0.22	0.15	0.08
Symptoms														
High fever	8863 (83.2%)	18260 (63.2%)	9067 (61.6%)	7571 (58.0%)	<0.001	508 (60.3%)	2371 (54.7%)	1092 (58.1%)	588 (64.6%)	<0.001	0.06	0.13	0.12	0.08
Cough	4967 (46.7%)	15272 (52.8%)	8288 (56.3%)	6971 (53.4%)	<0.001	411 (48.8%)	2431 (56.1%)	1171 (62.3%)	584 (64.2%)	<0.001	0.08	0.16	0.14	0.08
Odynophagia	4816 (42.5%)	15627 (54.1%)	6220 (42.3%)	2066 (15.8%)	<0.001	316 (37.5%)	2099 (48.4%)	780 (41.5%)	223 (24.5%)	<0.001	0.07	0.13	0.13	0.11
Headache	1643 (15.5%)	10893 (37.7%)	4213 (28.6%)	1143 (8.8%)	<0.001	227 (27.0%)	1903 (43.9%)	637 (33.9%)	104 (11.4%)	<0.001	0.14	0.17	0.15	0.09
Malaise	1926 (18.1%)	11312 (39.1%)	5790 (39.4%)	4701 (36.0%)	<0.001	118 (14%)	1304 (30.1%)	635 (33.8%)	283 (31.1%)	<0.001	0.06	0.12	0.11	0.06
Anosmia	105 (1.0%)	1569 (5.4%)	476 (3.2%)	94 (0.7%)	<0.001	69 (8.2%)	975 (22.5%)	250 (13.3%)	26 (2.9%)	<0.001	0.66	0.62	0.52	0.28
Myalgia	790 (7.4%)	7835 (27.1%)	3647 (24.8%)	1431 (11.7%)	<0.001	68 (8.1%)	991 (22.9%)	429 (22.8%)	114 (12.5%)	<0.001	0.09	0.13	0.12	0.08
Low fever	443 (4.2%)	1507 (5.2%)	645 (4.4%)	528 (4.0%)	<0.001	66 (7.8%)	396 (9.1%)	178 (9.5%)	48 (5.3%)	0.001	0.15	0.26	0.28	0.09

	Total population					Infected population					Ratio between infected population and total population			
Diarrhea	1221 (11.5%)	3163 (10.9%)	1537 (10.4)	943 (7.2%)	<0.001	57 (6.8%)	302 (7.0%)	159 (8.5%)	62 (6.8%)	0.163	0.05	0.10	0.10	0.09
Dysgeusia	105 (1.0%)	1290 (4.5 %)	408 (2.8%)	86 (0.6%)	<0.001	46 (5.5)	718 (16.6%)	187 (9.9%)	18 (2.0%)	<0.001	0.44	0.56	0.46	0.21
Abdominal pain	1015 (9.5%)	2288 (7.9%)	1185 (8.1%)	931 (7.13%)	<0.001	45 (5.3%)	178 (4.1%)	94 (5.0%)	46 (5.1%)	0.203	0.04	0.08	0.08	0.05
Tachypnea	1349 (12.7%)	3890 (13.5%)	3310 (22.5%)	5698 (43.7%)	<0.001	34 (4.0%)	330 (7.6%)	254 (13.5%)	281 (30.9%)	<0.001	0.03	0.08	0.08	0.05
Vomiting	1286 (12.1%)	1697 (5.9%)	803 (5.5%)	778 (6.0%)	<0.001	34 (4.0%)	127 (2.9%)	55 (2.9%)	31 (3.4%)	0.339	0.03	0.07	0.07	0.04
Dyspnea	943 (8.9%)	3143 (10.9%)	2624 (17.8%)	4624 (35.4%)	<0.001	33 (3.9%)	302 (7.0%)	217 (11.5%)	234 (25.7%)	<0.001	0.03	0.10	0.08	0.05
Food refusal	859 (8.1%)	1352 (4.7%)	818 (5.6%)	1341 (10.3%)	<0.001	31 (3.7%)	124 (2.9%)	74 (3.9%)	72 (7.9%)	<0.001	0.04	0.09	0.09	0.05
Arthralgia	432 (4.1%)	5172 (17.9%)	2578 (17.5%)	1214 (9.3%)	<0.001	23 (2.7%)	578 (13.3%)	280 (14.9%)	83 (9.1%)	<0.001	0.05	0.11	0.11	0.07
Conjunctival injection	315 (3.0%)	1161 (4.0%)	523 (3.6%)	289 (2.2%)	<0.001	11 (1.3%)	115 (2.7%)	48 (2.6%)	22 (2.4%)	0.144	0.03	0.10	0.09	0.08
Respiratory failure	275 (2.6%)	905 (3.1%)	1133 (7.7%)	2918 (22.4%)	<0.001	11 (1.3%)	100 (2.3%)	103 (5.5%)	135 (14.8%)	<0.001	0.04	0.11	0.09	0.05
Chest pain	252 (2.4%)	2472 (8.6%)	1462 (9.9%)	1084 (8.3%)	<0.001	10 (1.2%)	218 (5.0%)	131 (7.0%)	55 (6.0%)	<0.001	0.04	0.09	0.09	0.05
Irritability	170 (1.6%)	413 (1.4%)	303 (2.1%)	1029 (7.9%)	<0.001	8 (1.0%)	35 (0.8%)	19 (1.0%)	35 (3.8%)	<0.001	0.05	0.08	0.06	0.03
UAMB	674 (6.3%)	414 (1.4%)	555 (3.8%)	1626 (12.5%)	<0.001	7 (0.8%)	19 (0.4%)	26 (1.4%)	60 (6.6%)	<0.001	0.01	0.05	0.05	0.04
Seizures	189 (1.7%)	135 (0.5%)	99 (0.7%)	123 (0.9%)	<0.001	5 (0.6%)	6 (0.1%)	2 (0.1%)	3 (0.3%)	0.030	0.03	0.04	0.02	0.02
Mental confusion	63 (0.6%)	173 (0.6%)	209 (1.4%)	984 (7.5%)	<0.001	2 (0.2%)	13 (0.3%)	11 (0.6%)	32 (3.5%)	<0.001	0.03	0.08	0.05	0.03

UAMB – use of accessory muscles for breathing; **low fever** – low grade fever (37.5-37.9°C); **high fever** – high grade fever ($\geq 38^\circ\text{C}$).

Figure 2. Association between symptoms and COVID-19 in different age groups

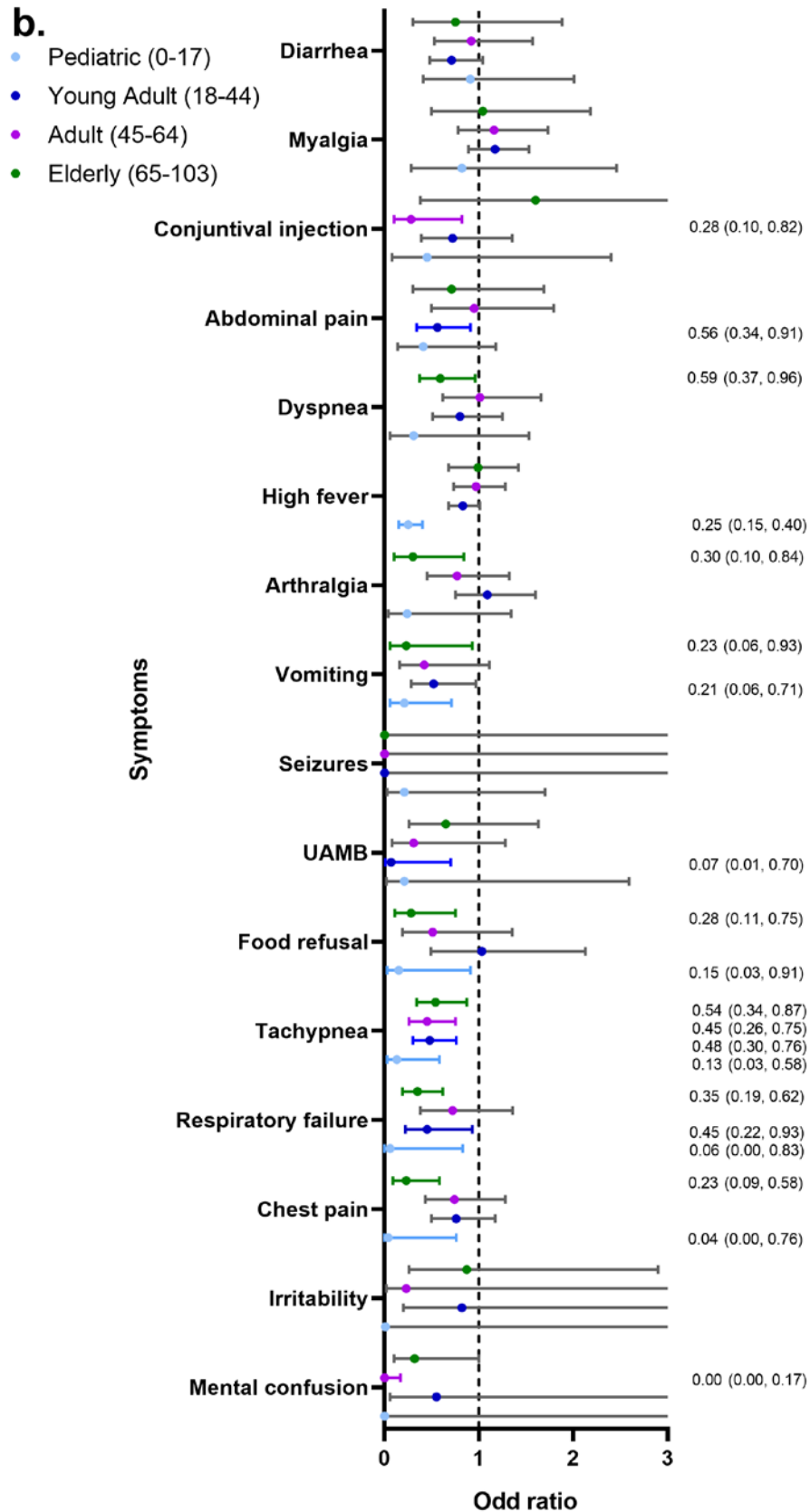


a. Symptoms with positive association with COVID-19 in the different age groups (years).

b. Symptoms with negative (or non-significant) association with COVID-19 in the different age groups (years).

Odds ratio values are reported with their respective 95% confidence interval in those symptoms that presented statistical significance ($p < 0.05$). Gray lines in confidence intervals indicate that there is no statistical significance, colored lines represent statistical significance.

UAMB - use of accessory muscles for breathing; **low fever** - low grade fever ($37.5-37.9^{\circ}\text{C}$); **high fever** - high grade fever ($\geq 38^{\circ}\text{C}$).



The main symptoms associated with COVID-19 among the different groups were anosmia, dysgeusia, headache, cough, low fever, malaise and odynophagia (Figure 2). However, not all of them presented a positive association in the different age groups.

In the pediatric group the symptoms with a positive association with COVID-19 were anosmia (OR=188.89, $p<0.001$), dysgeusia (OR=4.76, $p=0.040$), and headache (OR=2.08, $p=0.007$), while chest pain, high fever, food refusal, tachypnea, vomiting and respiratory failure had a negative association.

The rest of the symptoms did not show a statistically significant association with COVID-19, but some of them had a significant interaction with other symptoms (eFile 4, supplementary material). Significantly higher ORs were found with the combination of myalgia and respiratory failure (OR=2.29, $p=0.028$), seizures and vomiting (OR=7.78, $p=0.027$), and dyspnea and myalgia (OR=28.63, $p=0.010$).

In the young adult group, the symptoms with a positive association with COVID-19 were anosmia (OR=10.36, $p<0.001$), dysgeusia (OR=4.08, $p<0.001$), low fever (OR=1.72, $p=0.002$) and headache (OR=1.69, $p<0.001$), while abdominal pain, tachypnea, vomiting, respiratory failure, UAMB, and odynophagia had a negative association (Figure 2). In addition, significantly higher ORs were found with the combination of mental confusion and tachypnea (OR=61.36, $p=0.009$), and mental confusion and odynophagia (OR=50.0, $p=0.009$).

In the adult group, the symptoms with a positive association with COVID-19 were anosmia (OR=9.69, $p<0.001$), low fever (OR=2.50, $p<0.001$), dysgeusia (OR=2.43, $p=0.005$) and malaise (OR=1.41, $p=0.034$), while mental confusion, conjunctival injection, malaise, tachypnea and odynophagia had a negative association (Figure 2). In addition, significantly higher ORs were found with the combination of anosmia and vomiting (OR=133.43, $p=0.014$), and diarrhea and irritability (OR=12.91, $p=0.032$).

In the elderly group, the symptoms with a positive association with COVID-19 were odynophagia (OR=1.66, $p=0.035$), and cough

(OR=1.62, $p=0.010$), while arthralgia, mental confusion, dyspnea, chest pain, food refusal, tachypnea, vomiting, respiratory failure and odynophagia had a negative association (Figure 2). In addition, significantly higher ORs were found with the combination of irritability and UAMB (OR=6.75, $p=0.002$), and abdominal pain and chest pain (OR=1.52, $p=0.001$).

The symptoms associated with COVID-19 (anosmia, dysgeusia, headache, cough, low fever, malaise and odynophagia) presented differences in their diagnostic capacity in the different age groups (Table 2). In all groups the most specific symptoms were anosmia and dysgeusia. However, the most sensitive symptoms were cough and odynophagia for the pediatric, young adult and adult groups; and cough and malaise for the elderly group. In addition, some of the symptoms associated with COVID-19 presented differences according to sex in the different age groups in the infected population. In the pediatric group, the females presented 1.47 ($p=0.017$) higher odds than the males of presenting headaches. In the group of young adults, females had 1.35 ($p=0.002$), 1.41 ($p<0.001$) and 1.35 ($p<0.001$) higher odds of suffering from dysgeusia, odynophagia and headache than males. In the adult group, females had 1.45 ($p=0.021$) and 1.27 ($p=0.011$) higher odds of suffering from anosmia and odynophagia than males; and in the elderly group females had 2.63 ($p=0.040$) higher odds of suffering from anosmia than males.

Logistic regression predictive model

All tested models presented similar characteristics regarding area under the curve, AIC value and predictive value. The model with lower AIC value (35029) presented an AUC=0.72. The model considers 16 symptoms (1=present, 0=absent), sex (1=male, 0=female), and three quantitative variables: age, age squared and the number of symptoms (NS). Table 3 shows the regression coefficient of each variable, with its corresponding p-value. Cross validation did not show significant differences between the model calculated with the training set and 100 randomly chosen partitions.

Table 2. Diagnostic characteristics of the main symptoms associated with COVID-19

Variable	Pediatric (0-17 years)				Young adult (18-44 years)				Adult (45-64 years)				Elderly (65-103 years)			
	Sensitivity	Specificity	+LR	-LR	Sensitivity	Specificity	+LR	-LR	Sensitivity	Specificity	+LR	-LR	Sensitivity	Specificity	+LR	-LR
Anosmia	8.2 (6.3,10.1)	99.6 (99.5,99.8)	22.3 (15.0,33.2)	0.92 (0.90,0.94)	22.5 (21.2,23.7)	97.6 (97.4,97.8)	9.3 (8.4,10.3)	0.79 (0.78,0.81)	13.3 (11.8,14.8)	98.2 (98.0,98.4)	7.5 (6.3,9.0)	0.88 (0.87,0.90)	2.9 (1.8,3.9)	99.4 (99.3,99.6)	5.1 (3.3,8.0)	0.98 (0.97,0.99)
Dysgeusia	5.5 (3.9,7.1)	99.4 (99.2,99.6)	9.1 (6.2,13.3)	0.95 (0.94,0.97)	16.6 (15.4,17.7)	97.7 (97.5,97.9)	7.1 (6.4,7.9)	0.85 (0.84,0.87)	9.9 (8.6,11.2)	98.3 (98.1,98.5)	5.8 (4.8,7.0)	0.92 (0.90,0.93)	2.0 (1.1,2.9)	99.4 (99.3,99.6)	3.5 (2.1,5.9)	0.99 (0.98,0.99)
Low fever	36.3 (33.6,39.0)	96.2 (95.8,96.5)	9.5 (8.4,10.7)	0.66 (0.63,0.69)	9.1 (8.3,10.0)	95.5 (95.2,95.7)	2.0 (1.8,2.3)	0.95 (0.94,0.96)	9.5 (8.1,10.8)	96.4 (96.0,96.7)	2.6 (2.2,3.1)	0.94 (0.93,0.95)	5.3 (3.8,6.7)	96.0 (95.7,96.4)	1.3 (1.0,1.8)	0.99 (0.97,1.00)
Headache	27.0 (23.9,30.0)	85.6 (84.9,86.3)	1.9 (1.7,2.1)	0.85 (0.82,0.89)	43.9 (42.4,45.4)	63.4 (62.8,64.0)	1.2 (1.1,1.3)	0.88 (0.86,0.91)	33.9 (31.7,36.0)	72.1 (71.4,72.9)	1.2 (1.1,1.3)	0.92 (0.89,0.95)	11.4 (9.4,13.5)	91.4 (90.9,91.9)	1.3 (1.1,1.6)	0.97 (0.95,0.99)
Malaise	14.0 (11.6,16.4)	81.6 (80.8,82.3)	0.8 (0.6,0.9)	1.05 (1.02,1.08)	30.1 (28.7,31.4)	59.3 (58.7,59.9)	0.7 (0.7,0.8)	1.18 (1.15,1.21)	33.8 (31.6,35.9)	59.8 (59.0,0.61)	0.8 (0.7,0.9)	1.11 (1.07,1.15)	31.1 (28.1,34.1)	63.6 (62.7,64.4)	0.9 (0.8,0.9)	1.08 (1.04,1.13)
Odynophagia	37.5 (34.2,40.9)	54.1 (53.1,55.1)	0.8 (0.7,0.9)	1.15 (1.09,1.22)	48.4 (46.9,49.9)	44.9 (44.3,45.6)	0.88 (0.85,0.91)	1.15 (1.11,1.19)	41.5 (39.3,43.7)	57.6 (56.8,58.4)	1.0 (0.9,1.0)	1.02 (0.98,1.06)	24.5 (21.7,27.3)	84.8 (84.2,85.4)	1.6 (1.4,1.8)	0.89 (0.86,0.92)
Cough	48.8 (45.4,52.3)	53.6 (52.6,54.6)	1.1 (1.0,1.2)	0.96 (0.89,1.02)	56.1 (54.6,57.5)	47.7 (47.1,48.4)	1.1 (1.0,1.2)	0.92 (0.89,0.95)	62.3 (60.0,64.5)	44.5 (43.7,45.4)	1.1 (1.0,1.2)	0.85 (0.80,0.90)	64.2 (61.0,67.3)	47.4 (46.5,48.2)	1.2 (1.1,1.3)	0.76 (0.69,0.83)

+LR – positive likelihood ratio; -LR – negative likelihood ratio.

Table 3. Coefficients of the logistic regression model

Variable	Estimate	Std error	Odds ratio	P value
Anosmia	2.24000	0.07	173.78	<0.001
Headache	1.64000	0.15	43.65	<0.001
Dysgeusia	1.27000	0.08	18.62	<0.001
Cough	1.07000	0.12	11.75	<0.001
Low fever	0.99900	0.07	9.98	<0.001
Dysgeusia :Tachypnea	0.74300	0.22	5.53	<0.001
Anosmia : Conjunctival injection	0.73600	0.34	5.45	0.030
Dyspnea : Conjunctival injection	0.67700	0.22	4.75	0.002
Diarrhea: Chest pain	0.60300	0.18	4.01	<0.001
Myalgia	0.43800	0.06	2.74	<0.001
High fever : Tachypnea	0.41900	0.09	2.62	<0.001
Food refusal	0.26900	0.09	1.86	0.003
Arthralgia	0.26800	0.07	1.85	<0.001
High fever	0.24500	0.05	1.76	<0.001
Malaise	0.23300	0.05	1.71	<0.001
Dyspnea	0.18100	0.07	1.52	0.008
Diarrhea	0.14800	0.08	1.41	0.051
Conjunctival injection	0.08890	0.11	1.23	0.423
Age	0.03420	0.01	1.08	<0.001
Age : Sex (M)	0.01810	0.00	1.04	<0.001
Chest pain	0.01540	0.08	1.04	0.845
Age : NS	0.01490	0.00	1.03	<0.001
Age ² :Headache	0.00026	0.00	1.00	0.002
Age ² :Cough	0.00015	0.00	1.00	0.009
Age:Age ²	0.00001	0.00	1.00	<0.001
Age ² :NS	-0.00012	0.00	1.00	<0.001
Age ² :Sex (M)	-0.00020	0.00	1.00	<0.001
Age ²	-0.00113	0.00	1.00	<0.001
Age:Cough	-0.01410	0.01	0.97	0.009
Age:Headache	-0.03390	0.01	0.92	<0.001
Sex (M)	-0.18100	0.10	0.66	0.066
Tachypnea	-0.33600	0.09	0.46	<0.001
Seizures	-0.58700	0.29	0.26	0.041
NS	-0.79500	0.06	0.16	<0.001
(Intercept)	-2.14000	0.12	0.01	<0.001

Low fever – low grade fever (37.5-37.9°C); **high fever** – high grade fever ($\geq 38^\circ\text{C}$); **NS** – number of symptoms.

Based on the ROC curve of the model (Figure 3), a cut-off line of 0.08839 was chosen to obtain a sensitivity of 80% (95% CI: 78-82) and a specificity of 46% (95% CI: 45-47). In addition, the model presented positive and negative predictive values of 17% and 95% respectively, a positive likelihood ratio of 1.48 (95% CI: 1.42-1.54) and a negative likelihood ratio of 0.43 (95% CI: 0.38-0.49).

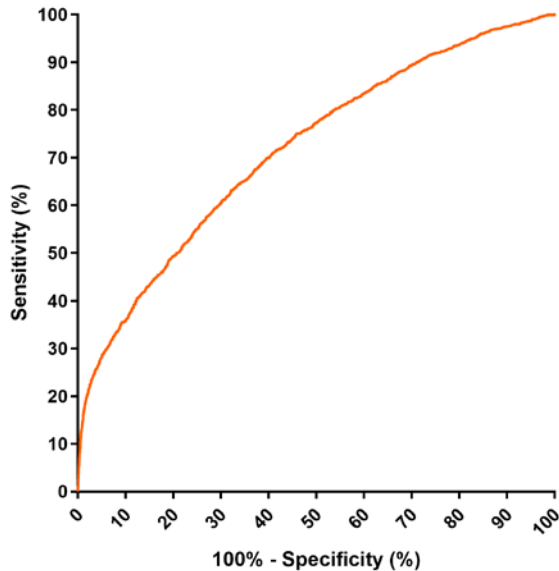


Figure 3. ROC curve of the logistic regression model based on 16 symptoms (anosmia, dysgeusia, low fever, dyspnea, diarrhea, myalgia, food refusal, arthralgia, high fever, malaise, conjunctival injection, chest pain, cough, headache, tachypnea, seizures), sex, age, age squared, and number of symptoms.

The variables with the highest coefficients were anosmia (estimate coefficient: 2.24), headache (estimate coefficient: 1.64), dysgeusia (estimate coefficient: 1.27), cough (estimate coefficient: 1.07), and low fever (estimate coefficient: 0.99). In addition, the interactions with the highest coefficients were dysgeusia with tachypnea, anosmia with conjunctival injection, dyspnea with conjunctival injection and diarrhea with chest pain. On the other hand, the variable “age squared” captures the way in which the ratio rate of positive cases changes with age (Figure 1). However, as observed by the value of its

coefficient (estimate coefficient: -0.00113), its contribution is very low. An alternative model without the variable age squared is described in Supplementary material (eFile 5, supplementary material). It can also be observed that the number of symptoms (estimate coefficient: -0.00113), had a negative contribution in the equation resulting in decreases in the likelihood of COVID-19 infection as the number of symptoms increases.

Discussion

The analysis of this large cohort allows an age-based characterization of the symptoms associated with COVID-19 at presentation, in the search for an improved clinical based decision strategy for the use of diagnostic tests and case management. Besides vaccine roll-out, in order to slow the spread of SARS-CoV-2, early detection, isolation, and the implementation of a robust system to trace contacts is required.⁴ Among the non-severe symptomatic cases, non-specific symptoms shared with other acute upper-respiratory infections hamper proper case identification and implementation of rapid isolation measures. In contrast to most acute upper-respiratory infections (including many respiratory coronaviruses), COVID-19 poses a special challenge in requiring specific identification for clinical and public health management. To facilitate the early detection of patients with COVID-19, we analyzed the symptoms of 67318 individuals studied with RT-PCR, divided into four groups based on age.

The most frequently reported symptoms among case series of COVID-19 include fever, cough, dyspnea and shortness of breath.^{13,27,28} Other symptoms include sore throat, odynophagia, anosmia, dysgeusia, diarrhea, nausea, vomiting, myalgia, headache and confusion.^{14,27,29} We found that in the present database the most frequent symptoms were cough, high fever, odynophagia, headache, malaise and myalgia, cough, high-grade fever, and odynophagia. The frequency of the symptoms showed significant differences between the different age groups. However, our analysis revealed that the symptoms positively associated with COVID-19 in the different groups were

some combination of these seven symptoms: anosmia, dysgeusia, headache, cough, low fever, malaise and odynophagia. In addition, the sensitivity and specificity of these seven symptoms showed variations between the different age groups. Anosmia and dysgeusia were the most specific symptoms for COVID-19. However, the odds ratio decreased with age, with the pediatric group having the highest predictive value (anosmia OR=188, dysgeusia OR=4.76, $p=0.040$), the adult group the lowest (anosmia OR=9.69, dysgeusia OR=2.43, $p<0.005$) and no statistical significance in the elderly group (anosmia OR=1.39 $p=0.718$, dysgeusia OR=3.59 $p=0.184$). The high value of the OR of anosmia in the pediatric group is due to the fact that this symptom presented a low frequency (8.2%) in this age range, but 66% of the cases with anosmia were COVID-19 cases. The decrease in the value of the OR of anosmia as age increases, is due to the fact that anosmia is more frequent in young COVID-19 patients than in elderly patients.¹⁴ It should be noted that studies carried out in Europe reveal frequencies of anosmia and dysgeusia greater than 60%,^{14,29} while in our study they presented frequencies lower than 23%. Despite the relatively low frequency, these symptoms have the highest predictive capacity. This is in accordance with other authors stating that anosmia and dysgeusia have a positive predictive value of 77% each of them and 83% if they are together.²⁹ Although anosmia can appear in other respiratory infections, it has been suggested that the lack of accompanying rhinitis or nasal swelling is more typically associated with COVID-19;³⁰ we were not able to test this hypothesis because the questionnaire does not contain rhinitis as a variable.

Cough was the most frequent and most sensitive symptom in each of the four age groups; however, a significant association with COVID-19 was only identified in the elderly group (OR=1.62). Similar results were obtained with a database from the United States, also using multivariate logistic regression. where individuals with cough and a median age of 45 years did not show significant association with COVID-19.³¹

The second most frequent symptom in all age groups was high fever. Fever has been reported in

numerous studies as a frequent symptom of COVID-19; however, we found that although fever is highly prevalent in many infections, low-grade fever appears to be specific to COVID-19. This could be observed in the group of young adults and adults who presented significant odds ratios for low-grade fever (37.5-37.9°C) (OR=1.72, OR=2.5). While high fever ($\geq 38^\circ\text{C}$) presented a negative association with individuals positive for SARS-CoV-2 in the pediatric group, and had no statistical significance in the other groups.

Another group of symptoms, including respiratory symptoms, showed a negative (or non-significant) association with a positive RT-PCR for SARS-CoV-2, therefore suggesting alternative diagnoses (Figure 2b). This is in line with the findings of a study from United Kingdom and United States,³¹ where anosmia presented the highest OR, while respiratory symptoms (chest pain, abdominal pain and dyspnea) presented the lowest values or lacked statistical significance.

Our analysis also revealed that some symptoms interact with each other, modifying their ORs. It should be noted that^{29,30} combinations of nonspecific symptoms resulted in high ORs such as dyspnea and myalgia (OR=28.63) in the pediatric group, mental confusion and tachypnea (OR=61.36) in the young adult group, diarrhea and irritability (OR=12.91) in the adult group and irritability and UAMB (OR=6.75) in the elderly group. This highlights that many symptoms are not independent, and that certain combinations may be more specific of COVID-19.

Middle-aged men appear to be more likely to be infected with COVID-19 than women. This could be observed in our study in the group of young adults (OR=1.36) and has been reported in studies with a median age of 50 years.³² In addition, the presentation of symptoms was different between males and females, where females showed a higher prevalence of anosmia, dysgeusia, headache, and odynophagia in other studies.¹⁴ This was also observed in our study, where females had higher odds of presenting these symptoms than males in the different age groups.

The diagnosis of COVID-19 is complex, since multiple factors such as age, sex and the interaction between different symptoms have to be taken into account. Therefore, we also developed a model based on symptoms at presentation, sex and age. Our model presents better diagnostic characteristics than individual symptoms, with an AUC=0.72, a sensitivity of 80% and a specificity of 46%. The main variables of this model were anosmia, headache, dysgeusia cough and low fever among other symptoms and symptom interactions (Table 3). Age, age squared, and sex had lower coefficients in the model. Although the age-dependent variability in the proportion of RT-PCR positive results in our study population might reflect differences in exposure due to school closure and an active protection of the elderly, the model achieves comparable predictive values in all age groups. Furthermore, our model achieved similar diagnostic characteristics with another symptom-based regression model.³¹ That model, which evaluated data from the United States and the United Kingdom presented an AUC=0.76;³¹ therefore, considering our results and those of studies carried out in those countries,³¹ the regression models based on symptoms achieve an AUC greater than 0.70. However, in view of the burden of morbidity and mortality in high-risk groups, models to predict the likelihood of infection should be interpreted with caution and not be used to limit the corresponding testing of groups at increased risk of severe disease.

This study was conducted with patients diagnosed between April and May 2020, prior to the influenza season in Argentina; this should also place our results in the adequate context of the alternative diagnoses that could modify the specificity of the symptoms based on the type and prevalence of these agents; in the current case, with most cases occurring in Buenos Aires and its metropolitan area, which was affected by its largest recorded dengue outbreak and a significant reduction of influenza-like illnesses (probably related to the lockdown measures) during that period.³³

This study has limitations; first the reference diagnostic technique used in this study was RT-PCR which has a sensitivity of 56-83%.⁹

Therefore, the sensitivity and specificity measures of the model are affected by that diagnostic accuracy. While suffering from the lack of a highly sensitive diagnostic standard, identifying clinical characteristics that raise the pretest probability of the infection could help interpretations of the RT-PCR result when combined with predictive models, epidemiological factors and complementary diagnostic methods like chest imaging.³⁴ Second, in our analysis, a significantly higher probability of a positive RT-PCR was identified if the test was performed when the interval between symptoms initiation and sampling was between 4 and 15 days. Third, another limitation derives from the circumstance of the data collection, which has not been designed as a prospective research tool but rather as a surveillance system with no monitoring system and therefore subject to errors in data capture. Finally, all symptoms were recorded by the physicians at initial presentation, without any updates for symptoms appearing later.

Conclusions

In summary, this symptoms-based analysis of a cohort of suspected COVID-19 cases identified a group of symptoms (anosmia, dysgeusia, headache, cough, low-grade fever, malaise and odynophagia) with significant association with a positive RT-PCR in different age groups. The presentation of symptoms showed significant differences between the different age groups. In those younger than 65 years, anosmia and dysgeusia were the symptoms most frequently associated with COVID-19, while in the elderly (65-103 years), cough and odynophagia were the symptoms most frequently associated. In addition, low-grade fever instead of high-grade fever was associated with COVID-19 in the young adult and adult age groups (18-64 years); there were also interactions between different symptoms, which means that there were certain combinations, which although appearing at low frequency, were highly associated to COVID-19. Finally, our findings show that a regression model based on multiple factors (age, sex, interaction between symptoms) can be used as an

accessory diagnostic method or for the rapid screening of suspected COVID-19 cases.

Authors' contributions statement: PEF contributed to: conceptualization, methodology, formal analysis, writing, original draft preparation, review and editing. JAP contributed to: conceptualization, methodology, formal analysis, review and editing. MIS contributed to: data curation, methodology, formal analysis, review and editing. CV contributed to: methodology, review and editing. ROC contributed to: methodology, review and editing. AJK contributed to: supervision, conceptualization review and editing. JPA contributed to: supervision, conceptualization review and editing. All authors read and approved the final version of the manuscript.

Conflicts of interest: All authors – none to declare.

Funding: None to declare.

References

- Zhu N, Zhang D, Wang W, et al. A novel coronavirus from patients with pneumonia in China, 2019. *N Engl J Med.* 2020;382:727-33. <https://doi.org/10.1056/NEJMoa2001017>
- Lu R, Zhao X, Li J, et al. Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *Lancet.* 2020;395:565-74. [https://doi.org/10.1016/S0140-6736\(20\)30251-8](https://doi.org/10.1016/S0140-6736(20)30251-8)
- Zhou P, Yang XL, Wang XG, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature.* 2020;579:270-3. <https://doi.org/10.1038/s41586-020-2012-7>
- World Health Organization. 2020. Novel coronavirus (2019-nCoV) situation report - 12. 1 February 2020. Accessed on: 04 April 2020. Available at: https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200201-sitrep-12-ncov.pdf?sfvrsn=273c5d35_2.
- Johns Hopkins Coronavirus Resource Center. COVID-19 dashboard by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU). Accessed on: 05 May 2020. Available at: <https://coronavirus.jhu.edu/map.html>.
- Ali I, Ali S. Why may COVID-19 overwhelm low-income countries like Pakistan? *Disaster Med Public Health Prep.* 2020;1-5. <https://doi.org/10.1017/dmp.2020.329>
- Ai T, Yang Z, Hou H, et al. Correlation of chest CT and RT-PCR testing for coronavirus disease 2019 (COVID-19) in China: a report of 1014 cases. *Radiology.* 2020;296:E32-40. <https://doi.org/10.1148/radiol.2020200642>
- Xie J, Ding C, Li J, et al. Characteristics of patients with coronavirus disease (COVID-19) confirmed using an IgM-IgG antibody test. *J Med Virol.* 2020;92:2004-10. <https://doi.org/10.1002/jmv.25930>
- Kokkinakis I, Selby K, Favrat B, Genton B, Cornuz J. [Covid-19 diagnosis: clinical recommendations and performance of nasopharyngeal swab-PCR]. *Rev Med Suisse.* 2020;16:699-701.
- Wölfel R, Corman VM, Guggemos W, et al. Virological assessment of hospitalized patients with COVID-2019. *Nature.* 2020;581:465-9. <https://doi.org/10.1038/s41586-020-2196-x>
- Espejo AP, Akgun Y, Al Mana AF, et al. Review of current advances in serologic testing for COVID-19. *Am J Clin Pathol.* 2020;154:293-304. <https://doi.org/10.1093/ajcp/aqaa112>
- Deeks JJ, Dinnes J, Takwoingi Y, et al. Antibody tests for identification of current and past infection with SARS-CoV-2. *Cochrane Database Syst Rev.* 2020;6:CD013652. <https://doi.org/10.1002/14651858.CD013652>
- Popov GT, Baymakova M, Vaseva V, Kundurzhiev T, Mutafchiyski V. Clinical characteristics of hospitalized patients with COVID-19 in Sofia, Bulgaria. *Vector Borne Zoonotic Dis.* 2020;20:910-5. <https://doi.org/10.1089/vbz.2020.2679>
- Lechien JR, Chiesa-Estomba CM, Place S, et al. Clinical and epidemiological characteristics of 1420 European patients with mild-to-moderate coronavirus disease 2019. *J Intern Med.* 2020;288:335-44. <https://doi.org/10.1111/joim.13089>
- Guan WJ, Ni ZY, Hu Y, et al. Clinical characteristics of coronavirus disease 2019 in China. *N Engl J Med.* 2020;382:1708-20. <https://doi.org/10.1056/NEJMoa2002032>
- Yang R, Gui X, Xiong Y. Comparison of clinical characteristics of patients with asymptomatic vs symptomatic coronavirus disease 2019 in Wuhan, China. *JAMA Netw Open.* 2020;3:e2010182. <https://doi.org/10.1001/jamanetworkopen.2020.10182>
- Mehta NS, Mytton OT, Mullins EWS, et al. SARS-CoV-2 (COVID-19): what do we know about children? A systematic review. *Clin Infect Dis* 2020;71:2469-79. <https://doi.org/10.1093/cid/ciaa556>
- Liu K, Chen Y, Lin R, Han K. Clinical features of COVID-19 in elderly patients: a comparison with young and middle-aged patients. *J Infect.* 2020; 80:e14-8. <https://doi.org/10.1016/j.jinf.2020.03.005>
- Li R, Tian J, Yang F, et al. Clinical characteristics of 225 patients with COVID-19 in a tertiary hospital near Wuhan, China. *J Clin Virol.* 2020;127:104363. <https://doi.org/10.1016/j.jcv.2020.104363>
- Datos Argentina - Casos COVID-19. 2021. Accessed on: 19 May 2021. Available at: https://datos.gob.ar/dataset/salud-covid-19-casos-registrados-republica-argentina/archivo/salud_fd657d02-a33a-498b-a91b-2ef1a68b8d16.
- InfoLEG - Ministerio de Justicia y Derechos Humanos - Argentina. 2021. Derecho de acceso a la información pública. Ley 27275. Objeto. Excepciones. Alcances. Accessed on: 19 May 2021. Available at: <http://servicios.infoleg.gob.ar/infolegInternet/anexos/265000-269999/265949/norma.htm>.

22. Argentina.gob.ar. Protección de los datos personales. Ley 25.326. Accessed on: 19 May 2021. Available at: <https://www.argentina.gob.ar/normativa/nacional/ley-25326-64790/actualizacion>.
23. Chang TH, Wu JL, Chang LY. Clinical characteristics and diagnostic challenges of pediatric COVID-19: a systematic review and meta-analysis. *J Formos Med Assoc.* 2020;119:982-9. <https://doi.org/10.1016/j.jfma.2020.04.007>
24. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2013.
25. Tuszynski J. Package 'caTools'. 2014.
26. Chongsuvivatwong V. epiDisplay: Epidemiological Data Display Package. 2018.
27. Wu YC, Chen CS, Chan YJ. The outbreak of COVID-19: An overview. *J Chinese Med Assoc.* 2020;83:217-20. <https://doi.org/10.1097/JCMA.0000000000000270>
28. Yang J, Zheng Y, Gou X, et al. Prevalence of comorbidities and its effects in patients infected with SARS-CoV-2: a systematic review and meta-analysis. *Int J Infect Dis.* 2020;94:91-5. <https://doi.org/10.1016/j.ijid.2020.03.017>
29. Zayet S, Klopfenstein T, Mercier J, et al. Contribution of anosmia and dysgeusia for diagnostic of COVID-19 in outpatients. *Infection.* 2021;49:361-5. <https://doi.org/10.1007/s15010-020-01442-3>
30. Zubair AS, McAlpine LS, Gardin T, Farhadian S, Kuruvilla DE, Spudich S. Neuropathogenesis and neurologic manifestations of the coronaviruses in the age of coronavirus disease 2019. *JAMA Neurol.* 2020;77:1018-27. <https://doi.org/10.1001/jamaneurol.2020.2065>
31. Menni C, Valdes AM, Freidin MB, et al. Real-time tracking of self-reported symptoms to predict potential COVID-19. *Nat Med.* 2020;26:1037-40. <https://doi.org/10.1038/s41591-020-0916-2>
32. Liu R, Han H, Liu F, et al. Positive rate of RT-PCR detection of SARS-CoV-2 infection in 4880 cases from one hospital in Wuhan, China, from Jan to Feb 2020. *Clin Chim Acta.* 2020;505:172-5. <https://doi.org/10.1016/j.cca.2020.03.009>
33. Ministerio de Salud Argentina. 2020. Boletín integral de vigilancia. Edición semanal. 499. Accessed on: 07 July 2020. Available at: https://www.argentina.gob.ar/sites/default/files/biv_49_9_se23.pdf.
34. Woloshin S, Patel N, Kesselheim AS. False negative tests for SARS-CoV-2 infection - challenges and implications. *N Engl J Med.* 2020;383:e38. <https://doi.org/10.1056/NEJMp2015897>

Please cite this article as:

Fleitas PE, Paz JA, Simoy MI, Vargas C, Cimino RO, Krolewiecki AJ, Aparicio JP. Clinical diagnosis of COVID-19. A multivariate logistic regression analysis of symptoms of COVID-19 at presentation. *GERMS.* 2021;11(2):221-237. doi: 10.18683/germs.2021.1259