scientific reports

Check for updates

OPEN

Genetic variants regulating the immune response improve the prediction of COVID-19 severity provided by clinical variables

Pablo Delgado-Wicke[1,19], Sara Fernández de Córdoba-Oñate[2,16,17,19], Emilia Roy-Vallejo[3], Estíbaliz Alegría-Carrasco[1], Diego A. Rodríguez-Serrano[4], Amalia Lamana[5], Nuria Montes[2,6,18], Ana Nicolao-Gómez[1], Rosa Carracedo-Rodríguez[1], Ana Marcos-Jiménez[7], Paula Díaz-Fernández[7], José M. Galván-Román[3], Laura Rabes-Rodríguez[3], Marta Sanz-Alba[3], Jesús Álvarez-Rodríguez[3], Almudena Villa-Martí[3], Carlos Rodríguez-Franco[3], Gonzalo Villapalos-García[8,9], Pablo Zubiaur[8,9], Francisco Abad-Santos[8,9], Ignacio de los Santos[3,11], Rosa P. Gomariz[5], Rosario García-Vicuña[2,10], Cecilia Muñoz-Calleja[7,10,11], Isidoro González-Álvaro[2,19]✉, Elena Fernández-Ruiz[1,10,19]✉ & PREDINMUN-COVID Group*

The characteristics of the host are crucial in the final outcome of COVID-19. Herein, the influence of genetic and clinical variants in COVID-19 severity was investigated in a total of 1350 patients. Twenty-one single nucleotide polymorphisms of genes involved in SARS-CoV-2 sensing as Toll-like-Receptor 7, antiviral immunity as the type I interferon signalling pathway (*TYK2*, *STAT1*, *STAT4*, *OAS1*, *SOCS*) and the vasoactive intestinal peptide and its receptors (*VIP/VIPR1,2*) were studied. To analyse the association between polymorphisms and severity, a model adjusted by age, sex and different comorbidities was generated by ordinal logistic regression. The genotypes rs8108236-AA (OR 0.12 [95% CI 0.02–0.53]; $p = 0.007$) and rs280519-AG (OR 0.74 [95% CI 0.56–0.99]; $p = 0.03$) in *TYK2*, and rs688136-CC (OR 0.7 [95% CI 0.5–0.99]; $p = 0.046$) in *VIP*, were associated with lower severity; in contrast, rs3853839-GG in *TLR7* (OR 1.44 [95% CI 1.07–1.94]; $p = 0.016$), rs280500-AG (OR 1.33 [95% CI 0.97–1.82]; $p = 0.078$) in *TYK2* and rs1131454-AA in *OAS1* (OR 1.29 [95% CI 0.95–1.75]; $p = 0.110$) were associated with higher severity. Therefore, these variants could influence the risk of severe COVID-19.

The clinical heterogeneity of COVID-19 ranges from asymptomatic infection to an exacerbated inflammatory response[1,2]. There are several host factors that can affect the severity of COVID-19 as age, sex and several comorbidities including hypertension, diabetes or heart disease[3,4]. However, early in the evolution of the

[1]Molecular Biology Unit, Hospital Universitario La Princesa and Health Research Institute (IIS-Princesa), Madrid, Spain. [2]Rheumathology Department, Hospital Universitario La Princesa (IIS-Princesa), Madrid, Spain. [3]Internal Medicine Department, Hospital Universitario La Princesa (IIS-Princesa), Madrid, Spain. [4]Intensive Care Unit, Hospital Universitario La Princesa (IIS Princesa), Madrid, Spain. [5]Cell Biology Department, Facultad de Biología, Universidad Complutense de Madrid, Madrid, Spain. [6]Pharmaceutical and Health Sciences Department, Faculty of Pharmacy, Universidad San Pablo-CEU, Boadilla del Monte, Spain. [7]Immunology Department, Hospital Universitario La Princesa (IIS-Princesa), Madrid, Spain. [8]Clinical Pharmacology Department, Hospital Universitario La Princesa (IIS-Princesa), Madrid, Spain. [9]Pharmacology Department, Faculty of Medicine, Universidad Autónoma de Madrid, Madrid, Spain. [10]Medicine Department, Faculty of Medicine, Universidad Autónoma de Madrid, Madrid, Spain. [11]Centro de Investigación Biomédica en Red en Enfermedades Infecciosas (CIBERINFEC), Instituto de Salud Carlos III, Madrid, Spain. [16]Present address: Hematology Department, Hospital General Universitario Gregorio Marañón, Madrid, Spain. [17]Present address: Instituto de Investigación Sanitaria Gregorio Marañón (IiSGM), Madrid, Spain. [18]Present address: Methodology Department, Hospital Universitario La Princesa (IIS-Princesa), Madrid, Spain. [19]These authors contributed equally: Pablo Delgado-Wicke, Sara Fernández de Córdoba-Oñate, Isidoro González-Álvaro and Elena Fernández-Ruiz. *A list of authors and their affiliations appears at the end of the paper. ✉email: isidoro.ga@ser.es; efruiz@salud.madrid.org

pandemic it was evidenced that those cases with severe forms of COVID-19 were related with an exacerbated immune response to the virus, leading to an hyperinflammatory status and, in the worst cases, a devastating effect on the lungs denominated Acute Respiratory Distress Syndrome (ARDS)[5], thrombosis and multiple organ failure[6,7].

The innate immune system plays an essential role in the antiviral response and rapidly senses viral infection through Toll-like receptors (TLRs) triggering type I interferon (IFN) production[8]. Among TLRs, the endosomal TLR7 recognizes single strand RNA viruses as SARS-CoV-2[9–11]. The effectiveness of the early immune response depends on the capability of cells to induce a competent expression of a variety of interferon-stimulated genes (ISGs) involved in the orchestration of the antiviral response and immune-regulation[12,13]. This process needs a fine coordination led by receptors and cellular mediators, as the Janus kinase and Signal transducers and activators of transcription (JAK-STAT) signalling pathway, whose participation is essential and whose deregulation has already been related to susceptibility to infections and the development of autoimmune diseases[14–16]. In addition, the vasoactive intestinal peptide (VIP) and its G protein-coupled receptors VPAC1 and VPAC2 modulate the immune response in several immune-mediated inflammatory disorders (IMID)[17–19] and viral infections[20], and its relationship with severe COVID-19 is an active area of research. A patented formulation of VIP RLF-100 (Aviptadil) is now under clinical trials to evaluate its use for COVID-19-related ARDS. Preliminary results showed that this treatment increases life expectancy by optimizing oxygenation and controlling the failure caused by the COVID-19 induced cytokine storm[21].

Genome-wide association studies (GWAS) have tried to identify genetic factors related to COVID-19 severity, and single nucleotide polymorphisms (SNPs) of several loci integrated in the pathway of virus entry into the cells have been reported[22]. Among them, previous work of our group has validated the transmembrane serine protease 2 (TMPRSS2 rs75603675) as a predictor of the severity of hospitalized patients[23]. Moreover, TMPRSS2 rs713400 was associated with SARS-CoV-2 viremia and a poor prognosis[24]. Furthermore, different gene pathways of activation and regulation of the immune response have been assessed, and missense variants either in the viral RNA sensor TLR7 or in the IFN α/β-receptor (IFNAR1/2) have been associated to severe cases of COVID-19[25–28]. Autosomal-recessive deficiencies in the intracellular mediators of the IFNAR1 signalling pathway, as tyrosine kinase 2 (TYK2) or STAT family members, favour intracellular bacterial and viral infections through impaired cytokine response[29–31]. Consequently, missense genetic variants of these genes participate in COVID-19 severity[28,32,33]. However, genetic variants leading to high expression of TYK2 are associated to critical COVID-19 disease[34] and the inhibition of this tyrosine kinase has been proposed as a therapeutic target in influenza A virus infections aggravated by secondary bacterial pneumonia[35].

Because of the disparity of observed findings, additional confirmatory and exploratory studies are warranted. Therefore, in this observational study we validated previously reported alleles and evaluated new SNP candidates in *TLR7*, *TYK2* and downstream interferon-stimulated genes (ISGs) as *STAT1, STAT4, OAS1* (2'-5' oligoadenylate synthetase 1) and *SOCS1* (suppressor of cytokine signalling 1), as well as *VIP, VIPR1* and *VIPR2* genes, determining their potential worsening or protective role in the outcome of COVID-19. For this purpose, first, a clinical model was developed with the characteristics and comorbidities associated with COVID-19 severity in our cohort (age, sex, cancer, dementia, diabetes, hypertension, obesity and people living with HIV). Then, each SNP was added to the clinical model and evaluated and, finally, those variants with a significant association were selected for a combined clinical and genetic model. We found that genetic variants in *TYK2* (rs8108236-AA and rs280519-AG) and *VIP* (rs688136-CC) had a protective role, while other variants in *TYK2* (rs280500-AG), *TLR7* (rs3853839-GG) and *OAS1* (rs1131454-AA) increased the likelihood of developing greater COVID-19 severity.

Results

Sociodemographic and clinical variables associated with COVID-19 severity in our population

A total of 1,440 patients were genotyped. Twenty-six of them were excluded from the statistical analysis due to genotyping failure (more than 5% missing genotypes), while 64 were excluded because of missing relevant clinical data. Therefore, a total of 1,350 patients were included in the study, with a mean age of 63.3 years (SD 16.7), 55.6% were male and 82.1% white non-Hispanic. The main comorbidities detected in the population were hypertension (41%) and dyslipidaemia (37%). Regarding severity (see Materials and Methods for definition), 31.4% of patients were classified as having mild, 45.5% moderate and 23.1% severe disease. There were significant differences between severity groups in terms of age ($p < 0.001$), male sex ($p < 0.001$), prevalence of hypertension ($p < 0.001$), dyslipidaemia ($p < 0.001$), diabetes mellitus ($p < 0.001$), obesity ($p = 0.002$), dementia ($p < 0.001$), cancer ($p = 0.001$) and people living with HIV (LHIV) ($p < 0.001$), as shown in Table 1. Treatment administered and analytical data collected during hospitalization are listed in Table S1.

When the multivariate analysis was performed to establish the clinical variables predicting severity, only some of them remained statistically significant. Among these variables, the interaction between sex and age was included to improve model adjustment. This interaction indicated that the risk increased in older age groups and was higher in males than in females. The risk of severe clinical outcome was also increased to different extents by other known comorbidities such as hypertension, obesity, dementia, diabetes (being higher for patients with organ damage) and cancer (being higher for metastatic cancer). In contrast, living with HIV exhibited a protective role notably decreasing COVID-19 severity. This multivariate model was used to adjust the association between genetic variants and COVID-19 severity (Table 2).

Genotyping and association analysis with clinical variables

Genotype analysis was performed for 21 SNPs within genes involved in the intracellular sensing of SARS-CoV-2 as *TLR7*, the IFN-I signalling pathway (*TYK2, STAT1, STAT4, OAS1, SOCS1*) and the VIP/VPAC axis. The studied SNPs were not in linkage disequilibrium (LD) and their minor allele frequency (MAF) was > 0.01

Variable	n	Mild, n = 424[a]	Moderate, n = 614[a]	Severe, n = 312[a]	P[b]
Sex x Age (years)	1350				< 0.001
Female < 45		69 (16%)	19 (3.1%)	4 (1.3%)	
Female 45–70		118 (28%)	130 (21%)	32 (10%)	
Female > 70		32 (7.5%)	129 (21%)	67 (21%)	
Male < 45		51 (12%)	42 (6.8%)	9 (2.9%)	
Male 45–70		123 (29%)	177 (29%)	94 (30%)	
Male > 70		31 (7.3%)	117 (19%)	106 (34%)	
Race/ethnicity	1349				0.3
White, non-Hispanic		345 (81%)	497 (81%)	266 (85%)	
White, Hispanic		74 (17%)	109 (18%)	40 (13%)	
Afro-descendent		1 (0.2%)	0 (0%)	1 (0.3%)	
Asian		4 (0.9%)	7 (1.1%)	5 (1.6%)	
Missing		0	1	0	
Hypertension	1350				< 0.001
Yes		94 (22%)	269 (44%)	178 (57%)	
Diabetes mellitus	1350				< 0.001
No		394 (93%)	500 (81%)	237 (76%)	
Without organ damage		26 (6.1%)	102 (17%)	55 (18%)	
With organ damage		4 (0.9%)	12 (2.0%)	20 (6.4%)	
Dyslipidaemia	1350				< 0.001
Yes		102 (24%)	250 (41%)	147 (47%)	
Obesity	1350				0.002
Yes		39 (9.2%)	89 (14%)	56 (18%)	
Dementia	1350				< 0.001
Yes		5 (1.2%)	25 (4.1%)	25 (8.0%)	
Cancer	1350				0.001
No		419 (99%)	594 (97%)	291 (93%)	
Without metastasis		3 (0.7%)	15 (2.4%)	14 (4.5%)	
With metastasis		2 (0.5%)	5 (0.8%)	7 (2.2%)	
People living with HIV	1350				< 0.001
Yes		25 (5.9%)	4 (0.7%)	0 (0%)	

Table 1. Demographic and clinical data according to COVID-19 severity scale. [a]n (%). [b]Pearson's Chi-squared test; Fisher's exact test.

(Table S2). In the case of *TYK2*, due to its upstream position in the IFN-I pathway, seven SNPs covering the gene were analysed (Fig S1).

To determine the association between each SNP and COVID-19 severity, the 21 SNPs were included separately in the clinical model by multivariate analysis and *TYK2* rs8108236, rs280519 and rs280500, *OAS1* rs1131454, *TLR7* rs3853839 and *VIP* rs688136 remained significant. The predicted probability of severity per genotype of significant SNPs in this model is shown in Fig. 1. This analysis showed that the rs280519-GA, rs8108236-AA and rs688136-CC genotypes were protective compared to each of their major alleles (Odds ratio, OR = 0.786/0.171/0.682, respectively) (Fig. 1, upper panels), while carrying rs1131454-AA, rs3853839-GG or rs280500-AG was considered a risk factor (OR = 1.323/1.478/1.104, respectively) (Fig. 1, lower panels).

Then, the combined effect or influence of these SNPs and clinical variables on COVID-19 severity was analysed (Fig. 2). Although clinical variables strongly influenced the disease outcome, showing higher OR compared to genetic variants, *TYK2* rs8108236-AA, rs280519-AG and *VIP* rs688136-CC genotypes were associated with a lower COVID-19 severity (OR < 1) and *TLR7* rs3853839-GG, *TYK2* rs280500-AG and *OAS1* rs1131454-AA genotypes favoured a worse disease outcome (OR > 1). Among them, the *TLR7* rs3853839-GG genotype had the highest OR value (OR = 1.44). This model resulting from including these variants was statistically different from the model that only included clinical variables ($p < 0.001$). The model combining both genetic variants and clinical variables provided a better explanation of data variability (R^2 Nagelkerke from 0.230 to 0.250) and improved the prediction of disease severity (R^2 McFadden from 0.106 to 0.117).

Discussion

Due to a global effort, the main knowledge on the effect of specific demographic and clinical characteristics on COVID-19 severity has been established[36]. Herein the modified WHO COVID ordinal outcomes scale, which stratifies mild, moderate and severe patients, was employed to analyse both protective and risk variables. Other

Variable	OR	95% CI	*p*-value
Interaction of sex and age (comparator: females under 45 y.o.)			
Female 45–70 y.o	3.76	2.25–6.5	< 0.001
Female > 70 y.o	9.98	5.74–17.85	< 0.001
Male < 45 y.o	3.6	1.97–6.73	< 0.001
Male 45–70 y.o	7.25	4.37–12.41	< 0.001
Male > 70 y.o	14.87	8.55–26.61	< 0.001
People living with HIV	0.08	0.02–0.21	< 0.001
Diabetes (comparator: no)			
Without organ damage	1.2	0.88–1.64	0.253
With organ damage	2.31	1.17–4.69	0.018
Dementia	1.85	1.07–3.2	0.027
Obesity	1.6	1.17–2.19	0.004
Hypertension	1.45	1.13–1.86	0.003
Cancer (comparator: no)			
Without metastasis	1.72	0.87–3.45	0.121
With metastasis	3.61	1.23–11.02	0.02

Table 2. Clinical model with demographic and clinical variables of the patients. n: 1350; AIC: 2587.679; R^2 McFadden: 0.106; R^2 Nagelkerke: 0.230. *CI* confidence interval, *OR* Odds ratio *y.o.* years old.

studies focused mainly on the severe states of the disease have selected additional sub-classifications[37]. Our study corroborates that age and sex are the two characteristics with the highest influence on COVID-19 severity as has been described[3,4]. Regarding comorbidities, the presence of oncological disease with metastasis, diabetes with organ damage, dementia, obesity and hypertension, in decreasing order of importance, significantly influence the risk of severe COVID-19 complications in our model, as previously reported[2,4].

The finding that LHIV is an independent protective factor in our cohort is remarkable. In contrast, the largest study on the effect of COVID-19 on people LHIV showed that HIV infection was an independent risk factor for poor prognosis[38]. However, this study recognized that having an undetectable viral load and being on antiretroviral treatment (ART) greatly reduced the risk of worse disease onset, which has been validated in other Spanish LHIV cohorts[39]. Additionally, this study also indicated that other patient features, such as age and health quality, are factors that notably affect COVID-19 severity. Patients LHIV included herein presented a lower average age (49.9 ± 11.6 years) compared to those in the general population (63.8 ± 16.6 years) and presented a very high rate of immune virologic control, since almost all of them received ART treatment. Thus, we interpret the detected protective effect as a consequence of a greater supervision, concomitant ART treatment, and good immunological state of the patients prior to COVID-19 infection.

Heterogeneity in clinical response to SARS-CoV-2 that could not be explained by these clinical variables led us to study the influence of certain SNPs on the severity of infection and outcome in Spanish patients, focusing on the impact of genetic variants of genes involved in the IFN-I signalling pathway and the VIP/VPAC axis. We found that the TLR7 rs3853839-GG genotype was the greatest risk factor for disease severity, in agreement with recently described data[40,41]. This variant is located in the 3' untranslated region (UTR) and increases TLR7 transcripts in peripheral blood mononuclear cells (PBMC) from carriers with systemic lupus erythematosus (SLE)[42], HCV-infected patients[43] and COVID-19 patients[44]. In addition, rare TLR7 missense variants have been described in 4 young male patients with severe COVID-19 whose PBMCs displayed impaired type I and II IFN responses[26] and Asano and colleagues reported TLR7 deficiency in 1% of men under 60 years with life-threatening COVID-19[45]. Moreover, Fallerini et al. described that TLR7 loss-of-function variants were present in 2% of young male patients thus establishing *TLR7* as the most important susceptibility gene for life-threatening disease[46].

Recently, we have described a close interaction between TLR7 expression and the IFN-I signalling pathway mediated through TYK2[47]. Among the *TYK2* SNPs evaluated herein, rs8108236-AA and rs280519-AG appear to be protective while rs280500-AG seems to increase the risk of COVID-19 severity. *TYK2* rs8108236-AA genotype is rare and was only found in nine patients (Table S3). This is a splicing variant located in the intronic position chr19:10355156 that could exert an influence on expression changes. Additionally, *TYK2* rs280519 is located 7 bp away from the start of intron 11 in a highly conserved interspecies region[48], suggesting an important role in the function of the protein. In this case, the AG and AA genotypes are both present at equal frequency in the European population. Other variants assessed in our model, reported previously to have a protective effect in autoimmune diseases, such as *TYK2* rs12720270[48,49], rs12720356 and rs34536443[50,51] were not found to be significant to predict COVID-19 severity, although the latter is a missense variant described as a risk allele for mycobacteria and viral infection[30] and was associated to hospitalization due to SARS-CoV-2 infection in another study[32]. Similarly, severity is not influenced by the rs2304256 variant in our setting, in contrast with the poor outcome found in COVID-19 Brazilian patients, especially among female and non-white populations[52]. In our

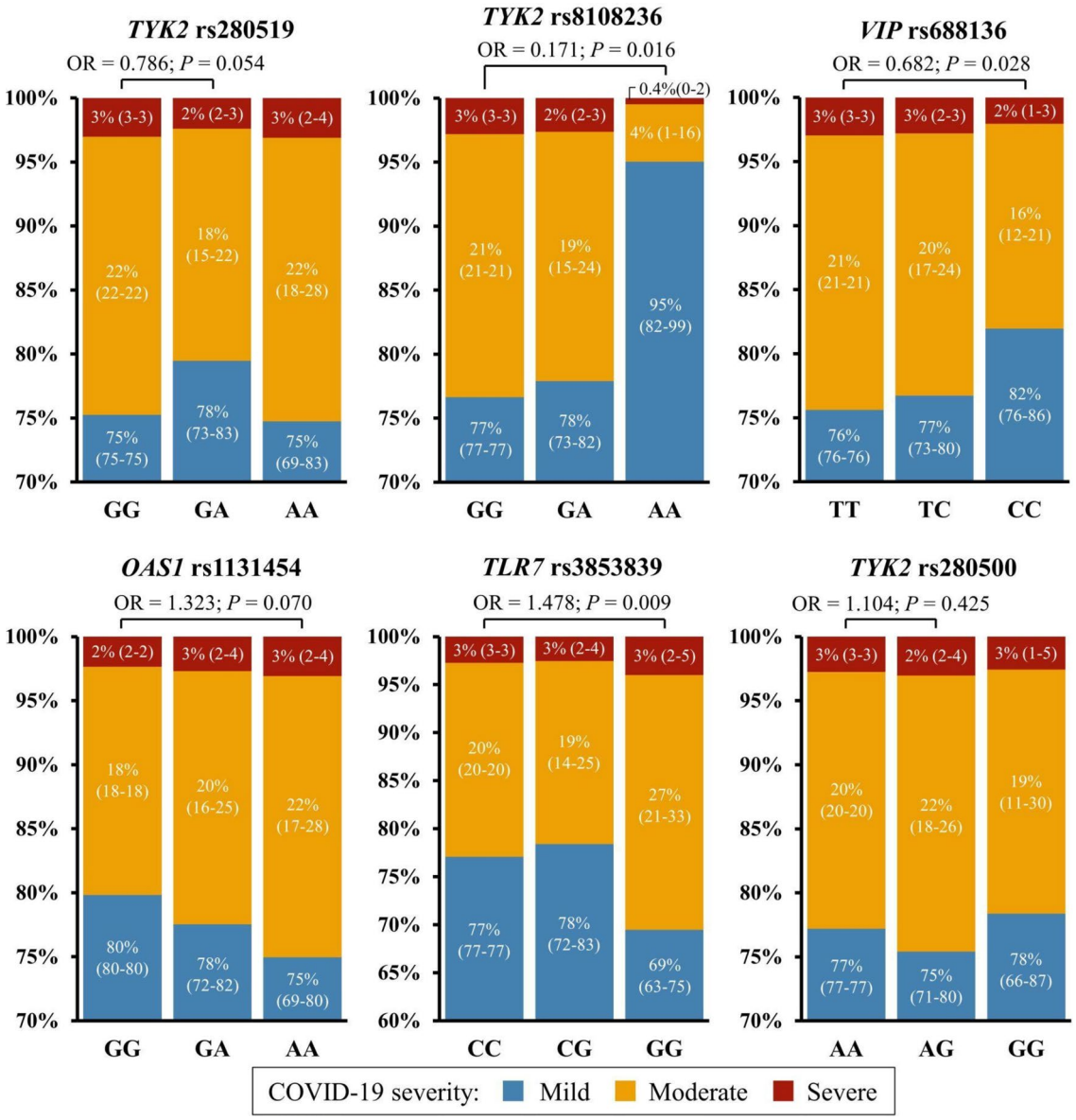

Fig. 1. Predicted probability of COVID-19 severity for each significant SNP in the clinical model. Numbers inside each graph bar indicate the predicted probability of COVID-19 severity as percentage (95% confidence interval) for each genotype: *OAS1* rs1131454, *TLR7* rs3853839, *TYK2* rs280500, rs280519 and rs8108236, and *VIP* rs688136. Upper horizontal lines indicate significant *p*-values and Odds ratios from the multivariable analysis in which each SNP was included separately in the clinical model.

research, we found that *TYK2* rs280500-AG increased the risk of COVID-19 severity, although the effect of this intronic variant on protein function is not known and it has not been previously associated with viral infection.

Downstream TYK2 in the IFN-I signalling pathway is OAS1, involved in the early antiviral response by degrading viral RNA[53]. We found the missense variant rs1131454-AA associated with disease severity, in agreement with previous reports relating this variant to COVID-19 hospitalization[54,55] or increased risk for Alzheimer's disease[56]. In fact, this variant appears associated with this neurodegenerative disease (OR 2.58 [95% CI 1.22–5.46]; $p = 0.013$) in our cohort. Moreover, *OAS1* rs1131454-AA achieved significant association with COVID-19 severity when correcting the model with dementia and other comorbidities, ensuring the independent effect of this SNP on clinical outcome. *OAS1* rs2660 was previously associated with a lower risk of SARS-CoV-2 viremia[24], but we did not find an association with disease outcome.

Altogether, these results reinforce the importance of timing in the host type I IFN response to control viral infection[57]; if this response is poor in the early stages of infection due to missense or impaired-function variants in *TLR7*, *TYK2* or/and *OAS1* it could favour virus evasion mechanisms and result in severe disease[1,25,58]. However, the persistence of an elevated IFN-I secretion when TLR7 and TYK2 expression is increased could induce inflammatory tissue damage, also leading to a worse outcome. Then, the fine-tuning of TLR7 and TYK2 expression is crucial to avoid triggering IFN-I mediated pathological processes.

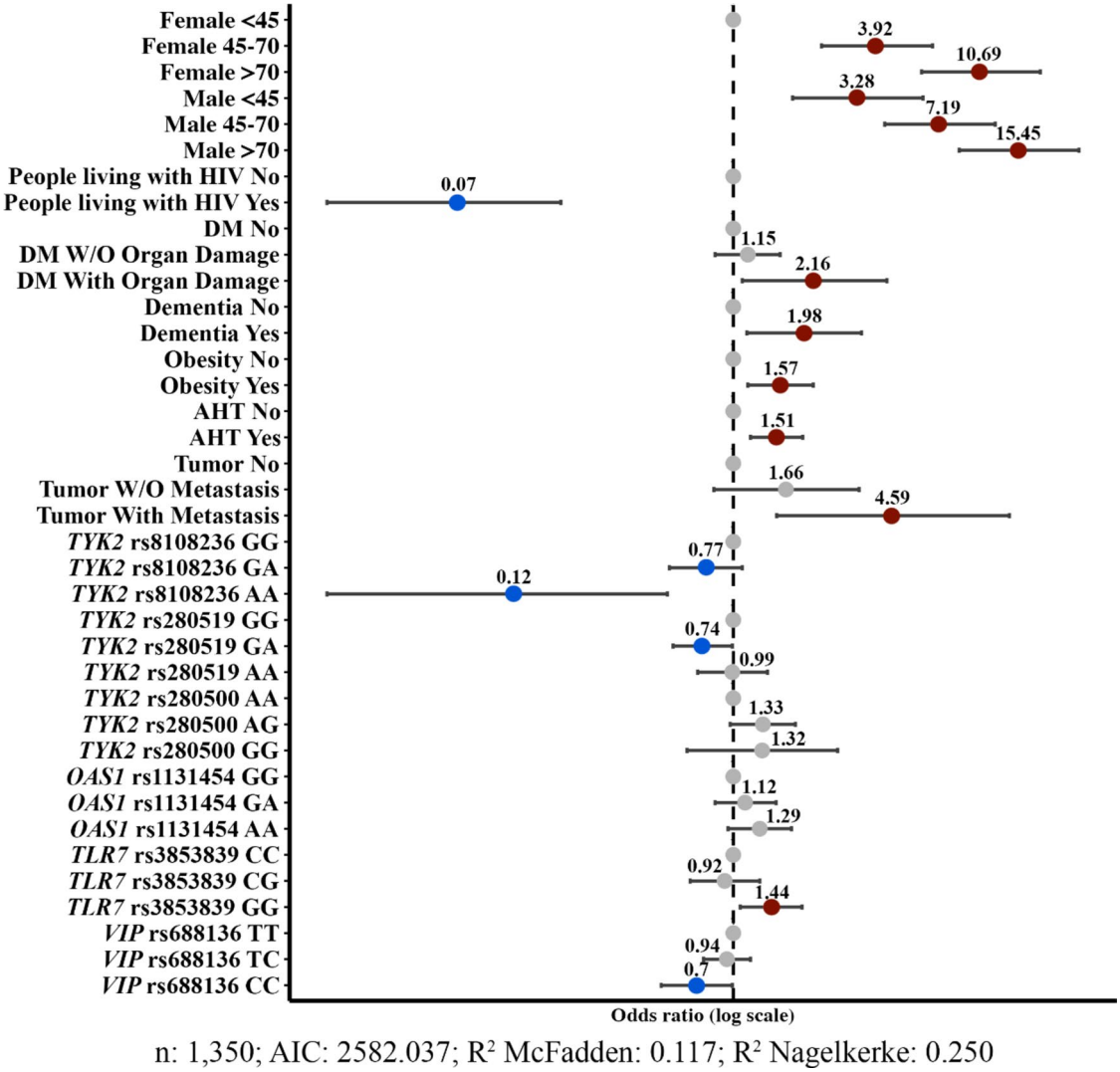

Fig. 2. Forest plot of the final multivariable model with clinical and genetic variables. The Odds ratios (dots) and 95% confidence intervals (horizontal bars) for all variables are depicted. ORs = 1 correspond to reference conditions and are represented by grey dots, as well as non-significant conditions. Conditions that significantly increase the probability of developing severe COVID-19 are represented by red dots, while those conditions significantly associated with the development of mild COVID-19 are represented in blue ($p < 0.05$). *HIV* human immunodeficiency virus, *DM* Diabetes mellitus, *AHT* arterial hypertension.

Regarding *VIP*, we have shown that patients carrying the rs688136-CC genotype had a lower probability to develop severe COVID-19. Previous studies of our group described that the expression of this variant, located in the 3' UTR region, is higher in homozygous patients with early arthritis[19]. Elevated serum VIP levels correlated with a better outcome and response to treatments in these patients[59]. Lower serum VIP levels are present in patients with severe COVID-19[60], therefore, high serum VIP could protect from a worst outcome in both IMID and COVID-19.

In conclusion, genetic variants in *TYK2* (rs8108236 and rs280519) and *VIP* (rs688136) were associated with a better outcome of COVID-19. In contrast, *TLR7* rs3853839, *TYK2* rs280500 and *OAS1* rs1131454 were identified as risk factors predicting higher disease severity. All these results improved the severity predictive model based on the patient's clinical features and comorbidities, helping to explain the variability in COVID-19 outcome within our population.

Materials and methods

Study design and population

This is a retrospective observational study including patients with COVID-19 treated at the University Hospital La Princesa from March 29th, 2020 to September 14th, 2021, before the vaccination campaign started in Spain (see results for detailed description of number of patients). Patients were included in the study if COVID-19 infection was confirmed by real-time polymerase chain reaction (RT-PCR) on nasopharyngeal samples or by serological testing, and if they were older than 18 years and had given informed consent. Blood samples of

admitted patients were collected during their hospitalization while those treated as outpatients were recruited when the pandemic situation did not pose a risk to their health.

Probes and genotyping

Total DNA was extracted from peripheral blood using MagNA Pure 2.0 and MagNA Pure LC DNA Isolation Kit (Roche Life Science, Basel, Switzerland). Concentration and integrity were measured with NanoDrop ND-1000 (Thermo Fisher Scientific). Genotyping was performed at Parque Científico of Autónoma University of Madrid (UAM, Campus Cantoblanco, Madrid) using QuantStudio 12 k Flex, TaqMan™ Genotyping Master Mix and TaqMan™ customized 384 plates (all from Applied Biosystems, Thermo Fisher Scientific, Walthman, MA, USA), using individual probes, duplicate samples and negative controls to verify assay´s accuracy. Allelic discrimination was automatically defined by the TaqMan™ SNP Genotyping App (within Thermo Fisher Connect™, Applied Biosystems™ ANALYSIS SOFTWARE, Genotyping Analysis Module, version 4.1 was employed), based on allele-specific fluorescence. The assessed SNPs and their predesigned TaqMan™ probes and primer sequences are detailed in Table S2.

Variables

Data were extracted retrospectively from electronic medical records and included in a database in a pseudo-anonymized/codified way by removing all identifiable information to ensure the privacy of the patients.

All variables collected can be classified into six groups: demographic data, COVID-19-related clinical data, comorbidities, pharmacological treatment before and during the hospitalization and laboratory parameters. Laboratory parameters and COVID-19 pre-treatments were collected on admission day and according to pre-existing pathologies, respectively (Table S4). Additionally, variables were registered during patient hospitalization period or, in the case of outpatients, at the time of their visit to the emergency department.

The main outcome considered in the study was COVID-19 severity, assessed following a modified version of the 8-point World Health Organization (WHO) Ordinal Scale (WOS) to obtain the worst outcome for each patient (Table S5). According to this scoring system, a categorical outcome variable was elaborated by pooling patients in three levels: mild for 1, 2 and 3 WOS score, moderate for 4, and severe for 5, 6 and 7. In addition, in order to establish a more detailed comparison, among comorbidity variables, two of them were considered categorical instead of binary: diabetes (0: diabetes absence; 1: diabetes without complications; 2: diabetes with end organ damage) and tumors (0: tumor absence; 1: non metastatic solid tumor; 2: metastatic solid tumor).

Statistical analysis

Statistical analyses were performed using Stata 14.0 for Windows (StataCorp LP, College Station, TX, USA). Those quantitative variables with non-normal distribution were represented as median and interquartile range (IQR), and the Mann Whitney or Kruskal–Wallis tests were used to analyse significant differences. Quantitative variables with gaussian distribution were described as mean and standard deviation. Qualitative variables were described as proportions, and the χ^2 test was used to compare categorical variables.

In order to identify those variables associated with variation of COVID-19 severity we fitted an ordered logistic regression analysis (command ologit of Stata), considering the dependent variable the WHO-severity scale grouped in Mild (levels 1 to 3), Moderate (level 4) and Severe (levels 5 to 7) as previously described[61]. The first model was performed by adding all sociodemographic variables, comorbidities and treatments previous to SARS-CoV-2 infection that achieved a p-value < 0.15 in the bivariate analysis (Table S6). The final clinical model was reached through backward stepwise removal of variables with p-value > 0.15, based on the Akaike information criterion (AIC)[62]. Once the best model was obtained, the different SNPs were forced in the model in order to determine whether there was an association with COVID-19 severity or not, using as reference the global major allele established in the Ensembl genome database. All those SNPs that achieved a p-value < 0.15 with this approach were included in the final model combining clinical and genetic variables, which was constructed as previously described for the clinical model. Finally, we evaluated whether the clinical model and the final model (with clinical and genetic variables) were different by AIC, R^2 and ANOVA.

Data availability

The dataset from this publication has been deposited to the Zenodo database and assigned the identifier 10245797.

Received: 14 December 2023; Accepted: 28 August 2024
Published online: 05 September 2024

References

1. Hadjadj, J. *et al.* Impaired type I interferon activity and inflammatory responses in severe COVID-19 patients. *Science* **369**, 718–724 (2020).
2. 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 72 314 cases from the Chinese Center for disease control and prevention. *JAMA* **323**, 1239–1242 (2020).
3. Zhang, J. J. Y., Lee, K. S., Ang, L. W., Leo, Y. S. & Young, B. E. Risk factors for severe disease and efficacy of treatment in patients infected with COVID-19: A systematic review, meta-analysis, and meta-regression analysis. *Clin. Infect Dis.* **71**, 2199–2206 (2020).
4. Zhou, F. *et al.* Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: A retrospective cohort study. *Lancet* **395**, 1054–1062 (2020).
5. Siddiqi, H. K. & Mehra, M. R. COVID-19 illness in native and immunosuppressed states: A clinical-therapeutic staging proposal. *J. Heart Lung Transplant.* **39**, 405–407 (2020).

6. Tsai, P.-H. *et al.* Clinical manifestation and disease progression in COVID-19 infection. *J. Chin. Med. Assoc.* **84**, 3–8 (2021).
7. Páramo, J. A. Inflammatory response in relation to COVID-19 and other prothrombotic phenotypes. *Reumatol. Clin. (Engl. Ed.)* **18**, 1–4 (2022).
8. Yang, L. *et al.* Potential role of IFN-α in COVID-19 patients and its underlying treatment options. *Appl. Microbiol. Biotechnol.* **105**, 4005–4015 (2021).
9. Diebold, S. S., Kaisho, T., Hemmi, H., Akira, S. & Reis e Sousa, C. Innate antiviral responses by means of TLR7-mediated recognition of single-stranded RNA. *Science* **303**, 1529–1531 (2004).
10. Lund, J. M. *et al.* Recognition of single-stranded RNA viruses by Toll-like receptor 7. *Proc. Natl. Acad. Sci. USA* **101**, 5598–5603 (2004).
11. van der Sluis, R. M. *et al.* TLR2 and TLR7 mediate distinct immunopathological and antiviral plasmacytoid dendritic cell responses to SARS-CoV-2 infection. *EMBO J.* **41**, e109622 (2022).
12. Crosse, K. M., Monson, E. A., Beard, M. R. & Helbig, K. J. Interferon-stimulated genes as enhancers of antiviral innate immune signaling. *J. Innate Immun.* **10**, 85–93 (2018).
13. Schoggins, J. W. Interferon-stimulated genes: What Do they all do?. *Annu. Rev. Virol.* **6**, 567–584 (2019).
14. Gonciarz, M., Pawlak-Buś, K., Leszczyński, P. & Owczarek, W. TYK2 as a therapeutic target in the treatment of autoimmune and inflammatory diseases. *Immunotherapy* **13**, 1135–1150 (2021).
15. Hu, X., Li, J., Fu, M., Zhao, X. & Wang, W. The JAK/STAT signaling pathway: From bench to clinic. *Signal Transduct. Target Ther.* **6**, 402 (2021).
16. Muromoto, R., Shimoda, K., Oritani, K. & Matsuda, T. Therapeutic advantage of Tyk2 inhibition for treating autoimmune and chronic inflammatory diseases. *Biol. Pharm. Bull.* **44**, 1585–1592 (2021).
17. Gomariz, R. P. *et al.* An overview of VPAC receptors in rheumatoid arthritis: Biological role and clinical significance. *Front. Endocrinol.* **10**, 729 (2019).
18. Martínez, C. *et al.* A Clinical approach for the use of VIP axis in inflammatory and autoimmune diseases. *Int. J. Mol. Sci.* **21**, 65 (2019).
19. Lamana, A. *et al.* VIP/VPAC axis expression in immune-mediated inflammatory disorders: Associated miRNA signatures. *Int. J. Mol. Sci.* **23**, 8578 (2022).
20. He, Y., Zang, N. & Liu, E.-M. Vasoactive intestinal peptide: A potential target for antiviral therapy. *Sheng Li Xue Bao* **74**, 419–433 (2022).
21. Mukherjee, T. *et al.* Anticipated pharmacological role of Aviptadil on COVID-19. *Environ. Sci. Pollut. Res. Int.* **29**, 8109–8125 (2022).
22. Colona, V. L., Vasiliou, V., Watt, J., Novelli, G. & Reichardt, J. K. V. Update on human genetic susceptibility to COVID-19: Susceptibility to virus and response. *Hum. Genom.* **15**, 57 (2021).
23. Villapalos-García, G. *et al.* Transmembrane protease serine 2 (TMPRSS2) rs75603675, comorbidity, and sex are the primary predictors of COVID-19 severity. *Life Sci. Alliance* **5**, e202201396 (2022).
24. Roy-Vallejo, E. *et al.* Occurrence of SARS-CoV-2 viremia is associated with genetic variants of genes related to COVID-19 pathogenesis. *Front Med* **10**, 1215246 (2023).
25. Blanco-Melo, D. *et al.* Imbalanced host response to SARS-CoV-2 drives development of COVID-19. *Cell* **181**, 1036-1045.e9 (2020).
26. van der Made, C. I. *et al.* Presence of genetic variants among young men with severe COVID-19. *JAMA* **324**, 663–673 (2020).
27. Zhang, Q. *et al.* Inborn errors of type I IFN immunity in patients with life-threatening COVID-19. *Science* **370**, eabd4570 (2020).
28. Zhang, Q. *et al.* Recessive inborn errors of type I IFN immunity in children with COVID-19 pneumonia. *J. Exp. Med.* **219**, e20220131 (2022).
29. Dupuis, S. *et al.* Impaired response to interferon-alpha/beta and lethal viral disease in human STAT1 deficiency. *Nat. Genet.* **33**, 388–391 (2003).
30. Kreins, A. Y. *et al.* Human TYK2 deficiency: Mycobacterial and viral infections without hyper-IgE syndrome. *J. Exp. Med.* **212**, 1641–1662 (2015).
31. Ott, N., Faletti, L., Heeg, M., Andreani, V. & Grimbacher, B. JAKs and STATs from a clinical perspective: Loss-of-function mutations, gain-of-function mutations, and their multidimensional consequences. *J. Clin. Immunol.* **43**, 1326–1359 (2023).
32. COVID-19 Host Genetics Initiative. Mapping the human genetic architecture of COVID-19. *Nature* **600**, 472–477 (2021).
33. Kousathanas, A. *et al.* Whole-genome sequencing reveals host factors underlying critical COVID-19. *Nature* **607**, 97–103 (2022).
34. Pairo-Castineira, E. *et al.* Genetic mechanisms of critical illness in COVID-19. *Nature* **591**, 92–98 (2021).
35. Berg, J. *et al.* Tyk2 as a target for immune regulation in human viral/bacterial pneumonia. *Eur. Respir. J.* **50**, 1601953 (2017).
36. Tian, W. *et al.* Predictors of mortality in hospitalized COVID-19 patients: A systematic review and meta-analysis. *J. Med. Virol.* **92**, 1875–1883 (2020).
37. Bautista-Becerril, B. *et al.* High expression levels of miR-21-5p in younger hospitalized COVID-19 patients are associated with mortality and critical disease. *Int. J. Mol. Sci.* **24**, 10112 (2023).
38. Bertagnolio, S. *et al.* Clinical features of, and risk factors for, severe or fatal COVID-19 among people living with HIV admitted to hospital: Analysis of data from the WHO Global Clinical Platform of COVID-19. *Lancet HIV* **9**, e486–e495 (2022).
39. Del Amo, J. *et al.* Incidence and severity of COVID-19 in HIV-positive persons receiving antiretroviral therapy: A cohort study. *Ann. Intern. Med.* **173**, 536–541 (2020).
40. Martínez-Gómez, L. E. *et al.* Implication of myddosome complex genetic variants in outcome severity of COVID-19 patients. *J. Microbiol. Immunol. Infect.* **56**, 939–950 (2023).
41. Naushad, S. M., Mandadapu, G., Ramaiah, M. J., Almajhdi, F. N. & Hussain, T. The role of TLR7 agonists in modulating COVID-19 severity in subjects with loss-of-function TLR7 variants. *Sci. Rep.* **13**, 13078 (2023).
42. Shen, N. *et al.* Sex-specific association of X-linked Toll-like receptor 7 (TLR7) with male systemic lupus erythematosus. *Proc. Natl. Acad. Sci. USA* **107**, 15838–15843 (2010).
43. Liao, T.-L. *et al.* Hepatitis C virus-induced exosomal microRNAs and toll-like receptor 7 polymorphism regulate B-cell activating factor. *mBio* **12**, e0276421 (2021).
44. El-Hefnawy, S. M. *et al.* COVID-19 susceptibility, severity, clinical outcome and Toll-like receptor (7) mRNA expression driven by TLR7 gene polymorphism (rs3853839) in middle-aged individuals without previous comorbidities. *Gene Rep.* **27**, 101612 (2022).
45. Asano, T. *et al.* X-linked recessive TLR7 deficiency in ~1% of men under 60 years old with life-threatening COVID-19. *Sci. Immunol.* **6**, eabl4348 (2021).
46. Fallerini, C. *et al.* Association of Toll-like receptor 7 variants with life-threatening COVID-19 disease in males: Findings from a nested case-control study. *Elife* **10**, e67569 (2021).
47. Bodega-Mayor, I. *et al.* Tyrosine kinase 2 modulates splenic B cells through type I IFN and TLR7 signaling. *Cell Mol. Life Sci* **81**, 199 (2024).
48. Cunninghame Graham, D. S., Akil, M. & Vyse, T. J. Association of polymorphisms across the tyrosine kinase gene, TYK2 in UK SLE families. *Rheumatology (Oxford)* **46**, 927–930 (2007).
49. Hellquist, A. *et al.* Evidence for genetic association and interaction between the TYK2 and IRF5 genes in systemic lupus erythematosus. *J. Rheumatol.* **36**, 1631–1638 (2009).
50. Eyre, S. *et al.* High-density genetic mapping identifies new susceptibility loci for rheumatoid arthritis. *Nat. Genet.* **44**, 1336–1340 (2012).

51. Diogo, D. *et al.* TYK2 protein-coding variants protect against rheumatoid arthritis and autoimmunity, with no evidence of major pleiotropic effects on non-autoimmune complex traits. *PLoS ONE* **10**, e0122271 (2015).
52. Dieter, C. *et al.* Polymorphisms in ACE1, TMPRSS2, IFIH1, IFNAR2, and TYK2 genes are associated with worse clinical outcomes in COVID-19. *Genes* **14**, 29 (2022).
53. Choi, U. Y., Kang, J.-S., Hwang, Y. S. & Kim, Y.-J. Oligoadenylate synthase-like (OASL) proteins: Dual functions and associations with diseases. *Exp. Mol. Med.* **47**, e144 (2015).
54. Tanimine, N. *et al.* Identification of aggravation-predicting gene polymorphisms in coronavirus disease 2019 patients using a candidate gene approach associated with multiple phase pathogenesis: A study in a Japanese City of 1 million people. *Crit. Care Explor.* **3**, e0576 (2021).
55. Banday, A. R. *et al.* Genetic regulation of OAS1 nonsense-mediated decay underlies association with COVID-19 hospitalization in patients of European and African ancestries. *Nat. Genet.* **54**, 1103–1116 (2022).
56. Magusali, N. *et al.* A genetic link between risk for Alzheimer's disease and severe COVID-19 outcomes via the OAS1 gene. *Brain* **144**, 3727–3741 (2021).
57. Park, A. & Iwasaki, A. Type I and Type III interferons—Induction, signaling, evasion, and application to combat COVID-19. *Cell Host Microbe* **27**, 870–878 (2020).
58. Zhang, X. *et al.* Viral and host factors related to the clinical outcome of COVID-19. *Nature* **583**, 437–440 (2020).
59. Seoane, I. V. *et al.* Vasoactive intestinal peptide gene polymorphisms, associated with its serum levels, predict treatment requirements in early rheumatoid arthritis. *Sci. Rep.* **8**, 2035 (2018).
60. Temerozo, J. R. *et al.* VIP plasma levels associate with survival in severe COVID-19 patients, correlating with protective effects in SARS-CoV-2-infected cells. *J. Leukoc. Biol.* **111**, 1107–1121 (2022).
61. Su, Y. *et al.* Multi-omics resolves a sharp disease-state shift between mild and moderate COVID-19. *Cell* **183**, 1479-1495.e20 (2020).
62. Akaike, H. A new look at the statistical model identification. *Sel. Pap. Hirotugu Akaike* https://doi.org/10.1007/978-1-4612-1694-0_16 (1974).

Acknowledgements

We thank Dr. Manuel Gómez (IIS-Princesa) for editing the manuscript and the Fundación de Investigación Biomédica (FIB) administrative staff for support. This study was funded with grants: "Fondos Supera COVID19" by Banco Santander and CRUE to CM-C and RPG; RD21/0002/0027, PI18/0371 and PI21/00526 to IG-Á; PI19/00096 and PI22/00428 to EF-R and RICORS RD21/0002/0004 to RPG from Ministerio de Ciencia e Innovación (Instituto de Salud Carlos III, ISCIII), co-funded by European Regional Development Fund (ERDF) "A way to make Europe"; and co-financed by the Community of Madrid (CAM) through the COVID 2019 Aid to IdeS. The work of ER-V and AM-J was funded by Rio-Hortega grants CM19/00149 and CM19/00254, respectively, from the Ministerio de Ciencia e Innovación (Instituto de Salud Carlos III, ISCIII) and co-funded by The ERDF "A way to make Europe"; GV-G and PD-F were co-financed by ISCIII and the European Social Fund (PFIS grants FI20/00090 and FI19/00092, respectively). PZ was financed by Universidad Autónoma de Madrid, Margarita Salas contract, grants for the requalification of the Spanish university system. EA-C and AN-G were financed by INVESTIGO (2022-C23.I01.P03.S0020-0000031) and INVESTIGO CAM (09-PIN1-00015.6/2022), by Ministerio de Ciencia e Innovación and CAM respectively, both financed by the European Union's Recovery, Transformation and Resilience Plan and NextGenerationEU.

Author contributions

P.D.-W., S.C.-O.: conceptualization, writing—original draft, data curation, investigation, formal analysis, methodology, visualization, writing—review and editing. E.R.-V.: conceptualization, writing—original draft, data curation, investigation, formal analysis, methodology, software, validation, visualization, writing—review and editing. E.A.-C., A.N.-G., F.A.-S., A.M.-J.: data curation, investigation, visualization, writing—review and editing. A.L.: conceptualization, writing—original draft, methodology, visualization, writing—review and editing. N.M.: formal analysis, methodology, software, validation, visualization, writing—review and editing. D.R.-S., R.C.-R., P.D.-F., J.G.-R., L.R.-R., M.S.-A., J.A.-R., A.V.-M., C.R.-F., G.V.-G., P.Z.: data curation, investigation, visualization. I.S.: funding acquisition, project administration, resources, supervision, writing—review and editing. R.P.G., C.M.-C., R.G.-V.: conceptualization, funding acquisition, project administration, resources, supervision, writing—review and editing. I.G.-Á.: conceptualization, writing—original draft, formal analysis, methodology, software, validation, funding acquisition, project administration, resources, supervision, visualization, writing—review and editing. E.F.-R.: conceptualization, writing—original draft, funding acquisition, project administration, resources, supervision, visualization, writing—review and editing.

Competing interests

TFA-S has been consultant or investigator in clinical trials sponsored by the following pharmaceutical companies: Abbott, Alter, Aptatargets, Chemo, FAES, Farmalider, Ferrer, Galenicum, GlaxoSmithKline, Gilead, Italfarmaco, Janssen-427 Cilag, Kern, Normon, Novartis, Servier, Teva and Zambon. IG-Á reports personal fees from Lilly and Sanofi; personal fees and non-financial support from BMS; personal fees and non-financial support from Abbvie; research support, personal fees and non-financial support from Roche Laboratories; research support from Gebro Pharma; non-financial support from MSD, Pfizer and Novartis, not related to the submitted work. RGV declares educational or research grants for her institution from Abbvie, Lilly, Janssen, MSD, Novartis, Sanofi and UCB; consultancies/speaking personal fees from Abbvie, Biogen, MSD, Pfizer, Sandoz and UCB; non-financial support from Abbvie, Janssen, Lilly, MSD, Novartis, Pfizer and UCB, all outside the present work. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Ethics

his study was approved by the Research Ethics Committee of University Hospital La Princesa, Madrid (register number 4070, March 30th 2020), and it was carried out following the ethical principles of the Declaration of Helsinki. All patients were informed about the study and gave oral or written consent to participate, which

was registered in their electronic clinical chart. Due to the COVID-19 pandemic emergency, oral consent was accepted as proposed by the AEMPS (Agencia Española de Medicamentos y Productos Sanitarios, The Spanish Agency for Medicines and Medical Devices) in April 2020. Due to the periodical medical supervision, the written consent of the people living with HIV was subsequently obtained. This article was written following the STREGA (STrengthening the REporting of Genetic Association studies) guidelines (Table S7).

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-024-71476-2.

Correspondence and requests for materials should be addressed to I.G. or E.F.-R.

Reprints and permissions information is available at www.nature.com/reprints.

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

PREDINMUN-COVID Group

Carmen Suárez Fernández[3], Ana Barrios[3], Jesús Sanz[3], Pedro Casado[3], Ángela Gutiérrez[3], Azucena Bautista[3], Pilar Hernández[3], Nuria Ruiz Giménez[3], Berta Moyano[3], Paloma Gil[3], María Jesús Delgado[3], Pedro Parra[3], Beatriz Sánchez[3], Carmen Sáez[3], Marta Fernández-Rico[3], Cristina Arévalo-Román[3], Marianela Ciudad[3], Santos Castañeda[2], Irene Llorente[2], Eva G. Tomero[2], Noelia García-Castañeda[2], Miren Uriarte[2], Laura Cardeñoso[12], Leticia Fontán García-Rodrigo[12], Diego Domingo García[12], Teresa Alarcón-Cavero[12], María Auxiliadora Semiglia Chong[12], Ainhoa Gutiérrez-Cobos[12], Nelly D. Zurita-Cruz[12], Francisco Sánchez-Madrid[7], Enrique Martín-Gayo[7], Ildefonso Sánchez-Cerrillo[7], Pedro Martínez-Fleta[7], Celia López-Sanz[7], Ligia Gabrie[7], Luciana del Campo-Guerola[7], Reyes Tejedor[7], Julio Ancochea[13], Elena García-Castillo[13], Elena Ávalos[13], Ana Sánchez-Azofra[13], Tamara Alonso[13], Carolina Cisneros[13], Claudia Valenzuela[13], Francisco J. García-Pérez[13], Rosa M. Girón[13], Javier Aspa[13], Celeste Marcos[13], M. del Perpetuo Socorro Churruca[13], Enrique Zamora[13], Adrián Martínez[13], Mar Barrio-Mayo[13], Rosalina Henares-Espi[13], Rosa Méndez[14], David Arribas[14], Marta Chicot-Llano[4], Begoña González[4], Begoña Quicios[4], Pablo Patiño[4], Marina Trigueros[4], Cristina Dominguez-Peña[4], David Jiménez-Jiménez[4], Pablo Villamayor[4], Alfonso Canabal[4], Rafael de la Cámara[15], Javier Ortiz[15] & Isabel Iturrate[15]

[12]Microbiology Department, Hospital Universitario La Princesa (IIS-Princesa), Madrid, Spain. [13]Pneumology Department, Hospital Universitario La Princesa (IIS-Princesa), Madrid, Spain. [14]Anesthesiology Department, Hospital Universitario La Princesa (IIS-Princesa), Madrid, Spain. [15]Hematology Department, Hospital Universitario La Princesa (IIS-Princesa), Madrid, Spain.