

Genetic prediction of ICU hospitalization and mortality in COVID-19 patients using artificial neural networks

Panagiotis G. Asteris¹ | Eleni Gavriilaki²  | Tasoula Touloumenidou² | Evaggelia-Evdoxia Koravou² | Maria Koutra² | Penelope Georgia Papayanni² | Alexandros Pouleres³ | Vassiliki Karali⁴ | Minas E. Lemonis¹ | Anna Mamou¹ | Athanasia D. Skentou¹ | Apostolia Papalexandri² | Christos Varelas² | Fani Chatzopoulou⁵ | Maria Chatzidimitriou⁶ | Dimitrios Chatzidimitriou⁵ | Anastasia Veleni⁷ | Evdoxia Rapti⁸ | Ioannis Kioumis⁹ | Evaggelos Kaimakamis¹⁰ | Milly Bitzani¹⁰ | Dimitrios Boumpas⁴ | Argyris Tsantes⁸ | Damianos Sotiropoulos² | Anastasia Papadopoulou² | Ioannis G. Kalantzis¹¹ | Lydia A. Vallianatou¹² | Danial J. Armaghani¹³ | Liborio Cavaleri¹⁴ | Amir H. Gandomi¹⁵ | Mohsen Hajihassani¹⁶ | Mahdi Hasanipanah¹⁷ | Mohammadreza Koopialipour¹⁸ | Paulo B. Lourenço¹⁹ | Pijush Samui²⁰ | Jian Zhou²¹ | Ioanna Sakellari² | Serena Valsami²² | Marianna Politou²²  | Styliani Kokoris⁸  | Achilles Anagnostopoulos²

¹Computational Mechanics Laboratory, School of Pedagogical and Technological Education, Athens, Greece

²Hematology Department, BMT Unit, G Papanicolaou Hospital, Thessaloniki, Greece

³Biomedical Laboratories, Athens, Greece

⁴Rheumatology and Clinical Immunology Unit, University General Hospital 'Attikon', Athens, Greece

⁵Microbiology Department, Aristotle University of Thessaloniki, Greece

⁶Biomedical Sciences Alexander Campus International Hellenic University, Thessaloniki, Greece

⁷Infectious Disease Committee, G Papanicolaou Hospital, Thessaloniki, Greece

⁸Laboratory of Hematology and Hospital Blood Transfusion Department, University General Hospital 'Attikon', NKUA, Medical School, Athens, Greece

⁹Respiratory Failure Department, G Papanicolaou Hospital, Aristotle University of Thessaloniki, Thessaloniki, Greece

¹⁰1st Intensive Care Unit, G Papanicolaou Hospital, Thessaloniki, Greece

¹¹Gastroenterology Department, Hellenic Red Cross Hospital, Athens, Greece

¹²Cath and EP Laboratory/Covid Intensive Care Unit, Hellenic Red Cross Hospital, Athens, Greece

¹³Department of Civil Engineering, Faculty of Engineering, University of Malaya, Kuala Lumpur, Malaysia

¹⁴Department of Civil, Environmental, Aerospace and Materials Engineering, University of Palermo, Palermo, Italy

¹⁵Faculty of Engineering & Information Technology, University of Technology Sydney, Ultimo, NSW, Australia

¹⁶Department of Mining Engineering, Faculty of Engineering, Urmia University, Urmia, Iran

¹⁷Institute of Research and Development, Duy Tan University, Da Nang, Vietnam

¹⁸Faculty of Civil and Environmental Engineering, Amirkabir University of Technology, Tehran, Iran

¹⁹Department of Civil Engineering, ISISE, University of Minho, Guimarães, Portugal

²⁰Department of Civil Engineering, National Institute of Technology Patna, Patna, Bihar, India

²¹School of Resources and Safety Engineering, Central South University, Changsha, China

Panagiotis G. Asteris, Eleni Gavriilaki, Styliani Kokoris and Achilles Anagnostopoulos contributed equally to this work.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2022 The Authors. *Journal of Cellular and Molecular Medicine* published by Foundation for Cellular and Molecular Medicine and John Wiley & Sons Ltd.

²²Hematology Laboratory - Blood Bank, Aretaieion Hospital, School of Medicine, NKUA, Athens, Greece

Correspondence

Eleni Gavriilaki, Hematology Department
– BMT Unit, G. Papanicolaou Hospital,
Exochi, 57010, Thessaloniki; Greece.
Email: elenicelli@yahoo.gr

Funding Information

Our research was supported by independent investigator-driven grants (Prefecture of Macedonia and Pfizer Pharmaceuticals).

Abstract

There is an unmet need of models for early prediction of morbidity and mortality of Coronavirus disease-19 (COVID-19). We aimed to a) identify complement-related genetic variants associated with the clinical outcomes of ICU hospitalization and death, b) develop an artificial neural network (ANN) predicting these outcomes and c) validate whether complement-related variants are associated with an impaired complement phenotype. We prospectively recruited consecutive adult patients of Caucasian origin, hospitalized due to COVID-19. Through targeted next-generation sequencing, we identified variants in complement factor H/*CFH*, *CFB*, *CFH-related*, *CFD*, *CD55*, *C3*, *C5*, *CFI*, *CD46*, *thrombomodulin/THBD*, and *A Disintegrin and Metalloproteinase with Thrombospondin motifs (ADAMTS13)*. Among 381 variants in 133 patients, we identified 5 critical variants associated with severe COVID-19: rs2547438 (*C3*), rs2250656 (*C3*), rs1042580 (*THBD*), rs800292 (*CFH*) and rs414628 (*CFHR1*). Using age, gender and presence or absence of each variant, we developed an ANN predicting morbidity and mortality in 89.47% of the examined population. Furthermore, *THBD* and *C3a* levels were significantly increased in severe COVID-19 patients and those harbouring relevant variants. Thus, we reveal for the first time an ANN accurately predicting ICU hospitalization and death in COVID-19 patients, based on genetic variants in complement genes, age and gender. Importantly, we confirm that genetic dysregulation is associated with impaired complement phenotype.

KEYWORDS

artificial intelligence, complement, complement inhibition, COVID-19, genetic susceptibility, SARS-CoV2

1 | INTRODUCTION

Coronavirus disease-19 (COVID-19) caused by severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) has led to unprecedented morbidity and mortality worldwide.¹ Although vaccination against SARS-CoV-2 has positively impacted the course of this pandemic,² the unmet need of reducing morbidity and mortality due to severe COVID-19, especially in special populations, remains. Accumulating evidence suggests that SARS-CoV-2 induces a vicious cycle of immune dysfunction, endothelial injury, complement activation and microangiopathy. In particular, severe COVID-19 is a multisystemic vascular disease characterized by endothelial dysfunction.^{3,4} Therefore, an improved understanding of endothelial dysfunction and complement activation is of utmost importance.

In this context, several groups worldwide have shown evidence of complement activation in experimental and clinical studies of severe COVID-19.⁵⁻¹³ Based on the paradigm of genetic susceptibility in complement-mediated disorders or complementopathies,¹⁴ our group and other researchers have suggested genetic susceptibility identifying complement genetic variants in COVID-19 patients.¹⁵⁻¹⁷

In parallel, complement inhibitors have been safe and effective in severe COVID-19 during the first wave.¹⁸ Encouraging results have been reported in case series for terminal complement inhibition with eculizumab,¹⁹ C3 inhibition with the AMY-101,²⁰ C1 inhibition with conestat alpha²¹ and lectin pathway inhibition with narsoplumab.²² The comparison of AMY-101 to eculizumab suggested a broader involvement of C3 in thromboinflammation.²³ Based on promising data, randomized controlled trials are ongoing for AMY-191 and eculizumab in severe COVID-19 (NCT04346797 and NCT04395456). Interim analysis is still pending on the paused phase 3 study of the long-acting terminal complement inhibitor ravulizumab. Nevertheless, several issues need to be considered for the wider use of complement inhibitors in COVID-19, including the complex setting of inflammatory responses in COVID-19, the cost and the limitation of drug accessibility. All require proper selection of patients that would potentially benefit from complement inhibition.

Taking into account the constantly evolving COVID-19 landscape due to vaccination and viral mutations, there is an unmet clinical need of a prediction tool based on robust variables. Therefore, we aimed initially to identify critical complement-related genetic

variants predicting severe COVID-19. Then, for the first time in COVID-19 patients, we sought to develop an artificial neural network (ANN) that predicts morbidity and mortality incorporating genetic variants and defining characteristics such as age and gender, but also to validate that the identified complement-related variants are associated with an impaired complement phenotype who might benefit from complement inhibitors.

2 | METHODS

2.1 | Artificial neural networks

Development of artificial neural networks (ANNs) is based on the concept of the biological neural network of the human brain and was initially used for medicine research purposes to simulate strongly non-linear relationships between numerous input and output parameters.²⁴⁻³⁰ ANN models were subsequently introduced into the wider context of engineering disciplines,^{31,32} which significantly enriched the mathematical background underpinning ANN. At present, ANNs are widely employed on an interdisciplinary interface between various disciplines.¹⁷ A few studies employing ANNs have also emerged focussing on COVID-19 prediction problems.³³⁻⁴⁰

The basic building block of ANNs is the artificial neuron, which is a mathematical model mimicking the behaviour of the biological neuron (Figure 1). Information is passed onto the artificial neuron as an input parameter and is then processed using a mathematical function to derive an output which determines the behaviour of the neuron (similar to the fire-or-not situation of the biological neuron). Before the information enters the neuron, it is weighted in order to simulate the random nature of the biological neuron. A group of such neurons comprise an ANN, similar to the structure of biological neural networks. To define an ANN, (i) the architecture of the ANN, (ii) the training algorithm used during the ANN's training stage and (iii) the mathematical functions underpinning the mathematical model are required.

The architecture of the ANN defines how the artificial neurons are organized and how the information flows within the network. For example, if the neurons are organized in more than one layer, then the network is called a multilayer ANN. The training stage can

be considered as a function minimization problem, in which the optimum weight values need to be determined by minimizing an error function.

2.2 | Patient population

We prospectively recruited consecutive adult patients of Caucasian origin, hospitalized due to COVID-19 in our referral centres (G Papanicolaou, Attikon Hospital) (April-December 2020). RT-PCR (reverse-transcriptase polymerase chain reaction) confirmed SARS-COV2 infection. Patients' history and course were documented by treating physicians following patients up to discharge or death. Our study was approved by Institutional Review Boards (IRBs) of referral centres and conducted according to the Declaration of Helsinki.

We studied 133 COVID-19 patients, 80 hospitalized in COVID-19 general ward (GW-not in ICU) and 53 hospitalized in intensive care units (ICU). Figure 2 presents patients characteristics according to age, gender and disease severity.

2.3 | Genetic studies

Genetic studies were performed as previously described.¹⁷ Briefly, peripheral blood samples were used to isolate genomic DNA that underwent next-generation sequencing (NGS, MiniSeq, Illumina, San Diego, California) analysis with our customized gene panel for complement factor *H/CFH*, *CFB*, *CFH-related*, *CFD*, *CD55*, *C3*, *C5*, *CFI*, *CD46*, *thrombomodulin/THBD*, and *A Disintegrin and Metalloproteinase with Thrombospondin motifs (ADAMTS13)*. DesignStudio (Illumina, San Diego, California) was used for probes design, in order to cover all exonic regions spanning 15 bases into introns (coverage 98%).

2.4 | Functional assessment of complement activation

Plasma was isolated from EDTA tubes collected at hospitalization for non-ICU patients or at ICU admission for ICU patients and stored

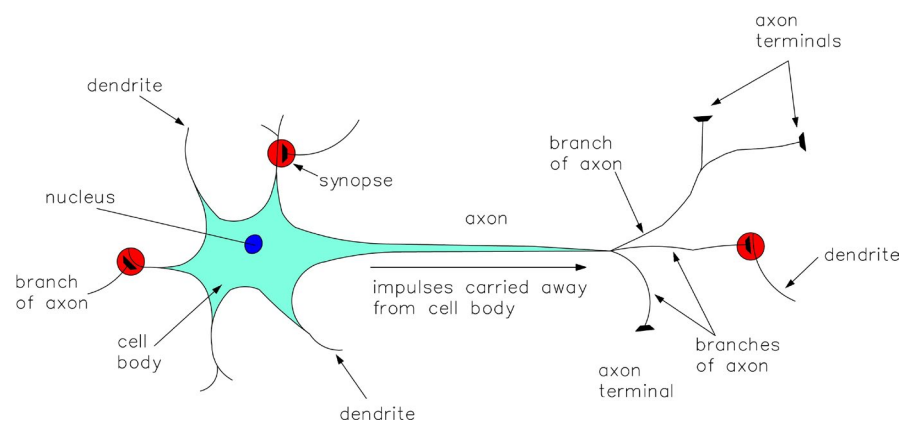


FIGURE 1 Typical biological neuron structure. The artificial neural network mathematical models have an analogous simplified architecture resembling the structure of neurons in the biological prototype

Severe	Age of Infected Patients	All Patients	Male Patients	Female Patients
ICU & not in ICU & Died	Up to 65	76	51	25
	Over 65	57	36	21
	All Ages	133	87	46
ICU	Up to 65	30	20	10
	Over 65	23	19	4
	All Ages	53	39	14
not in ICU	Up to 65	46	31	15
	Over 65	34	17	17
	All Ages	80	48	32
Died	Up to 65	9	5	4
	Over 65	17	11	6
	All Ages	26	16	10

FIGURE 2 Study population characteristics categorized by age, gender and infection severity (requiring or not requiring intensive care unit (ICU), mortality)

immediately at -80°C . Taking into account the difficulties of functional complement assessment in clinical laboratories,⁴¹ we studied markers that would be potentially accessible in everyday clinical practice or the setting of a clinical trial. Therefore, we measured THBD (R&D, Bio-Techne), C3a (Invitrogen, Thermo Fisher Scientific) and C5a (Affymetrix, Bender Medsystems) using commercially available ELISA assays. Statistical comparison of continuous variables between groups was performed using the statistical program SPSS 23.0 (IBM SPSS Statistics for Windows, Version 23.0.; IBM Corp). Descriptive statistics were described using median (interquartile range) or mean \pm standard deviation for continuous variables according to normality, and frequency for categorical variables. Continuous variables were compared using t test or Mann-Whitney, according to normality.

2.5 | Criteria for the optimum database filtering

Given the essentially infinite number of available variant combinations (in excess of yotta = 10^{24}), advanced computational analysis tools were required to correlate the effect of variant combinations with the severity of COVID-19 infection. To reduce the computational complexity associated with the significant number of available variants, a pre-defined set of criteria was used to identify the minimum number of required variants, which are mainly associated with severe COVID-19 infection.

Criterion I: Variants which are present in more than 90% or less than 10% of COVID-19 infected patients are considered as non-associated with COVID-19 severity.

Criterion II: From the remaining available variant combinations, single variant types associated with severe complications of COVID-19 infection are investigated. This is achieved using the following II.I and II.II Criteria.

Criterion II.I: Variants are identified, which are present in the 90% of COVID-19 patients requiring intensive care unit (ICU) and at the same time are not present in the 90% of patients not requiring ICU hospitalization.

Criterion II.II: Variants are identified, which are not present in the 90% of COVID-19 patients requiring ICU and at the same time are present in the 90% of COVID-19 patients not requiring ICU. Criterion II.II is essentially the reverse of Criterion II.I.

This constraint was driven by the computational requirements during the training and development of ANN models. In this study, different architectures of ANNs were trained and developed and the main stages are summarized in the results section, later in the text.

3 | RESULTS

3.1 | Genetic analysis

We identified a total of 381 variants, ranging from 40 to 101 per patient (mean value 71 and standard deviation 12). The database of variants is presented as supplementary material in the supplemental excel file entitled variants Database of 133 COVID-19 Patients.

Using the pre-defined set of criteria, a total of 381 variants are reduced to 5 critical variants, which are mainly associated with severe COVID-19 infection (morbidity and mortality), as shown in Table 1:

TABLE 1 Prediction of COVID-19 infection severity using the 5 critical variants obtained from criteria I and II

Ranking	Rs	Gene	Position	Percentage prediction						Criterion			
				In ICU (53)			Not in ICU (80)				In ICU and not in ICU (133)		
				All	Male	Female	All	Male	Female		All	Male	Female
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
5	rs2547438	C3	6718078	75.47	79.49	64.29	35.00	27.08	46.88	51.13	50.57	52.17	II
2	rs2250656	C3	6718534	62.26	71.79	35.71	52.50	52.08	53.13	56.39	60.92	47.83	II
1	rs1042580	THBD	23027621	49.06	56.41	28.57	66.25	62.50	71.88	59.40	59.77	58.70	II
4	rs800292	CFH	196642233	37.74	38.46	35.71	65.00	56.25	78.13	54.14	48.28	65.22	I
3	rs414628	CFHR1	196801078	62.26	64.10	57.14	52.50	45.83	62.50	56.39	54.02	60.87	I

Note: Criterion I: If variant (0) is present, then no ICU is required, whereas if variant (1) is not present, then ICU is required. Criterion II: If variant (0) is present, then ICU is required, whereas if variant (1) is not present, then ICU is not required. THBD: thrombomodulin; CFH: complement factor H; CFHR1: complement factor H-related; ICU: intensive care unit.

rs2547438 (C3), rs2250656 (C3), rs1042580 (THBD), rs800292 (CFH) and rs414628 (CFHR1). Variant characteristics are shown in detail in Supplementary Table 1. Interestingly, Figure 3 shows that variants satisfying Criterion II are by 15% more present in male than in female patients and the reverse.

The number of variant combinations is defined as follows:

$$\text{Number of patterns} = \sum_{i=1}^{npv} \binom{npv}{i} = \sum_{i=1}^{npv} \frac{npv!}{i!(npv-i)!} \quad (1)$$

where npv is the number of variants and npv is the number of variant patterns.

In this study, $npv = 381$ number of variants were investigated for a 133 COVID-19 patient sample.

3.2 | Development of ANNs

The database used in this research comprised of 133 data sets, with each data set containing 7 input parameters (age, gender and 5 parameters indicating the presence or absence of each of the 5 critical variants). Figure 4 illustrates a statistical analysis of the selected input parameters.

The 133 data sets were divided into three separate sets. Specifically, 89 of the 133 (66.92%) data sets were designated as training data sets, 22 of the 133 (16.54%) as validation data sets, while the remaining 22 (16.54%) data sets were used as testing data sets. The assignment of patients to the various data sets was algorithmic and pseudo-random in nature, in order to prevent any potential bias. It was ensured that the composition of the final data sets in terms of input parameters was representative of the entire population.

For the training and the development of ANNs, it is necessary to define a number of parameters, related to their architecture and algorithmic implementation. Rather than relying on expertise or trial-and-error approaches, a number of alternatives were examined for these parameters, leading to the training and development of a plethora of possible ANN model implementations. Supplementary Table 2 presents all the parameters considered for the development of multiple alternative ANN architectures in the present study. ANNs with one or two hidden layers were developed and trained. The number of neurons per hidden layer ranged from 1 to 30, with an incremental step of 1. Ten different initial values of weights and biases were applied for each architecture. Two functions, the mean square error (MSE) and sum square error (SSE) functions, were used as cost functions, during the training and validation process. Ten functions were used as transfer or activation functions. Considering that the implementation of 10 different transfer function results in 1000 different ANNs, different dual combinations of the 10 transfer functions were investigated. During the training stage of ANNs, the 133 data sets were used with and without normalization techniques. When normalization of

Nr	rs	Gene	Position	Variant frequency based on Sex difference (%)		
				Female	Male	
100	rs2547438	C3	6718078			19,34
117	rs1042580	THBD	23027621			17,72
264	rs551397	CFH	196642072			15,29
265	rs800292	CFH	196642233			15,29
102	rs2250656	C3	6718534			15,14
332	rs4230	CFH	196801042			15,14
333	rs414628	CFHR1	196801078			15,14
377	rs7144	CD46	207967719			14,64
18	rs17030	C3	6677989			14,09
85	rs1042580	C3	6710782			14,09
317	rs442759	CFHR1	196794519			13,99
319	rs437504	CFHR1	196795741			13,99
320	rs424925	CFHR1	196795761			13,99
322	rs424722	CFHR1	196795849			13,99
324	rs399507	CFHR1	196796184			13,99
325	rs438781	CFHR1	196796240			13,99
327	rs408519	CFHR1	196799433			13,87
373	rs1962149	CD46	207956559			13,49
305	rs400344	CFHR3	196757392			13,22
244	rs685523	ADAMTS13	136310908			12,92
335	rs390679	CFHR1	196801190			12,84
337	rs401332	CFHR1	196801321			12,84
220	NA	ADAMTS13	136295288			12,79
24	rs2277984	C3	6679511			11,92
26	rs2277983	C3	6679563			11,92
41	rs11569511	C3	6690571			11,27
10	rs10506	CFD	863356			10,57
256	rs3094374	ADAMTS13	136323248			10,37
43	NA	C3	6690742			10,12
44	rs11569508	C3	6690746			10,12
45	rs3745567	C3	6690771			10,12
46	rs11569507	C3	6690805			10,12
50	rs11569492	C3	6693273			10,12
51	rs2241389	C3	6693317			10,12

FIGURE 3 Variant frequency distribution based on gender. The variants satisfying Criterion II are 15% more present in male than in female patients and the reverse (red and blue bars denote higher percentage variant presence in female and male patients respectively)

the data was conducted, the minmax normalization technique in the range [0.10, 0.90] and [-1.00, 1.00] Zscore normalization technique were implemented.

The Levenberg-Marquardt algorithm was used for the training of the ANNs. Overall, the alternative parameters, examined for the ANN development phase, resulted in the creation and evaluation of 75.000.000 different ANN models, which were ranked according to the prediction accuracy of the COVID-19 severity.

The architecture of the ANN with the highest prediction accuracy is presented in Figure 5; it consists of 2 hidden layers with 17 neurons in the first hidden layer and 11 neurons in the second hidden layer. The optimum ANN predicts successfully COVID-19 severity (not ICU, ICU or mortality) in 89.47% of the study population, which corresponds to 90.8% male and 86.96% female patients (Figure 6).

3.3 | Functional assessment of complement activation

THBD values were significantly increased in patients requiring ICU hospitalization compared to non-ICU patients (median 2.3, interquartile range [1.6] versus 1.4 [0.79] ng/ml, $p = 0.025$) and patients harbouring the rs1042580 (THBD) variant ($p = 0.032$). Similarly, C3a values were significantly increased in patients requiring ICU hospitalization compared to non-ICU patients (410 [14.1] versus 312 [19.1] ng/ml, $p = 0.035$) and trended towards a significant increase in patients harbouring one of the two critical C3 alleles (rs2547438 or rs2250656, $p = 0.092$). Despite increased levels of C5a in ICU patients, this difference did not reach statistical significance (72.1 [7.2] versus 43.4 [11.3] ng/ml, $p = 0.244$).

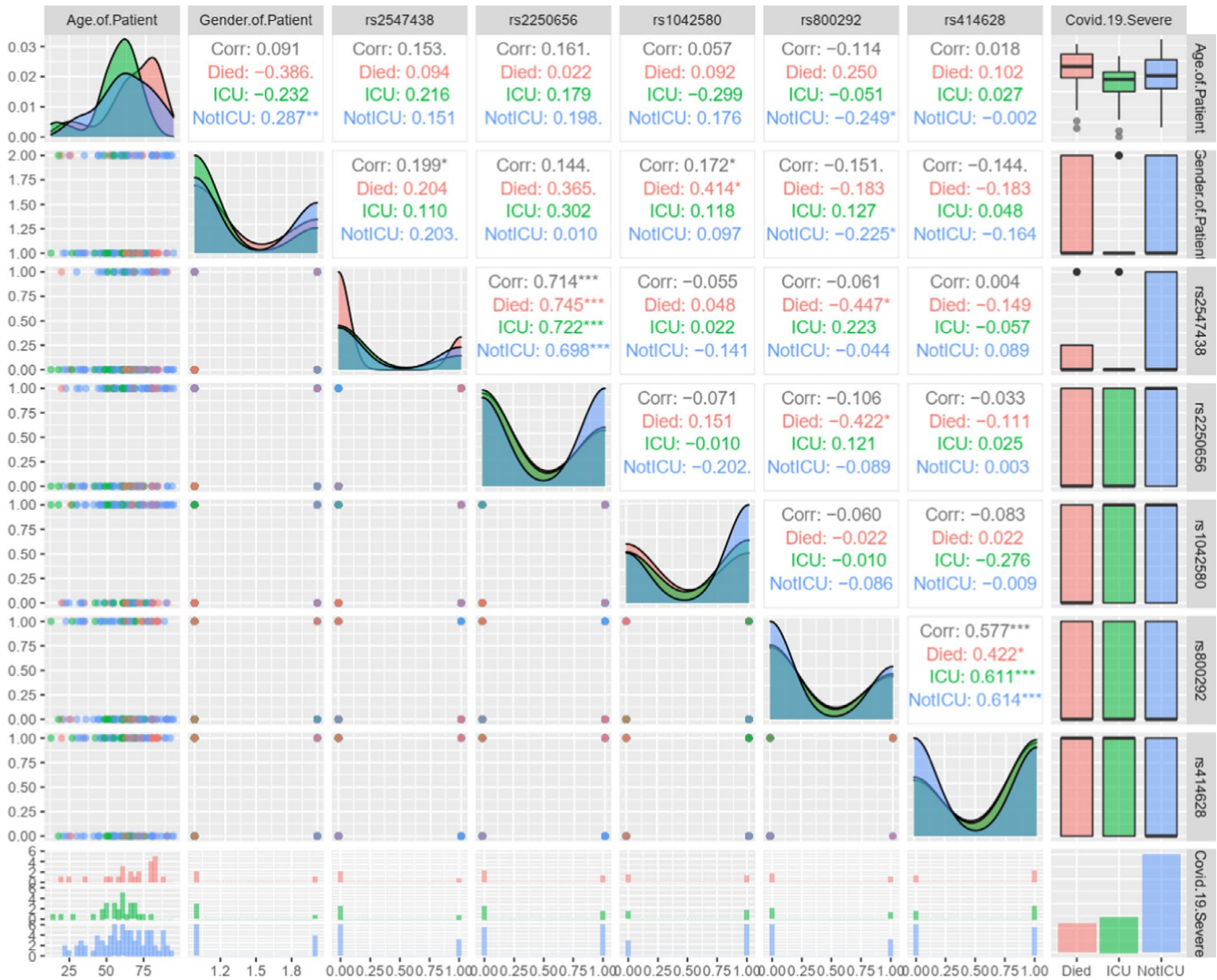


FIGURE 4 Scatterplot matrix of COVID-19 samples for the seven input parameters. Each input parameter is plotted against every other one, with different plot types drawn depending on the pair combination. Especially, the diagonal facets show the distributions of values for each parameter, ignoring all others. Along the right column, box plots show the distribution of each continuous parameter against the COVID-severe cases. The same is depicted in histogram form at the bottom row. Finally, dot plots show the relationships between parameter pairs

4 | DISCUSSION

We reveal for the first time an ANN able to accurately predict ICU hospitalization and death in COVID-19 patients, based on genetic variants in complement genes, age and gender. Importantly, we also confirm that genetic dysregulation is associated with an impaired complement phenotype.

Complement-related genetic variants have been investigated in limited COVID-19 studies.¹⁵⁻¹⁷ Valenti et al. have recently identified a variant (rs11385942) predisposing to severe COVID-19 that was associated with increased complement activation.¹⁵ Of note, this variant is not mapped in any complement gene coding loci, but it may be indirectly associated with enhanced complement activation during systemic inflammation and organ damage.⁴² Complement variants have been documented by genetic and transcriptional analysis, reporting 23 study-wide significant

single-nucleotide polymorphisms (SNPs) in 12 complement genes.¹⁶ An integrative analysis of molecular and functional data revealed four SNPs in three human complement genes (*C3*, *C4BPA*, *C5AR1*) that encode for missense polymorphic variants associated with viral susceptibility.⁴³

Considering that these analyses were not able to provide tools for disease severity prediction that could be helpful in clinical practice or a clinical trial setting, our group recently developed an algorithm identifying complement-related variants in *C3*, *CFH* and *THBD* that predict COVID-19 severity.¹⁷ Nevertheless, the prediction rate of this logical algorithm reached values above 80% only in patients not requiring ICU hospitalization and did not incorporate basic features associated with morbidity and mortality, such as age and gender. In the present study, we improved this logical algorithm in order to identify both ICU and non-ICU patients. Using the updated algorithm, we identified variants in complement-related genes (*C3*, *CFH*,

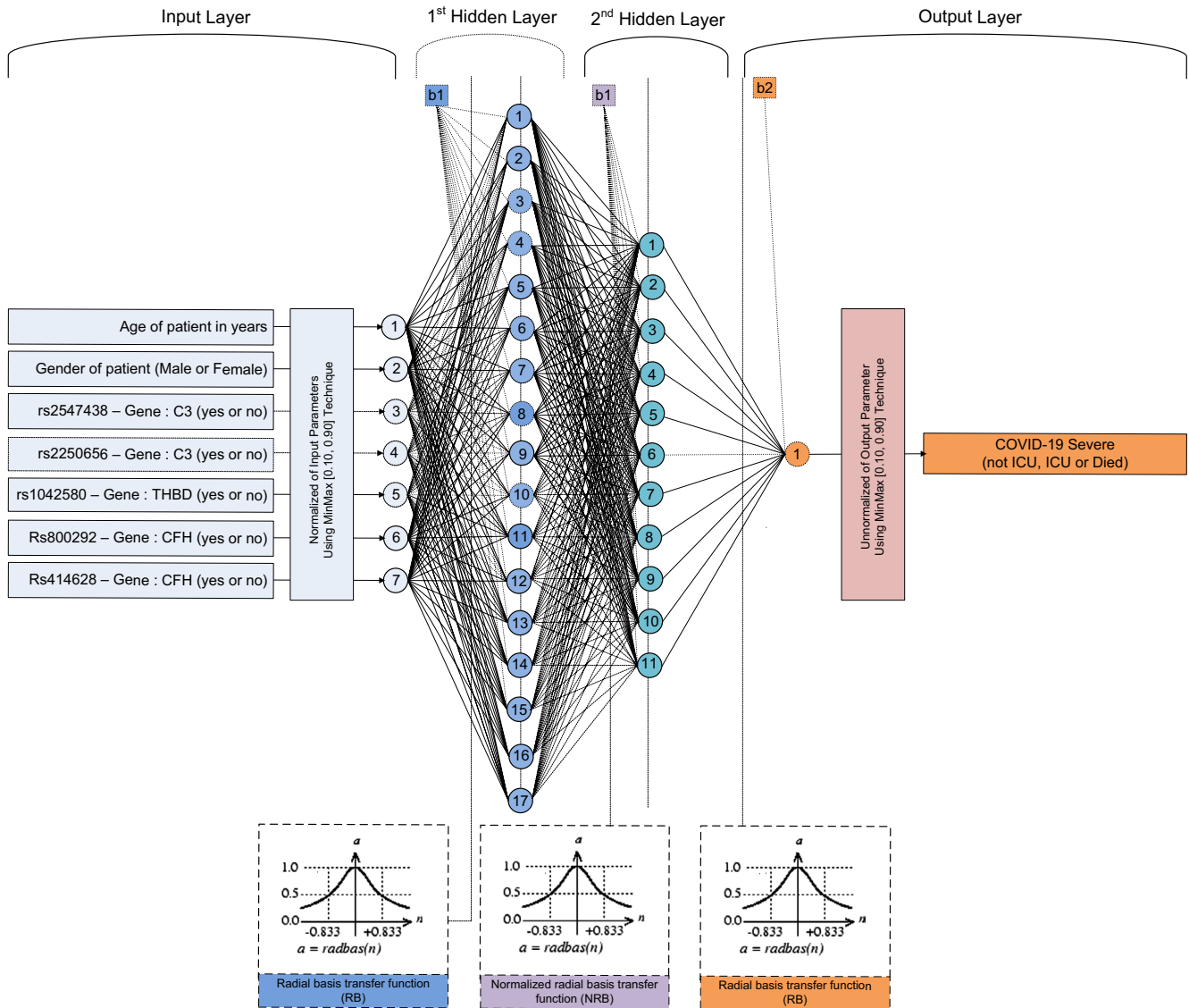


FIGURE 5 Architecture of the optimum ANN model. The input layer consists of seven parameters (input) which are age and gender of patient, and five crucial variants, while the output layer consists of a single parameter (output) which is the infection severity (requiring or not requiring intensive care unit, fatality). The two hidden layers consist of 17 and 11 neurons respectively. To the bottom of the schematic, the optimum transfer functions for the first hidden layer, second hidden layer and output layer are the radial basis transfer function, normalized radial basis transfer function and radial basis transfer function respectively

CFRH and THBD), among a targeted panel for complement-related genes, known to be dysregulated in complement-related disorder. Two of the five critical variants were also the ones that composed the initial algorithm. Based on the five critical variants derived from the updated algorithm, we further implemented an ANN incorporating age and gender. This tool is able to predict not only morbidity but also mortality in COVID-19 patients.

Indeed, a tool for early and individualized prediction of morbidity and mortality is a largely unmet clinical need not only in this COVID-19 pandemic but also in several clinical settings. In COVID-19, ANNs have been only used to predict outbreaks, and cases, deaths and hospital bed occupancy in the short term.^{44,45} However, in the era of precision medicine, individualized genetic data are able to further implement predictions of morbidity and mortality at a very early time point. This

prediction has an added value when it concerns early identification of patients that would potentially benefit from a safe and effective therapeutic approach.⁴⁶ In our case of a targeted complement-related genetic panel, individualized use of complement inhibitors could be of significant benefit in COVID-19 and other complement-related disorders. Therefore, the present ANN may serve as a prototype to further expand applications of precision medicine in this field.

The ANN presented in our study also incorporates the effect of age and gender on COVID-19 morbidity and mortality. Indeed, male COVID-19 patients show increased morbidity and mortality.⁴⁷ In addition, higher complement activity has been observed in healthy male versus female subjects.⁴⁸ In line with our previous analysis, the present study confirms differences in complement-related germline variants between male and female COVID-19 patients. Other factors

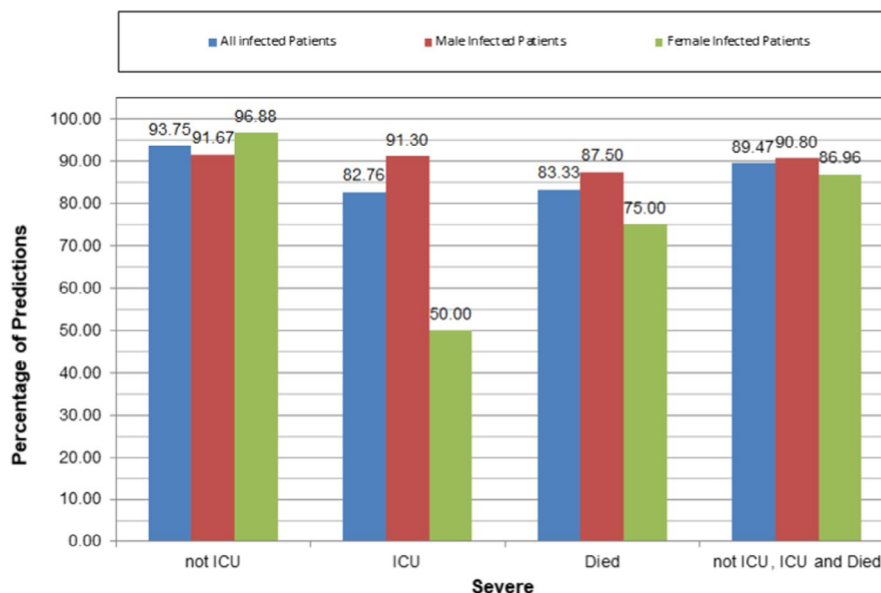


FIGURE 6 Percentage predictions of COVID-19 severity based on the optimum ANN model. The first group of three column bars represents the correct predictions of the ANN model for patients that did not require ICU treatment (achieving more than 93% successful predictions). The percentage of successful predictions is better for the female patients (more than 96%), while for male ones the respective percentage is 91.67%. The last group of three column bars represents the correct predictions of the ANN model for all cases of patient infection severity (requiring or not requiring intensive care unit and died). The percentage of successful predictions is more than 90% for male patients and close to 90% for female ones, while the combined percentage is 89.47%. The remaining groups of column bars report the respective results for the patients requiring ICU and those who have died

that may be also associated with differences in morbidity and mortality, such as differences in socioeconomic factors or comorbidities, are also incorporated in the algorithm by the addition of age and gender.

Our study is also the first to show an association between genotype and phenotype in complement dysregulation of COVID-19 patients. Despite difficulties of functional complement assessment in clinical laboratories,⁴¹ our results are in line with other functional studies in COVID-19 showing increased THBD and C3a levels in patients with severe disease.^{49,50} More than that, we also associated impaired levels with the presence of variants recognized as critical in the present study. These data highlight the potential of complement-related variants to predict complement dysregulation and presumably benefit from complement inhibition. The present study also focusses on alternative pathway variants, since this is directly activated by SARS-COV-2.⁷ In terms of complement inhibition, different pathways have been studied. In spite of encouraging results with different inhibitors,¹⁹⁻²² randomized studies have failed to show positive results yet. These could be attributed to several reasons, including study design, patient populations and choice of an appropriate inhibitor based on mechanistic evaluations.⁵¹ Therefore, broadening complement inhibition in COVID-19 requires well-designed studies, as well as appropriate patient selection that would overcome not only methodological but also practical issues of cost and availability during the pandemic. Our study has some limitations. Despite the prospective patient recruitment, our findings cannot be considered as definitive, indicating the need for future studies. In addition, our patient population is rather small, and therefore, the suggested ANN needs to be further validated in other real-world cohorts.

5 | CONCLUSIONS

An artificial neural network model was developed, trained and evaluated targeted to the prediction morbidity and mortality ratios of COVID-19 patients. A number of complement-related genetic variants, associated with severe COVID-19, were identified and used as inputs to the model, together with patient's age and gender. Using a sample of 133 patients, the developed ANN model was found capable to successfully predict COVID-19 severity in 89.47% of the study population.

In conclusion, germline complement-related genetic variants along with age and gender predict morbidity and mortality in COVID-19 patients. Given that vaccinations and viral mutations constantly change the landscape of COVID-19, a prediction tool based on such robust variables is of high importance in the future of this pandemic. Additionally, such a prediction tool is also expected to significantly contribute to better selection of patients that would benefit from targeted complement inhibition, considering the clinical phenotype associated with these variants. Last but not least, this novel approach of artificial intelligence paves the way for future application in additional clinical entities.

ACKNOWLEDGEMENT

We would like to thank the biologist Maria Spachidou for her technical assistance.

CONFLICTS OF INTEREST

E.G. has consulted for Omeros Cooperation and is supported by the ASH Global Research Award. Remaining authors declare no competing financial interest.

AUTHOR CONTRIBUTIONS

Panagiotis Asteris: Conceptualization (equal); Formal analysis (equal); Software (equal); Supervision (equal); Writing – original draft (equal). **Eleni Gavriilaki:** Conceptualization (equal); Data curation (equal); Funding acquisition (equal); Resources (equal); Writing – original draft (equal). **Tasoula Touloumenidou:** Investigation (equal); Methodology (equal). **Evaggelia-Evdoxia Koravou:** Methodology (equal). **Maria Koutra:** Methodology (equal). **Penelope Georgia Papayanni:** Methodology (equal). **Alexandros Pouleres:** Investigation (equal). **Vassiliki Karali:** Investigation (equal). **Minas Lemonis:** Software (equal). **Anna Mamou:** Software (equal). **Athanasia Skentou:** Software (equal). **Apostolia Papalexandri:** Resources (equal). **Christos Varelas:** Investigation (equal). **Fani Chatzopoulou:** Methodology (equal). **Maria Chatzidimitriou:** Methodology (equal). **Dimitrios Chatzidimitriou:** Validation (equal). **Anastasia Veleni:** Investigation (equal). **Evdoxia Rapti:** Methodology (equal). **Ioannis Kioumis:** Investigation (equal). **Evangelos Kaimakamis:** Investigation (equal). **Milly Bitzani:** Investigation (equal). **Dimitrios Boumpas:** Supervision (equal). **Argyris Tsantes:** Supervision (equal). **Damianos Sotiropoulos:** Supervision (equal). **Anastasia Papadopoulou:** Data curation (equal); Methodology (equal). **Ioannis Kalantzis:** Software (equal). **Lydia Vallianatou:** Software (equal). **Danial Armaghani:** Software (equal). **Liborio Cavaleri:** Software (equal). **Amir Gandomi:** Software (equal). **Mohsen Hajihassani:** Software (equal). **Mahdi Hasanipناه:** Resources (equal). **Paulo Lourenco:** Software (equal). **Pijush Samui:** Software (equal). **Jian Zhou:** Software (equal). **Ioanna Sakellari:** Supervision (equal); Writing – review & editing (equal). **Serena I. Valsami:** Methodology (equal); Writing – review & editing (equal). **Marianna Politou:** Project administration (equal); Writing – review & editing (equal). **Styliani Kokoris:** Validation (equal); Visualization (equal); Writing – review & editing (equal). **Achilles Anagnostopoulos:** Validation (equal); Visualization (equal); Writing – review & editing (equal).

DATA AVAILABILITY STATEMENT

Data will be immediately available upon communication.

ORCID

Eleni Gavriilaki  <https://orcid.org/0000-0002-8883-8208>

Marianna Politou  <https://orcid.org/0000-0002-7226-6452>

Styliani Kokoris  <https://orcid.org/0000-0002-1635-7309>

REFERENCES

- Kabanova A, Gavriilaki E, Pelzer BW, Brunetti L, Maiques-Diaz A. Effect of the COVID-19 pandemic on laboratory and clinical research: a testimony and a call to action from researchers. *Hemasphere*. 2020;4(6):e499.
- Rubin EJ, Longo DL. SARS-CoV-2 vaccination - an ounce (actually, much less) of prevention. *N Engl J Med*. 2020;383(27):2677-2678.
- Gavriilaki E, Brodsky RA. Severe COVID-19 infection and thrombotic microangiopathy: success does not come easily. *Br J Haematol*. 2020;189(6):e227-e230.
- Gavriilaki E, Anyfanti P, Gavriilaki M, Lazaridis A, Douma S, Gkaliagkousi E. Endothelial dysfunction in COVID-19: lessons learned from coronaviruses. *Curr Hypertens Rep*. 2020;22(9):63.
- Gao T, Hu M, Zhang X, et al. Highly pathogenic coronavirus N protein aggravates lung injury by MASP-2-mediated complement over-activation. *medRxiv*. 2020:2020.03.29.20041962.
- Magro C, Mulvey JJ, Berlin D, et al. Complement associated microvascular injury and thrombosis in the pathogenesis of severe COVID-19 infection: a report of five cases. *Transl Res*. 2020;220:1-13.
- Yu J, Yuan X, Chen H, Chaturvedi S, Braunstein EM, Brodsky RA. Direct activation of the alternative complement pathway by SARS-CoV-2 spike proteins is blocked by factor D inhibition. *Blood*. 2020;136(18):2080-2089.
- Holter JC, Pischke SE, de Boer E, et al. Systemic complement activation is associated with respiratory failure in COVID-19 hospitalized patients. *Proc Natl Acad Sci USA*. 2020;117(40):25018-25025.
- Busch MH, Timmermans S, Nagy M, et al. Neutrophils and contact activation of coagulation as potential drivers of Covid-19. *Circulation*. 2020;142(18):1787-1790.
- Skendros P, Mitsios A, Chrysanthopoulou A, et al. Complement and tissue factor-enriched neutrophil extracellular traps are key drivers in COVID-19 immunothrombosis. *J Clin Investig*. 2020;130(11):6151-6157.
- Yu J, Gerber GF, Chen H, et al. Complement dysregulation is associated with severe COVID-19 illness. *Haematologica*. 2021. [Online ahead of print]. doi:10.3324/haematol.2021.279155
- Satyam A, Tsokos MG, Brook OR, Hecht JL, Moulton VR, Tsokos GC. Activation of classical and alternative complement pathways in the pathogenesis of lung injury in COVID-19. *Clin Immunol*. 2021;226:108716.
- Cugno M, Meroni PL, Gualtierotti R, et al. Complement activation and endothelial perturbation parallel COVID-19 severity and activity. *J Autoimmun*. 2021;116:102560.
- Gavriilaki E, Brodsky RA. Complementopathies and precision medicine. *J Clin Investig*. 2020;130(5):2152-2163.
- Valenti L, Griffini S, Lamorte G, et al. Chromosome 3 cluster rs11385942 variant links complement activation with severe COVID-19. *J Autoimmun*. 2021;117:102595.
- Ramlall V, Thangaraj PM, Meydan C, et al. Immune complement and coagulation dysfunction in adverse outcomes of SARS-CoV-2 infection. *Nat Med*. 2020;26(10):1609-1615.
- Gavriilaki E, Asteris PG, Touloumenidou T, et al. Genetic justification of severe COVID-19 using a rigorous algorithm. *Clin Immunol*. 2021;226:108726.
- Risitano AM, Mastellos DC, Huber-Lang M, et al. Complement as a target in COVID-19? *Nat Rev Immunol*. 2020;20(6):343-344.
- Diurno F, Numis FG, Porta G, et al. Eculizumab treatment in patients with COVID-19: preliminary results from real life ASL Napoli 2 Nord experience. *Eur Rev Med Pharmacol Sci*. 2020;24(7):4040-4047.
- Mastaglio S, Ruggeri A, Risitano AM, et al. The first case of COVID-19 treated with the complement C3 inhibitor AMY-101. *Clin Immunol*. 2020;215:108450.
- Urwiler P, Moser S, Charitos P, et al. Treatment of COVID-19 with Conestat Alfa, a regulator of the complement, contact activation and Kallikrein-Kinin system. *Front Immunol*. 2020;11:2072.
- Rambaldi A, Gritti G, Mico MC, et al. Endothelial injury and thrombotic microangiopathy in COVID-19: Treatment with the lectin-pathway inhibitor narsoplimab. *Immunobiology*. 2020;9:152001.
- Mastellos DC, Pires da Silva BGP, Fonseca BAL, et al. Complement C3 vs C5 inhibition in severe COVID-19: early clinical findings reveal differential biological efficacy. *Clin Immunol*. 2020;220:108598.
- Wu Y, Giger ML, Doi K, Vyborny CJ, Schmidt RA, Metz CE. Artificial neural networks in mammography: application to decision making in the diagnosis of breast cancer. *Radiology*. 1993;187(1):81-87.

25. Kononenko I. Machine learning for medical diagnosis: history, state of the art and perspective. *Artif Intell Med.* 2001;23(1):89-109.
26. Mazurowski MA, Habas PA, Zurada JM, Lo JY, Baker JA, Tourassi GD. Training neural network classifiers for medical decision making: the effects of imbalanced datasets on classification performance. *Neural Netw.* 2008;21(2-3):427-436.
27. Ciresan DC, Giusti A, Gambardella LM, Schmidhuber J. Mitosis detection in breast cancer histology images with deep neural networks. *Med Image Comput Comput Assist Interv.* 2013;16(Pt 2):411-418.
28. Anthimopoulos M, Christodoulidis S, Ebner L, Christe A, Mougiakakou S. Lung pattern classification for interstitial lung diseases using a deep convolutional neural network. *IEEE Trans Med Imaging.* 2016;35(5):1207-1216.
29. Shen D, Wu G, Suk HI. Deep learning in medical image analysis. *Annu Rev Biomed Eng.* 2017;21(19):221-248.
30. Ribeiro AH, Ribeiro MH, Paixao GMM, et al. Automatic diagnosis of the 12-lead ECG using a deep neural network. *Nat Commun.* 2020;11(1):1760.
31. Gandomi AH, Fridline MM, Roke DA. Decision tree approach for soil liquefaction assessment. *Scientific World J.* 2013;2013:346285.
32. Asteris PG, Tsaris AK, Cavaleri L, et al. Prediction of the fundamental period of infilled RC frame structures using artificial neural networks. *Comput Intell Neurosci.* 2016;2016:5104907.
33. Niazkar HR, Niazkar M. Application of artificial neural networks to predict the COVID-19 outbreak. *Glob Health Res Policy.* 2020;5(1):50.
34. Upadhyay AK, Shukla S. Correlation study to identify the factors affecting COVID-19 case fatality rates in India. *Diabetes Metab Syndr.* 2021;15(3):993-999.
35. Kafieh R, Arian R, Saeedzadeh N, et al. COVID-19 in Iran: forecasting pandemic using deep learning. *Comput Math Methods Med.* 2021;2021:6927985.
36. Rahimi I, Gandomi AH, Asteris PG, Chen F. Analysis and Prediction of COVID-19 Using SIR, SEIQR, and Machine Learning Models: Australia, Italy, and UK Cases. *Information.* 2021;12(3):109.
37. Niazkar HR, Niazkar M. Application of artificial neural networks to predict the COVID-19 outbreak. *Glob Health Res Policy.* 2020;5(1):50.
38. Çaparoglu ÖF, Ok Y, Tutam M. To restrict or not to restrict? Use of artificial neural network to evaluate the effectiveness of mitigation policies: a case study of Turkey. *Chaos, Solitons Fractals.* 2021;151:111246.
39. Gómez-Herrera S, Sartori Jeunon Gontijo E, Enríquez-Delgado SM, Rosa AH. Distinct weather conditions and human mobility impacts on the SARS-CoV-2 outbreak in Colombia: Application of an artificial neural network approach. *Int J Hyg Environ Health.* 2021;238:113833.
40. Asteris PG, Douvika MG, Karamani CA, et al. A novel heuristic algorithm for the modeling and risk assessment of the COVID-19 pandemic phenomenon. *Comput Model Eng Sci.* 2020;125(2):815-828. 10.32604/cmesci.2020.013280
41. Gavriilaki E, Anagnostopoulos A, Mastellos DC. Complement in thrombotic microangiopathies: unraveling Ariadne's thread into the labyrinth of complement therapeutics. *Front Immunol.* 2019;10:337.
42. Bianco C, Baselli G, Malvestiti F, et al. Genetic insight into COVID-19-related liver injury. *Liver Int.* 2021;41(1):227-229.
43. Tsiftoglou SA. SARS-CoV-2 associated Complement genetic variants possibly deregulate the activation of the Alternative pathway affecting the severity of infection. *Mol Immunol.* 2021;135:421-425.
44. Braga MB, Fernandes RDS, Souza GN Jr, et al. Artificial neural networks for short-term forecasting of cases, deaths, and hospital beds occupancy in the COVID-19 pandemic at the Brazilian Amazon. *PLoS One.* 2021;16(3):e0248161.
45. Mollalo A, Rivera KM, Vahedi B. Artificial neural network modeling of novel coronavirus (COVID-19) incidence rates across the continental United States. *Int J Environ Res Public Health.* 2020;17(12):4204.
46. Gavriilaki E, Peffault de Latour R, Risitano AM. Advancing therapeutic complement inhibition in hematologic diseases: PNH and beyond. *Blood.* 2021. [Online ahead of print]. doi:10.1182/blood.2021012860
47. Peckham H, de Gruijter NM, Raine C, et al. Male sex identified by global COVID-19 meta-analysis as a risk factor for death and ICU admission. *Nat Commun.* 2020;11(1):6317.
48. Gaya da Costa M, Poppelaars F, van Kooten C, et al. Age and sex-associated changes of complement activity and complement levels in a healthy Caucasian population. *Front Immunol.* 2018;9:2664.
49. Goshua G, Pine AB, Meizlish ML, et al. Endotheliopathy in COVID-19-associated coagulopathy: evidence from a single-centre, cross-sectional study. *Lancet Haematol.* 2020;7(8):e575-e582.
50. de Nooijer AH, Grondman I, Janssen NAF, et al. Complement activation in the disease course of coronavirus disease 2019 and its effects on clinical outcomes. *J Infect Dis.* 2021;223(2):214-224.
51. Mastellos DC, Skendros P, Calado RT, Risitano AM, Lambris JD. Efficacy matters: broadening complement inhibition in COVID-19. *Lancet Rheumatol.* 2021;3(2):e95.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Asteris PG, Gavriilaki E, Touloumenidou T, et al. Genetic prediction of ICU hospitalization and mortality in COVID-19 patients using artificial neural networks. *J Cell Mol Med.* 2022;26:1445-1455. doi:[10.1111/jcmm.17098](https://doi.org/10.1111/jcmm.17098)