Predicting the severity of disease progression in COVID-19 at the individual and population level: A mathematical model

Narendra Chirmule[1], Pradip Nair[2], Bela Desai[3],

Ravindra Khare[1], Vivek Nerurkar[4], Amitabh Gaur[5],

[1] SymphonyTech Biologics, Philadelphia, Pennsylvania, USA; [2] Biocon, Bangalore, Karnataka, India; [3] NanoCellect Biomedical, Inc., San Diego, California, USA, [4] Department of Tropical Medicine, Medical Microbiology and Pharmacology, John A. Burns School of Medicine, University of Hawaii at Manoa, Honolulu, Hawaii, USA, [5] Innovative Assay Solutions LLC, San Diego, California, USA.

Address Correspondence to:

Narendra Chirmule at Narendra.Chirmule@symphonytech.com or

Amitabh Gaur at agaursd@gmail.com

Key words: innate, interferon, cytokine-storm, lymphopenia, neutralizing-antibodies, viral factors, modeling, prediction

18 Conflict of Interest Statement.

19 Narendra Chirmule and Ravindra Khare are employed by SymphonyTech Biologics Inc, and
20 own stock in the company; the opinions of Pradip Nair do not represent those of Biocon; Bela
21 Desai, Vivek Nerurkar and Amitabh Gaur do not have any conflict of interest.

22

23

ABSTRACT

The impact of COVID-19 disease on health and economy has been global, and the magnitude of devastation is unparalleled in modern history. Any potential course of action to manage this complex disease requires the systematic and efficient analysis of data that can delineate the underlying pathogenesis. We have developed a mathematical model of disease progression to predict the clinical outcome, utilizing a set of causal factors known to contribute to COVID-19 pathology such as age, comorbidities, and certain viral and immunological parameters. Viral load and selected indicators of a dysfunctional immune response, such as cytokines IL-6 and IFNα, which contribute to the cytokine storm and fever, parameters of inflammation d-dimer and ferritin, aberrations in lymphocyte number, lymphopenia, and neutralizing antibodies were included for the analysis. The model provides a framework to unravel the multi-factorial complexities of the immune response manifested in SARS-CoV-2 infected individuals. Further, this model can be valuable to predict clinical outcome at an individual level, and to develop strategies for allocating appropriate resources to mitigate severe cases at a population level.

INTRODUCTION

The COVID-19 pandemic caused by infection with SARS-CoV-2 was officially announced in March 2020 by the CDC and WHO [1, 2]. As of this publication, more than 100 million infections and over 2.6 million deaths have been reported worldwide. Majority of the subjects have asymptomatic infections. The rate of fatality is disproportionately high in the elderly and patients with comorbidities such as diabetes, cardiac disease, and kidney disease [3, 4]. The consequences of the pandemic are fraught with potential loss of lives, social and economic distress, and the uncertainty of disease progression because of variable individual pathogenesis.

A unique and dysregulated immune response has been shown to be a hallmark of COVID-19 [5-9]. **Figure 1** schematically depicts the cascade of events that contribute to the progression of disease. Mathematical models have been utilized by several investigators to understand the mechanisms of disease pathogenesis, immune pathways involved and course of viral infections [10, 11]. In this article, we have proposed a predictive model that utilizes the levels of clinical and laboratory parameters to determine the severity of clinical outcomes ranging from asymptomatic to mild, moderate, severe, and critical disease states. The proposed model can be useful to predict clinical outcome at the individual-level and develop efficient and effective treatment strategies to manage public health challenges at the population-level.

The questions the model attempts to answer are: i) At an individual level, what is the probability of an individual infected with SARS-CoV-2, given the clinical signs and laboratory values on various days, likely to progress to severe disease, and ii) At a population level, what are the prioritized clinical and laboratory parameters that are most likely to contribute to progression to severe disease. We have used a multiple regression based model to predict

severity of the outcome of COVID-19. To evaluate the combinatorics that are not observed in the sample, we have applied resampling methods based on Monte Carlo simulation.

METHODS

Development of a Simulated Dataset

A simulated data set of 45 individual subjects was created with 15 subjects assumed to be asymptomatic, 15 with moderate disease, and 15 with severe COVID-19 [12, 13]. The simulated values for the viral and immune parameters were generated using data from clinical reports published in the last year for each of the selected parameters. **Table 1** provides the ranges and the related references for the values for all parameters and **Figure 2** shows the box-and-whisker plots for the distribution of the values for each parameter.

Data Modeling

We have applied Multiple Linear Regression approach to the simulated data set for COVID-19 subjects generated and analyzed to understand the impact of each of the parameters on the outcome of disease severity. We chose a multiple regression model since both, the outcomes and predictors, were numeric. We used regression models to establish a predictive transfer function and evaluated significance of results. In this model, the relationship between independent variables (**x1, x2…x_n**) with dependent variable (**y**) can be visualized by the equation, **y=f (x1, x2…x_n).** This is the transfer function that is derived through analysis. The validity of the model was established using 'Goodness of Fit' and ANOVA. The statistical significance of the model was tested by evaluating residuals and F Ratio in one-way ANOVA, based on the criteria of p

<0.05 and goodness-of-fit with Adj. RSq >90%. The assumption for this analysis was that each of the parameters was independent. However, in cases where factual patient datasets will be subjected to this type of analysis, there may be multi-co-linearity within the parameters that should be rationalized using dimensionality reduction methods [14, 15].

Since the model may not exhibit multiple combination of parameters in the limited dataset of 45 subjects, we have used resampling methods using Monte Carlo simulation to achieve a better density of combinations. The simulation was applied for resampling of the transfer function with 2000 runs, where a convergence was achieved after multiple runs. The simulation was performed in order to understand the impact of possible parameter combinations on clinical outcomes. Monte Carlo simulation uses random variates from selected range of values to model the impact of progression of events leading to outcomes.

Data analysis using training and testing data sets

Model building involved partitioning the data set into 'training' and 'testing' sets. We apportioned 70% of the data to train the model and used the remaining 30% to test the model, using random selection algorithms. Following development of the model, we analyzed a set of test data to compare predicted versus observed results to validate the model. The regression model generated a prediction formula as follows:

Outcome = -36.898 - 0.020 AGE + 0.894 COMORBID - 0.048 Viral Load - 0.004 IFNα + 0.444 Fever - 0.003 IL6 + 0.271 D-Dimer + 0.000 Ferritin - 0.000 Lymphocyte Count - 0.037 O_2 saturation - 2.57034e-006 NAB

The linear coefficients of the prediction equation determined the weights of each parameter to predict the clinical outcome.

Estimation of the coefficients of input parameters

The modeling approach was based on utilizing clinical and laboratory parameters to fit the regression models. Since direct comparison of regression coefficients was not necessary, and interactions in factors were not considered on account of assumption of independence of factors, we chose to leave the factor-data in the original scales.

Rationale for the parameters included in the analysis

The input parameters selected for this model, which requires cause (clinical and laboratory parameters) and effect (clinical outcome) relationships, were based on the data reported in recent scientific publications. **Figure 1** shows the schematic representation of the stage of disease progression and parameters associated with the increasing severity of diseases. The following parameters were chosen:

1. **Comorbidities:** Though the precise mechanism(s) of disease progression in patients with comorbidities has yet to be elucidated, pre-existing conditions such as diabetes, cancer, neurological, cardiac and lung and kidney disease have been reported to contribute towards severity of COVID-19 [16, 17]. The simulated data for comorbidity was generated using an arbitrary range of 1 to 4, where 1 represented a healthy individual and 4 represented an individual with a severe co-morbidity.

2. **Age**: A range of 18 to 100 years was utilized for generating the mock data set. The assumption used in generating the data was that disease progression was directly proportional to age [17]. Reports of certain rare pathogenic conditions in children, e.g.,

Kawasaki disease [18, 19], have not been considered in the current model. Reports indicate that majority of children infected with SARS-CoV-2 are asymptomatic [19].

3. **Viral load**: SARS-CoV-2 infects individuals through the nasopharyngeal pathway. This infection is the cause of all subsequent effects. Viral load is measured by reverse-transcriptase quantitative PCR (RT-qPCR), which detects viral RNA from nasopharyngeal swabs [20]. The test relies on multiple cycles of RNA amplification to produce detectable amount of RNA in the mixed nucleic acid sample, reflected in the Cycle-time (Ct) value, which is defined as the number of cycles necessary to detect the virus. A Ct value of less than 20 is considered a high viral load while a Ct value of 35 and higher indicates a lower level or near absence of viral infection [20]. Viral load in patients is dependent on various factors, including number of ACE2 and TMPRSS2 receptors [21, 22], comorbidities, cytokines, number of viral particles at infection, and the overall immune health status of the patients [23-26]. Viral loads have been demonstrated to have a direct correlation with severity of disease and mortality in COVID-19 [27, 28].

4. **Cytokine Storm**: High viral loads evoke defensive mechanisms that can induce inflammation leading to a dysregulated innate immune response that could result in a cytokine storm characterized by fever-inducing levels of cytokines such as IL6, IFNα, IL1β and CXCL-10 [27, 29-33]. CXCL-10, interestingly was also found to be indicative of severe outcomes in patients affected by the SARS CoV1 outbreak in 2002 [34]. Cytokine storm has been implicated in contributing to pulmonary immunopathology, leading to severe clinical disease and mortality. In this model, we have included levels of IFNα and IL6 obtained from the published data.

5. **Systemic Inflammation**: Laboratory based parameters indicating inflammation in the serum, such as D-Dimer and Ferritin, have been shown to lead to a reduction in blood oxygen saturation levels, reflecting inadequate oxygenation in the lungs [35, 36].

6. **Lymphopenia:** Viral infection can lead to marked lymphopenia that can affect both CD4+ and CD8+ T cells [3, 28, 36]. Lymphopenia, reflected by significantly reduced CD4 and CD8 T cells in peripheral blood, is likely due to sequestration and cell death and reflected by significantly reduced CD4 and CD8 T cells in peripheral blood, has been reported in moderate and severe COVID-19 patients. In addition, antigen specific CD8 Cytotoxic T lymphocyte (CTL) responses have been detected approximately a week following viral infection, and the magnitude of the response was observed to have protective or damaging effects [37].

7. **Neutralizing antibodies**: Neutralizing antibodies bind to specific surface receptors on infectious agents such as viruses and toxins, reducing or eliminating their ability to exert harmful effects on cells. SARS-CoV-2 infected individuals generate a robust and long-lasting neutralizing antibody response, and plasma from convalescent COVID-19 patients has been used for treatment of severe disease with some success [38, 39]. It has recently been reported that neutralizing antibodies to SARS-CoV-2 can predict severity and survival, with higher titers being associated with severe disease in some instances [40].

.

RESULTS

We evaluated multiple approaches to develop mathematical models using parameters that can predict the progression of disease. Candidate parameters were selected from mechanistic understanding of the process of pathogenesis of COVID-19 to evaluate their possible impact on the clinical outcome. Regression models utilize data to build predictive models. Hypotheses are examined and confirmed with pre-determined statistical confidence and inferential power. These models incorporate all the experimental variability in the data set. Since the models contained numeric factors and numeric ordinal outcomes, we utilized methods of Multiple Linear Regression [41]. In this approach, we used the simulated data set from COVID-19 affected subjects, organized, and analyzed it to understand the variability of each of the parameters.

Regression modeling approach

The data set was parsed into training and testing partitions using methods of randomization. The validity of the model was based on goodness-of-fit of R Sq. > 90% and ANOVA, [p value <0.05] and a consequent F Ratio **(Table 2a and 2b).** These statistical results confirmed acceptable degree of predictability of the model.

Following this multiple-regression analysis, we conducted 2,000 bootstrap samplings using the predicted coefficients and random variates from chosen intervals of parameters. The assumption for this analysis was that each of the parameters were independent variables. The coefficients of each parameter were determined by using multiple regression analyses, which is the multiplier to the parameter value in a linear regression equation. The inclusion of all the variables in analysis ensures their contribution to the model [41]. However, analysts applying this model in the future may, at their judgment, evaluate statistical significance of regression coefficients. Parameters

that are not significant maybe excluded using step wise regression. In our analysis, results based on training dataset predictors matched with those from the test dataset confirming an acceptable degree of predictability of the model. We invite the readers of this article to contact us to analyze the predictive potential of the model using their data.

Monte Carlo simulation

To determine the factors that contribute to the clinical outcome at the population level, Monte Carlo simulation was performed on a sample set of laboratory and clinical parameters covering the full range, from asymptomatic to severe disease, of outcomes (stages 1-7 [Figure 2]) [12, 13]. The histogram and cumulative data (**Figure 3a)** show the distribution of asymptomatic to severe outcomes. The Tornado chart (**Figure 3b**) shows the sensitivity of parameter to the outcome in the selected range.

The predictive model

Based on the correlation coefficient of the parameters and the outcome from the training data set, we developed a model using the prediction equation. **Table 4** shows the process of predicting the outcome. When the numerical values of the individual parameters for each patient are entered into the columns, the model predicts the outcome. The validation of the model will require data from patients and clinical trials. The goal of this exercise was to develop a model that can be used to predict the outcome in a large number of patients.

DISCUSSION

We have evaluated multiple regression analysis for mathematically modeling the course of COVID-19 to predict clinical outcome. The premise of this model is that quantitatively measured clinical and laboratory parameters involved in the pathogenesis of disease progression can be mathematically mapped to a multiple-regression model. COVID-19 is initiated by infection of the subject with SARS-CoV-2 with subsequent replication in the epithelial cells of the lung. The factors that contribute to the viral load include number of cells that express the ACE2 and other receptors, and inflammatory cytokines. Comorbidities contribute towards a more serious disease progression. Virus infection of antigen presenting cells, such as dendritic cells, macrophages, and other cell types including endothelial cells, results in activation of biochemical signals, which lead to secretion of a battery of cytokines that include IL1β and IL-6. The viral infection as well as inflammatory cytokines cause fever and an increase in serum inflammatory factors such as D-Dimer and Ferritin. Induction of an inflammatory response contributes to reduction of the total numbers of lymphocytes from circulation. The inflammation results in a loss of lung function (e.g., reduction in blood-oxygen levels), cardiac function (blood pressure) and can culminate in multi-organ failure.

Subjects with a normal immune response can generally mount an adequate innate and adaptive response to the virus. These individuals clear the virus by generating adaptive T cell responses and neutralizing antibodies. Subjects with comorbid conditions can have compromised immune function which could result in dysfunctional activation of inflammatory responses, leading to worse clinical outcomes.

Selection of the parameters that were included in the model building process was influenced by their perceived significance from current research reports. This list of factors is by no means complete and it is expected that in due course a more comprehensive list will emerge. This report provides a basis for creating a tool, independent of the number and type of parameters, that could find utility in predicting the disease outcome using those parameters.

Viral Load. Association of viral load and progression of diseases has been reported for several viral infections [42-44]. Viral load in COVID-19 is measured by qRT-PCR of SARS-CoV-2 using primers for the spike gene [43]. The correlation of high viral load with severity of disease progression has been extensively demonstrated. The systemic dissemination of the virus has been associated with expression of the ACE2 receptor on endothelial cells [21]]. Comorbid conditions could enhance the expression of receptors and enable distribution of virus, thereby enhancing the viral load, which can result in progression of disease.

IFNα. The critical role of Type I interferons in innate and adaptive immunity, leading to both protective and pathogenic responses, has been reported in the case of several viral and bacterial infections [45]. SARS-CoV-2 infection has been shown to result in a diverse range of effects on Type I immune responses. Most patients elicit a strong IFNα response along with a battery of inflammatory cytokines, some of which progress to a cytokine storm [46, 47]. Specific blocking of the type I mediated signal transduction by various proteins of SARS-CoV-2 has been demonstrated [48]. A remarkably high proportion of male subjects experiencing severe or critical COVID-19 disease expressed an inability to produce sufficient levels of IFNα due to various types of errors in the IFN genes. Curiously, majority of the male subjects possessed circulating IFNα autoantibodies that had the ability to neutralize the endogenously produced cytokine,

thereby effectively reducing the available IFNα. The discovery of these two mechanisms for lowering IFNα levels underscores its relevance in controlling the progression of disease in individuals infected with the SARS-CoV-2 [49].

D-Dimer. D-Dimer is routinely measured in clinical situations because its levels correlate with serious underlying conditions including venous thromboembolism, cancer and sepsis [48]. In the case of COVID-19 patients, introduction of the virus brings about infection-induced inflammatory alterations leading to coagulopathy. Lungs being the target of SARS-CoV-2, acute injury to the lung as well as multi organ failure have been caused by the virus-induced cascade of the inflammatory pathway. In an early study on 41 COVID-19 patients, those with severe disease had higher levels of D-Dimer along with high levels of IL-8, TNFα and IL-2R [31]. Male patients were found to have higher levels of IL-6, IL-2R, Ferritin and other markers of inflammation compared to female. High levels of IL-6 showed a statistically significant correlation with severe disease in a retrospective study as well [27]. One can hypothesize that such patients would likely benefit from anticoagulation therapy.

Ferritin. A high level of ferritin, measure of stored iron, was found to be associated with severe disease in COVID-19 patients and was linked to high fatality rates in a 72 patient prospective study [33, 50, 51]. In another study on 39 patients, those with mild COVID-19 symptoms had lower levels of ferritin while those with moderate or severe symptoms expressed higher levels of ferritin [50].

Lymphopenia. Loss of lymphocytes after viral infections has been associated with severe disease. The mechanisms involved in lymphodepletion can been implicated to be due to cell death, cytokine storm and/or redistribution of lymphocyte populations [3, 33, 37]. In this model,

we have utilized lymphopenia as a measure of severity of disease progression. Loss of immune function could result in several potential mechanisms of pathogenesis including autoimmunity, hyperactivation, increased susceptibility to infections and organ dysfunction.

Neutralizing Antibodies. Induction of neutralizing antibodies directed to the receptor-binding domain of the spike protein is critical for restricting entry of the virus into the cells and has been one of the central tenets of a protective immune response. In this model, we have used a range of IgG titers to spike protein for the simulated data set [52]. However, the role of neutralizing antibodies induced in a large proportion of subjects following natural infection is still being studied [53]. Some subjects do not elicit strong antibody responses. Sub-optimal levels of antibodies may catalyze generation of virus mutants [54]. Neutralizing antibodies to the virus have generally not correlated with reduced severity of disease in the primary infection. In addition, it will be interesting to decipher the role of pre-existing antibodies reported recently in the modulation of disease and its impact on vaccination regimens. Thus, the mechanisms involved in the induction of antibodies, the repertoire and diversity of responses, and effects on protection versus progression, remains to be clearly established.

The predictive model can have multiple applications, such as forecasting the percentage of the population that will progress to severe disease in each geography, enabling logistics planning for hospital beds, health care providers and personal-protective safety equipment. Analysis of the coefficient of correlations of parameters with outcome of disease may provide clues to a better understanding of the mechanism of action of disease pathogenesis. The model can predict the probability of disease progression at an individual level, based on parameter data, and can be used to understand the effect and impact of therapeutic interventions. The predictive

model can be utilized to analyze large amounts of data to develop algorithms for personalized treatment regimens.

In summary, we have developed a probabilistic model that can be utilized to predict progression of disease following infection with SARS-CoV-2. This model was developed using simulated data based on published levels of COVID-19 related clinical and laboratory parameters and provides an approach to predicting the outcome of disease. Validation of the model will require existing data and the clinical outcomes of patients. Prediction of disease progression can be highly valuable at an individual as well as population level.

TABLES LEGENDS:

Table 1. Ranges of values for the parameters used for developing the simulated dataset for the mathematical model. The range of comorbidities was assigned arbitrary nominal value between 1 to 4, with 1 being healthy, and 4 having multiple health-conditions (e.g., diabetes, cancer etc.). The age ranges in the model were 18-100 years.

Table 2A. Statistical analysis of coefficients for each parameter based on the multiple regression analysis. The Table shows the regression statistics. The statistical terms are: DF (degrees of freedom); SeqSS (Sequential Sum Square); AdjSS (Adjusted Sum Square), AdjMS (Adjusted Mean Squares), F ratio, p value.

Table 2B. **Coefficient and standard error for parameters**. The coefficients for each parameter were determined by using multiple regression analyses, which is the multiplier to the parameter value in a linear regression equation. The values of the coefficients of each of the parameters are shown using ANOVA. The table shows the standard error, t and p values. The p value denotes statistical significance to the outcome.

Table 3. The Ranges for the Monte Carlo Simulation.

Table 4. The Prediction of outcome based on observed and predicted values. The values of the parameters for each of the seven subjects are entered in columns, upon running of the model. The predicted values are calculated in numerical values in a range of 1-7, with 1 being asymptomatic, and 7 most severe.

FIGURE LEGENDS:

Figure 1. Schematic representation of the progression of disease. The width of the triangles denotes increase in levels of viral load (purple), cytokine storm (blue), and anti-inflammatory symptoms (green); blue arrows denote T and B cell responses.

Figure 2. Box-and-whisker plots of the simulated data. The figures show the visual representation of the summary, which includes median (Q2/50th percentile); first quartile (Q1/25th percentile); third quartile (Q3/75h percentile); interquartile range in whiskers, maximum and outliers.

Figure 3a. **Histogram from Monte Carlo Simulation**. 2,000 bootstrap samplings were generated using the predicted coefficients from the linear regression analysis, from the intervals of parameters. The minimum and maximum values for each of the parameters was set to the levels. The distribution of the severity of outcome is in this frequency histogram. The values on the x axis denote the disease severity, and y axis denote frequency of the population in each level of clinical outcome.

Figure 3b. The Tornado Chart. The tornado chart shows the influence of each of the parameters on the outcome. The positive values correlate positively towards the severity of disease, and negative values towards asymptomatic disease.

ACKNOWLEDGEMENTS

VRN is partly supported by a grant (P30GM114737) from the Centers of Biomedical Research Excellence, National Institute of General Medicine, National Institutes of Health.

350 **REFERENCES**

351

352 1. CDC. Coronavirus 2020 [31 October, 2020]. Available from:
353 https://www.cdc.gov/coronavirus/2019-ncov/index.html.

354 2. WHO. Coronavirus 2020 [31 October, 2020]. Available from: https://covid19.who.int/.

355 3. Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of patients infected
356 with 2019 novel coronavirus in Wuhan, China. Lancet. 2020;395(10223):497-506. Epub
357 2020/01/28. doi: 10.1016/s0140-6736(20)30183-5. PubMed PMID: 31986264; PubMed Central
358 PMCID: PMCPMC7159299.

359 4. Kim GU, Kim MJ, Ra SH, Lee J, Bae S, Jung J, et al. Clinical characteristics of
360 asymptomatic and symptomatic patients with mild COVID-19. Clin Microbiol Infect. 2020.
361 Epub 2020/05/04. doi: 10.1016/j.cmi.2020.04.040. PubMed PMID: 32360780.

362 5. Kuri-Cervantes L, Pampena MB, Meng W, Rosenfeld AM, Ittner CAG, Weisman AR, et
363 al. Comprehensive mapping of immune perturbations associated with severe COVID-19. Sci
364 Immunol. 2020;5(49). Epub 2020/07/17. doi: 10.1126/sciimmunol.abd7114. PubMed PMID:
365 32669287; PubMed Central PMCID: PMCPMC7402634.

366 6. Kuri-Cervantes L, Pampena MB, Meng W, Rosenfeld AM, Ittner CAG, Weisman AR, et
367 al. Immunologic perturbations in severe COVID-19/SARS-CoV-2 infection. bioRxiv. 2020.
368 Epub 2020/06/09. doi: 10.1101/2020.05.18.101717. PubMed PMID: 32511394; PubMed Central
369 PMCID: PMCPMC7263541.

370 7. Vabret N, Britton GJ, Gruber C, Hegde S, Kim J, Kuksin M, et al. Immunology of
371 COVID-19: Current State of the Science. Immunity. 2020. Epub 2020/06/09. doi:
372 10.1016/j.immuni.2020.05.002. PubMed PMID: 32505227; PubMed Central PMCID:

PMCPMC7200337 Sciences, Primevax, Novartis, Array BioPharma, Roche, Avidea, Boeringer Ingelheim, Rome Therapeutics, Roswell Park, and the Parker Institute for Cancer Immunotherapy. N.B. receives research support from the Parker Insitute, Novocure, Celldex, Genentech, Oncovir, and Regeneron. M.M. serves as an advisor/board member for Celsius, Pionyr, Compugen, Myeloids and Innate pharma and ad hoc for Takeda. M.M. receives research support from Regeneron, Takeda, and Genentech. A.M. has equity in Gilead Sciences and Regeneron Pharmaceuticals.

8. Zheng HY, Zhang M, Yang CX, Zhang N, Wang XC, Yang XP, et al. Elevated exhaustion levels and reduced functional diversity of T cells in peripheral blood may predict severe progression in COVID-19 patients. Cell Mol Immunol. 2020;17(5):541-3. Epub 2020/03/24. doi: 10.1038/s41423-020-0401-3. PubMed PMID: 32203186; PubMed Central PMCID: PMCPMC7091621.

9. Rydyznski Moderbacher C, Ramirez SI, Dan JM, Grifoni A, Hastie KM, Weiskopf D, et al. Antigen-Specific Adaptive Immunity to SARS-CoV-2 in Acute COVID-19 and Associations with Age and Disease Severity. Cell. 2020;183(4):996-1012.e19. Epub 2020/10/05. doi: 10.1016/j.cell.2020.09.038. PubMed PMID: 33010815; PubMed Central PMCID: PMCPMC7494270.

10. Norton KA, Gong C, Jamalian S, Popel AS. Multiscale Agent-Based and Hybrid Modeling of the Tumor Immune Microenvironment. Processes (Basel). 2019;7(1). Epub 2019/02/01. doi: 10.3390/pr7010037. PubMed PMID: 30701168; PubMed Central PMCID: PMCPMC6349239.

11. Foy BH, Gonçalves BP, Higgins JM. Unraveling Disease Pathophysiology with Mathematical Modeling. Annu Rev Pathol. 2020;15:371-94. Epub 2020/01/25. doi: 10.1146/annurev-pathmechdis-012419-032557. PubMed PMID: 31977295.

12. Zhang C, Qin L, Li K, Wang Q, Zhao Y, Xu B, et al. A Novel Scoring System for Prediction of Disease Severity in COVID-19. Front Cell Infect Microbiol. 2020;10:318. Epub 2020/06/26. doi: 10.3389/fcimb.2020.00318. PubMed PMID: 32582575; PubMed Central PMCID: PMCPMC7292148.

13. Cadegiani FA, Zimerman RA, Campello de Souza B, McCoy J, Pereira ECRA, Gustavo Wambier C, et al. The AndroCoV Clinical Scoring for COVID-19 Diagnosis: A Prompt, Feasible, Costless, and Highly Sensitive Diagnostic Tool for COVID-19 Based on a 1757-Patient Cohort. Cureus. 2021;13(1):e12565. Epub 2021/01/14. doi: 10.7759/cureus.12565. PubMed PMID: 33437562; PubMed Central PMCID: PMCPMC7793341.

14. Booth AL, Abels E, McCaffrey P. Development of a prognostic model for mortality in COVID-19 infection using machine learning. Mod Pathol. 2020:1-10. Epub 2020/10/18. doi: 10.1038/s41379-020-00700-x. PubMed PMID: 33067522; PubMed Central PMCID: PMCPMC7567420.

15. Ko J, Baldassano SN, Loh PL, Kording K, Litt B, Issadore D. Machine learning to detect signatures of disease in liquid biopsies - a user's guide. Lab Chip. 2018;18(3):395-405. Epub 2017/12/02. doi: 10.1039/c7lc00955k. PubMed PMID: 29192299; PubMed Central PMCID: PMCPMC5955608.

16. Pinto BGG, Oliveira AER, Singh Y, Jimenez L, Gonçalves ANA, Ogava RLT, et al. ACE2 Expression is Increased in the Lungs of Patients with Comorbidities Associated with

Severe COVID-19. medRxiv. 2020. Epub 2020/06/09. doi: 10.1101/2020.03.21.20040261. PubMed PMID: 32511627; PubMed Central PMCID: PMCPMC7276054.

17. Bonafe M, Prattichizzo F, Giuliani A, Storci G, Sabbatinelli J, Olivieri F. Inflamm-aging: Why older men are the most susceptible to SARS-CoV-2 complicated outcomes. Cytokine Growth Factor Rev. 2020. Epub 2020/05/12. doi: 10.1016/j.cytogfr.2020.04.005. PubMed PMID: 32389499.

18. Feldstein LR, Rose EB, Horwitz SM, Collins JP, Newhams MM, Son MBF, et al. Multisystem Inflammatory Syndrome in U.S. Children and Adolescents. N Engl J Med. 2020;383(4):334-46. Epub 2020/07/01. doi: 10.1056/NEJMoa2021680. PubMed PMID: 32598831; PubMed Central PMCID: PMCPMC7346765.

19. CDC. COVID-19 in Children and Teens 2020 [10 Jan 2020]. Available from: https://www.cdc.gov/coronavirus/2019-ncov/daily-life-coping/children/symptoms.html.

20. Afzal A. Molecular diagnostic technologies for COVID-19: Limitations and challenges. J Adv Res. 2020;26:149-59. Epub 2020/08/25. doi: 10.1016/j.jare.2020.08.002. PubMed PMID: 32837738; PubMed Central PMCID: PMCPMC7406419.

21. Ratajczak MZ, Bujko K, Ciechanowicz A, Sielatycka K, Cymer M, Marlicz W, et al. SARS-CoV-2 Entry Receptor ACE2 Is Expressed on Very Small CD45(-) Precursors of Hematopoietic and Endothelial Cells and in Response to Virus Spike Protein Activates the Nlrp3 Inflammasome. Stem Cell Rev Rep. 2020:1-12. Epub 2020/07/22. doi: 10.1007/s12015-020-10010-z. PubMed PMID: 32691370; PubMed Central PMCID: PMCPMC7370872.

22. Lukassen SA-O, Chua RL, Trefzer T, Kahn NC, Schneider MA, Muley T, et al. SARS-CoV-2 receptor ACE2 and TMPRSS2 are primarily expressed in bronchial transient. Embo J. 2020;39(10):e105114 LID - 10.15252/embj.20105114 [doi] LID - e105114.

439 23. Hirano T, Murakami M. COVID-19: A New Virus, but a Familiar Receptor and Cytokine
440 Release Syndrome. Immunity. 2020;52(5):731-3. Epub 2020/04/24. doi:
441 10.1016/j.immuni.2020.04.003. PubMed PMID: 32325025; PubMed Central PMCID:
442 PMCPMC7175868.
443 24. Hoffmann M, Kleine-Weber H, Schroeder S, Krüger N, Herrler T, Erichsen S, et al.
444 SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically. Cell.
445 2020;181(2):271-80.e8 LID - S0092-8674(20)30229-4 [pii] LID - 10.1016/j.cell.2020.02.052
446 [doi].
447 25. Finkel Y, Mizrahi O, Nachshon A, Weingarten-Gabbay S, Morgenstern D, Yahalom-
448 Ronen Y, et al. The coding capacity of SARS-CoV-2. Nature. 2020. Epub 2020/09/10. doi:
449 10.1038/s41586-020-2739-1. PubMed PMID: 32906143.
450 26. Wang K, Chen W, Zhang Z, Deng Y, Lian JQ, Du P, et al. CD147-spike protein is a
451 novel route for SARS-CoV-2 infection to host cells. Signal Transduct Target Ther.
452 2020;5(1):283. Epub 2020/12/06. doi: 10.1038/s41392-020-00426-x. PubMed PMID: 33277466;
453 PubMed Central PMCID: PMCPMC7714896.
454 27. Zhou Z, Ren L, Zhang L, Zhong J, Xiao Y, Jia Z, et al. Heightened Innate Immune
455 Responses in the Respiratory Tract of COVID-19 Patients. Cell Host Microbe. 2020;27(6):883-
456 90.e2. Epub 2020/05/15. doi: 10.1016/j.chom.2020.04.017. PubMed PMID: 32407669; PubMed
457 Central PMCID: PMCPMC7196896.
458 28. To KK, Tsang OT, Leung WS, Tam AR, Wu TC, Lung DC, et al. Temporal profiles of
459 viral load in posterior oropharyngeal saliva samples and serum antibody responses during
460 infection by SARS-CoV-2: an observational cohort study. Lancet Infect Dis. 2020;20(5):565-74.

461 Epub 2020/03/28. doi: 10.1016/s1473-3099(20)30196-1. PubMed PMID: 32213337; PubMed
462 Central PMCID: PMCPMC7158907.
463 29. Azkur AK, Akdis M, Azkur D, Sokolowska M, van de Veen W, Brüggen MC, et al.
464 Immune response to SARS-CoV-2 and mechanisms of immunopathological changes in COVID-
465 19. Allergy. 2020. Epub 2020/05/13. doi: 10.1111/all.14364. PubMed PMID: 32396996;
466 PubMed Central PMCID: PMCPMC7272948.
467 30. Catanzaro M, Fagiani F, Racchi M, Corsini E, Govoni S, Lanni C. Immune response in
468 COVID-19: addressing a pharmacological challenge by targeting pathways triggered by SARS-
469 CoV-2. Signal Transduction and Targeted Therapy. 2020;5(1):84. doi: 10.1038/s41392-020-
470 0191-1.
471 31. Chi Y, Ge Y, Wu B, Zhang W, Wu T, Wen T, et al. Serum Cytokine and Chemokine
472 Profile in Relation to the Severity of Coronavirus Disease 2019 in China. J Infect Dis.
473 2020;222(5):746-54. Epub 2020/06/21. doi: 10.1093/infdis/jiaa363. PubMed PMID: 32563194;
474 PubMed Central PMCID: PMCPMC7337752.
475 32. Picchianti Diamanti A, Rosado MM, Pioli C, Sesti G, Laganà B. Cytokine Release
476 Syndrome in COVID-19 Patients, A New Scenario for an Old Concern: The Fragile Balance
477 between Infections and Autoimmunity. Int J Mol Sci. 2020;21(9). Epub 2020/05/14. doi:
478 10.3390/ijms21093330. PubMed PMID: 32397174; PubMed Central PMCID:
479 PMCPMC7247555.
480 33. Grifoni A, Weiskopf D, Ramirez SI, Mateus J, Dan JM, Moderbacher CR, et al. Targets
481 of T Cell Responses to SARS-CoV-2 Coronavirus in Humans with COVID-19 Disease and
482 Unexposed Individuals. Cell. 2020;181(7):1489-501.e15. Epub 2020/05/31. doi:

10.1016/j.cell.2020.05.015. PubMed PMID: 32473127; PubMed Central PMCID: PMCPMC7237901.

34. Ernst D, Bolton G, Recktenwald D, Cametron MJ, Danesh A, Persad D, et al. Bead Based Flow Cytometric Assays : A Multiplex Assay Platform with Applications in Diagnostic Microbiology. In: Advanced Techniques in Diagnostic Microbiology , Eds : Y.W. Tang and C.W. Stratton. Springer. 2006. p. 427-43.

35. Bassoli C, Oreni L, Ballone E, Foschi A, Perotti A, Mainini A, et al. Role of serum albumin and proteinuria in patients with SARS-CoV-2 pneumonia. Int J Clin Pract. 2020:e13946. Epub 2020/12/18. doi: 10.1111/ijcp.13946. PubMed PMID: 33332691.

36. Payán-Pernía S, Gómez Pérez L, Remacha Sevilla Á F, Sierra Gil J, Novelli Canales S. Absolute Lymphocytes, Ferritin, C-Reactive Protein, and Lactate Dehydrogenase Predict Early Invasive Ventilation in Patients With COVID-19. Lab Med. 2020. Epub 2020/12/19. doi: 10.1093/labmed/lmaa105. PubMed PMID: 33336243.

37. Chen Z, John Wherry E. T cell responses in patients with COVID-19. Nat Rev Immunol. 2020;20(9):529-36. Epub 2020/07/31. doi: 10.1038/s41577-020-0402-6. PubMed PMID: 32728222; PubMed Central PMCID: PMCPMC7389156 Roche, Pieris, Elstar and Surface Oncology. E.J.W. has a patent licensing agreement on the PD1 pathway with Roche/Genentech. E.J.W. is a founder of Arsenal Biosciences. Z.C. declares no competing interests.

38. Wajnberg A, Amanat F, Firpo A, Altman DR, Bailey MJ, Mansour M, et al. Robust neutralizing antibodies to SARS-CoV-2 infection persist for months. Science. 2020;370(6521):1227-30. Epub 2020/10/30. doi: 10.1126/science.abd7728. PubMed PMID: 33115920.

505 39. Bloch EM, Shoham S, Casadevall A, Sachais BS, Shaz B, Winters JL, et al. Deployment
506 of convalescent plasma for the prevention and treatment of COVID-19. J Clin Invest.
507 2020;130(6):2757-65. Epub 2020/04/08. doi: 10.1172/jci138745. PubMed PMID: 32254064;
508 PubMed Central PMCID: PMCPMC7259988.
509 40. Legros V, Denolly S, Vogrig M, Boson B, Siret E, Rigaill J, et al. A longitudinal study of
510 SARS-CoV-2-infected patients reveals a high correlation between neutralizing antibodies and
511 COVID-19 severity. Cell Mol Immunol. 2021;18(2):318-27. Epub 2021/01/08. doi:
512 10.1038/s41423-020-00588-2. PubMed PMID: 33408342; PubMed Central PMCID:
513 PMCPMC7786875.
514 41. Box G, E., Hunter W, Hunter S. Statistics for Experimenters: design, discovery and
515 innovation: John Wiley and Sons; 2005.
516 42. Piyathilake CJ, Badiga S, Burkholder GA, Harada S, Raper JL. The accuracy of HPV
517 genotyping in isolation and in combination with CD4 and HIV viral load for the identification of
518 HIV-infected women at risk for developing cervical cancer. Cancer Med. 2021. Epub
519 2021/02/20. doi: 10.1002/cam4.3785. PubMed PMID: 33605553.
520 43. Bland J, Kavanaugh A, Hong LK, Kadkol SS. Development and Validation of Viral Load
521 Assays to Quantitate SARS-CoV-2. J Virol Methods. 2021:114100. Epub 2021/02/19. doi:
522 10.1016/j.jviromet.2021.114100. PubMed PMID: 33600849.
523 44. Romeo F, Louge Uriarte E, Delgado SG, González-Altamiranda E, Pereyra S, Morán P,
524 et al. Effect of bovine viral diarrhea virus on subsequent infectivity of bovine gammaherpesvirus
525 4 in endometrial cells in primary culture: an in vitro model of viral co-infection. J Virol Methods.
526 2021:114097. Epub 2021/02/19. doi: 10.1016/j.jviromet.2021.114097. PubMed PMID:
527 33600847.

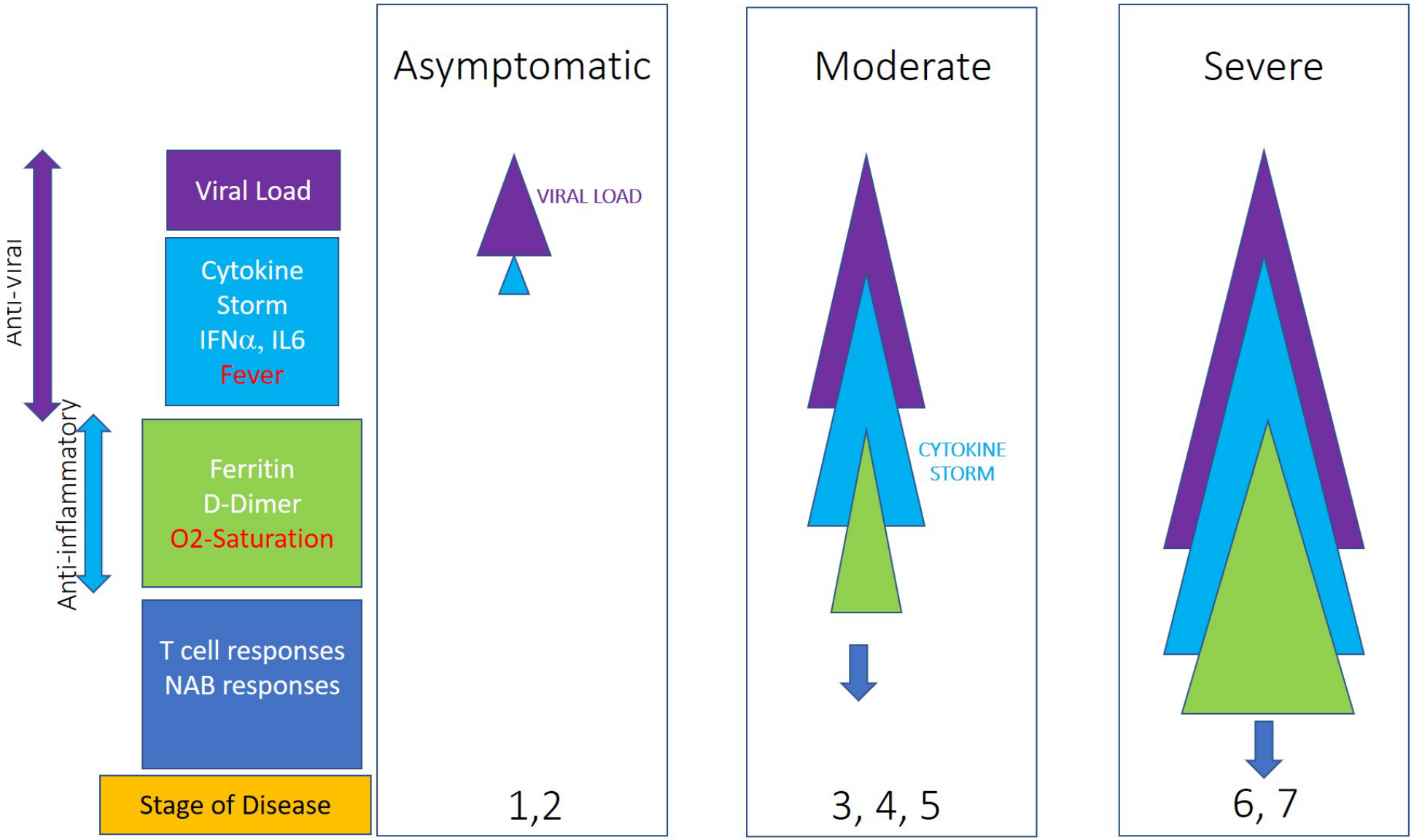
528 45. McNab F, Mayer-Barber K, Sher A, Wack A, O'Garra A. Type I interferons in infectious
529 disease. Nat Rev Immunol. 2015;15(2):87-103. Epub 2015/01/24. doi: 10.1038/nri3787. PubMed
530 PMID: 25614319; PubMed Central PMCID: PMCPMC7162685.
531 46. Buszko M, Park JH, Verthelyi D, Sen R, Young HA, Rosenberg AS. The dynamic
532 changes in cytokine responses in COVID-19: a snapshot of the current state of knowledge. Nat
533 Immunol. 2020;21(10):1146-51. Epub 2020/08/29. doi: 10.1038/s41590-020-0779-1. PubMed
534 PMID: 32855555.
535 47. Cardone M, Yano M, Rosenberg AS, Puig M. Lessons Learned to Date on COVID-19
536 Hyperinflammatory Syndrome: Considerations for Interventions to Mitigate SARS-CoV-2 Viral
537 Infection and Detrimental Hyperinflammation. Front Immunol. 2020;11:1131. Epub 2020/06/24.
538 doi: 10.3389/fimmu.2020.01131. PubMed PMID: 32574265; PubMed Central PMCID:
539 PMCPMC7272717.
540 48. Vabret N, Britton GJ, Gruber C, Hegde S, Kim J, Kuksin M, et al. Immunology of
541 COVID-19: current state of the science. Immunity. 2020. doi: 10.1016/j.immuni.2020.05.002.
542 PubMed PMID: PMC7200337.
543 49. Zhang Q, Bastard P, Liu Z, Le Pen J, Moncada-Velez M, Chen J, et al. Inborn errors of
544 type I IFN immunity in patients with life-threatening COVID-19. Science. 2020;370(6515).
545 Epub 2020/09/26. doi: 10.1126/science.abd4570. PubMed PMID: 32972995.
546 50. Rusu D, Blaj M, Ristescu I, Patrascanu E, Gavril L, Lungu O, et al. Outcome Predictive
547 Value of Serum Ferritin in ICU Patients with Long ICU Stay. Medicina (Kaunas). 2020;57(1).
548 Epub 2020/12/31. doi: 10.3390/medicina57010001. PubMed PMID: 33375016; PubMed Central
549 PMCID: PMCPMC7822040.

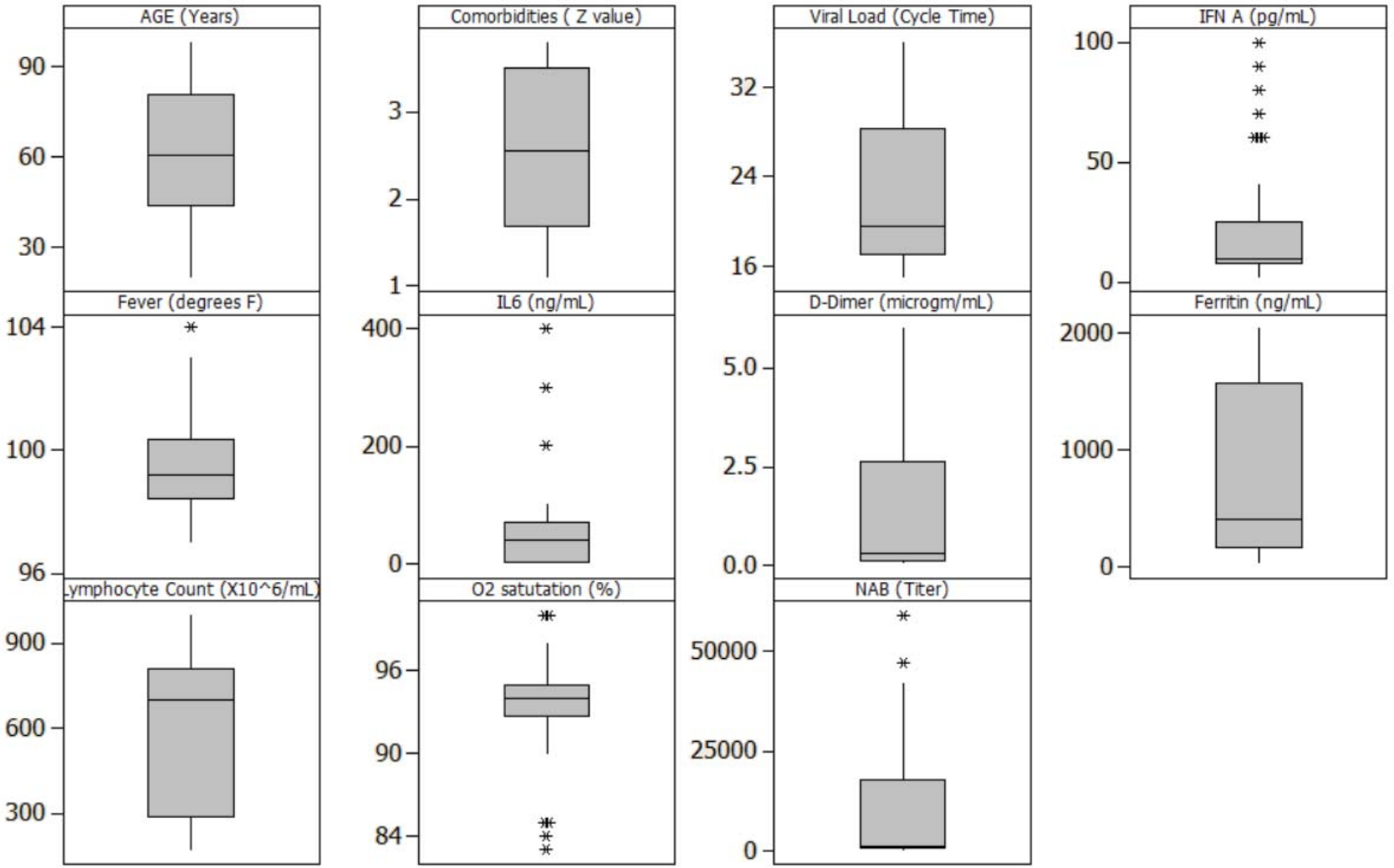
550 51. Dahan S, Segal G, Katz I, Hellou T, Tietel M, Bryk G, et al. Ferritin as a Marker of
551 Severity in COVID-19 Patients: A Fatal Correlation. Isr Med Assoc J. 2020;22(8):494-500. Epub
552 2020/11/26. PubMed PMID: 33236582.
553 52. Mateus J, Grifoni A, Tarke A, Sidney J, Ramirez SI, Dan JM, et al. Selective and cross-
554 reactive SARS-CoV-2 T cell epitopes in unexposed humans. Science. 2020;370(6512):89-94.
555 Epub 2020/08/06. doi: 10.1126/science.abd3871. PubMed PMID: 32753554; PubMed Central
556 PMCID: PMCPMC7574914.
557 53. Gentili M, Nachohen A. Surprising effects of antibodies in severe COVID. Nature. 2021.
558 54. Poeschla E. Neutralizing SARS-CoV-2. Elife. 2020;9. Epub 2020/12/16. doi:
559 10.7554/eLife.64496. PubMed PMID: 33320086; PubMed Central PMCID: PMCPMC7738179.

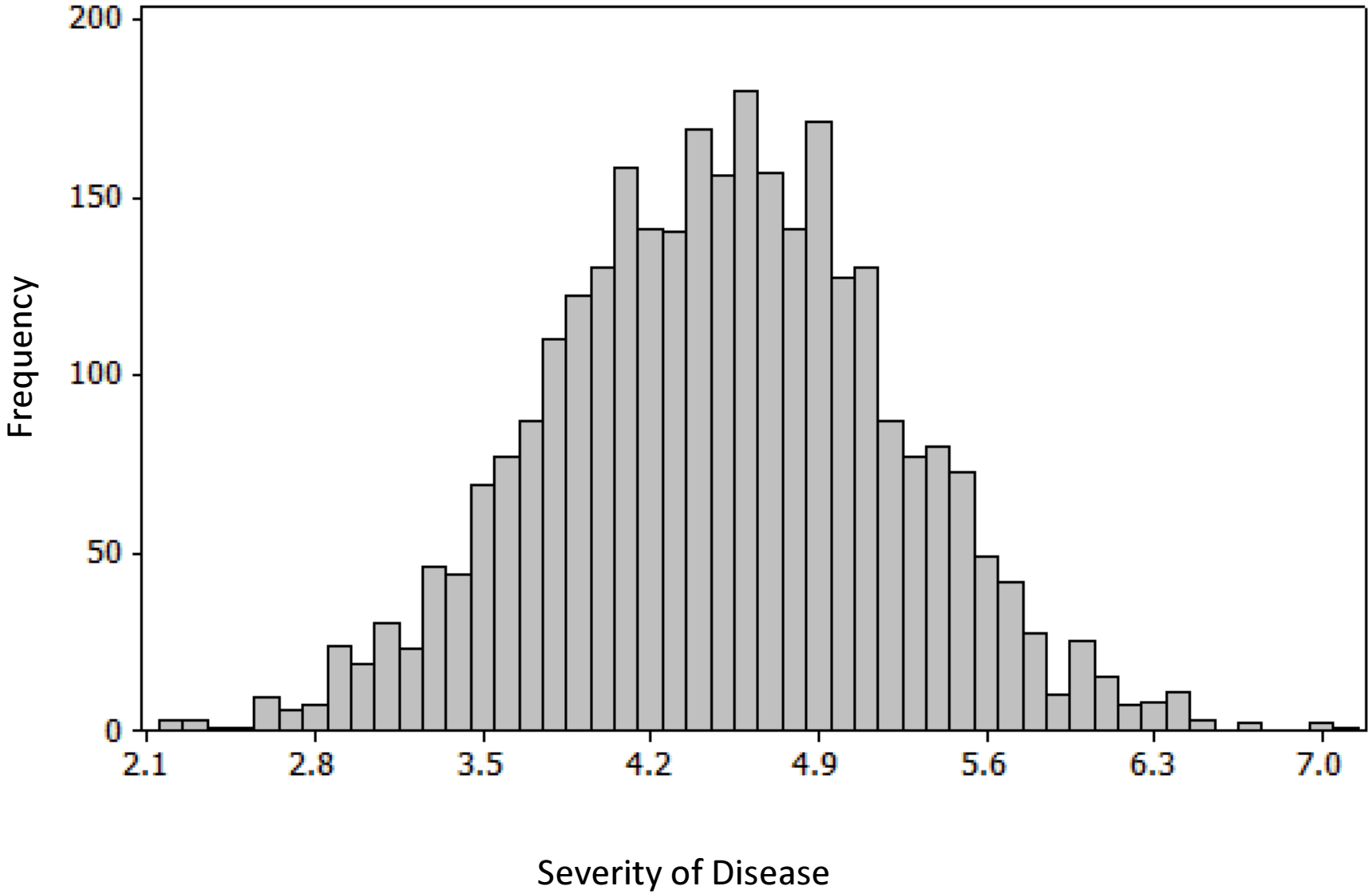
560 55. Qian, Z. et al. (2013) Innate immune response of human alveolar type II cells infected with
561 severe acute respiratory syndrome-coronavirus. Am J Respir Cell Mol Biol 48 (6), 742-8.
562 56. Kronbichler, A. et al. (2020) Asymptomatic patients as a source of COVID-19 infections: A
563 systematic review and meta-analysis. Int J Infect Dis 98, 180-186. PMID: 32562846;
564 PMCID: PMC7832751; DOI: 10.1016/j.ijid.2020.06.052

565 57. Zhao, J. et al. (2020) Antibody responses to SARS-CoV-2 in patients of novel coronavirus
566 disease 2019. LID - 10.1093/cid/ciaa344 [doi] LID - ciaa344. Clin Infect Dis.

567







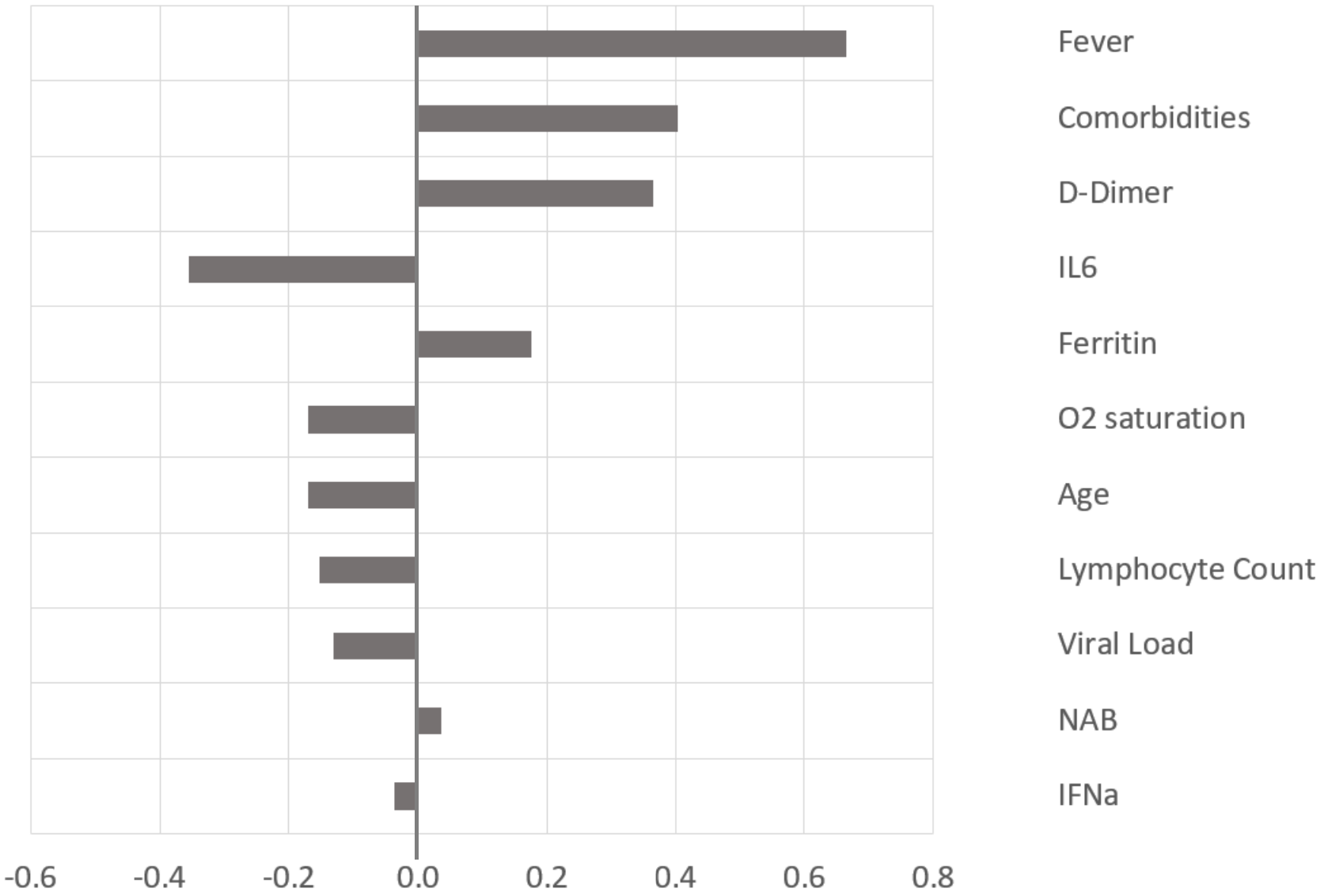


Table 1. Ranges of values for the parameters used for developing the simulated dataset for the mathematical model.

Parameter	Unit	Reference	COVID-19 Ranges	COVID-19 Ranges	COVID-19 Ranges
			Asymptomatic	Moderate	Severe
Viral Load	Cycle Time	[20]	>28	20-15	22-16
IFNα	pg/mL	[45, 46]	<10	10-100	10-2
Fever	°F	[55, 56]	97-98.6	98.6-100	100-104
D-Dimer	μg/mL	[31]	<0.1	0.15-0.62	0.5-9.3
Ferritin	ng/L	[31]	20-200	286-1,275	1,400-2,000
Oxygen Saturation	%	[3, 56]	95-100	85-94	60-84
IL-6	ng/mL	[45, 46]	<1	19-76	19-430
Lymphocyte count	x 10^6/mL	[27]	>785	588-785	169-415
NAB	Titer	[57]	1,000-45,000	200-20,000	500-60,000

Table 2A. Statistical analysis of coefficients for each parameter based on the multiple regression analysis. The Table shows the regression statistics. The statistical terms are: DF (degrees of freedom); SeqSS (Sequential Sum Square); AdjSS (Adjusted Sum Square), AdjMS (Adjusted Mean Squares), F ratio, p value.

Source	DF	SeqSS	AjdSS	AdjMS	F	p
Regression	11	13.172	139.172	12.652	10.259	0
AGE	1	116.579	0.095	0.095	0.817	0.337
COMORBID	1	10.085	0.175	0.175	1.5	0.231
Viral Load	1	0.172	0.392	0.392	3.357	0.078
IFNα	1	1.037	0.136	0.135	1.161	0.291
Fever	1	8.159	0.378	0.378	3.238	0.083
IL6	1	0.808	0.172	0.171	1.469	0.236
D-Dimer	1	1.511	0.444	0.444	3.799	0.062
Ferritin	1	0.574	0.098	0.098	0.839	0.368
Lymphocyte Count	1	0.018	0.039	0.039	0.334	0.568
Oxygen saturation	1	0.91	0.133	0.133	1.141	0.295
NAB	1	0.039	0.039	0.039	0.337	0.566

Table 2B. **Coefficient and standard error for parameters**. The coefficients for each parameter were determined by using multiple regression analyses, which is the multiplier to the parameter value in a linear regression equation. The values of the coefficients of each of the parameters are shown using ANOVA. The table shows the standard error, t and p values. The p value denotes statistical significance to the outcome.

Term	Coefficient	Standard Error	t	p
Constant	-36.898	24.867	-1.484	0.150
AGE	-0.021	0.023	-0.904	0.374
COMORBID	0.894	0.730	1.225	0.232
Viral Load	-0.048	0.026	-1.832	0.078
IFN A	-0.005	0.005	-1.077	0.291
Fever	0.444	0.247	1.799	0.084
IL6	-0.003	0.003	-1.212	0.236
D-Dimer	0.271	0.139	1.949	0.062
Ferritin	0.000	0.001	0.916	0.368
Lymphocyte Count	-0.001	0.001	-0.578	0.568
Oxygen saturation	-0.038	0.036	-1.068	0.295
NAB	0.000	0.000	-0.580	0.567
Error	26.000	3.039	3.039	0.117
Total	37.000	142.211		

Table 3. The Ranges for the Monte Carlo Simulation.

Parameters	Range				Parametric
	Max		Min		
Age	**66**		**20**		43
Comorbidities	**3.6**		**1.2**		2.4
Viral Load	**30**		**18**		24
IFNa	**48**		**5**		26.5
Fever	**104**		**97.7**		100.85
IL6	**390**		**2**		196
D-Dimer	**6**		**0.02**		3.01
Ferritin	**1775**		**30**		902.5
Lymphocyte Count	**983**		**150**		566.5
O2 saturation	**100**		**81**		90.5
NAB	**1580**		**2**		791
					4.50 **Simulated Rank**

Table 4. The Prediction of outcome based on observed and predicted values.

Variable	Subject 1	Subject 2	Subject 3	Subject 4	Subject 5	Subject 6	Subject 7
Age	20	52	55	55	62	73	80
Comorbidity	1.1	2.2	2.3	2.4	2.6	3.2	3.5
Viral Load	36	19	16	15	17	21	17
IFNa	7	20	50	60	40	9	5
Fever	98.6	99.1	99.8	100	99.4	100.1	100
IL6	1	30	60	70	50	50	90
D-Dimer	0.05	0.25	0.40	0.45	0.35	1.5	4.0
Ferritin	30	350	800	1275	500	1500	1800
Lymphocyte Count	800	740	620	580	640	340	200
Oxygen Saturation	100	95	83	85	90	95	94
NAB	2	20	150	200	100	220	400
Calculated Outcome Rank	0.723	3.299	4.008	4.237	3.228	4.916	6.427
PREDICTED	1	3	4	4	3	5	6
OBSERVED	1	3	4	4	4	5	6