Predictors of COVID-19 Hospital Treatment Outcome

Ryszard Tomasiuk1
Jan Dabrowski2
Jolanta Smykiewicz3
Magdalena Wiacek4

1Department of Medicine, Faculty of Medical and Health Sciences, The University of Technology and Humanities, Radom, Poland; 2Department of Pharmacology, National Medicines Institute, Warsaw, Poland; 3Central Medical Diagnostic Laboratory, Dr. Tytus Chałubiński Specialist Hospital in Radom, Radom, Poland; 4Department of Physiotherapy, Faculty of Medical and Health Sciences, The University of Technology and Humanities, Radom, Poland

Background: There are more than 228,394,572 confirmed cases and 4,690,186 confirmed deaths caused by COVID-19 worldwide. The magnitude of the COOVID-19 pandemic has stimulated research on the treatment and diagnosis of COVID-19 patients.

Objective: In this report, a battery of specific parameters was used to develop a model that allows prediction of the outcome of the COVID-19 treatment. These parameters are C-reactive protein, procalcitonin, fibrinogen, D-dimers, immature granulocytes, and interleukin-6.

Methods: The study was carried out on a sample of N = 49 survivors (22 men, 27 women) and 83 deceased patients (62 men, 21 women). The distribution of means and differences in means of the parameters studied between survivors and deceased patients were evaluated using the bootstrap method.

Results: A mathematical model that allows for the prediction of hospitalization outcome was obtained using the Naive Bayes model. The results demonstrated a statistically significant difference between survivors and deceased patients in all parameters studied. A mathematical model employing a battery of parameters provided a 97% precision in predicting the outcome of hospitalization.

Conclusion: This study showed that the cross-correlation of survivability with absolute levels of C-reactive protein, procalcitonin, fibrinogen, D-dimers, immature granulocytes, and interleukin-6 could be used successfully in the hospital setting as a diagnostic tool.

Keywords: Covid-19, biological markers, C-reactive protein, procalcitonin, fibrinogen, D-dimers, immature granulocytes, interleukin-6

Introduction
The world has been plagued by coronavirus disease (COVID-19) for the past 2 years. Globally, there are 228,394,572 confirmed cases with 4,690,186 confirmed deaths.1 Coronaviruses are part of the Coronaviridae subfamily2 and belong to the family of positive-sense RNA viruses responsible for respiratory diseases in mammals and birds.3 In the Coronaviridae subfamily, there are three groups. The viruses of Groups 1 and 2 have only mammalian hosts, whereas the viruses of Group 3 have only been found in birds.4

Methods for diagnosing COVID-19 include 1) nucleic acid amplification test (NAAT), 2) the serological test, and 3) the hematologic test.5 Hematologic disorders can be divided into typical, most common, and hematologic abnormalities in conjunction with coagulopathy.6 For example, a typical abnormality of COVID-19 disease is high levels of C-reactive protein (CRP)7 and D-dimer.8 Furthermore, serum procalcitonin (PCT)9 and fibrinogen (Fg)10 also have diagnostic value for patients with COVID-19.
Studies on COVID-19, a detrimental influence on medical services induced by the Covid-19 outbreak, indicate the need for robust and quick methods to predict Covid-19 mortality. Among such, one may distinguish study of Castelnuovo et al, who employed a machine learning tool to predict cardiovascular risk factors on Covid-19 mortality. Furthermore, machine-supported decision increases the speed of decisions on Covid-19 treatment, which is extremely important because Covid-19 outburst delayed the diagnosis of other diseases.

The purpose of this report is to evaluate the relative differences in CRP, PCT, Fg, D-dimers, immature granulocytes (IG), and interleukin-6 (IL-6) in subjects who recovered or died from COVID-19 and to provide a robust mathematical model for the prediction of COVID-19 mortality and treatment outcome. The mathematical model used for the estimation of COVID-19 rendered mortality was the Naive Bayes classification. Consequently, this report is among a few that use mathematical modelling to test a wide range of blood markers to predict the outcome of the disease.

Methods

The study was performed accordingly to World Medical Association (WMA) declaration of Helsinki. Data were collected on the date of death or discharge from the hospital. Institutional Ethics Clearance (IEC) granted by the Regional Ethics Committee of Medical Chamber in Gdańsk, Poland, was obtained for this study: KB-29/21. Additionally, each person provided a signed informed consent form.

Study Subjects

The data used in this study included blood tests from patients admitted to Dr. Tytus Chałubiński Specialist Hospital in Radom, Poland. In this study, a total of N = 132 subjects were used: N = 49 survivors (22 men, 27 women) and N = 83 deceased patients (62 men, 21 women).

Experimental Methods

CRP protein concentration was measured using an in vitro immunoturbidimetric assay (Tina-quant C-reactive protein IV) in serum. It is a latex particle-enhanced immunoturbidimetric assay that consists of TRIS buffer with bovine serum albumin (BSA) with preservatives R2 and latex particles coated with mouse anti-CRP glycine buffer and mouse immunoglobulins with preservative. Human CRP was agglutinated with latex particles covered with anti-CRP monoclonal antibodies, and the precipitate was measured by turbidimetry.

PCT concentration was determined in vitro using the Elecsys BRAHMS PCT serum assay. The test was carried out in three steps. The first step was to incubate a complex composed of sample antigen, biotinylated PCT-specific monoclonal antibodies, and PCT-specific monoclonal antibodies labeled with a ruthenium complex. The second step involved adding labeled microparticles to streptavidin to bind the complex to the solid phase using the affinity of biotin and streptavidin. The third step was to transfer the reaction mixture to the measuring chamber, where the microparticles were attracted to the electrode surface by a magnet. The unbound particles were processed using the ProCell/ProCell M method. Electrochemiluminescence and photon emission were induced by the applied voltage and measured using a photomultiplier. The results were quantified by reference to the calibration curve.

Fg was quantified using Dia-FIB. In principle, the method allows measuring the clotting time after adding a high concentration of thrombin to the diluted plasma. Plasma fibrinogen concentration is inversely correlated with clotting time.

The concentration of D-dimers was measured using the photometric Dia-D-DIMER method. Dia-D-DIMER is an immunoturbidimetric test reinforced with latex particles. The method uses the binding of the sample to antibodies directed at D-dimers and the coating of latex particles. The concentration of D-dimers was determined by an antigen–antibody photometric reaction.

IG analysis was performed automatically using a Sysmex XN-550 multiparameter hematology analyzer from blood samples obtained by venipuncture in EDTA vacutainer tubes.

IL-6 concentration was measured in serum collected from separation gel tubes. The procedure was a four-step process that included 1) first incubation: 30 μL of a sample were incubated with a biotinylated monoclonal IL6-specific antibody, 2) second incubation: After the addition of a monoclonal IL6-specific antibody labeled with a ruthenium complex and streptavidin-coated microparticles, the antibodies form a sandwich complex with the antigen of the sample, 3) the reaction mixture was aspirated into the measuring cell where the microparticles are magnetically captured on the surface of the electrode. The unbound substances were then removed with Procell/ProCell M. The application of a voltage to the
electrode then induced chemiluminescent emission, which was measured by a photomultiplier, and 4) the results were determined by a calibration curve that is instrument-specific generated by a 2-point calibration and a master curve.

Statistical Analysis
The distribution of means and differences in means between survivors and deceased patients were tested using a bootstrap test consisting of 10,000 repeats with replacement.\(^\text{18}\)

The Naive Bayes-supervised nonlinear classification algorithm was implemented using the R language.\(^\text{19}\) The characteristics used in the construction of the model were following: age, PRC, PCT, Fg, D-dimers, IG%, and IL-6. The data set was divided into training and testing sets, where the number of cases in the training data was 71 and the number of cases in the testing data was 45. The data set was divided into two classes: survivors and deceased patients. The scheme of the model construction procedure, including the structure of the input database, is illustrated in Figure 1.

Results
Figure 2 provides an age distribution of survivors and deceased patients. The means of CRP, PCT, Fg, D-dimers, IG, and IL-6 stratified by hospitalization outcome are collected in Table 1 and shown in Figures 3A–C to 8A–C.

Analysis of data collected in Table 1 shows that survivors were defined by three times higher CRP levels, two times higher PCT levels, six times higher D-dimers levels, five times higher in IG% and three times higher IL-6 levels than deceased subjects. Twofold lower Fg levels were observed between the respective groups.

Figure 3A shows that the mean distribution of CRP in survivors included normal (<4.1 mg/L),\(^\text{20}\) high probability (>4.1 mg/L),\(^\text{21}\) and high severity (>8.7 mg/L)\(^\text{22}\) of the COVID-19 infection reference range. Figure 3B shows that the distribution of CRP concentration among deceased subjects fell within the high-severity range of COVID-19 infection. Figure 3C reveals a statistically significant difference in mean CRP levels between survivors and deceased subjects.

Figure 4A shows that the mean distribution of PCT among survivors was in a significant but moderate systemic inflammatory response, a likely range of sepsis and severe sepsis, and a high-risk range of developing organ dysfunction.\(^\text{23}\) Figure 4B shows that the distribution of PCT in the deceased patient group ranged from a nonprobable systemic infection to severe sepsis and a high risk of developing organ dysfunction. Finally,
Figure 4C shows statistically significant differences in mean PCT between survivors and deceased subjects.

Figure 5A and B show that the distribution of Fg means among survivors and deceased groups was in the normal range (>4 g/L). Furthermore, there was a statistically significant difference in the means of Fg between survivors and deceased patients, Figure 5C.

The distribution of the mean concentration of D dimers among survivors and the deceased group, Figure 6A and B, respectively, was above the normal range (>250 ng/mL). Figure 6C shows statistically significant differences in the mean of D dimers between survivors and deceased subjects.

Figure 7A shows that the IG distribution in survivors encompassed the normal range (<1%) and above the normal range (≥1%) range. The distribution of the mean IG% among the deceased patients included levels above only the normal range, Figure 7B. Figure 7C shows a statistically significant difference in the means of IG% between survivors and deceased subjects.

Figure 8A reveals that the distributions of IL-6 among survivors encompassed values within (≤24 pg/mL) and above the normal range (>24 pg/mL). The deceased groups covered values above the normal range, Figure 8B. Moreover, a statistically significant difference in the mean levels of IL-6 between survivors and deceased subjects is observed, Figure 8C.

Discussion

To the best of our knowledge, the present study is among the first to prove the usefulness of a combination of a battery of blood test parameters and a machine learning model to predict the outcome of COVID-19 treatment.

Several reports showed that C-reactive protein (CRP) levels increase significantly in response to injury, infection, and inflammation and participate in the inflammation process. Furthermore, CRP can recognize self- and foreign molecules by pattern recognition. Therefore, its mode of action is different from that of immunoglobulins, which recognize only distinct antigenic epitopes. A recent study also demonstrated CRP usability as a predictor of the severity of COVID-19. This report confirmed a correlation between survival status and CRP levels in subjects infected with COVID-19.

A three-times higher level of CRP defined patients who died from COVID-19 than subjects who survived infection and disease: 17.7 vs 6.4 mg/L (Figure 3A–C). PCT, the precursor of calcitonin, is the hormone responsible for calcium homeostasis. Its levels increase substantially in sepsis, systemic infections, and severe inflammation. A recent study also showed that PCT concentration is significantly higher in patients with COVID-19. This observation is in agreement with the results of this report (Table 1, Figure 4A–C). Therefore, patients who died of COVID-19 were defined by a mean PCT equal to 4.3 ng/mL, while the PCT level equal to 2.19 ng/mL defined patients who survived COVID-19.
Figure 3 Histogram of CRP levels distribution in (A) subjects who survived COVID-19 infection, (B) deceased subjects, and (C) comparison of differences in means between survived and deceased subjects. * - denotes $P < 0.05$.

Figure 4 Histogram of PCT levels distribution in (A) subjects who survived COVID-19 infection, (B) deceased subjects, and (C) comparison of differences in means between survived and deceased subjects. * - denotes $P < 0.05$. 
**Figure 5** Histogram of Fibrinogen levels distribution in (A) subjects who survived COVID-19 infection, (B) deceased subjects, and (C) comparison of differences in means between survived and deceased subjects. * - denotes $P < 0.05$.

**Figure 6** Histogram of D-dimers levels distribution in (A) subjects who survived COVID-19 infection, (B) deceased subjects, and (C) comparison of differences in means between survived and deceased subjects. * - denotes $P < 0.05$.  

https://doi.org/10.2147/IJGM.S334544

DovePress

For personal use only.
Figure 7 Histogram of IG% levels distribution in (A) subjects who survived COVID-19 infection, (B) deceased subjects, and (C) comparison of differences in means between survived and deceased subjects. * - denotes $P < 0.05$.

Figure 8 Histogram of IL6 levels distribution in (A) subjects who survived COVID-19 infection, (B) deceased subjects, and (C) comparison of differences in means between survived and deceased subjects. * - denotes $P < 0.05$. 
This study also found an inverse correlation between Fg and D-dimers as a function of the severity of COVID-19. Deceased subjects were defined by two times lower levels of Fg and six times higher levels of D-dimers compared to patients who recovered from COVID-19 infection. Therefore, the levels of the Fg and D dimers were equal to 4.5 g/L and 11,118 ng/mL, vs 8.4 g/L and 1833 ng/mL for the subjects deceased and survived, respectively (Figures 5A–C and 6A–C).

Some studies showed a reciprocal relationship between inflammation severity and IG number. Furthermore, elevated levels of IG were also discovered among COVID-19 patients. The results of this study confirmed a statistically significant increase in IG% levels among patients who died of COVID-19. Thus, a mean IG% equal to 4.5 defined deceased subjects, while survivors are defined by a mean IG% equal to 0.8 (Figure 7A–C).

IL-6 synthesized by an organism in response to infection, tissue injury, or inflammation, was also used as a predictor of the severity of COVID-19. This study confirmed this observation and found a direct correlation between IL-6 concentration and COVID-19 survival status. Therefore, patients who died of COVID-19 were defined by a mean IL-6 concentration equal to 146.2 pg/mL, while survivors were defined by a mean IL-6 concentration equal to 42.4 pg/mL (Figure 8A–C).

This study focuses on the problem of predicting standard government-approved COVID-19 treatment versus mortality. However, it must be noted that other approaches of COVID-19 treatment using, for example, heparin were also studied. Furthermore, the means of education on vaccination against COVID-19 should also be taken into account for the prevention and treatment of COVID-19. The relationship described above allows us to conclude that all parameters investigated differentiate the survivability of patients with COVID-19. Therefore, they could be the appropriate parameters for mathematical models predicting recovery after COVID-19 infection and disease.

Naive Bayes (NB) is a popular machine learning technique that offers a highly effective probabilistic classifier with a solid mathematical foundation and has been shown to perform well in medical diagnosis. Furthermore, a recent study has shown that it is suitable for the detection of patients with COVID-19. Using our findings, we attempted to build a mathematical model using the NB algorithm to predict the survivability of patients with COVID-19. However, the model presented in this study differs from the model previously proposed by Mansour et al. Thus, the following numerical characteristics (concentrations) were used in the construction of the model: CRP, PCT, Fg, D-dimers, IL-6, and %IG, as well as the chronological age of a patient. Consistent application of these parameters allowed us to elucidate a model that allowed us to predict the outcomes of COVID-19 with 97% precision.

The limitation of this study is the size of the study sample. However, the provided model can be easily updated for much larger datasets.

Conclusions
The results of the study show the direct usability of serum C-reactive protein concentrations, procalcitonin, fibrinogen, D-dimers, immature globulin percentage, and interleukin 6 to evaluate the probability of survival of COVID-19 patients in the hospital environment. The obtained Naive Bayes model may be used per se for prediction of the COVID-19 outcome or further tuned by incorporating additional parameters to increase its predictive power. In brief, inclusion of laboratory-derived data on C-reactive protein, procalcitonin, fibrinogen, D-dimer, immature globulin percentage, and interleukin 6 to the provided model predicts the probable outcome of the treatment. Although some may view the presumption on treatment outcome as unethical, it is the tool that can be used to intensify treatment of change of treatment to change predicted unsuccessful outcome.

Data Sharing Statement
Model and the data are available upon request.

Disclosure
The authors report no conflicts of interest in this work.

References

10254 https://doi.org/10.2147/IJGM.S334544
DovePress
International Journal of General Medicine 2021:14


