Systematic Review and Meta-Analysis

Medicine®

OPEN

An artificial neural network model to predict the mortality of COVID-19 patients using routine blood samples at the time of hospital admission

Development and validation study

Ju-Kuo Lin, MD[a,b], Tsair-Wei Chien, MBA[c], Lin-Yen Wang, MD[d,e,f], Willy Chou, MD[g,h,*]

Abstract

Background: In a pandemic situation (e.g., COVID-19), the most important issue is to select patients at risk of high mortality at an early stage and to provide appropriate treatments. However, a few studies applied the model to predict in-hospital mortality using routine blood samples at the time of hospital admission. This study aimed to develop an app, name predict the mortality of COVID-19 patients (PMCP) app, to predict the mortality of COVID-19 patients at hospital-admission time.

Methods: We downloaded patient records from 2 studies, including 361 COVID-19 patients in Wuhan, China, and 106 COVID-19 patients in 3 Korean medical institutions. A total of 30 feature variables were retrieved, consisting of 28 blood biomarkers and 2 demographic variables (i.e., age and gender) of patients. Two models, namely, artificial neural network (ANN) and convolutional neural network (CNN), were compared with each other across 2 scenarios using

1. raw laboratory versus normalized data and
2. training vs testing datasets (n = 361 and n = 106/361 ≅ 30%) to verify the model performance (e.g., sensitivity [SENS], specificity [SPEC], and area under the receiver operating characteristic curve [AUC]).

An app for predicting the mortality of COVID-19 patients was developed using the model's estimated parameters for the prediction and classification of PMCP at an earlier stage. Feature variables and prediction results were visualized using the forest plot and category probability curves shown on Google Maps.

Results: We observed that

1. the normalized dataset gains a relatively higher AUC(>0.9) when compared to that(<0.9) in the raw-laboratory dataset based on training data,
2. the normalized dataset in ANN yielded a high AUC of 0.96 that that(=0.91) in CNN based on testing data, and
3. a ready and available app, where anyone can access the model to predict mortality, for PMCP was developed in this study.

Conclusions: Our new PMCP app with ANN model accurately predicts the mortality probability for COVID-19 patients. It is publicly available and aims to help health care providers fight COVID-19 and improve patients' classifications against treatment risk.

Editor: Flavio Palmieri.

This study was waived due to data from publicly free website in articles.

All data used in this study are available in Supplemental Contents.

The authors have no funding and conflicts of interests to disclose.

Supplemental Digital Content is available for this article.

The datasets generated during and/or analyzed during the current study are publicly available.

[a] *Department of Ophthalmology, Chi-Mei Medical Center, Yong Kang, Tainan City, Taiwan,* [b] *Department of Optometry, Chung Hwa University of Medical Technology, Jen-Teh, Tainan City, Taiwan,* [c] *Department of Medical Research, Chi-Mei Medical Center, Tainan, Taiwan,* [d] *Department of Pediatrics, Chi-Mei Medical Center, Tainan, Taiwan,* [e] *Department of Childhood Education and Nursery, Chia Nan University of Pharmacy and Science, Tainan, Taiwan,* [f] *School of Medicine, College of Medicine, Kaohsiung Medical University, Kaohsiung, Taiwan,* [g] *Department of Physical Medicine and Rehabilitation, Chung San Medical University Hospital, Taichung, Taiwan,* [h] *Department of Physical Medicine and Rehabilitation, Chi Mei Medical Center, Tainan, Taiwan.*

[*] *Correspondence: Willy Chou, Department of Childhood Education and Nursery, Chia Nan University of Pharmacy and Science, Tainan 700, Taiwan (e-mail: ufan0101@ms22.hinet.net).*

How to cite this article: Lin JK, Chien TW, Wang LY, Chou W. An artificial neural network model to predict the mortality of COVID-19 patients using routine blood samples at the time of hospital admission: Development and validation study. Medicine 2021;100:28(e26532).

Received: 28 December 2020 / Received in final form: 14 June 2021 / Accepted: 15 June 2021

http://dx.doi.org/10.1097/MD.0000000000026532

Abbreviations: AI = artificial intelligence, LDH = lactate dehydrogenase, PMCP = predict the mortality of COVID-19 patients.

Keywords: app, artificial neural network, convolutional neural network, Google Maps, predict the mortality of COVID-19 patients, receiver operating characteristic curve

Highlights

- We applied an artificial neural network (ANN) to design an app to help health care providers fight COVID-19.
- The ANN and CNN were compared to each other in model performance to PMCP at hospital-admission time.
- Feature variables and prediction results were visualized using the forest plot and category probability curves shown on Google Maps, which is novel and innovative never seen before in the literature.

1. Introduction

COVID-19 is a highly contagious infection caused by SARS-CoV2.[1–3] As of June 10, 2021, COVID-19 cases and deaths are approaching 1.75 billion and 3.78 million, respectively, worldwide in more than 191 countries/regions.[4]

1.1. A prognostic prediction model is required

When a new disease (e.g., COVID-19) starts to spread, many questions emerge.[5,6] One of the most frequently asked questions is how high is the case fatality rate (CFR)[5] and what is the model to predict patients at risk of high mortality.[1,7] If we ignore the classification of patients at high risk with proper treatments, the condition of patients might rapidly deteriorate.[1] Numerous studies addressed that deceased patients in the COVID-19 epidemic initially had mild symptoms and then suddenly transitioned to a critical stage, or leading to death.[8–10] For example, over 75% of deceased patients in Italy presented mild symptoms(e.g., fever, dyspnea, and cough) at admission to the hospital.[3] In the COVID-19 pandemic, the shortage of resources and medical staff causes big problems (e.g., high case fatality rate [5]) in the health-care system.[1] As such, a prognostic model to predict mortality at an earlier stage is required for the development.

1.2. Traditional solutions and modern prediction models

Traditional solutions to classify the mortality of COVID-19 patients (PMCP) merely

1. predicted the mortality of individual patients more than 10 days to survival or death (i.e., at least 10 days from admission to hospital to death or discharge),[2]
2. compared model accuracies without an app provided to readers,[11–13]
3. applied machine learning techniques with professional software[14–17] instead of deep learning on the popular and friendly-use MS Excel (Microsoft Corp.),[18–20] and
4. displayed the prediction probability and the classification (e.g., death or survival) only with a statice measure (or diagram)[1,14] rather than a dashboard that can be practiced to interpret the prediction result on their own.

Accordingly, artificial intelligence (AI), helping healthcare providers efficiently and effectively classify COVID-19 patients is critical and of importance.

Furthermore, many types of prediction methods that can be used in prediction mortality on COVID-19, such as Logistic regression, Naïve Bayes, Decision trees, Random Forests, Gradient tree boosting,[21–30] and other artificial neural networks(e.g., a Feed forward Neural Network, a Radial Basis Function Neural Network, a Multilayer Perceptron, a Recurrent Neural Network, a Modular Neural Network, or a Sequence-To-Sequence Model).[31] None of the research applied artificial neural network (ANN) or convolutional neural network (CNN) to develop an app used for PMCP at the time of hospital admission.

Many studies have extracted feature variables and developed AI prediction models of mortality.[1,2,14–20] Some[1,2,11] applied blood samples to make a prediction of mortality for COVID-19 patients. However, none provided an app (or web application) with dynamical visualizations of prediction results.

1.3. Raw-laboratory and normalized data of feature variables

In machine learning, data normalization instead of raw data has been discussed in the literature.[32] Similar data dealt with different data formats(e.g., raw-laboratory and normalized data) were studid.[1,2] Although using normalization methods to remove various technical biases was suggested, there are no previous studies evaluating the impacts of normalization on disease diagnosis.[32] Feature variables might be seen based on either the original observed scores[2] or the corresponding normalized data[1] in machine learning. We are motivated to investigate whether the PMCP performance would be different using raw or normalized data.

1.4. Study objectives

Therefore, in this study, we aimed to

1. develop an ANN model to predict the mortality of COVID-19 patients at hospital-admission time, and
2. design an app, name PMCP app, on a public website so that all patients and medical staff could predict mortality of COVID-19 patients on their own.

2. Methods

We downloaded patient records from two studies, including 361 COVID-19 patients in Wuhan, China,[2] and 106 COVID-19 patients in 3 Korean medical institutions.[1] A total of 30 feature variables were retrieved, consisting of 28 blood biomarkers and 2 demographic variables (i.e., age and gender) of patients.[2]

All data[33] (Supplemental Content 1, http://links.lww.com/MD2/A274) used in this study were downloaded from the deposited datasets in the previous studies,[1,2] which means the study is not necessary for ethical approval according to the regulation promulgated by the Taiwan Ministry of Health and Welfare.

2.1. Feature variables

We replaced the missing data (i.e., 4.9% and 31.6% in training and testing sets,[1] respectively) with the mean value of each biomarker for training and testing data sets. Next, the mean and standard deviation (SD) were computed for each variable from the training set. Normalized data (i.e., mean = 0, SD = 1) were applied to the training and testing data sets[1] using the formula (=[observed scores – mean]/standard deviation [SD]). We then added 2 more features (i.e., age, and gender) to the 28 biomarkers and trained our AI model using 30 features.[2]

Forest plots[34,35] were drawn to present these 30 features. The standard mean difference (SMD) method was utilized to compare the differences in variables alone (like *t*-test) using the forest plot.

The Chi-Squared test was conducted to assess the heterogeneity between variables. The forest plots (confidence interval [CI] plot) were drawn to display the effect estimates and their CIs for each study.

2.2. Two models and 2 scenarios

Two models, namely, artificial neural network (ANN) and convolutional neural network (CNN), were compared with each other across 2 scenarios using

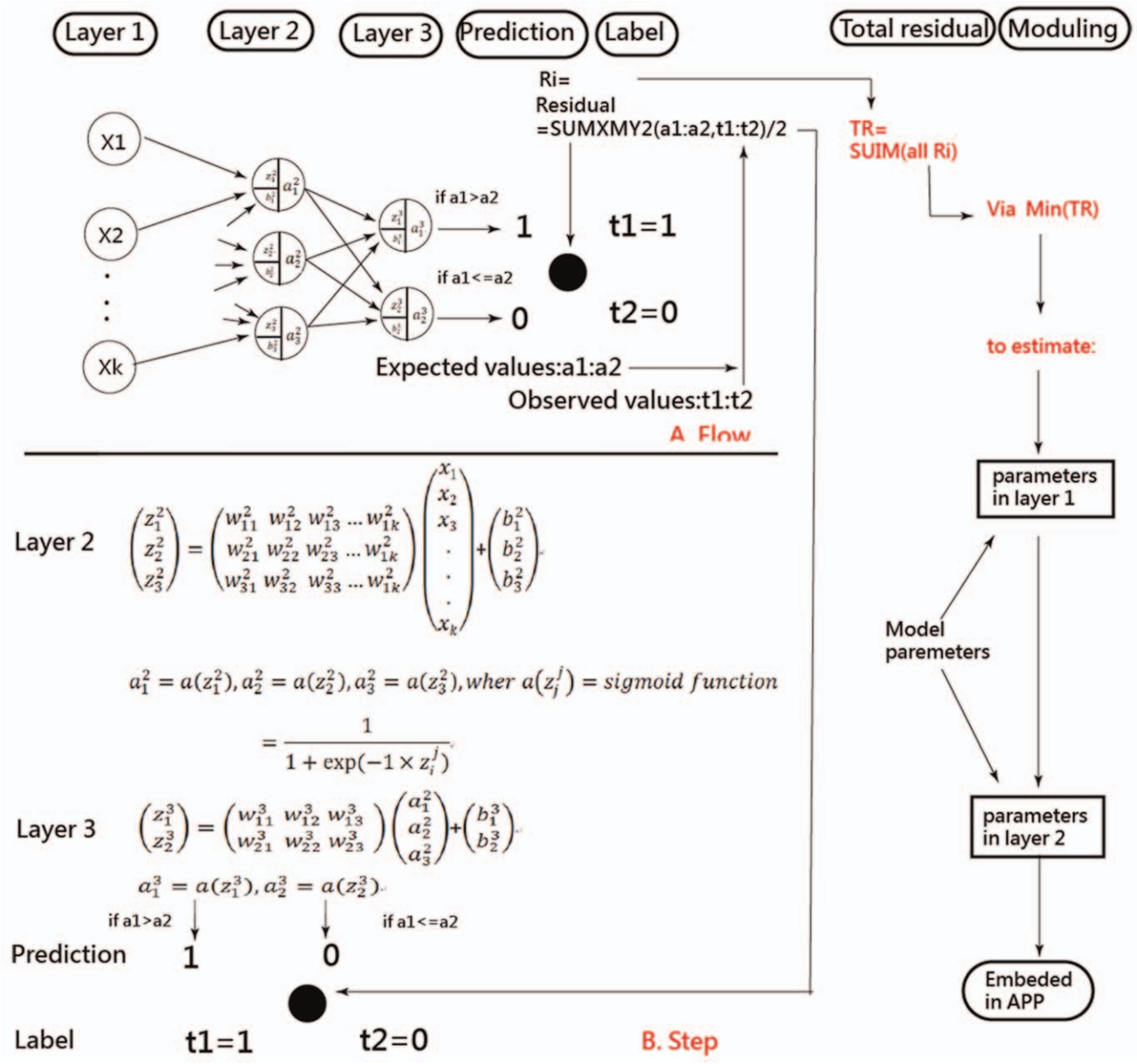

Figure 1. Process of estimating parameters in the ANN model.

1. raw laboratory versus normalized data and
2. training vs testing datasets (n = 361 and n = 106/361≅30%) to verify the model performance (e.g., sensitivity [SENS], specificity [SPEC], and area under the receiver operating characteristic curve [AUC]).

First, the 361-patient training and 106-patient testing datasets came from the studies,[1,2] where the training set was used to predict the testing set. Second, the accuracies (e.g., SENS, SPEC, and AUC) in training data used to predict the testing set were verified. The data are shown in reference.[33]

The artificial neural network (ANN) and convolutional neural network (CNN) were analyzed using the 2 scenarios (i.e., data types and model sets) previously mentioned.

The CNN has traditionally been performed on Microsoft (MS) Excel (Microsoft Corp.),[18–20] whereas the ANN has not been performed on MS Excel in the past. As illustrated in Figure 1, the ANN process involves data input in Layer 1 where the data are combined with 2 types of parameters and run through the sigmoid function algorithms in Layers 2 and 3. Finally, as shown on the right side and bottom of Figure 1, the prediction model was deemed complete when the total residuals were minimized through the MS Excel function of SUMXMY2 and Solver add-in.

2.3. Tasks for performing ANN and CNN

2.3.1. Task 1: Comparison of model performance on 2 datasets and 2 scenarios. Accuracy was determined by observing the high AUC along with indicators of SENS, SPEC, balanced accuracy, and accuracy in both models.

Comparisons of balanced accuracies across 2 scenarios were made in both ANN and CNN models based on the medians in respective models.

2.3.2. Task 2: Comparison of prediction models using the Weka software. To better understand the effectiveness and efficacy of the ANN and CNN models, several machine learning algorithms in the Weka software[36] were used to compare the ANN and CNN models and evaluate the high indicators of SENS, SPEC, balanced accuracy, and AUC in 2 data types (i.e., raw and normalized data) and 2 types of the model mentioned in the previous section.

2.3.3. Task 3: Developing an App for PMCP. An app, named PMCP app, for early PMCP, was designed and developed. Model parameters were embedded in the computer module. The results of the classification (i.e., death and survival) would instantly appear on smart phones. The visual representation with binary (i.e., deceased and survival) categorical probabilities are shown on a dashboard displayed on Google Maps.[18–20]

2.4. Statistical tools and data analysis

An author-made MS Excel VBA module was applied to draw the forest plot and perform the ANN and CNN algorithms. The significance level of Type I error was set at 0.05. ANN and CNN were performed on MS Excel as well.

A visual representation of the classification was plotted using two curves based on the Rasch model.[37] The study flowchart and the ANN modeling process are shown in Figure 2 and Supplemental Content 1, http://links.lww.com/MD2/A275.

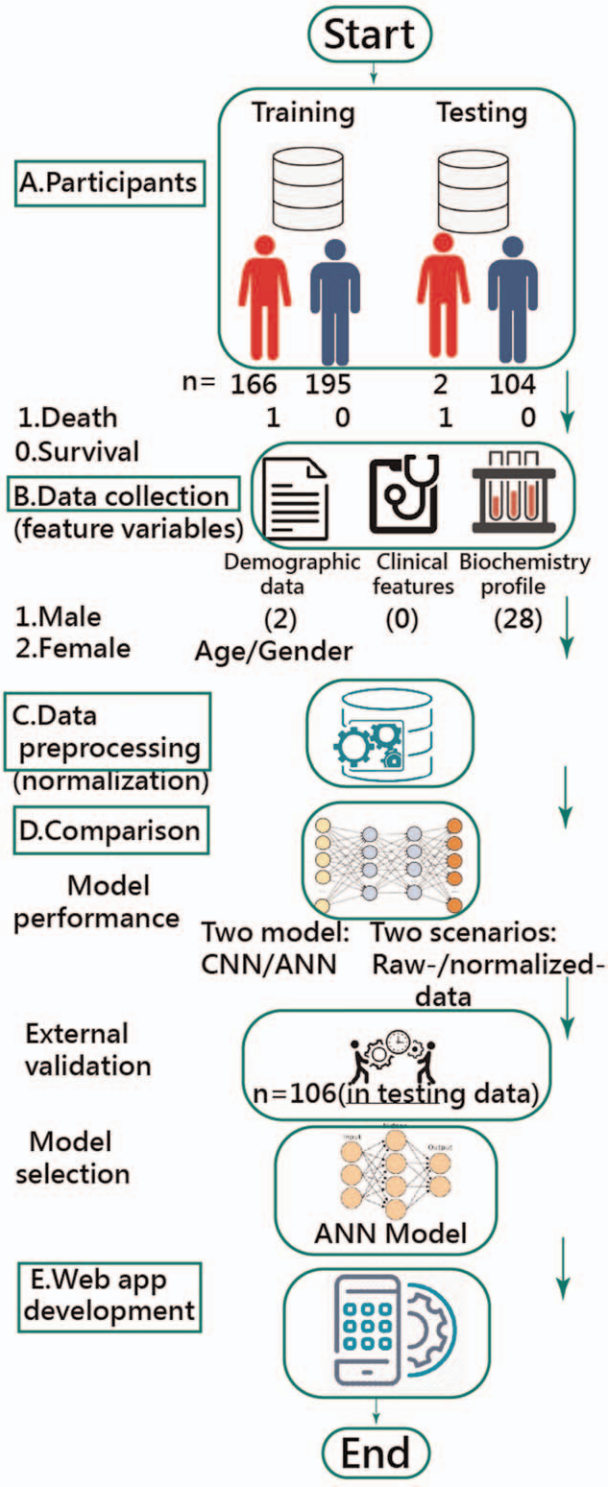

Figure 2. Study flowchart.

3. Results

3.1. Four hundred sixty seven patients in training/testing data and 30 feature variables

For training data, we used the blood test results obtained from 375 COVID-19 patients collected between January 10, 2020, and February 24, 2020, in Tongji Hospital, Wuhan, China.[2] Of

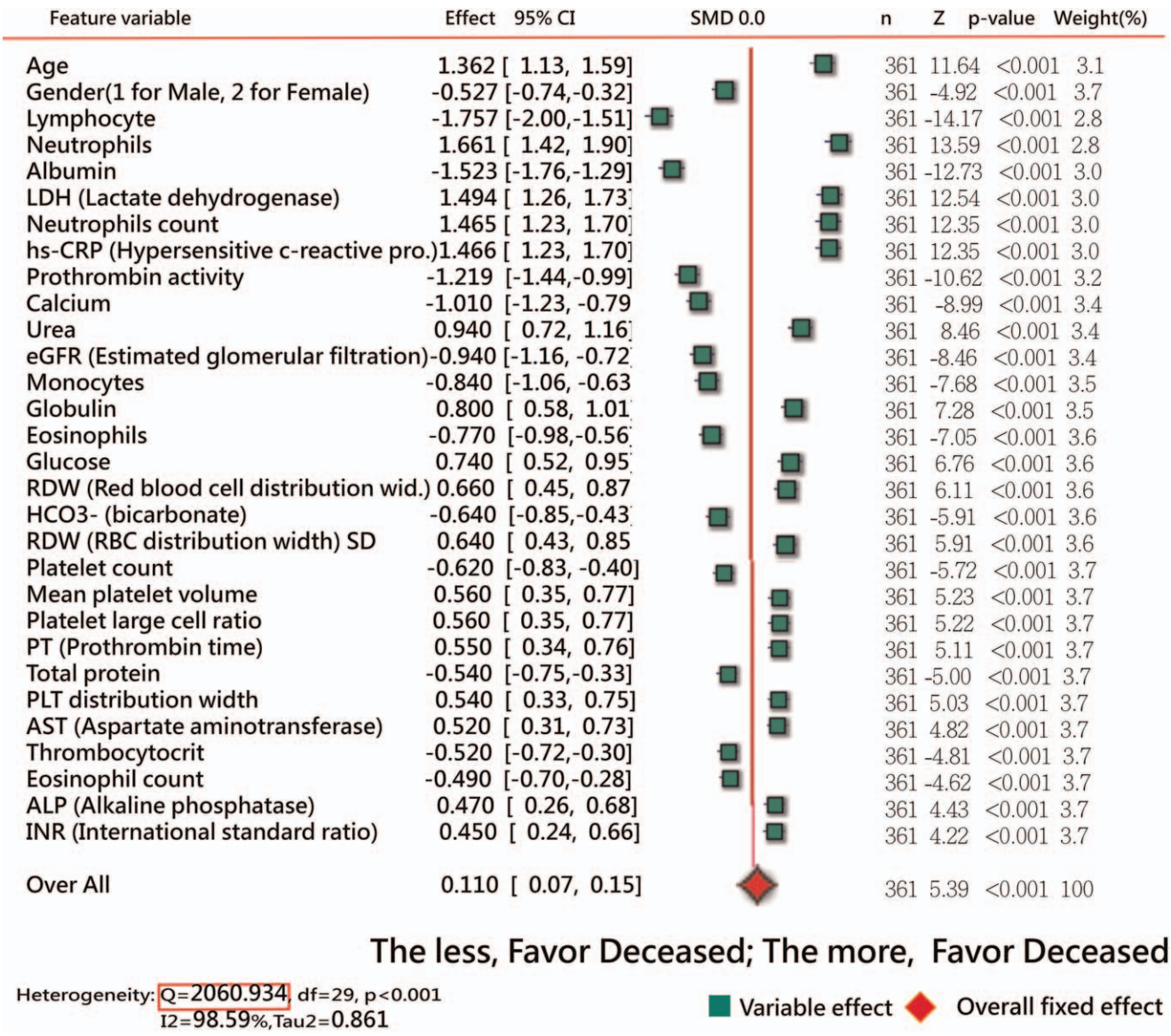

Figure 3. Thirty feature variables used in this study using the forest plot to display.

them, 14 patients without a blood test within 1 day after the hospital admission were excluded, and 361 patients (212 males, 58.7%; 149 females, 41.3%; mean age 58.9 years, SD = 16.5) were included with information of age, gender, mortality outcome, and results of blood tests obtained within 24 hours after hospital admission. For testing data, we downloaded 106-patient data (similar to that in the training set) between February 2020 and July 2020 from 3 Korean medical institutions.

The 30 feature variables from the previous study[1] were displayed using the forest plot in Figure 3, where the age showing the older with a high probability of death on the right-hand side due to older, and the gender on the left-hand side, indicating the male (i.e., coded 1 vs female coded 2) with a tendency toward death due to a lower value. As such, we can easily discriminate which variables with the tendency toward death are positively (the more) to the right panel or negatively (the less) to the left part. The highest Z-score (=−14.17) is at the variable of Lymphocyte and the least (=4.22) at INT (International standard ratio).

The Q-statistic is 2,060,934 with degrees of freedom = 29 ($P < .001$), indicating that the SMD of the 30 variables is significantly different. Readers are invited to click on the links[38] and examine the detail of the feature variables on the online forest plot.

3.2. Comparison of model performance in 2 types of data formats

3.2.1. AUC equals 0.96 in ANN using normalized data. Three parts are separated in Table 1, including (A) raw data, (B) normalized data, and (C) the models from the previous study.[1] We can see that

1. the normalized dataset gains a higher AUC(>0.9) when compared to that (<0.9) in the raw-laboratory dataset when we referred to 95% CIs of AUC in CNN and ANN (see the footnote beneath Table 1),
2. the normalized dataset in ANN yielded a high AUC of 0.96 that that(=0.91) if 0.05 CIs of AUC was considered.

Table 1

Comparison of model accuracy using normalized training data (n = 361).

Model	Sensitivity	Specificity	Accuracy	Balanced accuracy	AUC*
A. Raw data					
CNN	0.90	0.86	0.88	0.88	0.85
ANN	0.95	0.88	0.91	0.92	0.89
Forest (J48)	0.99	0.99	0.99	0.99	0.99
Random forest	1.00	1.00	1.00	1.00	1.00
Random tree	1.00	1.00	1.00	1.00	1.00
REPT tree	0.88	0.91	0.89	0.90	0.88
BayesNet	0.92	0.92	0.92	0.92	0.90
Naïve Bayes	0.78	0.95	0.87	0.87	0.86
Logistic	0.94	0.92	0.93	0.93	0.91
SMO	0.89	0.91	0.09	0.90	0.88
Median				0.92	
B. Normalized data					
CNN	0.93	0.93	0.92	0.93	0.91
ANN	0.95	0.98	0.97	0.97	0.96
Forest (J48)	0.99	0.97	0.98	0.98	0.97
Random forest	1.00	1.00	1.00	1.00	1.00
Random tree	1.00	1.00	1.00	1.00	1.00
REPT tree	0.92	0.93	0.93	0.93	0.91
BayesNet	0.92	0.92	0.92	0.92	0.90
Naïve Bayes	0.80	0.92	0.87	0.86	0.84
Logistic	0.93	0.92	0.93	0.93	0.91
SMO	0.87	0.93	0.91	0.90	0.88
Median				0.93	
C. Ko et al[1]					
Random forest	0.89	0.89	0.89	0.89	0.87
Deep neural network	0.91	0.93	0.92	0.92	0.90
EDRnet	0.92	0.93	0.93	0.93	0.91
Median				0.92	

* The 95% CIs of AUC (=0.9) = $1.96 \times \sqrt{0.9 \times 0.1}/361 = 0.03$ for the training dataset.

3.2.2. AUC equals 0.80 in ANN using normalized data. Similarly, we examine the model performances using the training data to predict the testing data in Table 2. The CNN has a higher balanced accuracy (=0.86) than that in ANN (=0.82) based on the raw-laboratory data. In contrast, the ANN has a higher balanced accuracy (=0.80) than that in CNN (=0.73) based on the normalized data.

3.2.3. Model accuracies in ANN and CNN compared to other models. Basically, all those models have equivalently equal balanced accuracies, see the median of 0.92, 0.93, 0.92 (at the bottoms in these 3 panels of Table 1), and the footnote of AUC 95% CIs beneath Table 1. The reason for explaining the difference in AUC (=0.87 and 1.0) using Random forest methods is worthy of investigation in the discussion section due to similar data used in our study and the previous research.[1]

The median (=0.86) in the normalized data is higher than that (=0.84) in the raw-laboratory data, but the ANN has a higher balanced accuracy (=0.82) in raw-laboratory data than that (=0.80) in normalized data if the 95% CIs were ignored.

If the 95% CIs of AUC are considered in the normalized data, the ANN has 0.72 and 0.88 in the range of AUC, equal to the CNN ranged between 0.65 and 0.81, and other models, but below the EDRnet with the boundary limits between 0.92 and 0.99 reported in the previous study.[1]

3.3. Web app development for PMCP used in this study

The interface of the PMCP app is shown on the left-hand side of Figure 4. Readers are invited to click on the links[39] and interact with the PMCP app. Notably, all 53 model parameters are embedded in the 30-item ANN model. Once the responses after clicking on the icon of the test[39] are submitted, the app generates the result (shown on the right-hand side of Fig. 4) as a classification of either possible death or survival on smart phones.

An example in which the patient scored with a high probability (0.96) of survival is shown on the right-hand side of Figure 4. The curve starts from the top-left corner to the bottom-right corner. The sum of the probabilities of death and survival is 1.0. The odds ratio can be calculated using the formula p/[1−p] (= 0.04/0.96 = 0.04), indicating that this COVID19 patient has a high probability of survival within the later treatments in hospitalization.

4. Discussion

4.1. Principal findings

We observed that

1. the normalized dataset gains a higher AUC(>0.9) when compared to that(<0.9) in the raw-laboratory dataset based on training data,
2. the normalized dataset in ANN yielded a high AUC of 0.96 that that(=0.91) in CNN on testing data, and
3. a ready and available PMCPapp was developed for anyone who can practice the model to predict mortality of COVID-19 patients on their own.

4.2. What this finding adds to what we already knew

The sudden increase in COVID-19 cases drives high pressure to healthcare services and requires a rapid, accurate, and early

Table 2

Model comparison using training data to predict testing data (n = 106).

Model	True negative	False-positive	False-negative	True positive	Sensitivity	Specificity	Accuracy	Balanced Accuracy
A.Raw data								
CNN	75	29	0	2	1	0.72	0.73	0.86
ANN	66	38	0	2	1	0.63	0.64	0.82
Forest(J48)	82	22	1	1	0.5	0.78	0.78	0.64
Random forest	88	16	0	2	1	0.85	0.85	0.93
Random tree	78	26	1	1	0.5	0.75	0.75	0.63
REPT tree	51	53	0	2	1	0.49	0.50	0.75
BayesNet	85	19	0	2	1	0.82	0.82	0.91
Naïve Bayes	85	19	0	2	1	0.82	0.82	0.91
Logistic	22	82	1	1	0.5	0.21	0.23	0.36
SMO	92	12	0	2	1	0.89	0.89	0.95
Median								0.84
B. Normalization data								
CNN	48	56	0	2	1	0.46	0.47	0.73
ANN	62	42	0	2	1	0.6	0.60	0.80
Forest(J48)	81	23	1	1	0.5	0.78	0.77	0.64
Random forest	82	22	0	2	1	0.79	0.79	0.90
Random tree	88	16	0	2	1	0.85	0.85	0.93
REPT tree	73	31	0	2	1	0.7	0.71	0.85
BayesNet	86	18	0	2	1	0.83	0.83	0.92
Naïve Bayes	86	18	0	2	1	0.83	0.83	0.92
Logistic	75	29	0	2	1	0.72	0.73	0.86
SMO	91	10	1	1	0.5	0.9	0.9	0.70
Median								0.86
C. From Ko et al.								
XGBoost	80	24	0	2	1	0.77	0.77	0.88
AdaBoost	81	23	0	2	1	0.78	0.78	0.89
Random forest	87	17	0	2	1	0.84	0.84	0.92
Deep neural network	95	9	1	1	0.5	0.91	0.91	0.71
DNN + XGBoost	80	24	0	2	1	0.77	0.77	0.88
DNN + AdaBoost	96	8	1	1	0.5	0.92	0.92	0.71
Yan et al model	36	68	0	2	1	0.35	0.36	0.67
EDRnet	95	9	0	2	1	0.91	0.92	0.96
Median								0.88

*The 95% CIs of AUC (=0.9) = $1.96 \times \sqrt{0.9 \times 0.1}/106 = 0.057$ for the testing dataset.

clinical assessment of the disease severity in the hospital.[2] The highest Z-score(=−14.17) is at Lymphocyte, the fourth at lactate dehydrogenase (LDH) LDH=12.54, and the sixth at high-sensitivity C-reactive protein (hs-CRP) (=12.35), see Figure 3.

Numerous studies have shown that the disease progression of COVID-19 is not only associated with lymphocyte,[2,40,41] LDH,[2,7,42–45] and hs-CRP,[2,7,44,46–49] but also with other blood-based biomarkers, such as neutrophil counts(at the 5th in Fig. 3),[48,50,51] albumin(at the third in Fig. 3),[50,52,53] and prothrombin activity (at the 7th in Fig. 3).[50,54–56]

Similarly, Yan et al applied the mortality prediction model with 3 blood biomarkers (i.e., LDH, lymphocyte, and hs-CRP) to predict mortality with 90% accuracy.[2] Hu et al proposed 4 variables of age, hs-CRP, lymphocyte count, and d-dimer level to PMCP with AUC=0.881.[14] Ko et al developed an AI model, EDRnet, to PMCP using 28 blood biomarkers obtained within 24 hours after hospital admission along with 2 demographic data of age and gender (similar to the current study) with high sensitivity (100%), specificity (91%), and accuracy (92%).[1]

However, Yan et al[2] predicted mortality ten days before the occurrence of survival or death instead of blood biomarkers obtained within 24 hours. Instead, the mortality prediction at the time of admission can be substantially more informative for clinicians because the critical time regarding disease progression is 10 to 14 days from the onset of symptoms.[41,48,57]

Many studies compared model performance as we did in Tables[11–13] and performed machine learning using professional software[14–17] instead of the popular and friendly-use MS Excel (Microsoft Corp.)[18–20] as we did in Figure 1. A few[1,14] provided an app to display the prediction probability and the classification (e.g., death or survival) rather than an animation-type dashboard to interpret the prediction result as Figure 4 shown in this study.

4.3. Strengths of this study

Besides 3 features, including

1. ANN model performed in Microsoft Excel,
2. an app combined with ANN and PMCP displayed on a dashboard, and
3. the online forest plot used to present feature variables, in this study, other 2 are

1. many prediction models compared in a study using the same data to present the model accuracy, and
2. model accuracy compared between raw data and normalized data.

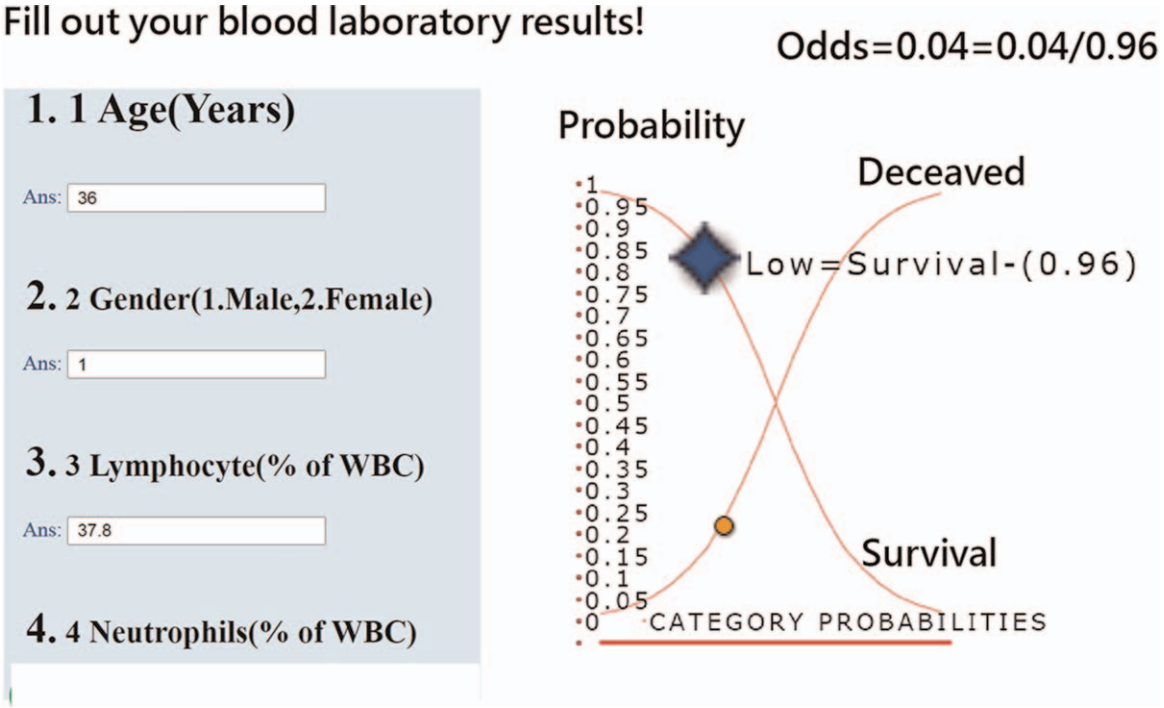

Figure 4. Snapshot of an app designed in this letter.

ANN[58,59] was performed on MS Excel, which has not been reported in the literature, but is easily understood by readers who are familiar with Microsoft Excel. The animation-type dashboard enable readers practice the app on their own with an easy understanding of the classification resultswe.

Furthermore, the online forest plot is unique and modern to display feature variables in comparison with each other,[38] unlike others commonly using a simple table with *P* value to present feature variables. It is novel and easy to understand

1. the direction of a feature toward death on either the left (with the less value) or right (with the more value) side, and
2. the importance of a feature with the Z-score.
3. For instance, Yan et al[2] plotted a decision rule using 3 key features and their thresholds to PMCP. In which, LDH$\leq$365 UI^{-1} denotes death (similar to the right side in Fig. 3), Lymphocyte$\leq$14.7% represents death (similar to the right side in Fig. 3), and hs-CRO$\geq$41.2 mgI^{-1} possibly stands for death (similar to the right side in Fig. 3).

The different types of algorithms for classification in machine learning[60,61] were compared each other, include logistic regression, support vector machine, naïve Bayes, random forest classification, ANN, CNN,[18–20] and *k*-nearest neighbor, by observing their 95% CIs. In addition, whether data in machine learning should be normalized has been evident in Tables 1 and 2, which provide an reference to the furture study related to the machine learning.

4.4. Implications and future work

A single-center usually attempts to increase model performance.[62] As expected, the model accuracies reached 90% and 0.881 for 2 studies,[2,14] respectively, in Tongji Hospital, Wuhan, China.[2] Surprisingly, Ko et al gained a higher (not lower as expected) accurate rate (92%, see Random forest at the bottom panel in Table 2) using data in China to predict the testing data in 3 Korean medical institutions.[1] Two questions were thus raised: why results of Random forest are different from 2 studies using similar data

1. in training datasets(e.g., 1 vs 0.89 in Table 1), and
2. in testing datasets (e.g., 0.90 vs 0.92 in Table 2) to externally validate the final predictive model.

We doubt the inconsistent result ($0.92 > 0.87$ in testing and training data, respectively[1]) that multicenter models reasonably exhibit poor performance,[62] which should be cautious and discussed in the future.

The ANN exhibited better accuracy than CNN in normalized data (Table 2). We have not seen others using the ANN approach to predict PMCP in the literature, which is a breakthrough in this study. We have also not noticed any article incorporating indicators of raw and normalized data to compare model performance, which is also required for verifications in the future.

More than 2114 articles searched using the keyword "artificial neural network" [Title] have been found in PubMed Central on December 26, 2020.[63] None of the articles used MS Excel to perform the ANN. The interpretations of the ANN concept and process, as well as the parameter estimations, are shown in Figure 1, Supplemental Content 1, http://links.lww.com/MD2/A275, and the app.[39] Readers can adjust the input values in the ANN model on their own and examine the differences in results.

In addition to the performance of the ANN model (i.e., Balanced accuracy = 0.96 and 0.80 with normalized raw data in training and testing data, respectively), we considered its generalizability in the future. As did in the previous study,[1] researchers are encouraged to apply the original training data[33]

to PMCP against the testing data of their own for COVD-19 patients in the future.

The categorical probability curves are shown in Figure 4. The binary categories (e.g., success and failure of an assessment in the psychometric field) have been frequently applied in health-related outcomes.[18–20,64–66] However, we are the first to provide the categorical probability curves of the PMCP dashboard displayed on Google Maps (Fig. 4).

4.5. Limitations and suggestions

Although our model was designed to be specific to COVID-19 patients, it does not work for patients of pregnant and breastfeeding women, patients younger than 18 years, recordings with data material less than 80% complete,[2] and patients without a blood test within 1 day after the hospital admission according to the definition in the previous study.[1]

Second, we have not discussed possible further improvement in predictive accuracy. Whether other feature variables (e.g., variables not shown in Fig. 3) should be applied to the ANN model to increase the accuracy rate is worth discussing. In the future, it would be useful to look for other variables that can improve the power of the PMCP model, such as underlying comorbidities.

Third, the study was performed using the ANN model. Whether other prediction models not illustrated in Table 1 have higher accuracy than the ANN model has yet to be investigated.

Fourth, the study patients in training data were taken from Tongji Hospital, Wuhan, China.[2] The model parameters estimated for PMCP are only suitable for the Chinese COVID-19 patients because geolocation is associated with socioeconomic status, which has been reported to be linked to the difference.[4]

Similarly, the testing data were from 3 Korean medical institutions.[1] The results in Table 2 cannot be generalized to other disparate regions.

Thus, a generalization of these PMCP findings (e.g., the model parameters) needs to be cautiously done because the sample only included patients with COVID-19 aged ≥18 years in China. Additional studies in other countries are required in the future to re-examine the feature variables that are similar to those used in this study (Supplemental Content 1: Abstract video of this study at https://youtu.be/rBVMehC7TgI).

5. Conclusion

In this study,

1. ANN is performed on MS Excel,
2. an online app is built to display the results using the visual dashboard on Google Maps, and
3. the categorical probability curves based on the Rasch model are combined with the ANN prediction model.

The novelty of the app with our ANN algorithm improves the accuracy of predicting PMCP up to AUC = 0.96 in the normalized training data. The integration of this app into the hospital information system is encouraged to the hospital, not limited to Korea as we did in this study.

Acknowledgments

We thank Enago (www.enago.tw) for the English language review of this manuscript.

Supplemental Content 1: Data source deposited at https://osf.io/n2vts/

Author contributions

Conceptualization: Tsair-Wei Chien.
Data curation: Tsair-Wei Chien.
Formal analysis: Ju-Kuo Lin, Tsair-Wei Chien.
Investigation: Lin-Yen Wang, Willy Chou.
Methodology: Tsair-Wei Chien.
Resources: Tsair-Wei Chien.
Software: Tsair-Wei Chien.
Supervision: Willy Chou.

References

[1] Ko H, Chung H, Kang WS, et al. An artificial intelligence model to predict the mortality of COVID-19 patients at hospital admission time using routine blood samples: development and validation of an ensemble model. J Med Internet Res 2020;22:e25442. doi: 10.2196/25442.
[2] Yan L, Zhang H, Goncalves J, et al. An interpretable mortality prediction model for COVID-19 patients. Nat Mach Intell 2020;2:283–8.
[3] COVID-19 Surveillance Group. Characteristics of COVID-19 Patients Dying in Italy: Report Based on Available Data on March 20th, 2020. Rome, Italy: Istituto Superiore di Sanità; 2020. 2020/12/25 available at https://www.epicentro.iss.it/coronavirus/bollettino/Report-COVID-2019_20_marzo_eng.pdf
[4] Johns Hopkins. COVID-19 Dashboard. 2020/12/25 available at https://gisanddata.maps.arcgis.com/apps/opsdashboard/index.html#/85320e2ea5424dfaaa75ae62e5c06e61
[5] Chang CS, Yeh YT, Chien TW, Lin JJ, Cheng BW, Kuo SC. The computation of case fatality rate for novel coronavirus (COVID-19) based on Bayes theorem: An observational study. Medicine (Baltimore) 2020;99:e19925. doi: 10.1097/MD.0000000000019925.
[6] Jen TH, Chien TW, Yeh YT, Lin JJ, Kuo SC, Chou W. Geographic risk assessment of COVID-19 transmission using recent data: An observational study. Medicine (Baltimore) 2020;99:e20774.
[7] Li C, Ye J, Chen Q, et al. Elevated lactate dehydrogenase (LDH) level as an independent risk factor for the severity and mortality of COVID-19. Aging (Albany NY) 2020;12:15670–81.
[8] Luo L, Xu M, Du M, et al. Early coagulation tests predict risk stratification and prognosis of COVID-19. Aging (Albany NY) 2020;12:15918–37.
[9] Lai C, Liu YH, Wang C, et al. Asymptomatic carrier state, acute respiratory disease, and pneumonia due to severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2): Facts and myths. J Microbiol Immunol Infect 2020;53:404–12.
[10] Sohrabi C, Alsafi Z, O'Neill N, et al. World Health Organization declares global emergency: a review of the 2019 novel coronavirus (COVID-19). Int J Surg 2020;76:71–6.
[11] Vaid A, Somani S, Russak AJ, et al. Machine learning to predict mortality and critical events in a cohort of patients with COVID-19 in New York City: model development and validation. J Med Internet Res 2020;22:e24018. doi: 10.2196/24018. PMID: 33027032; PMCID: PMC7652593.
[12] Wollenstein-Betech S, Silva AAB, Fleck JL, Cassandras CG, Paschalidis IC. Physiological and socioeconomic characteristics predict COVID-19 mortality and resource utilization in Brazil. PLoS One 2020;15:e0240346. doi: 10.1371/journal.pone.0240346.
[13] Yadaw AS, Li YC, Bose S, Iyengar R, Bunyavanich S, Pandey G. Clinical features of COVID-19 mortality: development and validation of a clinical prediction model. Lancet Digit Health 2020;2:e516–25.
[14] Hu C, Liu Z, Jiang Y, et al. Early prediction of mortality risk among patients with severe COVID-19, using machine learning. Int J Epidemiol 2020;dyaa171. doi: 10.1093/ije/dyaa171. Epub ahead of print. PMID: 32997743; PMCID: PMC7543461.
[15] Sánchez-Montañés M, Rodríguez-Belenguer P, Serrano-López AJ, Soria-Olivas E, Alakhdar-Mohmara Y. Machine learning for mortality analysis in patients with COVID-19. Int J Environ Res Public Health 2020;17:8386. doi: 10.3390/ijerph17228386. PMID: 33198392; PMCID: PMC7697463.

[16] Clift AK, Coupland CAC, Keogh RH, et al. Living risk prediction algorithm (QCOVID) for risk of hospital admission and mortality from coronavirus 19 in adults: national derivation and validation cohort study. BMJ 2020;371:m3731. doi: 10.1136/bmj.m3731. PMID: 33082154; PMCID: PMC7574532.
[17] Colombi D, Villani GD, Maffi G, et al. Qualitative and quantitative chest CT parameters as predictors of specific mortality in COVID-19 patients. Emerg Radiol 2020;27:701–10.
[18] Yan YH, Chien TW, Yeh YT, Chou W, Hsing SC. An app for classifying personal mental illness at workplace using fit statistics and convolutional neural networks: survey-based quantitative study. JMIR Mhealth Uhealth 2020;8:e17857. doi: 10.2196/17857. PMID: 32735232; PMCID: PMC7428910.
[19] Lee YL, Chou W, Chien TW, Chou PH, Yeh YT, Lee HF. An app developed for detecting nurse burnouts using the convolutional neural networks in microsoft excel: population-based questionnaire study. JMIR Med Inform 2020;8:e16528. doi: 10.2196/16528. PMID: 32379050; PMCID: PMC7243132.
[20] Ma SC, Chou W, Chien TW, et al. An app for detecting bullying of nurses using convolutional neural networks and web-based computerized adaptive testing: development and usability study. JMIR Mhealth Uhealth 2020;8:e16747. doi: 10.2196/16747. PMID: 32432557; PMCID: PMC7270851.
[21] Nelder JA, Wedderburn RWM. Generalized linear models. Journal of the Royal Statistical Society Series A (General) 1972;135:370.
[22] Cox DR. The regression analysis of binary sequences. J Royal Statistical Soc 2018;20:215–32.
[23] Szumilas M. Explaining odds ratios. J Can Acad Child Adolesc Psychiatry 2010;19:227–9.
[24] Bayes T, Price RLII. An essay towards solving a problem in the doctrine of chances. By the late Rev. Mr. Bayes, F. R. S. communicated by Mr. Price, in a letter to John Canton, A. M. F. R. S. Phil. Trans R Soc 1997;53:370–418.
[25] Domingos P, Pazzani M. On the optimality of the simple bayesian classifier under zero-one loss. Mach Learn 1997;29:103–30.
[26] Quinlan JR. Induction of decision trees. Mach Learn 1986;1:81–106.
[27] Breiman L. Sammut C, Webb GI. Random Forests. Encyclopedia of Machine Learning and Data Mining Boston, Massachusetts: Springer; 2017;32.
[28] Tin KH. Random decision forests. Proceedings of 3rd International Conference on Document Analysis and Recognition; 14-16 Aug; Montreal, Canada. 1995.
[29] Friedman JH. Stochastic gradient boosting. Computational Statistics Data Analysis 2002;38:367–78.
[30] Freund Y, Mason L. The Alternating Decision Tree Algorithm. Proceedings of the 16th International Conference on Machine Learning; 26-27 June; Ithaca, New York. 1999.
[31] Mehta A. Digital Vidya. 2019. Jan 25,.[2020-02-12] A Comprehensive Guide to Types of Neural Networks https://www.digitalvidya.com/blog/types-of-neural-networks/
[32] Han H, Men K. How does normalization impact RNA-seq disease diagnosis? J Biomed Inform 2018;85:80–92.
[33] Chien TW. Code, data, and results for this article. 2020/11/27 available at https://osf.io/n2vts/
[34] Hamling J, Lee P, Weitkunat R, Ambühl M. Facilitating meta-analyses by deriving relative effect and precision estimates for alternative comparisons from a set of estimates presented by exposure level or disease category. Stat Med 2008;27:954–70.
[35] Chen CJ, Wang LC, Kuo HT, Fang YC, Lee HF. Significant effects of late evening snack on liver functions in patients with liver cirrhosis: A meta-analysis of randomized controlled trials. J Gastroenterol Hepatol 2019;34:1143–52.
[36] Hall M, Frank E, Holmes G, Pfahringer B, Reutemann P, Ian H. The WEKA Data Mining Software: An Update. SIGKDD Explorations 2009; 11(1). 2020/11/20 available at https://www.kdd.org/exploration_files/p2V11n1.pdf
[37] Rasch G. Probabilistic Models for Some Intelligence and Attainment Tests. Chicago, US: University of Chicago Press; 1980.
[38] Chien TW. Using the forest plot to interpret the feature variables. 2020/12/25 at http://www.healthup.org.tw/gps/covidlab2020.htm
[39] Chien TW. PMCPapp used for input data. 2020/12/24 available at http://www.healthup.org.tw/irs/irsin_e.asp?type1=96
[40] Huang I, Pranata R. Lymphopenia in severe coronavirus disease-2019 (COVID-19): systematic review and meta-analysis. J Intensive Care 2020;8:36.
[41] Ruan Q, Yang K, Wang W, Jiang L, Song J. Clinical predictors of mortality due to COVID-19 based on an analysis of data of 150 patients from Wuhan, China. Intensive Care Med 2020;46:846–8.
[42] Wu MY, Yao L, Wang Y, et al. Clinical evaluation of potential usefulness of serum lactate dehydrogenase (LDH) in 2019 novel coronavirus (COVID-19) pneumonia. Respir Res 2020;21:171.
[43] Henry BM, de Oliveira MHS, Benoit S, Plebani M, Lippi G. Hematologic, biochemical and immune biomarker abnormalities associated with severe illness and mortality in coronavirus disease 2019 (COVID-19): A meta-analysis. Clin Chem Lab Med 2020;58:1021–8.
[44] Henry BM, Aggarwal G, Wong J, et al. Lactate dehydrogenase levels predict coronavirus disease 2019 (COVID-19) severity and mortality: a pooled analysis. Am J Emerg Med 2020;38:1722–6.
[45] Panteghini M. Lactate dehydrogenase: An old enzyme reborn as a COVID-19 marker (and not only). Clin Chem Lab Med 2020;58:1979–81.
[46] Wang L. C-reactive protein levels in the early stage of COVID-19. Med Mal Infect 2020;50:332–4.
[47] Wang D, Hu B, Hu C, et al. Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus-infected pneumonia in Wuhan, China. JAMA 2020;323:1061–9.
[48] Guan W, Ni Z, Hu Y, Liang W, Ou C, He J. China Medical Treatment Expert Group for Covid-19. Clinical characteristics of coronavirus disease 2019 in China. N Engl J Med 2020;382:1708–20.
[49] Tan C, Huang Y, Shi F, et al. C-reactive protein correlates with computed tomographic findings and predicts severe COVID-19 early. J Med Virol 2020;92:856–62.
[50] Long H, Nie L, Xiang X, et al. D-dimer and prothrombin time are the significant indicators of severe COVID-19 and poor prognosis. Biomed Res Int 2020;2020:6159720.
[51] Kong M, Zhang H, Cao X, Mao X, Lu Z. Higher level of neutrophil-to-lymphocyte is associated with severe COVID-19.
[52] Aziz M, Fatima R, Lee-Smith W, Assaly R. The association of low serum albumin level with severe COVID-19: A systematic review and meta-analysis. Crit Care 2020;24:255.
[53] Aziz M, Fatima R, Assaly R. Elevated interleukin-6 and severe COVID-19: a meta-analysis. J Med Virol 2020;92:2283–5.
[54] The Lancet Haematology. COVID-19 coagulopathy: an evolving story. Lancet Haematol 2020;7:e425.
[55] Arachchillage DRJ, Laffan M. Abnormal coagulation parameters are associated with poor prognosis in patients with novel coronavirus pneumonia. J Thromb Haemost 2020;18:1233–4.
[56] Bowles L, Platton S, Yartey N, et al. Lupus anticoagulant and abnormal coagulation tests in patients with COVID-19. N Engl J Med 2020;383:288–90.
[57] Qian G, Yang N, Ding F, et al. Epidemiologic and clinical characteristics of 91 hospitalized patients with COVID-19 in Zhejiang, China: a retrospective, multi-centre case series. QJM 2020;113:474–81.
[58] Falissard L, Morgand C, Roussel S, et al. A deep artificial neural network-based model for prediction of underlying cause of death from death certificates: algorithm development and validation. JMIR Med Inform 2020;8:e17125. doi: 10.2196/17125. PMID: 32343252; PMCID: PMC7218605.
[59] Frenkenfield J. Artificial neural network (ANN).2020/10/10 available at https://www.investopedia.com/terms/a/artificial-neural-networks-ann.asp
[60] Mellors BOL, Spear AM, Howle CR, Curtis K, Macildowie S, Dehghani H. Machine learning utilising spectral derivative data improves cellular health classification through hyperspectral infrared spectroscopy. PLoS One 2020;15:e0238647. Published 2020 Sep 15. doi:10.1371/journal.pone.0238647.
[61] Singh VK, Maurya NS, Mani A, Yadav RS. Machine learning method using position-specific mutation based classification outperforms one hot coding for disease severity prediction in haemophilia 'A' [published online ahead of print, 2020 Sep 11]. Genomics 2020;S0888-7543(20)30819-3. doi:10.1016/j.ygeno.2020.09.020.
[62] Mather JF, Fortunato GJ, Ash JL, Davis MJ, Kumar A. Prediction of pneumonia 30-day readmissions: a single-center attempt to increase model performance. Respir Care 2014;59:199–208.
[63] PMC. Articles with title in artificial neural network in Pubmed. 2020/12/26 available at https://pubmed.ncbi.nlm.nih.gov/?term=%E2%80%9Cartificial+neural+network%E2%80%9D+%5BTitle%5D

[64] Chien T, Lin W. Improving inpatient surveys: web-based computer adaptive testing accessed via mobile phone QR codes. JMIR Med Inform 2016;4:e8. doi: 10.2196/medinform.4313. https://medinform.jmir.org/2016/1/e8/.
[65] Hulin Cl , Drasgow F, Parsons C. Item Response Theory: Applications to Psychological Measurement. Homewood: Dow & Jones Irwin; 1983.
[66] Kim H, Hung WW, Paik MC, et al. Predictors and outcomes of unplanned readmission to a different hospital. Int J Qual Health Care 2015;27:513–9.