RESEARCH ARTICLE

Assessing the risk of dengue severity using demographic information and laboratory test results with machine learning

Sheng-Wen Huang[1]*, Huey-Pin Tsai[2,3], Su-Jhen Hung[1], Wen-Chien Ko[4], Jen-Ren Wang[2,3,5,6]*

1 National Mosquito-Borne Diseases Control Research Center, National Health Research Institutes, Tainan, Taiwan, 2 Department of Pathology, National Cheng Kung University Hospital, and College of Medicine, National Cheng Kung University, Tainan, Taiwan, 3 Department of Medical Laboratory Science and Biotechnology, National Cheng Kung University, Tainan, Taiwan, 4 Department of Internal Medicine, National Cheng Kung University Hospital, and College of Medicine, National Cheng Kung University, Tainan, Taiwan, 5 National Institute of Infectious Diseases and Vaccinology, National Health Research Institutes, Tainan, Taiwan, 6 Center of Infectious Disease and Signaling Research, National Cheng Kung University, Tainan, Taiwan

* joehuang@nhri.edu.tw (S-WH); jrwang@mail.ncku.edu.tw (J-RW)

Abstract

Background

Dengue virus causes a wide spectrum of disease, which ranges from subclinical disease to severe dengue shock syndrome. However, estimating the risk of severe outcomes using clinical presentation or laboratory test results for rapid patient triage remains a challenge. Here, we aimed to develop prognostic models for severe dengue using machine learning, according to demographic information and clinical laboratory data of patients with dengue.

Methodology/Principal findings

Out of 1,581 patients in the National Cheng Kung University Hospital with suspected dengue infections and subjected to NS1 antigen, IgM and IgG, and qRT-PCR tests, 798 patients including 138 severe cases were enrolled in the study. The primary target outcome was severe dengue. Machine learning models were trained and tested using the patient dataset that included demographic information and qualitative laboratory test results collected on day 1 when they sought medical advice. To develop prognostic models, we applied various machine learning methods, including logistic regression, random forest, gradient boosting machine, support vector classifier, and artificial neural network, and compared the performance of the methods. The artificial neural network showed the highest average discrimination area under the receiver operating characteristic curve (0.8324 ± 0.0268) and balance accuracy (0.7523 ± 0.0273). According to the model explainer that analyzed the contributions/co-contributions of the different factors, patient age and dengue NS1 antigenemia were the two most important risk factors associated with severe dengue. Additionally, co-existence of anti-dengue IgM and IgG in patients with dengue increased the probability of severe dengue.

Conclusions/Significance

We developed prognostic models for the prediction of dengue severity in patients, using machine learning. The discriminative ability of the artificial neural network exhibited good performance for severe dengue prognosis. This model could help clinicians obtain a rapid prognosis during dengue outbreaks. However, the model requires further validation using external cohorts in future studies.

Author summary

Dengue virus infects millions of people annually and is associated with a high mortality rate. When outbreaks occur, hospitals are often overcrowded with patients. Thus, novel approaches are required to accelerate patient triage for hospitalization, or further intensive care. Machine learning is being widely applied for resolving various problems, including medical diagnosis and outcome prediction. Here, we combined information from patients, including age, sex, and rapid virus test results, to develop a machine learning model for severe outcome prediction. The developed machine learning model displayed good performance for severe dengue disease prediction, and all information required for the model could be easily obtained. We also found that patients who were over 60 years old, who had detectable nonstructural protein-1 from dengue virus, or who had both detectable anti-dengue IgM and IgG antibodies in their sera, had a greater risk of progression to severe dengue. This study established a new approach to predict dengue disease outcomes by applying machine learning and defined the risk factors for severity prediction.

Introduction

Dengue virus (DENV) causes more than 90 million acute infection cases and 0.5 million fatalities worldwide each year [1]. Dengue disease is an acute febrile disease caused by the DENV, which is transmitted from mosquitos to humans [2]. Most patients present with acute dengue fever, and approximately 5–20% of patients progress to severe dengue with bleeding, plasma leakage, shock, organ failure, and even death [3]. Four serotypes of DENV, including DENV-1 to DENV-4, have recently circulated in tropical and subtropical regions around the world [4]. Although robust antibody responses have been detected in individuals who have recovered from primary DENV infections, these antibodies only have the capacity to prevent re-infection by the same serotype (homologous serotype). Thus, individuals remain susceptible to a second infection with a different serotype (heterologous serotype), and re-infection by heterologous serotypes is known to increase the risk of severe dengue disease through antibody-dependent enhancement (ADE) of DENV [5]. Thus, the antibody response to DENV infection is both beneficial and harmful to the host.

Virological and serological methods, including testing for viral RNA, DENV nonstructural protein 1 (NS1) antigenemia, anti-dengue IgM, and antigen-dengue IgG, have been widely applied in the diagnosis of DENV infection. DENV viremia occurs for 3–5 days prior to fever onset and continues for approximately 5 days into febrile illness [6]. During viremia, viral RNA and NS1 antigen can be detected in serum or plasma samples from infected patients. Among these markers, NS1 antigen is a widely applied marker in rapid diagnosis owing to its abundance along with viral RNA on the day of disease onset in patient serum. In addition to viral RNA and NS1 antigen, the presence of anti-dengue IgM and IgG antibodies is also

commonly evaluated following dengue viremia [6]. In primary DENV-infected patients, anti-dengue IgM antibodies gradually increase after the day of disease onset, and anti-dengue IgG antibodies increase after IgM antibody increases. During secondary DENV infection, anti-dengue IgM and IgG antibodies simultaneously increase after viremia [7]. Since viral RNA and NS1 antigenemia may not be detectable in patient's serum after the acute febrile phase, either the presence of anti-dengue IgM or the elevated amount of anti-dengue IgG antibodies can be used as markers to diagnose dengue infection.

After identifying individuals who are infected by DENV, evaluating their risk of developing severe dengue becomes the key factor for disease control. Clinical presentation combined with the results of laboratory blood tests have generally been used as markers of dengue severity prognosis [8]. According to clinical presentation, the World Health Organization (WHO) announced improved criteria in 2009 [8]; based on these criteria, patients can be divided into three categories, i.e., dengue fever without warning signs, dengue with warning signs, and severe dengue. The warning signs of mild dengue disease include high fever, severe headache, pain behind the eyes, muscle and joint pain, nausea, vomiting, swollen glands, and skin rash. Patients with severe dengue further present with severe abdominal pain, persistent vomiting, rapid breathing, bleeding gums, fatigue, restlessness, and hematemesis. In addition to clinical manifestations, other markers, such as high fever, platelet depletion, comorbidities, secondary infections, hemoconcentrations, rhabdomyolysis, prolonged prothrombin time, virus sero-types, and increased viral antigen NS1 levels, have been reported to be associated with increased severity in dengue patients [9–13]. These warning signs and identified risk factors have been applied for the evaluation of the severity of dengue; however, the time required to perform additional laboratory blood tests with advanced medical devices makes it difficult to rapidly triage patients with a high risk of severe dengue for further medical treatment or hospi-talization [14]. Hence, a more efficient prognostic tool is urgently required, particularly for application during major dengue outbreaks.

Artificial intelligence has recently attracted much attention in various fields of health and medicine. Different artificial intelligence and machine learning (ML) methods have been applied for various purposes, including image recognition, patient phenotyping, and outcome prediction for diseases such as cancer [15–19], cardiac arrest [20], Alzheimer's disease [21–23], respiratory diseases [24,25], rheumatic diseases [26], cornea and retinal diseases [27,28], gas-trointestinal diseases [29,30], and infectious diseases [31–35]. These studies revealed that artifi-cial intelligence has the capacity to assist clinicians in the disease diagnosis with high efficiency and accuracy. The advantages of artificial intelligence include improved medical treatment of patients and reduced duration of diagnosis after patients are examined using medical imaging or laboratory tests. However, very few artificial intelligence-based approaches or ML methods have been developed to predict dengue severity thus far [36]. Accordingly, in this study, we ret-rospectively established a rapid prognosis system for severe dengue using an ML approach according to rapid diagnostic test results and demographic characteristics of the patients. Since rapid diagnostic tests are performed when patients are suspected to be infected with DENV, these test results along with patient demographic information are available without the requirement of additional tests.

Methods

Ethics statement

All demographic and clinical data were anonymized and de-identified prior to analysis; thus, informed consent was waived. The waiver was approved by the Institutional Review Board of National Cheng Kung University Hospital (approval no. B-ER-107-224).

Demographic and clinical characteristics of patients with dengue

Demographic information and laboratory test results for patients with dengue were obtained from our previous study [37]. Suspected DENV-infected patients were enrolled at National Cheng Kung University Hospital between July and November 2015. Among these patients, 798 for whom all demographic information and laboratory test results were available, including disease severity, age, sex, viral RNA copies, NS1 antigen, anti-dengue IgM, and anti-dengue IgG, were enrolled in this study. Additional molecular tests for DENV serotyping were performed, and all patients in this study were confirmed to be infected with DENV-2. Patients were categorized as having mild or severe dengue by clinicians before hospital discharge (inpatients) or by the Taiwan CDC dengue surveillance system in the severe case report (outpatients), according to the 2009 WHO diagnostic criteria for severity. In brief, patients with mild dengue typically develop high-grade fever that is often accompanied by facial flushing, skin erythema, generalized body ache, myalgia, arthralgia, headache, anorexia, nausea, and vomiting, and some of these patients may have sore throat, injected pharynx, and conjunctival injection. The liver is often enlarged and tender after a few days of fever. The following criteria were used to categorize cases as severe dengue: severe plasma leakage, severe bleeding, or severe organ involvement.

Clinical specimens

Serum samples from patients with suspected DENV infection were collected at the Clinical Virology Laboratory of National Cheng Kung University Hospital (NCKUH) between July and November 2015. Samples were screened using one-step immunochromatographic Dengue DuoDengue NS1 Ag + Ab Combo assays (SD BIOLINE, Yongin, Republic of Korea) according to the manufacturer's instructions for qualitative antigen and antibody detection. For estimating DENV viral loads, viral RNA copies were amplified using LightMix dengue virus EC quantitative reverse transcription polymerase chain reaction, as previously described [37]. In brief, viral RNA was extracted from serum samples and an extraction control sample using QAIamp viral RNA mini kit (Qiagen, Venlo, Netherlands) or the automated extraction system, Lab-Turbo Virus mini Kit within the LabTurbo 48 Compact System (Taigen BioscienceCorp., Taipei, Taiwan). qRT-PCR analysis of viral loads was performed using LightMix dengue virus EC kit (qRT-PCR; TIB Molbiol, Berlin, Germany), which is capable of identifying all four dengue serotypes. cDNA was generated from the RNA samples using a FirstStrand cDNA Synthesis kit (Roche, Basel, Switzerland) and a DENV-specific primer included in the LightMix dengue virus EC kit. The qRT-PCR assay was then performed in a LightCycler 2.0 or LightCycler 480 II device (Roche, Basel, Switzerland), according to the manufacturer's instructions. The Light-Mix dengue virus EC kit provides cloned dengue DNA at concentrations of 10^1 to 10^6 copies/reaction as standards. The cycle number of the Crossing Point (Cp) of each sample was calculated automatically by the Second Derivative Maximum method (Automated (F'' max)) using the LightCycler software. A logarithmic transformation was applied to all resulting data. The amount of virus per sample (viral load) was reported in copies/reaction, as automatically generated by the LightCycler software.

Statistical and descriptive analysis

To identify the characteristics that contribute to disease severity, we analyzed the contribution of demographic information and laboratory test results of DENV-2 infected patients. Both age and viral RNA amount were compared using Student's t-test. Categorical characteristics, including sex, DENV NS1, anti-DENV IgG, and anti-DENV IgM test results, were compared with DENV disease severity using the Chi-squared test of independence. We further performed multivariate logistic regression (LR) to evaluate the association of demographic

information (age and sex) and clinical test results (NS1, IgM, IgG) with the prognosis of severe dengue using R version 3.4.4. To analyze the age variable more intuitively, we further categorized age into two groups. As the average age among patients was 55.8 years, we chose 50 years as a cut-off to categorize cases for multivariate logistic regression analysis. We determined the univariate and multivariate-adjusted odds ratio (OR) and 95% confidence interval (CI) for each factor to predict severity using LR. The factors that exhibited $P < 0.05$ from this analysis were considered significantly associated with severe dengue.

ML models and preprocessing

The ML models and preprocessing strategy were adapted from a previous report by Nanayak-kara et al. [20]. We used LR, support vector machine (SVM), random forest (RF), gradient boosting machine (GBM), and artificial neural network (ANN) methods, which are commonly used for binary classification questions in medicine, to establish patterns for distinguishing severe cases from mild cases [20]. Briefly, RFs first established multiple decision trees and generated various divisions of the data to retrieve an output [38]. Decision trees selected these divisions based on minimizing impurity. Similar to RFs, GBMs first trained the models in a gradual, additive, and sequential manner and collected weak decision trees [38]. Through a process of iteratively training new models to improve the weakly classified observations of the previous models, GBMs assembled these models together and made predictions. SVMs created a line or a hyperplane to divide the data into different classes to address the predictions [38]. Through imitating biological neurons of the human brain, the ANN method then established multiple hidden layers of neurons that link inputs to the output neuron [39]. Models were then constructed using open-source software and its libraries, including Scikit-learn 0.22.2 [40], Tensorflow 2.0.0 [41], and Python 3.6.10 [42].

To validate the prognosis performance of the models, we used a stratified 10-fold cross-validation approach with training and testing datasets. In this approach, the original sample was randomly partitioned into 10 equal size subsamples. Of the 10 subsamples, a single subsample was retained as the validation data for testing the model, and the remaining 9 subsamples were used as training data. The folds were selected so that each fold contains roughly the same proportions of severe and mild case labels. In a 90% and 10% training-testing set, each subsample contained 80 cases, including 14 severe dengue cases. We then repeated this process 10 times (the folds), with each of the 10 subsamples being used the validation data, exactly once. In each instance of partition, the remaining training data was used to develop the prognosis model. To fine-tune the parameters of the prognosis model, we used either grid or random hyperparameter searches to search for optimal hyperparameters for each model by another stratified 10-fold cross validation process using the training data. A broad range of hyperparameters was applied in the parameter search, and the area under the receiver operating characteristic curve (AUROC) was referred to as the optimization metric. The AUROC of the model was assessed using the logarithmic loss function. The age variable was normalized by 100 to yield values between 0 and 1. Feature normalization is beneficial for improving the numerical stability of the model and often reduces training time because similar ranges of values help to quickly converge the gradient descents during model training. The performance of the established models was validated by receiver operating characteristic (ROC) curves, which summarized the trade-off between the true- and false-positive rates, and the AUROC was calculated to validate the performance of the established models. We also compared the performance of the models balance accuracy values based on the average recall obtained on severe and mild case groups. The 10 AUROC and balance accuracy values from each fold were then averaged for model performance validation.

Model explanation methodology

To identify potentially relevant features on a per-patient basis, we assessed explainability using SHapley Additive exPlanations (SHAP). This method has been previously described in detail [43] and connects game theory with local explanations, uniting several previous methods. Briefly, SHAP generates a locally interpretable model for individual prediction from a complex model using an explainer method that combines the inputs together to evaluate the effects on the predictive model. Assessing how the ML model makes predictions is essential, particularly for predictive models with good performance; however, the ML model often achieved good performance by using complex models, which are difficult to explain. Thus, to explain the developed ANN model, we applied the SHAP model explainer [43], which is a unified framework to interpret predictions generated by our model. In the model explanation, we used a stratified holdout approach by randomly partitioning patients into subsets with 90% and 10% of training and testing datasets, with the same proportion of severe and mild cases in training and testing datasets. The training dataset was used to establish the ANN with hyperparameter tuning via stratified 10-fold cross validation. The testing dataset was analyzed using the SHAP explainer, which illustrated the output of ML models by assigning an importance value (the SHAP value) to each feature for prediction. This also included the identification of a new class of additive feature importance measures as well as theoretical results. The SHAP explainer could interpret for all predictions, whether each feature increased (SHAP value > 0) or decreased (SHAP value < 0) the potency for different classification results. Additionally, the SHAP explainer improved computational performance and consistency with higher local accuracy than other model explainer approaches [43]. The summary plot demonstrates how the SHAP values vary along with the feature values (age) and the designated values (NS1, anti-DENV IgM, and anti-DENV IgG: positive as 1 and negative as 0; sex: male as 1 and female as 0). The dependence plot reveals the combined effects of two features on the SHAP values. In the dependence plots, features that were most strongly associated with each other were paired as a "feature pair" for which the combined effects on disease severity prognosis were examined.

Results

Statistics and machine learning models using demographic information and laboratory test results

To develop a rapid prognostic model to predict severe dengue, we retrospectively analyzed the profiles of DENV-infected patients from the 2015 Taiwan outbreak. Among 798 recruited patients, 17.4% (138) of cases had severe dengue, and 82.6% (660) of cases had mild dengue. Patients with DENV were evaluated for presence of DENV NS1 antigen, anti-DENV IgG, anti-DENV IgM, and viral RNA in serum samples.

To identify the characteristics associated with disease severity that should be included as features of the prognostic model, we first analyzed association of the characteristics, including age, sex, DENV NS1 antigen, viral RNA amount, and anti-DENV IgG and IgM antibodies, with the severity of dengue (Table 1). The results indicated that patients with severe dengue were significantly older than patients with mild dengue (median age: 75 versus 55 years, respectively; $p < 0.001$). Additionally, more severe cases exhibited detectable DENV NS1 antigen (94.9% versus 74.2%, respectively; $p < 0.001$), anti-DENV IgG antibody (30.2% versus 23.6%, respectively; $p < 0.001$), and anti-DENV IgM antibody (25.8% versus 24.3%, respectively; $p = 0.008$) compared with mild cases in serum samples collected on day 1 when they sought medical advice in our hospital during acute febrile illness. In contrast, neither patient

Table 1. Demographic information and laboratory test results of DENV-2-infected patients. Age is reported as the median with interquartile range shown in square brackets. The viral RNA amount is reported as the mean with standard deviation in parentheses. For the remaining variables, the percentages of total number of patients are given in parentheses.

Characteristics	Severe dengue (n = 138)	Mild dengue (n = 660)	P value
Age (years)	75 [68–79]	55 [35–71]	< 0.001
Male sex	76 (54.6%)	342 (51.8%)	0.486
DENV NS1 positive	132 (94.9%)	490 (74.2%)	< 0.001
Viral RNA amount (copies/140 μL serum)	1.72×10^7 (4.93×10^7)	1.03×10^7 (3.97×10^7)	0.123
Anti-DENV IgG positive	42 (30.2%)	156 (23.6%)	< 0.001
Anti-DENV IgM positive	36 (25.8%)	161 (24.3%)	0.008

sex nor viral RNA amount in serum was significantly associated with disease severity. In summary, patient characteristics, including age, serum DENV NS1 antigen, and anti-DENV IgG antibody levels, were associated with disease severity in DENV-2-infected patients. This implies that characteristics may be good predictors of disease severity

To develop a prognosis model for rapid triage, we next excluded viral RNA amount, which result have to be retrieved with need for advanced medical instrument, and applied the patient demographic characteristics, and their laboratory test results, including age, sex, DENV NS1 antigen, anti-DENV IgG antibody, and anti-DENV IgM antibody for model development. First, we used the multivariate LR method to assess the impact of clinical and virological/serological laboratory characteristics on DENV disease severity (Table 2). Among the evaluated characteristics, age over 50 years (univariate OR: 7.22, 95% CI: 4.14–13.66; multivariate-adjusted OR: 7.58, 95% CI: 4.28–14.54; $p < 0.001$) and presence of NS1 (NS1 +) in the serum (univariate OR: 7.63, 95% CI: 3.60–19.74; multivariate-adjusted OR: 10.07, 95% CI: 4.61–26.58; $p < 0.001$), were the two major risk factors for severe DENV disease.

In addition to the multivariate LR method, we established ML models to predict the probability of patients progressing to severe dengue. The ML approach used diverse computational methods capable of incrementally building an accurate data model according to a measure of how well the model supported a given task, which may be applied for diagnosis classification in medicine. We thus utilized various ML methods, including LR, SVM, RF, GBM, and ANN, to establish models for severe dengue prognosis (Table 3). In addition to model development, we searched the hyperparameters to fine-tune each ML method for severity prognosis with high prognosis performance.

To evaluate whether the ANN model was not overfitting and thus generalized to unseen data, we used a 10-fold cross-validation scheme to assess the performance of the models (Fig 1 and Table 3). The performance values of the model are not single iterations; they are the averages across ten validations of the 10-fold cross-validation process. The ANN classified the clinical outcomes of patients with slightly higher AUROC (0.8324 ± 0.0268) and balance accuracy (0.7523 ± 0.0275), compared with the LR, SVM, RF, and GBM methods (Table 3). Thus, the

Table 2. Multivariate logistic regression model of clinical and laboratory diagnosis characteristics for dengue virus type 2 infection.

Characteristic	Severe (n = 138)	Mild (n = 660)	Univariate OR (95% CI)	P value	Multivariate-adjusted OR (95% CI)	P value
Age $\geq$ 50 years	377	125	7.22 (4.14–13.66)	< 0.001	7.58 (4.28–14.54)	< 0.001
DENV NS1 positive	176	622	7.63 (3.60–19.74)	< 0.001	10.07 (4.61–26.58)	< 0.001
Anti-DENV IgM positive	36	161	1.09 (0.71–1.65)	0.68	1.37 (0.82–2.28)	0.22
Anti-DENV IgG positive	42	156	1.41 (0.94–2.11)	0.09	1.15 (0.70–1.87)	0.57
Male	76	342	1.14 (0.79–1.65)	0.49	1.11 (0.75–1.65)	0.61

Table 3. Comparison of machine learning and deep learning models by 10-fold cross-validation for disease severity prediction in dengue virus type 2 infection.

Values	Logistic regression	Support vector machine	Random forest	Gradient boosting machine	Artificial neural network
Area under receiver operating characteristic curve (AUROC)	0.8152 ± 0.0718	0.8233 ± 0.0624	0.8187 ± 0.0543	0.8104 ± 0.0594	0.8324 ± 0.0268
Balance accuracy	0.5425 ± 0.0468	0.7485 ± 0.0688	0.5448 ± 0.0543	0.5867 ± 0.0712	0.7523 ± 0.0273

Values are shown as mean ± standard deviation.

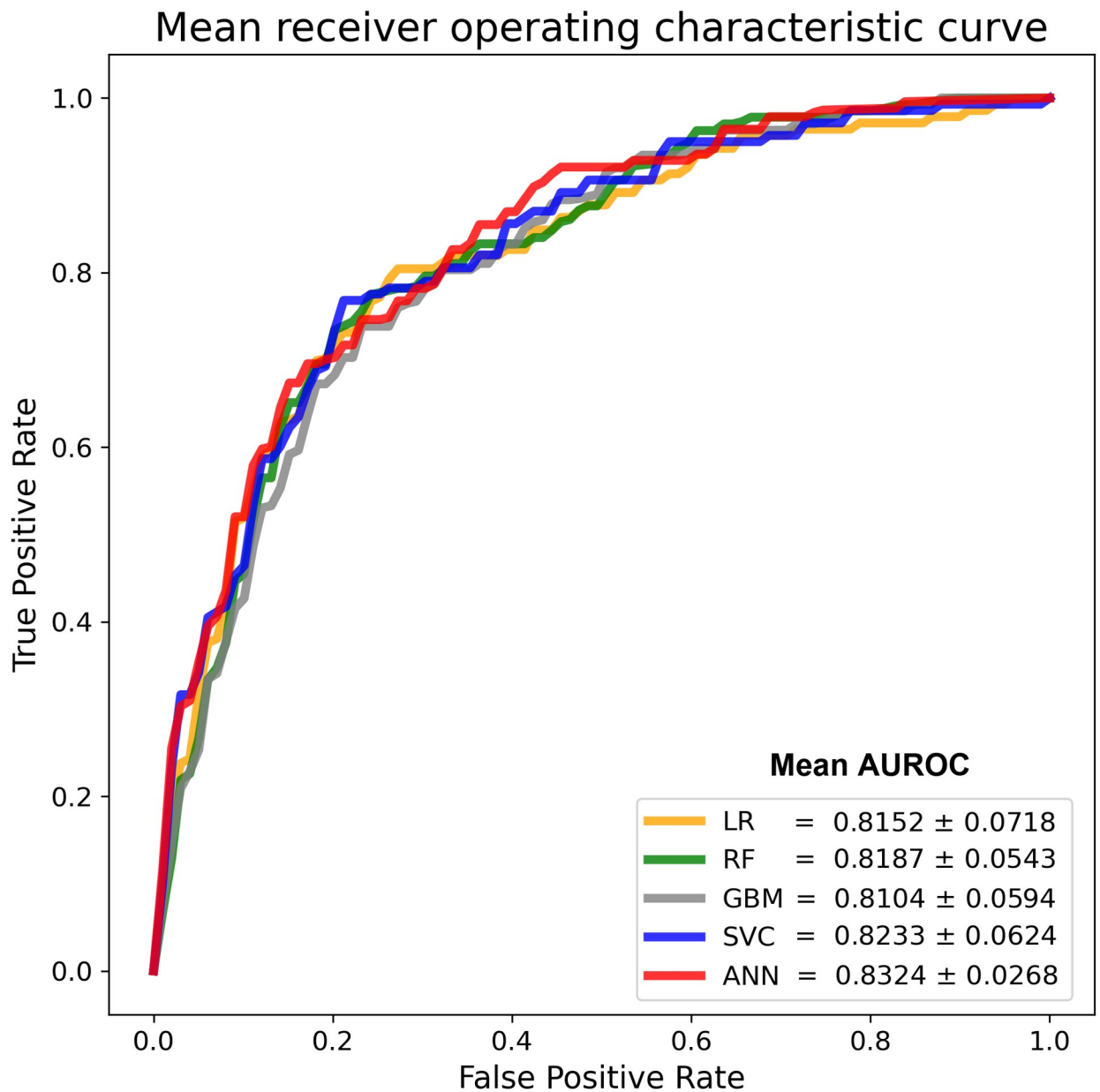https://doi.org/10.1371/journal.pntd.0008960.t003

Fig 1. Comparing receiver operating characteristic curves of machine learning methods with 10-fold cross-validation. Mean receiver operating characteristic curves of prognostic models established using machine learning methods are displayed in the indicated colors.

https://doi.org/10.1371/journal.pntd.0008960.g001

cross-validation results of our prognostic models for dengue disease severity revealed that utilization of the ANN method yielded a model with good prediction performance, as well as high AUROC and balance accuracy values. Ideally, a prognostic system should use readily available data for decision-making [20]. In the large 2015 Taiwan outbreak, we applied the rapid SD BIOLINE Dengue Duo kit for dengue diagnosis, and the results of NS1, IgM, and IgG were available within hours for patients in the emergency department of our hospital [37]. Since all of the characteristics used in the prognostic model, including age, sex, NS1 antigen, anti-DENV IgM, and anti-DENV IgG, could be evaluated quickly after rapid tests were performed, the ANN model that we developed can assist with and accelerate the triage of patients when hospitals are overcrowded during large outbreaks.

Model explainability for the prognostic ANN model

Since cross-validation results indicated that the ANN method exhibited a relatively better average performance compared with the other methods, we next analyzed the ANN model to determine how the models classified disease severity by utilizing the SHAP program. Similar to the statistic and multivariate LR model results (Tables 1 and 2), the importance plot of the SHAP program showed that age and NS1 antigen were more important features than sex, anti-DENV IgM, and anti-DENV IgG (Fig 2A), although the multivariate LR model indicated that NS1 antigen had a higher OR than patient ages (Table 2). Taken together, these results confirm the importance of these specific factors in DENV disease severity.

Next, we analyzed the contribution of each feature to the prediction using summary (Fig 2B) and dependence plots (Fig 3) of SHAP values. In the age panel of the summary plot, the SHAP value gradually increased as patient age increased (Figs 2B and 3A). In the independence plot of age, the SHAP values changed from negative to positive as the age increased to greater than 60 years, suggesting that age greater than 60 years increased the probability of the patient progressing to severe dengue (Fig 3A). NS1 antigen was another feature that positively contributed to the SHAP value (Figs 2B and 3B). In the summary plot, we observed that the appearance of NS1 antigen increased the probability of severe dengue as an outcome prediction when evaluated using the ANN model (Fig 2B). Meanwhile, SHAP values were not directly correlated with sex, anti-DENV IgM antibodies, or anti-DENV IgG antibodies (Figs 2B, 3C and 3D).

To further assess whether dual features cooperatively affected disease severity, we next determined the effects of the feature pairs on the SHAP value to evaluate the associations between features using dependence plots (Fig 3). In the dependence plots of age, NS1, and sex, we did not observe obvious association patterns within feature pairs; thus age, NS1, and sex were considered to be independent of the other features in the ANN model prognosis (Fig 3A–3C). In contrast, we found that IgM and IgG double-positive (IgM+/IgG+) was associated with positive SHAP values in the dependence plot of anti-DENV IgM antibody (IgM) and anti-DENV IgG antibody (IgG) (Fig 3D). IgM negative and IgG positive (IgM-/IgG+), or IgM positive and IgG negative (IgM+/IgG-) generally yielded negative SHAP values. (Fig 3D).

In addition to potential decision rules in the prediction, analyzing the characteristics of misclassified cases helped us to determine the limitations of the ANN model. We thus analyzed the 24 misclassified cases (Table 4), including two severe cases and 22 mild cases; these cases were incorrectly classified by the ANN model among the 80 cases of the 90:10 hold-out test dataset. The two false-negative cases were female, aged 56 and 63 years. They had NS1 antigenemia, with no anti-DENV IgM or IgG antibodies. Among the 22 false-positive cases, nine were male and 13 were female, 19 of whom were older than 60 years. Twenty-one false-positive cases were NS1 positive. In addition, two cases had NS1 antigenemia and anti-DENV IgG
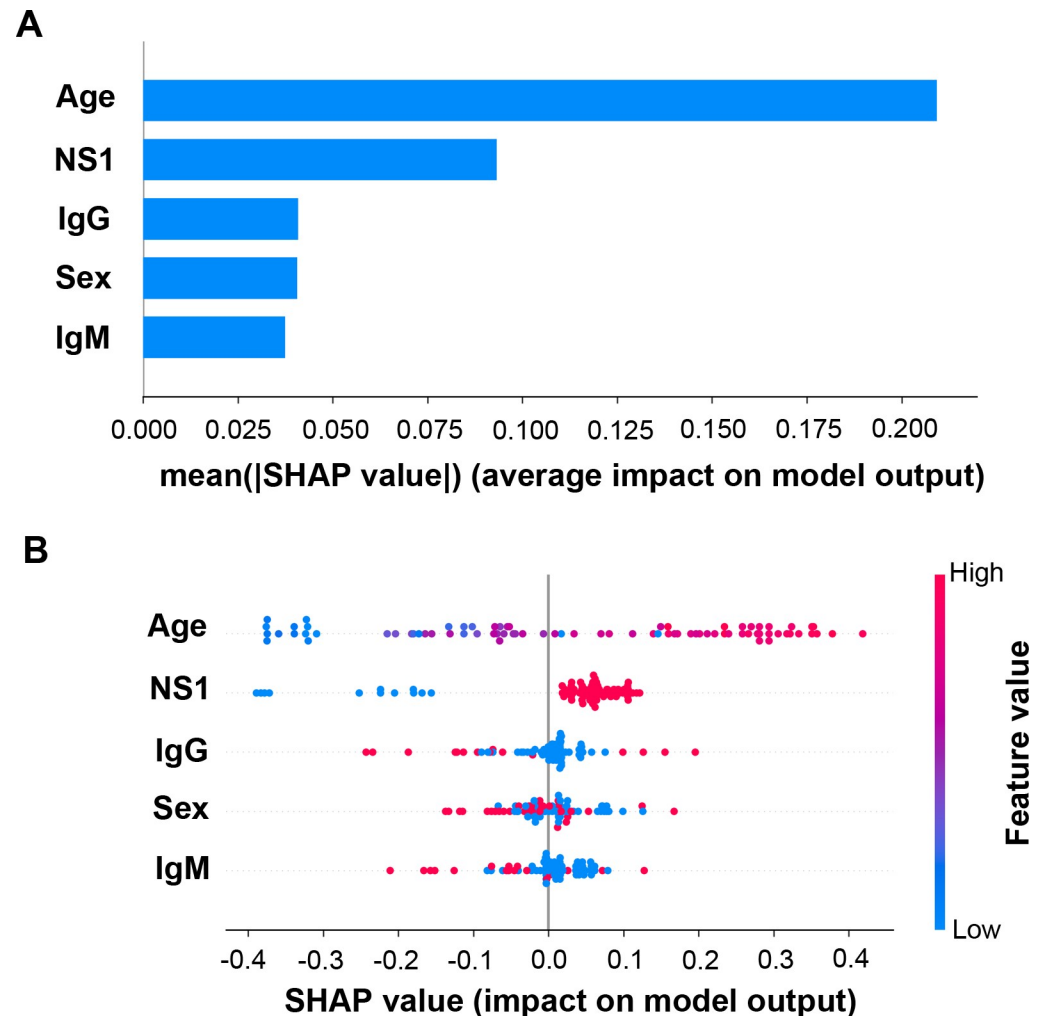
A

B

Fig 2. Importance plot and summary dot plot of the ANN model using SHAP explainer. (A) Importance plot depicting the features most effective at predicting dengue disease outcomes. (B) Summary dot plot mapping the effects of the indicated features on prediction outcomes, which assigns features to continuous color for age, and discrete color for other features. For NS1, anti-DENV IgM, and anti-DENV IgG, positive is red, and negative is blue; For sex, male is red, and female is blue.

antibodies; three cases simultaneously had NS1 antigenemia, anti-DENV IgM, and anti-DENV IgG antibodies. The calibration plot for the prediction is displayed in Fig 4. The figure shows under-prediction of severe dengue risk among the cases of the ANN model. Misclassifications, particularly the cases that were falsely classified as negative by the ANN model, should be noted. Such cases included two females with NS1 antigenemia but no anti-DENV IgM and IgG antibodies.

Discussion

Since DENV activity is dramatically increased during large outbreaks, a large number of patients suspected of having dengue will seek medical advice at hospitals, particularly from the emergency department. To help accelerate triage in overcrowded emergency rooms or out-patient clinics, we developed ML models for dengue severity prognosis according to demographic information (age and sex) and laboratory test results (NS1 antigen, anti-DENV IgM
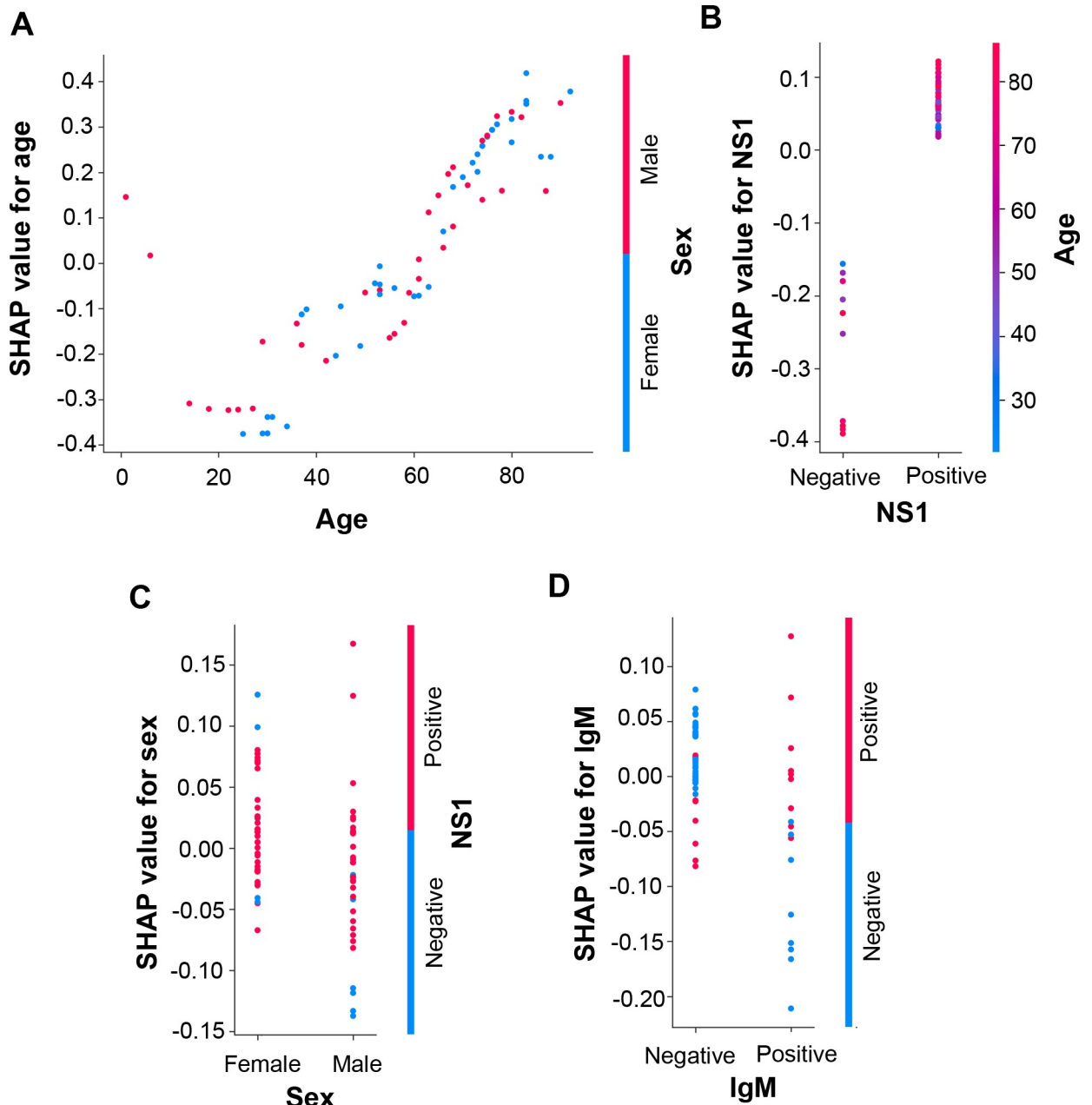
Fig 3. Dependence plots of the ANN model using SHAP explainer. Dependence plots of the ANN model showed the potential cooperative effects of feature pairs on prediction outcomes. For each feature, its most associated feature was chosen to be analyzed using dependence plots. The feature pairs included (A) age and sex, (B) NS1 and age, (C) sex and IgM, and (D) IgM and IgG.

antibody, and anti-DENV IgG antibody). The established ANN model predicted disease severity with good performance, and the SHAP explanation showed that patient age and the presence of NS1 antigen in serum samples were two major factors for severe dengue prediction. When we combined sex and anti-DENV IgM/IgG antibody in the serum, the ANN model could quickly predict severity using the NS1 antigen and anti-DENV IgM/IgG rapid tests. This method could enable clinicians to determine the prognosis of patients with dengue by analyzing retrieved information with a laptop, without the need for advanced medical instruments.

Table 4. Definitions of false-negative and false-positive severe dengue cases in the test dataset, using ANN model prediction.

	Severity	Sex	Age	NS1 antigen	Anti-DENV IgM	Anti-DENV IgG	Prediction probability of severe cases using the ANN model
False-negative cases							
	Severe	Female	56	Positive	Negative	Negative	0.3872
	Severe	Female	63	Positive	Negative	Negative	0.2694
False-positive cases							
	Mild	Female	72	Positive	Negative	Negative	0.6373
	Mild	Female	52	Positive	Positive	Positive	0.6311
	Mild	Female	83	Positive	Negative	Positive	0.7466
	Mild	Female	53	Positive	Positive	Positive	0.6272
	Mild	Female	83	Positive	Negative	Negative	0.8125
	Mild	Male	67	Positive	Negative	Negative	0.7621
	Mild	Male	77	Positive	Negative	Positive	0.5912
	Mild	Male	65	Positive	Negative	Negative	0.7412
	Mild	Male	75	Positive	Negative	Negative	0.7996
	Mild	Female	77	Positive	Negative	Negative	0.7504
	Mild	Male	63	Positive	Positive	Positive	0.8439
	Mild	Female	76	Positive	Negative	Negative	0.7349
	Mild	Male	1	Positive	Negative	Negative	0.9256
	Mild	Female	70	Positive	Negative	Negative	0.5919
	Mild	Male	61	Positive	Negative	Negative	0.5275
	Mild	Female	83	Positive	Negative	Negative	0.8125
	Mild	Female	80	Positive	Negative	Positive	0.6627
	Mild	Female	73	Positive	Negative	Negative	0.6639
	Mild	Female	73	Negative	Negative	Negative	0.5512
	Mild	Female	75	Positive	Negative	Negative	0.7159
	Mild	Male	68	Positive	Negative	Negative	0.7679
	Mild	Male	75	Positive	Negative	Negative	0.7996

Moreover, this approach can be applied in not only hospitals but also local clinics, which will help in identifying patients who may need further medical treatment or care.

According to the patient demographic information and laboratory test result dataset used in this study, patient age and NS1 antigenemia were the two factors that were most important for prediction. Regarding age, the SHAP values for severe dengue increased with increasing age and changed from negative to positive as age increased to greater than 60 years, suggesting that patients in this advanced age category had a higher probability of progressing to severe dengue. A potential explanation for this scenario is the gradual increase in seroprevalence rates with age in Taiwan [44]. According to a previous study, these rates in Taiwan increased significantly from 2.1% in the 30–39-year age group to 17.1% in the 60–69-year age group, and 50% in the 70–79-year age group [44]. The increasing seroprevalence rate from the younger group (30–39 years) to the older group (> 60 years) implied the low prevalence of DENV in Taiwan. Most individuals in the aged group were thought to be monotypically immune as they had been previously infected with DENV. Thus, the monotypically immune individuals in the aged group were at a high risk of secondary infection by heterologous DENV serotypes, which could induce ADE or immune mimicry in dengue infection and cause immunopathogenesis-related damage to endothelial cells, thereby causing severe plasma leakage [45–49]. This may explain the elevated potential for disease severity in the aged group. Moreover, since monotypically immune individuals were then exposed to infection by a homologous DENV serotype,
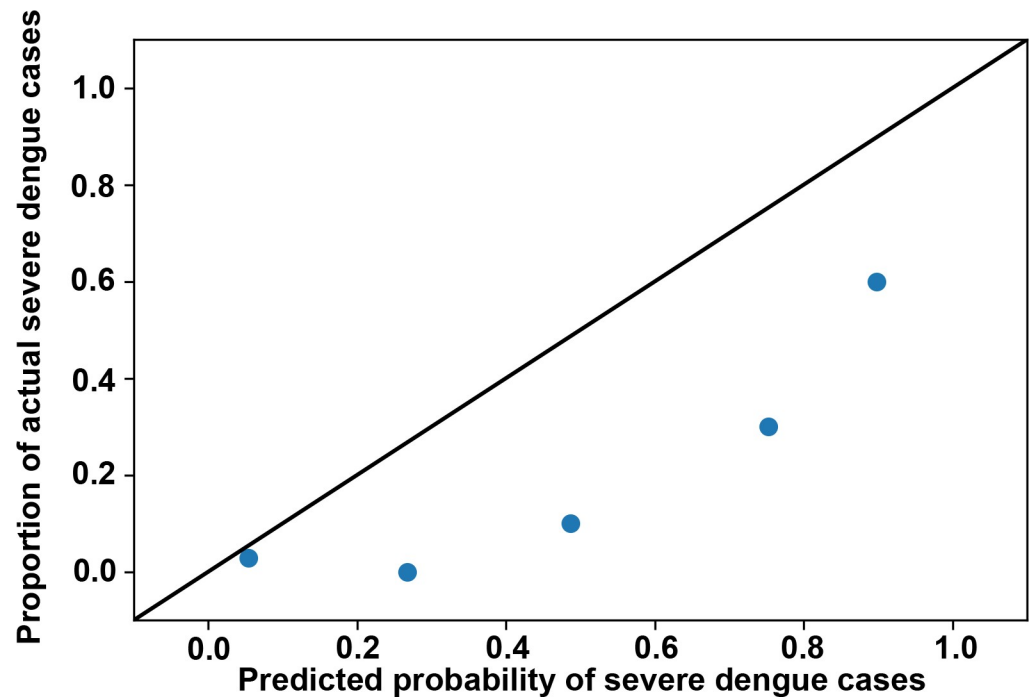
Fig 4. Calibration plot for the ANN model prediction. For the ANN model, the predicted probability of severe dengue cases is compared to the actual proportion of severe dengue cases among subjects in the 90–10 holdout test set. Perfect predictive ability is represented by the dashed diagonal line.

they may be asymptomatic or exhibit mild dengue disease owing to the protective effects of pre-existing anti-DENV IgG antibodies [5]. However, ADE with aberrant activation of cross-reactive T cells causes severe dengue, such as dengue hemorrhagic fever and dengue shock syndrome, when monotypically immune individuals are infected with heterologous DENV serotypes [45–49]. A similar explanation can be applied to the findings of the model explainer, which indicated that IgM+/IgG+ marginally increased disease severity, but IgM-/IgG+ or IgM+/IgG- reduced disease severity. Most specimens were collected on day 1 when they sought medical advice in clinics or emergency departments of our hospital during acute febrile illness. Therefore, the simultaneous elevation of IgM and IgG (IgM+/IgG+) in the acute febrile phase of infection implied that the patients were immune to monotypic DENV and were recently infected with a heterologous serotype, which could increase the risk of progression to severe disease. Considering that the antibody, which was developed from a monotypic immune response, could protect individuals from re-infection by the homologous serotype [5], the scenario of IgM-/IgG+ in patients during the acute infection phase implied that the patients had been infected by the homologous serotype virus, which stimulated a pre-existing subset of plasma cells or memory B cells to produce anti-DENV IgG antibodies against the virus. The scenario of IgM+/IgG- in patients with acute infection may imply that they were primarily infected with DENV and that their naïve B cells have begun to produce anti-DENV IgM antibodies against DENV infection. Hence, both IgM-/IgG+ and IgM+/IgG- likely indicated that the individuals would potentially not be re-infected by the heterologous serotype, which may have caused mild dengue disease. While the information of patients regarding primary or secondary infection was not available in our de-identified data, it might be worthwhile to investigate the association between the rapid IgM/IgG test results and primary/secondary infection among dengue patients in the future.

Another important risk factor was NS1 antigenemia. NS1 antigen has been widely applied as a diagnosis marker in dengue rapid tests and identified as a virulence factor in DENV infection [50,51]. By disrupting the endothelial glycocalyx, NS1 causes vascular leakage and increases endothelial permeability [50], which can lead to plasma leakage or hemorrhage. High amounts of NS1 antigenemia have been suggested as an essential factor for severe dengue, particularly in the endemic population [52–56]. Thus, these reports implied that NS1 causes the pathogenesis of dengue infection and is associated with disease severity. Similar to previous studies, the SHAP explainer for the developed ANN model showed high SHAP values for the NS1 antigenemia group, suggesting greater risk of severity among NS1 positive patients.

In addition to the characteristics analyzed in this study, several studies have reported other potential markers in patient serum, such as immune activation markers (interleukin [IL]-6, IL-10, interferon-γ, macrophage inhibitory factor, and C-C chemokine motif ligand-4), endothelial activation markers (increase levels of angiotensin [Ang]-2, von Willibrand factor, circulating vascular adhesion molecule, vascular endothelial growth factor [VEGF], and VEGF receptor [VEGFR] 1 and decrease levels of Ang-1, a disintegrin and metalloproteinase with a thrombospondin type 1 motif member 13, and VEGFRII), and biochemical markers (increased levels of lipopolysaccharide, aspartate aminotransferase [AST], and alanine aminotransferase [ALT] and decreased levels of lipids, inhibitor alpha inhibitor protein, and nitric oxide), as predictors of severe dengue disease [57]. In addition to the soluble markers in patient serum, recent investigations have identified biomarkers associated with severe dengue development via microarray-based analysis to analyze host gene expression in peripheral blood samples from patients with different disease severities [57–67]. By combining results retrieved from patients, including demographic characteristics, dengue warning signs, other symptoms or signs, and laboratory features associated with severe dengue, researchers have further applied scoring systems to evaluate the potential for severe dengue with high performance [36,58,68]. In contrast, our prediction system applied patient information, including demographic characteristics (age and sex) and laboratory test results (NS1 antigen and anti-DENV IgM/IgG antibodies) obtained from rapid tests. Dengue NS1 and anti-DENV IgM/IgG rapid tests have been applied in the laboratory diagnosis of dengue infection, particularly during outbreaks and epidemics [37]. Since these tests are potentially conducted to diagnose dengue infection in suspected cases at clinics, hospitals, and medical centers, NS1 antigen and anti-DENV IgM/IgG antibody test results can be rapidly available as diagnostic factors for decision-making, without the need for special medical devices to detect or quantify other markers in the blood. The data can then be analyzed using a laptop to determine severity prediction using the ANN model. Through the ANN prediction system, the risk of patients progressing to severe dengue can be rapidly estimated. Although our prognostic approach exhibited marginally lower performance in terms of AUROC values compared to those reported in previous studies [58,68], the developed ANN prognostic system could rapidly generate predictions within hours, including the durations of collecting blood, performing rapid tests, and running the data prediction model. In addition, we believed that more patients' blood test results and dengue disease histories may improve the sensitivity and decrease the false positive rate of this method. Thus, by combining demographic and laboratory test results, we can simultaneously diagnose DENV infection and predict disease outcomes using the ANN method without the requirement for additional medical laboratory instruments. This could facilitate the widespread use of this prognostic approach in hospitals and local clinics. Accordingly, by utilizing rapid dengue detection tests and the ML method, our ANN model system may improve and accelerate patient triage and therefore reduce morbidity and mortality during dengue outbreaks.

Nevertheless, several issues remain that must be explored when applying the ANN prognostic model. First, we only used a single type of rapid test kit in this study. A previous report showed that various commercial dengue rapid test kits exhibit differences in sensitivity and specificity [69]. Thus, the sensitivity, specificity, and accuracy of the ANN model will need to be further verified using different rapid test kits. Second, the age of patients with severe dengue varies among countries and regions. We recognize that our dataset included only Taiwanese patients. Most patients with severe disease in our study were older than 60 years; however, in southeastern Asia [60] and South America [56], the most severe cases are found in children. These differences in age distribution among severe cases in various endemic or outbreak regions may be the result of the prevalence of monotypic immune responses in the studied populations. Therefore, it is necessary to rebuild a new ML model when applying the prognostic system to a different population. Third, the predominant serotype or strain that appears during outbreaks of severe dengue may vary according to the region. In Taiwan, we faced two large outbreaks in 2014 and 2015. DENV-1 caused the first wave in 2014, with a mortality rate of 0.127% (15,732 cases and 20 fatalities); however, DENV-2 caused a large outbreak in 2015, with a mortality of 0.497% (43,784 cases with 218 fatalities), which was more than 3-fold higher than that of the 2014 DENV-1 outbreak. The different mortality rates between the two outbreaks, which were caused by different predominant strains in the same population, indicated that the virulence of different DENV serotypes may vary. Because our retrospective study used data retrieved from the DENV-2 outbreak in 2015 in Taiwan, our findings will need to be verified for other serotypes. Four, most severe dengue cases were observed in our hospital, the largest medical center of Tainan city, which was overcrowded during the 2015 outbreak. Although our hospital served the entire population of Taiwan, more severe cases might be recruited preferentially into this study, which might result in sampling bias. Finally, as a retrospective study, our findings still need to be validated in a prospective cohort before the prognostic model could be applied in the clinical setting.

In conclusion, we established an ML ANN model for dengue outcome prediction based on demographic information and patient laboratory test results. Upon examining the prognosis performance of the established models, we showed that the ML model developed by us using the ANN method demonstrated superior performance in severe dengue prognosis than established models. Therefore, this new prognostic model can assist physicians in evaluating the risk of patients progressing to severe dengue and can increase the efficiency of patient triage into hospitals, particularly in overcrowded medical facilities during dengue outbreaks. Since all information used in this model was available upon initial dengue infection diagnosis, the prognostic system may be widely applied by all medical institutes, particularly local clinics, for the prediction of dengue disease outcomes.

Acknowledgments

Author Contributions

Conceptualization: Sheng-Wen Huang, Su-Jhen Hung, Wen-Chien Ko.

Data curation: Huey-Pin Tsai, Wen-Chien Ko, Jen-Ren Wang.

Formal analysis: Sheng-Wen Huang, Huey-Pin Tsai, Su-Jhen Hung, Jen-Ren Wang.

Investigation: Sheng-Wen Huang, Jen-Ren Wang.

Methodology: Sheng-Wen Huang, Su-Jhen Hung.

Project administration: Sheng-Wen Huang.

Resources: Sheng-Wen Huang, Huey-Pin Tsai, Wen-Chien Ko.

Software: Sheng-Wen Huang.

Supervision: Sheng-Wen Huang, Jen-Ren Wang.

Validation: Sheng-Wen Huang, Su-Jhen Hung, Jen-Ren Wang.

Visualization: Sheng-Wen Huang, Su-Jhen Hung.

Writing – original draft: Sheng-Wen Huang.

Writing – review & editing: Sheng-Wen Huang, Huey-Pin Tsai, Su-Jhen Hung, Jen-Ren Wang.

References

1. Bhatt S, Gething PW, Brady OJ, Messina JP, Farlow AW, Moyes CL, et al. The global distribution and burden of dengue. Nature. 2013; 496:504507. https://doi.org/10.1038/nature12060 PMID: 23563266

2. Gubler DJ. The global emergence/resurgence of arboviral diseases as public health problems. Arch Med Res. 2002; 33:330–342. https://doi.org/10.1016/s0188-4409(02)00378-8 PMID: 12234522

3. Khursheed M, Khan UR, Ejaz K, Fayyaz J, Qamar I, Razzak JA. A comparison of WHO guidelines issued in 1997 and 2009 for dengue fever—single centre experience. J Pak Med Assoc. 2013; 63:670–674. PMID: 23901662

4. Roehrig JT. Antigenic structure of flavivirus proteins. Adv Virus Res. 2003; 59:141–175. https://doi.org/10.1016/s0065-3527(03)59005-4 PMID: 14696329

5. Rothman AL. Immunology and immunopathogenesis of dengue disease. Adv Virus Res. 2003; 60:397–419. https://doi.org/10.1016/s0065-3527(03)60010-2 PMID: 14689699

6. Hunsperger EA, Munoz-Jordan J, Beltran M, Colon C, Carrion J, Vazquez J, et al. Performance of dengue diagnostic tests in a single-specimen diagnostic algorithm. J Infect Dis. 2016; 214:836–844. https://doi.org/10.1093/infdis/jiw103 PMID: 26984143

7. Janis K. Immunology. New York, NY: W. H. Freeman and Co.; 1997.

8. World Health Organization. First WHO report on Neglected Tropical Diseases., Working to Overcome the Global Impact of Neglected Tropical Diseases. Geneva, Switzerland; 2010.

9. Srikiatkhachorn A, Green S. Markers of dengue disease severity. Curr Top Microbiol Immunol. 2010; 338:67–82. https://doi.org/10.1007/978-3-642-02215-9_6 PMID: 19802579

10. Nguyen TH, Nguyen TL, Lei HY, Lin YS, Le BL, Huang KJ, et al. Association between sex, nutritional status, severity of dengue hemorrhagic fever, and immune status in infants with dengue hemorrhagic fever. Am J Trop Med Hyg. 2005; 72:370–374. PMID: 15827272

11. Anders KL, Nguyet NM, Chau NV, Hung NT, Thuy TT, Lien le B, et al. Epidemiological factors associated with dengue shock syndrome and mortality in hospitalized dengue patients in Ho Chi Minh City, Vietnam. Am J Trop Med Hyg. 2011; 84:127–134. https://doi.org/10.4269/ajtmh.2011.10-0476 PMID: 21212214

12. Vaughn DW, Green S, Kalayanarooj S, Innis BL, Nimmannitya S, Suntayakorn S, et al. Dengue viremia titer, antibody response pattern, and virus serotype correlate with disease severity. J Infect Dis. 2000; 181:2–9. https://doi.org/10.1086/315215 PMID: 10608744

13. Balmaseda A, Hammond SN, Perez L, Tellez Y, Saborio SI, Mercado JC, et al. Serotype-specific differences in clinical manifestations of dengue. Am J Trop Med Hyg. 2006; 74:449–456. PMID: 16525106

14. Hadinegoro SR. The revised WHO dengue case classification: does the system need to be modified? Paediatr Int Child Health. 2012; 32 Suppl 1:33–38. https://doi.org/10.1179/2046904712Z.00000000052 PMID: 22668448

15. Xing L, Zhang X, Guo M, Zhang X, Liu F. Application of machine learning in developing a novelty five-pseudogene signature to predict prognosis of head and neck squamous cell carcinoma: a new aspect of "junk genes" in biomedical practice. DNA Cell Biol. 2020; 39:709–723. https://doi.org/10.1089/dna.2019.5272 PMID: 32045271

16. Syed-Abdul S, Firdani RP, Chung HJ, Uddin M, Hur M, Park JH, et al. Artificial intelligence based models for screening of hematologic malignancies using cell population data. Sci Rep. 2020; 10:4583.

17. Qian Y, Qiu Y, Li CC, Wang ZY, Cao BW, Huang HX, et al. A novel diagnostic method for pituitary adenoma based on magnetic resonance imaging using a convolutional neural network. Pituitary. 2020; 23:246–252. https://doi.org/10.1007/s11102-020-01032-4 PMID: 32062801

18. Pesapane F, Suter MB, Rotili A, Penco S, Nigro O, Cremonesi M, et al. Will traditional biopsy be substituted by radiomics and liquid biopsy for breast cancer diagnosis and characterisation? Med Oncol. 2020; 37:29. https://doi.org/10.1007/s12032-020-01353-1 PMID: 32180032

19. Milanez-Almeida P, Martins AJ, Germain RN, Tsang JS. Cancer prognosis with shallow tumor RNA sequencing. Nat Med. 2020; 26:188–192. https://doi.org/10.1038/s41591-019-0729-3 PMID: 32042193

20. Nanayakkara S, Fogarty S, Tremeer M, Ross K, Richards B, Bergmeir C, et al. Characterising risk of in-hospital mortality following cardiac arrest using machine learning: A retrospective international registry study. PLoS Med. 2018; 15(11):e1002709. https://doi.org/10.1371/journal.pmed.1002709 PMID: 30500816

21. Raamana PR, Rosen H, Miller B, Weiner MW, Wang L, Beg MF. Three-class differential diagnosis among Alzheimer disease, frontotemporal dementia, and controls. Front Neurol. 2014; 5:71. https://doi.org/10.3389/fneur.2014.00071 PMID: 24860545

22. Jin D, Wang P, Zalesky A, Liu B, Song C, Wang D, et al. Grab-AD: Generalizability and reproducibility of altered brain activity and diagnostic classification in Alzheimer's Disease. Hum Brain Mapp. 2020 May 4. https://doi.org/10.1002/hbm.25023 PMID: 32364666

23. Lee T, Lee H. Prediction of Alzheimer's disease using blood gene expression data. Sci Rep. 2020; 10:3485. https://doi.org/10.1038/s41598-020-60595-1 PMID: 32103140

24. Mekov E, Miravitlles M, Petkov R. Artificial intelligence and machine learning in respiratory medicine. Expert Rev Respir Med. 2020; 14:559–564. https://doi.org/10.1080/17476348.2020.1743181 PMID: 32166988

25. Kwon JM, Kim KH, Medina-Inojosa J, Jeon KH, Park J, Oh BH. Artificial intelligence for early prediction of pulmonary hypertension using electrocardiography. J Heart Lung Transplant. 2020; 39(8):805–814. https://doi.org/10.1016/j.healun.2020.04.009 PMID: 32381339

26. Hugle M, Omoumi P, van Laar JM, Boedecker J, Hugle T. Applied machine learning and artificial intelligence in rheumatology. Rheumatol Adv Pract. 2020; 4:rkaa005. https://doi.org/10.1093/rap/rkaa005 PMID: 32296743

27. Islam MM, Poly TN, Walther BA, Yang HC, Li YJ. Artificial intelligence in ophthalmology: a meta-analysis of deep learning models for retinal vessels segmentation. J Clin Med. 2020; 9(4):1018. https://doi.org/10.3390/jcm9041018 PMID: 32260311

28. Esporcatte LPG, Salomao MQ, Lopes BT, Vinciguerra P, Vinciguerra R, Roberts C, et al. Biomechanical diagnostics of the cornea. Eye Vis (Lond). 2020; 7(1):9. https://doi.org/10.1186/s40662-020-0174-x PMID: 32042837

29. Abadir AP, Ali MF, Karnes W, Samarasena JB. Artificial intelligence in gastrointestinal endoscopy. Clin Endosc. 2020; 53:132–141. https://doi.org/10.5946/ce.2020.038 PMID: 32252506

30. Jiang P, Lai S, Wu S, Zhao XM, Chen WH. Host DNA contents in fecal metagenomics as a biomarker for intestinal diseases and effective treatment. BMC Genomics. 2020; 21:348. https://doi.org/10.1186/s12864-020-6749-z PMID: 32393180

31. Jamal S, Khubaib M, Gangwar R, Grover S, Grover A, Hasnain SE. Artificial intelligence and machine learning based prediction of resistant and susceptible mutations in Mycobacterium tuberculosis. Sci Rep. 2020; 10:5487. https://doi.org/10.1038/s41598-020-62368-2 PMID: 32218465

32. Belfiore MP, Urraro F, Grassi R, Giacobbe G, Patelli G, Cappabianca S, et al. Artificial intelligence to codify lung CT in Covid-19 patients. Radiol Med. 2020; 125:500–504. https://doi.org/10.1007/s11547-020-01195-x PMID: 32367319

33. Huang Y, Cheng W, Zhao N, Qu H, Tian J. CT screening for early diagnosis of SARS-CoV-2 infection. Lancet Infect Dis. 2020; 20(9):P1010–1011. https://doi.org/10.1016/S1473-3099(20)30241-3 PMID: 32222164

34. Yu Q, Wang Y, Huang S, Liu S, Zhou Z, Zhang S, et al. Multicenter cohort study demonstrates more consolidation in upper lungs on initial CT increases the risk of adverse clinical outcome in COVID-19 patients. Theranostics. 2020; 10:5641–5648. https://doi.org/10.7150/thno.46465 PMID: 32373237

35. Li G, Xiong Z, Zhou H, Xie J, Chen W, Zhou M, et al. Value of CT application in the screening, diagnosis, and treatment of COVID-19. Zhong Nan Da Xue Bao Yi Xue Ban. 2020; 45:269–274.

36. Sippy R, Farrell DF, Lichtenstein DA, Nightingale R, Harris MA, Toth J, et al. Severity Index for Suspected Arbovirus (SISA): Machine learning for accurate prediction of hospitalization in subjects

suspected of arboviral infection. PLoS Negl Trop Dis. 2020:e0007969. https://doi.org/10.1371/journal.pntd.0007969 PMID: 32059026

37. Tsai HP, Tsai YY, Lin IT, Kuo PH, Chang KC, Chen JC, et al. Validation and application of a commercial quantitative real-time reverse transcriptase-PCR assay in investigation of a large dengue virus outbreak in Southern Taiwan. PLoS Negl Trop Dis. 2016; 10:e0005036. https://doi.org/10.1371/journal.pntd.0005036 PMID: 27732593

38. Hastie T, Tibshirani R, Friedman J. The Elements of Statistical Learning. New York, NY: Springer; 2017.

39. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015; 521:436–444. https://doi.org/10.1038/nature14539 PMID: 26017442

40. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: Machine learning in Python. Mach Learn Res. 2011; 12:2825–2830.

41. Abadi M, Agarwal A, Barham P, Brevdo E, Chen Z, Citro C, et al. TensorFlow: Large-scale machine learning on heterogeneous distributed systems. arXiv preprint arXiv. 2015;1603.04467.

42. Van Rossum G, Drake FL. Python 3 Reference Manual. Scotts Valley, CA: CreateSpace; 2009.

43. Lundberg SM, Lee SI. A unified approach to interpreting model predictions. In: Guyon I, Luxburg UV, Bengio S, Wallach H, Fergus R, Vishwanathan S, Garnett R, editors. Advances in neural information processing systems; 2017.

44. Tsai JJ, Liu CK, Tsai WY, Liu LT, Tyson J, Tsai CY, et al. Seroprevalence of dengue virus in two districts of Kaohsiung City after the largest dengue outbreak in Taiwan since World War II. PLoS Negl Trop Dis. 2018; 12(10):e0006879. https://doi.org/10.1371/journal.pntd.0006879 PMID: 30356316

45. Gordon A, Kuan G, Mercado JC, Gresh L, Aviles W, Balmaseda A, et al. The Nicaraguan pediatric dengue cohort study: incidence of inapparent and symptomatic dengue virus infections, 2004–2010. PLoS Negl Trop Dis. 2013; 7(9):e2462. https://doi.org/10.1371/journal.pntd.0002462 PMID: 24086788

46. Zivna I, Green S, Vaughn DW, Kalayanarooj S, Stephens HA, Chandanayingyong D, et al. T cell responses to an HLA-B*07-restricted epitope on the dengue NS3 protein correlate with disease severity. J Immunol. 2002; 168:5959–5965. https://doi.org/10.4049/jimmunol.168.11.5959 PMID: 12023403

47. Wang TT, Sewatanon J, Memoli MJ, Wrammert J, Bournazos S, Bhaumik SK, et al. IgG antibodies to dengue enhanced for FcgammaRIIIA binding determine disease severity. Science. 2017; 355:395–398. https://doi.org/10.1126/science.aai8128 PMID: 28126818

48. Katzelnick LC, Gresh L, Halloran ME, Mercado JC, Kuan G, Gordon A, et al. Antibody-dependent enhancement of severe dengue disease in humans. Science. 2017; 358:929–932. https://doi.org/10.1126/science.aan6836 PMID: 29097492

49. Rodriguez-Barraquer I, Cordeiro MT, Braga C, de Souza WV, Marques ET, Cummings DA. From re-emergence to hyperendemicity: the natural history of the dengue epidemic in Brazil. PLoS Negl Trop Dis. 2011; 5(1):e935. https://doi.org/10.1371/journal.pntd.0000935 PMID: 21245922

50. Puerta-Guardo H, Glasner DR, Harris E. Dengue virus NS1 disrupts the endothelial glycocalyx, leading to hyperpermeability. PLoS Pathog. 2016; 12(7):e1005738. https://doi.org/10.1371/journal.ppat.1005738 PMID: 27416066

51. Nunes PCG, Nogueira RMR, Heringer M, Chouin-Carneiro T, Damasceno Dos Santos Rodrigues C, de Filippis AMB, et al. NS1 antigenemia and viraemia load: potential markers of progression to dengue fatal outcome? Viruses. 2018; 10:326. https://doi.org/10.3390/v10060326 PMID: 29903980

52. Libraty DH, Young PR, Pickering D, Endy TP, Kalayanarooj S, Green S, et al. High circulating levels of the dengue virus nonstructural protein NS1 early in dengue illness correlate with the development of dengue hemorrhagic fever. J Infect Dis. 2002; 186:1165–1168. https://doi.org/10.1086/343813 PMID: 12355369

53. Libraty DH, Endy TP, Houng HS, Green S, Kalayanarooj S, Suntayakorn S, et al. Differing influences of virus burden and immune activation on disease severity in secondary dengue-3 virus infections. J Infect Dis. 2002; 185:1213–1221. https://doi.org/10.1086/340365 PMID: 12001037

54. Nunes PC, Sampaio SA, da Costa NR, de Mendonca MC, Lima Mda R, Araujo SE, et al. Dengue severity associated with age and a new lineage of dengue virus-type 2 during an outbreak in Rio De Janeiro, Brazil. J Med Virol. 2016; 88:1130–1136. https://doi.org/10.1002/jmv.24464 PMID: 27061403

55. Tricou V, Minh NN, Farrar J, Tran HT, Simmons CP. Kinetics of viremia and NS1 antigenemia are shaped by immune status and virus serotype in adults with dengue. PLoS Negl Trop Dis. 2011; 5(9):e1309. https://doi.org/10.1371/journal.pntd.0001309 PMID: 21909448

56. Allonso D, Meneses MD, Fernandes CA, Ferreira DF, Mohana-Borges R. Assessing positivity and circulating levels of NS1 in samples from a 2012 dengue outbreak in Rio de Janeiro, Brazil. PLoS One. 2014; 9(11):e113634. https://doi.org/10.1371/journal.pone.0113634 PMID: 25412084

57. John DV, Lin YS, Perng GC. Biomarkers of severe dengue disease—a review. J Biomed Sci. 2015; 22:83. https://doi.org/10.1186/s12929-015-0191-6 PMID: 26462910

58. Robinson M, Sweeney TE, Barouch-Bentov R, Sahoo MK, Kalesinskas L, Vallania F, et al. A 20-gene set predictive of progression to severe dengue. Cell Rep. 2019; 26:1104–1111. https://doi.org/10.1016/j.celrep.2019.01.033 PMID: 30699342

59. Devignot S, Sapet C, Duong V, Bergon A, Rihet P, Ong S, et al. Genome-wide expression profiling deciphers host responses altered during dengue shock syndrome and reveals the role of innate immunity in severe dengue. PLoS One. 2010; 5(7):e11671. https://doi.org/10.1371/journal.pone.0011671 PMID: 20652028

60. Hoang LT, Lynn DJ, Henn M, Birren BW, Lennon NJ, Le PT, et al. The early whole-blood transcriptional signature of dengue virus and features associated with progression to dengue shock syndrome in Vietnamese children and young adults. J Virol. 2010; 84:12982–12994. https://doi.org/10.1128/JVI.01224-10 PMID: 20943967

61. Loke P, Hammond SN, Leung JM, Kim CC, Batra S, Rocha C, et al. Gene expression patterns of dengue virus-infected children from nicaragua reveal a distinct signature of increased metabolism. PLoS Negl Trop Dis. 2010; 4(6):e710. https://doi.org/10.1371/journal.pntd.0000710 PMID: 20559541

62. Long HT, Hibberd ML, Hien TT, Dung NM, Van Ngoc T, Farrar J, et al. Patterns of gene transcript abundance in the blood of children with severe or uncomplicated dengue highlight differences in disease evolution and host response to dengue virus infection. J Infect Dis. 2009; 199:537–546. https://doi.org/10.1086/596507 PMID: 19138155

63. Nascimento EJ, Braga-Neto U, Calzavara-Silva CE, Gomes AL, Abath FG, Brito CA, et al. Gene expression profiling during early acute febrile stage of dengue infection can predict the disease outcome. PLoS One. 2009; 4(11):e7892. https://doi.org/10.1371/journal.pone.0007892 PMID: 19936257

64. Popper SJ, Gordon A, Liu M, Balmaseda A, Harris E, Relman DA. Temporal dynamics of the transcriptional response to dengue virus infection in Nicaraguan children. PLoS Negl Trop Dis. 2012; 6(12): e1966. https://doi.org/10.1371/journal.pntd.0001966 PMID: 23285306

65. Simmons CP, Popper S, Dolocek C, Chau TN, Griffiths M, Dung NT, et al. Patterns of host genome-wide gene transcript abundance in the peripheral blood of patients with acute dengue hemorrhagic fever. J Infect Dis. 2007; 195:1097–1107. https://doi.org/10.1086/512162 PMID: 17357045

66. Sun P, Garcia J, Comach G, Vahey MT, Wang Z, Forshey BM, et al. Sequential waves of gene expression in patients with clinically defined dengue illnesses reveal subtle disease phases and predict disease severity. PLoS Negl Trop Dis. 2013; 7(7):e2298. https://doi.org/10.1371/journal.pntd.0002298 PMID: 23875036

67. van de Weg CA, van den Ham HJ, Bijl MA, Anfasa F, Zaaraoui-Boutahar F, Dewi BE, et al. Time since onset of disease and individual clinical markers associate with transcriptional changes in uncomplicated dengue. PLoS Negl Trop Dis. 2015; 9(3):e0003522. https://doi.org/10.1371/journal.pntd.0003522 PMID: 25768297

68. Lee IK, Liu JW, Chen YH, Chen YC, Tsai CY, Huang SY, et al. Development of a simple clinical risk score for early prediction of severe dengue in adult patients. PLoS One. 2016; 11(5):e0154772. https://doi.org/10.1371/journal.pone.0154772 PMID: 27138448

69. Hunsperger EA, Yoksan S, Buchy P, Nguyen VC, Sekaran SD, Enria DA, et al. Evaluation of commercially available anti-dengue virus immunoglobulin M tests. Emerg Infect Dis. 2009; 15: 436–440. https://doi.org/10.3201/eid1503.080923 PMID: 19239758