Multivariable prediction model for predicting deaths in severe dengue cases

Saiful Safuan Md Sani (dcicantab5@icloud.com)
Hospital Kuala Lumpur https://orcid.org/0000-0001-9008-7506
Masliza Zaid
Department of Medicine, Hospital Sultanah Aminah, Ministry of Health Malaysia
Kar Nim Leong
Department of Medicine, Hospital Pulau Pinang, Ministry of Health Malaysia
Rossman Hawari
Emergency and Trauma Department, Hospital Sungai Buloh, Ministry of Health Malaysia
Anilawati Mat Jelani
Department of Medicine, Hospital Raja Perempuan Zainab II, Ministry of Health Malaysia
Marzilawati Abd-Rahman
Department of Medicine, Hospital Kuala Lumpur, Ministry of Health Malaysia
Ping Yee Ooh
Emergency and Trauma Department, Hospital Tuanku Ampuan Rahimah, Ministry of Health Malaysia
Anusha Shanmugarajoo
Department of Medicine, Hospital Tuanku Ampuan Rahimah, Ministry of Health Malaysia
Suresh Kumar
Department of Medicine, Hospital Sungai Buloh, Ministry of Health Malaysia
Mahiran Mustafa
Department of Medicine, Hospital Raja Perempuan Zainab II, Ministry of Health

Research article

Keywords: Dengue; severe; model; predict; deaths; phase

Posted Date: May 31st, 2019

DOI: https://doi.org/10.21203/rs.2.9960/v1

Abstract

Background Many predictive models have been developed to predict an outbreak, identify and stratify dengue but none in predicting mortality in severe dengue cases. To build a predictive model for deaths in severe dengue, a multicentre retrospective cohort study was conducted. Methods Patients with severe dengue based on WHO 2009 classification were studied. Demographic, clinical and laboratory data were collected at diagnosis of severe dengue. Penalised regression was used for variable selection and model-building. Ten-fold cross-validation with 1000 repeats was performed for internal validation. Results A cohort of 786 severe dengue cases including 35 deaths was analysed. Our model that predicts death in severe dengue cases comprises eight independent predictors: persistent diarrhoea, BMI, respiratory rate, platelet count, AST, serum bicarbonate, serum lactate and serum albumin. The AUROC is 89·6% with a sensitivity of 99·6%, specificity of 23·6%, positive predictive value of 96·6%, negative predictive value of 71·1%, positive likelihood ratio 1·45 and negative likelihood ratio 0·01. We also found that the proportion of patients that were in the febrile phase at diagnosis of severe dengue for the overall cohort, decompensated and compensated shock were 74·3%, 73% and 75·4%, respectively. Conclusions We developed a high performance dengue mortality prediction model comprising clinical and laboratory data and deployed an open access web-based tool (www.saifulsafuan.com/REPROSED2017E2) for any centre to utilise for local validation and found that a large majority of patients developed severe dengue during febrile phase.

Introduction

Dengue is a common and endemic infection in Malaysia. The number of recorded dengue cases in Malaysia has risen from 6543 to 120836 cases between 1995 to 2015.[1] The incidence rate had risen from 31·6/100000 population in the year 2000 to 396·4/100000 in 2015.[1] Expectedly, the number of deaths increased from 28 deaths in 1995 to 336 deaths in 2015.[1] The case fatality rate, however, remained at or more than 0·2% per year, for the exception of years 2011-2012.[1]

Malaysia sets a national target case fatality rate of <0·2%.[2] In a recent estimate on 12 countries in South East Asia, the region where Malaysia resides, over the years 2001-2010 annual dengue cases amounted to 2·9 million cases with 5906 deaths per year.[3]

In South East Asia, the estimated economic burden amounted to USD950 million [95% CI: USD610-1384 million] or USD1·65 [95%CI: USD1·06-2·41] per capita annually.[3] In 2010, USD73·5 million was spent by Malaysia for its National Dengue Vector Control Program alone or USD2·68 per capita. This represents 0·03% of Malaysia's gross domestic product.[4] Direct medical costs and costs related to productivity loss and premature mortality, was USD102·2 million in 2009.[4–6]

Modelling studies have been conducted but mostly in the prediction of dengue outbreak.[7-13] Several studies have looked into two clinical aspects of dengue management: predicting the identification of dengue in unselected cases of febrile illnesses,[14,15] and predicting the severity of the disease.[14,16,17,18]

However, there have only been three studies on predicting mortality in dengue.[19-21] Huang et al. utilised risk scoring.[19] The development of risk scores has been the favoured strategy in medicine for their relative ease of interpretation. However, risk scoring is not without weaknesses and have been deemed to have serious problems.[22] Risk scoring involves converting continuous predictors into categorical predictors. The conversion results in loss of granularity of information contained within continuous predictors. Md-Sani et al. faced the problem of having only 20 events (death).[20] Building a predictive model to predict death among severe dengue cases is not easy as dengue death, the event or outcome of interest, is actually uncommon - in Malaysia it is just above 0.2% of all dengue cases.[1] This poses a huge problem in that, in the usual approach of statistical modelling using multivariate logistic regression, 5-10 events per candidate predictor variable (EPV) are required. This ratio dictates the number of candidate predictors that may be simultaneously analysed in the multivariate model. When the EPV is less than this, the number of candidate predictors that may be simultaneously analysed to identify which among them are truly independent predictors is limited. Pinto et al. unfortunately employed WHO 1997 dengue classification instead of WHO 2009 schema, which makes local application in Malaysia difficult.[21]

If the prediction of an explosion of the global volume of dengue cases due to impending climate changes materialises,[23] the global community of clinicians caring for patients with severe dengue must be intelligently equipped. While literature in dengue on the identification of risk factors or predictors of dengue outbreak and disease severity are aplenty, the imperative ought to include to translate findings to the 'bedside'. Identification of risk factors should culminate in the building of prediction models that should next be deployed as software applications for clinical use at the bedside. Thus, we conducted a multicentre retrospective cohort study to build a predictive model that will predict mortality in severe dengue cases, and provide an open access 'bedside' prediction tool.

Methods

We conducted a multicentre retrospective cohort study with a universal sampling of patients with severe dengue who were admitted during the period 1st January 2017 until 31st December 2017. Our report is based on the Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD) 2015 guideline.[24]

The five participating centres were Hospital Raja Perempuan Zainab II, Hospital Sungai Buloh, Hospital Kuala Lumpur, Hospital Tuanku Ampuan Rahimah, and Hospital Sultanah Aminah. Patients were selected for inclusion if, 1) they were ≥15 years old, 2) the case fulfilled the study definition of severe dengue, and 4) the presence of dengue viral infection was confirmed via laboratory confirmation by non-structural protein 1 (NS1) antigen, or presence of dengue RNA by reverse transcription polymerase chain reaction (RT-PCR), or presence of high-titre level of immunoglobulin G (IgG), or positive immunoglobulin M (IgM) from an admission serum sample. Excluded patients were those who were pregnant and those who were admitted for transfer of care from non-participating centres. Patients with positive IgM-only immunology were also excluded if their clinical and laboratory blood results patterns were incongruous

with dengue disease course. High-titre IgG detects acute dengue infection at titres >1:2560 which therefore is highly specific for dengue infection and also identifies such infection as acute secondary dengue infection.

The definition of severe dengue in our study was based on World Health Organisation (WHO) 2009 definition but with adaptation for this study.[25] We defined severe dengue by presence of any one of the following: 1) decompensated shock due to severe plasma leakage, 2) compensated shock due to severe plasma leakage, 3) respiratory compromise due to plasma leakage, 4) severe bleeding that required intervention, or 5) severe organ involvement such as acute kidney injury defined by elevated serum creatinine above upper limit normal (according to gender-specific levels), severe hepatitis, myocarditis, or encephalopathy. Decompensated shock was defined by the presence of systolic blood pressure (SBP) less than 90 mmHg, mean arterial pressure (MAP) of less than 65 mmHg, or a drop in systolic blood pressure of more than 40 mmHg from a patient's known usual baseline readings. Compensated shock required signs of impaired peripheral perfusion, occurring in combination rather than singly, in the presence of systolic blood pressure of ≥90 mmHg. Severe hepatitis was defined as aspartate transaminase (AST) level >1,000 IU/L or alanine transaminase (ALT) level >1,000 IU/L.

Data extracted from records review were demographic, co-morbidity, clinical parameters at the time of diagnosis of severe dengue, serial laboratory investigations at the time of diagnosis of severe dengue and at 24 hours later (with their timings), nadir and peak values of selected laboratory investigations, serial treatment, time data and outcome data. A data collection proforma was used to ensure the integrity of data. Time data, in the format of dates and times, were: fever onset, admission, the time of diagnosis of severe dengue, the time of defervescence (exact start of temperature persistently <38°C), the time of occurrence of the outcome and times of all laboratory investigations. Fever onset was determined by careful history-taking, specified to the best estimated time to the nearest hour. If missing it was imputed as midday (12:00 pm). As such, the day of illness may be determined for all variables. The time of onset of severe dengue was taken as the time of diagnosis of severe dengue. Temperature was monitored 4-hourly and the temperature at diagnosis of severe dengue was determined from the temperature graph, extrapolated by the time of diagnosis of severe dengue.

We calculated the required sample size on the basis of comparison of the area under a receiver-operating characteristic (AUROC) curve with a null hypothesis value of 0·5, i.e. no discriminatory power. A sample size of 21 deaths was needed for this study for an AUROC of 80%, at the confidence level of 95% and a power of 95%.

Statistical Methods and Predictive Model Building

Our analyses were two pronged, one involved predictive model building and the other descriptive in nature. As our study was focused on building a predictive model that would predict mortality among severe dengue cases, we used selected variables that were collected at the onset of severe dengue as candidate predictors. There were 28 candidate predictors: age, body mass index (BMI), systolic blood

pressure, diastolic blood pressure, pulse rate, respiratory rate, white blood cell count, platelet count, haematocrit, aspartate transaminase, alanine transaminase, serum albumin, serum bicarbonate, serum lactate, serum creatinine, gender, presence of any comorbidity, presence of non-structural protein 1 antigen, presence of immunoglobulin M, presence of immunoglobulin G, persistent vomiting, persistent diarrhoea, presence of abdominal pain, presence of abdominal tenderness, presence of clinical fluid accumulation, occurrence of non-severe bleeding, day of onset of severe dengue, severe dengue presentation (presented as severe dengue or not). Persistent vomiting and persistent diarrhoea were quantified as three or more episodes within the 24-hour interval before admission. Clinical fluid accumulation refers to the presence of any ascites or pleural effusion by clinical or radiological detection. The outcome of interest was any death during the hospital stay.

We initially performed univariable and multivariable logistic regression. As the study outcome (event) rarely occurs with dengue and was expected to be similarly rare in our cohort, in the usual multivariable logistic regression, each candidate predictor was adjusted for age, gender, centre and time of onset of severe dengue. The rare outcome led to low events per variable which prevented all 28 candidate predictors from being simultaneously analysed in a model using usual multivariable logistic regression approach - the number of events restricts the number of variables to be analysed within a model. We then report the adjusted odds ratios in comparison to corresponding unadjusted values. This analysis was made on complete case basis.

In building our predictive model, we performed variable selection, model search, and internal validation. In order to select truly independent predictors of death, all 28 candidate predictors were simultaneously analysed in a single model using a technique called penalised regression.[26] Our prediction model was built using multivariable penalised logistic regression using the glmnet and caret packages in R.[27,28] Penalised or regularised regression was chosen as the method is known to be able to perform variable selection in the condition of low events per variable very well.[29] In penalised logistic regression, a penalty term is introduced into the algorithm such that non-significant regression coefficients undergo shrinkage towards zero while coefficients of significant or influential variables would be retained as non-zero. This is in contrast to the usual method of stepwise logistic regression to select independent predictors whereby variables are chosen based on the magnitude of the p-value. Penalised regression also reduces overfitting which is usual with logistic regression. Overfitting leads to the overoptimistic performance of a model.

We performed a model search for the best performing model by tuning different values of model hyperparameters α and λ of penalised regression. The best performing model was chosen by the highest AUROC. We then extracted the corresponding values of α and λ to build the final model. Model search and the final model underwent repeated k-fold cross-validation, using 1000 repeats and 10 folds that generated 10000 sample models. In this technique, resampling with replacement were made. Cross-validation further minimised overoptimistic performance and served as our internal validation method. Performance measures we reported were the area under the curve of the receiver-operating curve,

sensitivity, specificity, positive predictive value, negative predictive value, positive likelihood ratio, and negative likelihood ratio. We also report the calibration curve of the model.

In prediction model building, missing data within the candidate predictors were imputed using k-nearest neighbours imputation during model-building via the caret package in R.[27] Apart from serum lactate and serum bicarbonate, which had 16% and 11·6% missing data respectively, all other candidate predictors had less than 10% missing data. Since the proportion of missing data were less than 20%, we kept all the candidate predictors. Continuous candidate predictors were retained at their original scales.

Additionally, we also reported other variables to describe the characteristics of our cohort clearly. These descriptive variables include those which are related to laboratory investigations at 24 hours after the diagnosis of severe dengue, nadir and peak values of selected laboratory investigations, and variables related to management at 24-hourly time intervals. We described our cohort with descriptive and inferential analyses. Continuous variables were tested for normality with the Shapiro-Wilk test. As our data were mostly non-parametric, we used non-parametric analyses for data interrogation. Categorical variables were expressed as frequencies and percentages. Continuous variables were summarised as median and inter-quartile range (IQR). Wherever applicable, we reported the statistical parameter with its 95% confidence interval. All tests of significance were 2-sided and we took p-value <0·05 indicating statistical significance. All analyses were made using R, R Core Team (2016), R Foundation for Statistical Computing, Vienna, Austria.[30] We illustrated the overall study pipeline in Figure 1 and the internal validation step in Figure 2.

Figure 1. Overall study pipeline.

Figure 2. Internal validation scheme.

Results

The study cohort was composed of 786 laboratory-confirmed dengue cases of whom 35 (4·5%) died. NS1 antigen was positive in 643 (81·8%) and a small proportion, 54 patients (6·9%) had only positive IgM. Distribution of cases according to centres were: 366 cases (46·6%) from Hospital Kuala Lumpur, 217 cases (27·6%) from Hospital Tuanku Ampuan Rahimah, 103 cases (13·1%) from Hospital Sungai Buloh, 61 cases (7·8%) from Hospital Sultanah Aminah, and 39 cases (5·0%) from Hospital Raja Perempuan Zainab II. Compensated shock occurred in 44·9%, acute kidney injury in 27·1%, decompensated shock in 24%, carditis in 21·1%, severe hepatitis in 16·7%, respiratory compromise in 12·7%, severe bleeding in 7·8%, and encephalopathy in 3·2% of patients.

The median age was 30·3 years old, 57·6% were males, 31·3% had co-morbidity, and 13·4% had no warning signs. Among those without warning signs, 28·6% presented as severe dengue. Severe dengue was diagnosed upon presentation in 23·8% and 82·7% of patients were in febrile phase at admission. The proportion who were in febrile phase at diagnosis of severe dengue were 74·3%; by subset according to type of severe dengue the proportions were: for decompensated shock 73%, compensated shock 75·4%,

respiratory compromise due to fluid accumulation 55%, severe bleeding 45.9%, severe hepatitis 58.8%, acute kidney impairment 77.5%, encephalopathy 68% and carditis 62.7%. The median duration of febrile phase was 4·8 days, the median onset of severe dengue was day 4·0, and median day of admission was on day 3·6. The median length of stay was 4 days and median length of illness was 7·6 days. Inotropes were administered in 16·2%, invasive and assisted ventilation in 17·7%, renal replacement therapy in 2·7%, and blood product transfusion in 15·3%. Demographic and clinical characteristics are given in Table 1 and laboratory variables are given in Table 2. Univariate and multivariable analysis of selected variables are given in Table 3 (full univariate and multivariate analysis of all study variables including management related variables at different time intervals are provided in Supplement Table 1).

Table 1. Demographic and clinical variables.

Table 2. Laboratory variables.

Table 3. Univariate and multivariable analysis of selected variables for risk of death in severe dengue.

Our final model was a penalised logistic regression with elastic net penalty model or an elastic net regression model, defined by model hyperparameters $\alpha = 0{\cdot}1$ and $\lambda = 0{\cdot}01613622$. The model comprised eight predictors: the presence of persistent diarrhoea, BMI, respiratory rate, levels of platelet count, serum bicarbonate, serum lactate, serum albumin, and aspartate transaminase. Since the predictors were timed at diagnosis of severe dengue, the model, therefore, is to be used from this time point onwards. The model coefficients and intercept are given in Figure 3 and its calibration curve in Figure 4. During variable selection, centre and day of onset of severe dengue were found to be non-significant (coefficients underwent shrinkage to zero).

We report the performance of the model based on internal validation method of 1000 repeated 10-fold cross-validation that generated 10000 models in Table 4. The prediction model may be accessed at the following website: www.saifulsafuan.com/REPROSED2017E2.

Table 4. Performance measures of final model, penalised logistic regression with elastic net penalty model.

Figure 3. Coefficients of the final model, penalised logistic regression with elastic net penalty model.

Figure 4. Calibration curve of the final model, penalised logistic regression with elastic net penalty model.

Discussion

Making predictions in dengue research has gained momentum. Modelling studies in outbreak prediction utilised several predictive analytics which include ensemble methods, time series regression, and support vector machine.[7-13] Modelling studies involving clinical aspects of management of dengue - identifying and stratifying dengue - used decision trees, logistic regression, and structural equation models.[14-18] However, only 3 studies modelled mortality.[19-21] Huang et al. studied all patients with dengue that

included 34 deaths, identified five independent predictors of death: age >64 years old, diabetes mellitus, systolic blood pressure <90 mmHg, chronically bedridden, and haemoptysis.[19] However, this study used a scoring method and presence of any of the first three predictors only occurred in 162 (20·6%) cases of our cohort and the combination of all three in only 2 cases. Risk scoring has a prominent disadvantage of loss of information and resolution.[22] Moreover, it is imperative in clinical prediction that a predictor is sufficiently prevalent for the achievement of reasonable accuracy.[31] Md-Sani et al. examined severe dengue cases and predicted death at the onset of severe dengue similar to the current study but at a single centre.[20] Their multivariate analyses were limited by the number of events per variable (EPV) ratio, thus, the study employed the approach of using adjustment or controlling variables instead of assessing all candidate predictors simultaneously. Pinto et al. which had a large cohort with 61 severe dengue deaths, built a simple predictive model comprising only 4 categorical predictors: age (binary, cutoff age 55), haematuria, gastrointestinal bleeding and thrombocytopaenia (binary, cutoff platelet count 20,000 cells/mm^3).[21] In our study, we did not specifically identify the source of bleeding. However, assuming the variable warning sign of bleeding and variable severe bleeding represent haematuria and gastrointestinal bleeding, respectively, and applying their predictive model to our cohort, the presence of any of these predictors only occurred in 259 (33%) cases. The presence of all 4 predictors only occurred in 3 cases. Therefore, these models may not be adequate and accurate to prognosticate death in severe dengue. Our model kept variables in their original continuous attribute and persistent diarrhoea was the only categorical variable. Persistent diarrhoea occurred in 38.5% of our cohort, which is more prevalent than any of the predictors of Huang et al. and Pinto et al. Thus, because of this and the predominantly continuous attribute of predictors in our model, we expect a higher performance accuracy.

While there are other studies on mortality in dengue, these were studies examining the association with mortality in mostly unselected dengue cases and not prediction modelling studies.[32-39] Our study is the first modelling study, based on the latest and widely adopted WHO 2009 classification scheme, to model prediction of mortality on severe dengue cases. As mentioned previously, there have been many studies that built models for predicting identification and stratification in dengue. Our model complements this and completes the prediction aspect of clinical management.

We developed a predictive model which we have deployed as an open-access web-based tool that has very good performance and high certainty with an AUROC 89·6% (95% CI: 89·46-89·74). The model comprises 8 predictors, incorporated demographic, clinical, and laboratory variables. The model was developed through regression search for influential variables of the disease. We believe the model reflects all important underlying pathophysiological aspects of the disease represented by these variables and therefore led to its high performance. An interesting observation is that the model included the clinical symptom of persistent diarrhoea which is a common symptom in other viral hemorrhagic infection. We developed this model to be employed at the moment when severe dengue is diagnosed, which may reasonably be assumed as the time of onset of severe dengue. This is likely the time when outcome-determining pathophysiological processes become critical. This is the earliest and most appropriate choice of time to prognosticate a patient in terms of mortality in severe cases. We believe the selection of

point of time of prediction is crucial in the development of a dengue death prediction model. An earlier time point would be too early in which pathophysiological processes (which is yet to be elucidated with clarity) may not have reach outcome-determining significance. Selection of a later time point would be too late for the prediction model to be beneficial to change the outcome. By that argument and the fact that the model performed well, we postulate that investigating processes represented by these variables could elucidate the pathophysiology of dengue with better clarity. We also believe that the model may be used to track the progression of a patient through the course of illness and assists in guiding prognostication and this may be demonstrated through further studies.

The study additionally documented another interesting finding. We found that almost three-quarters of the cases were still in the febrile phase (study definition temperature >38°C) at diagnosis of severe dengue. This finding supports similar documentation in a previous study.[20] In fact, a similar proportion was found for those who had shock. This finding is different from what guidelines have assumed and depicted in which shock only occurs during the critical phase, i.e. upon defervescence and later.[2,25]

The limitation of our study is its retrospective design in which missing data was inevitable. However, only variables of utmost importance were included for variable selection. We included serum bicarbonate and serum lactate though these had higher missing proportions (11·6 and 16%, respectively) as we believe they play essential roles in determining not only if an infection progresses to a severe form but also the ultimate outcome of death.[20,40,41] Even though external validation of the model was not performed and will be a future study, we believe the repeated k-fold cross-validation algorithm we employed had ensured the robustness of the model for unseen data. A final limitation is that though our model may save costs due to its accessibility, it will require additional laboratory-related resources. Nevertheless, any severe dengue cases should be treated in a setting with adequate resources to implement evidence-based clinical practice and the laboratory predictors in our model are commercially available. Zakaria et al. demonstrated that the WHO 2009 dengue severity stratification scheme classifies more patients (4.6%) into the most severe form as compared to the previous WHO 1997 scheme (0.7%).[42] They highlighted that this may pose significant impact on hospital resources. Our model can potentially prioritise patients to local resources based on their risk of death.

In conclusion, we have developed a dengue mortality prediction model comprising clinical and laboratory data and deployed an open access web-based tool for any centre to utilise for local validation. The findings from this study would be valuable to the global community of clinicians who treat dengue, hopefully paving better and tailored clinical and resource management. In terms of research, the tool may be a useful yardstick, similar to how APACHE is useful to critical care medicine.[43,44]

Declarations

Ethics approval

The study was approved by the Medical Research and Ethics Committee (MREC), Ministry of Health of Malaysia (NMRR-16-2074-33141).

Consent for publication

Not applicable.

Availability of data

Datasets used and source codes to our analyses are given in supplementary files accompanying this article.

Competing interests

The authors declare that they have no potential conflicts of interest.

Funding

This work was supported by Medical Research Grant (NMRR-16-2074-33141), National Institutes of Health, Ministry of Health Malaysia.

Authors' contributors

SSM, MZ, RH, AMJ, PYO, SK, AS, and MM collected and contributed to data. SSM, MZ, AMJ, MA, SK, and MM designed the study. SSM analysed the data. SSM, KNL, MZ, and RH wrote the first draft of the manuscript. All authors performed data interpretation and critically reviewed the manuscript.

Acknowledgements

The authors thank Dr Siti Zubaidah A. Subki from Medical Development Division, Ministry of Health Malaysia, Dr Lee Jen Ven and Dr Duratul'ain M. Nazri from the Clinical Research Centre, Hospital Kuala Lumpur, for their commitment and diligence in their assistance running this study; and the Director General of Health Malaysia for his permission to publish this article.

Abbreviations

CI, confidence interval; AUROC, area under receiver operating characteristic curve.

References

1. Malaysian Remote Sensing Agency and National Disease Control Division, Ministry of Health Malaysia. Statistik. In: iDengue untuk komuniti. http://idengue.remotesensing.gov.my/idengue/content/statistik.pdf. Accessed 17 November 2018.

2. Ministry of Health Malaysia, Medical Development Division, Malaysia Health Technology Assessment Section (MaHTAS). Clinical Practice Guidelines: Management of Dengue Infection in Adults. 3rd Edition. Putrajaya: Ministry of Health Malaysia; 2015.

3. Shepard DS, Undurraga EA, Halasa YA. Economic and disease burden of dengue in Southeast Asia. PLoS Negl Trop Dis. 2013;7(2):e2055.

4. Packierisamy PR, Ng CW, Dahlui M, et al. Cost of dengue vector control activities in Malaysia. Am J Trop Med Hyg. 2015;93(5):1020–7.

5. Shepard DS, Undurraga EA, Lees RS, Halasa Y, Lum LC, Ng CW. Use of multiple data sources to estimate the economic cost of dengue illness in Malaysia. Am J Trop Med Hyg. 2012;87(5):796–805.

6. Shepard DS, Undurraga E, Lees R, Halasa YA, Lum LCS, Ng C. Errata. Am J Trop Med Hyg. 2013;88:606.

7. Ho CC, Ting CY, Raja DB. Using Public Open Data to Predict Dengue Epidemic: Assessment of Weather Variability, Population Density, and Land use as Predictor Variables for Dengue Outbreak Prediction using Support Vector Machine. Indian J Sci Technol. 2018;11(4).

8. Ahmad R, Suzilah I, Najdah WM, Topek O, Mustafakamal I, Lee HL. Factors determining dengue outbreak in Malaysia. PLoS One. 2018;13(2):e0193326..

9. Loshini T, Asirvadam VS, Dass SC, Gill BS. Predicting localised dengue incidences using ensemble system identification. In: 2015 International Conference on Computer, Control, Informatics and its Applications (IC3INA), 5-7 Oct 2015;Bandung, Indonesia. New York: IEEE; 2015. 6–11.

10. Buczak AL, Baugher B, Moniz LJ, Bagley T, Babin SM, Guven E. Ensemble method for dengue prediction. PLoS One. 2018;13(1):e0189988.

11. Morin CW, Monaghan AJ, Hayden MH, Barrera R, Ernst K. Meteorologically driven simulations of dengue epidemics in San Juan, PR. PLoS Negl Trop Dis. 2015;9(8):e0004002..

12. Yamana TK, Kandula S, Shaman J. Superensemble forecasts of dengue outbreaks. J R Soc Interface. 2016;13(123):20160410.

13. Ramadona AL, Lazuardi L, Hii YL, Holmner Å, Kusnanto H, Rocklöv J. Prediction of dengue outbreaks based on disease surveillance and meteorological data. PLoS One. 2016;11(3):e0152688.

14. Tanner L, Schreiber M, Low JG, et al. Decision tree algorithms predict the diagnosis and outcome of dengue fever in the early phase of illness. PLoS Negl Trop Dis. 2008;2(3):e196.

15. Cucunawangsih DB, Sungono V. Scoring Model to Predict Dengue Infection in the Early Phase of Illness in Primary Health Care Centre. Arch Clin Microbiol. 2015;6: 2.

16. Lee VJ, Lye DC, Sun Y, Leo YS. Decision tree algorithm in deciding hospitalisation for adult patients with dengue haemorrhagic fever in Singapore. Trop Med Int Health. 2009;14(9):1154–9.

17. Potts JA, Gibbons RV, Rothman AL, et al. Prediction of dengue disease severity among pediatric Thai patients using early clinical laboratory indicators. PLoS Negl Trop Dis. 2010;4(8):e769.

18. Carrasco LR, Leo YS, Cook AR, et al. Predictive tools for severe dengue conforming to World Health Organization 2009 criteria. PLoS Negl Trop Dis. 2014;8(7):e2972.

19. Huang CC, Hsu CC, Guo HR, Su SB, Lin HJ. Dengue fever mortality score: A novel decision rule to predict death from Dengue fever. J Infect. 2017;75(6):532–40.

20. Md-Sani SS, Md-Noor J, Han WH, et al. Prediction of mortality in severe dengue cases. BMC Infect Dis. 2018;18(1):232.

21. Pinto RC, de Castro DB, de Albuquerque BC, et al. Mortality predictors in patients with severe dengue in the State of Amazonas, Brazil. PLoS One. 2016;11(8):e0161884.

22. Hubbard D, Evans D. Problems with scoring methods and ordinal scales in risk assessment. IBM J Res Dev. 2010;54(3):2–1.

23. Hales S, De Wet N, Maindonald J, Woodward A. Potential effect of population and climate changes on global distribution of dengue fever: an empirical model. Lancet. 2002;360(9336):830–4.

24. Moons KG, Altman DG, Reitsma JB, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. Ann Intern Med. 2015;162(1):W1–73.

25. Department of Control of Neglected Tropical Diseases, World Health Organization. Epidemic, & Pandemic Alert. Dengue: guidelines for diagnosis, treatment, prevention and control. Geneva: World Health Organization; 2009.

26. James G, Witten D, Hastie T, et al. An introduction to statistical learning. New York: Springer; 2013; Chapter 6 Linear Model Selection and Regularization (Section 6.2 Shrinkage Methods): 214-227.

27. Friedman J, Hastie T, Tibshirani R. Regularization Paths for Generalized Linear Models via Coordinate Descent. Journal of Statistical Software. 2010;33(1):1-22.

28. Kuhn, M. Building Predictive Models in R Using the caret Package. Journal of Statistical Software. 2008;28(5):1-26.

29. Pavlou M, Ambler G, Seaman SR, et al. How to develop a more accurate risk prediction model when there are few events. BMJ. 2015;351:h3868.

30. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2016. URL https://www.R-project.org

31. Schummers L, Himes KP, Bodnar LM, Hutcheon JA. Predictor characteristics necessary for building a clinically useful risk prediction model: a simulation study. BMC Med Res Methodol. 2016;16(1):123.

32. Amâncio FF, Heringer TP, Fassy LB, et al. Clinical profiles and factors associated with death in adults with dengue admitted to intensive care units, Minas Gerais, Brazil. PLoS One. 2015;10(6):e0129046.

33. Werneck GL, Macias AE, Mascarenas C, et al. Comorbidities increase in-hospital mortality in dengue patients in Brazil. Mem Inst Oswaldo Cruz. 2018;113(8).

34. Almas A, Parkash O, Akhter J. Clinical factors associated with mortality in dengue infection at a tertiary care centre. Southeast Asian J Trop Med Public Health. 2010;41(2):333–40.

35. Khalil MA, Tan J, Khalil MA, Awan S, Rangasami M. Predictors of hospital stay and mortality in dengue virus infection-experience from Aga Khan University Hospital Pakistan. BMC Res Notes. 2014;7(1):473.

36. Karunakaran A, Ilyas WM, Sheen SF, Jose NK, Nujum ZT. Risk factors of mortality among dengue patients admitted to a tertiary care setting in Kerala, India. J Infect Public Health. 2014;7(2):114–20.

37. Chowdhury R, Pan K, Sarkar A, Chakrabarti S, Samanta T. Predictors of dengue mortality in a tertiary care hospital at kolkata: a cross-sectional study. Int J Med Sci Public Health. 2013;2(2):254–8.

38. Acharya V, Khan MF, Kosuru S, Mallya S. Predictors of mortality in adult patients with dengue: a study from South India. International Journal of Research in Medical Sciences. 2018;6(5):1605–10.

39. Singh J, Zeya MT, Dhir G, et al. Mortality and Severity Predictors of Dengue Fever. Scholars Journal of Applied Medical Sciences. 2014;2(6A):1958–61.

40. Thanachartwet V, Oer-Areemitr N, Chamnanchanunt S, et al. Identification of clinical factors associated with severe dengue among Thai adults: a prospective study. BMC Infect Dis 2015;15(1):420.

41. Shankar-Hari M, Phillips GS, Levy ML, et al. Developing a new definition and assessing new clinical criteria for septic shock: for the Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). JAMA. 2016;315(8):775–787.

42. Zakaria Z, Zainordin NA, Sim BL, et al. An evaluation of the World Health Organization's 1997 and 2009 dengue classifications in hospitalized dengue patients in Malaysia. The Journal of Infection in Developing Countries. 2014;8(07):869-75.

43. Bouch DC, Thompson JP. Severity scoring systems in the critically ill. Continuing Education in Anaesthesia, Critical Care & Pain. 2008;8(5):181-5.

44. Vincent JL, Moreno R. Clinical review: scoring systems in the critically ill. Critical care. 2010;14(2):207.

Tables

Table 1. Demographic and clinical variables.

	Died (N=35)		Survived (751)		
	n	Median (IQR) / %	n	Median (IQR) / %	*p*a value
Age (years)	35	42 (22·6)	751	30 (20)	·001
Gender (male)	20	57·1	433	57·7	NS
BMI (kg/m2)	31	28·1 (7·1)	686	24·7 (7·7)	·001
Systolic BP (mmHg)	35	105 (3·5)	751	111 (24)	NS
Diastolic BP (mmHg)	35	67 (25)	751	68 (18)	NS
Mean Arterial Pressure (mmHg)	35	81·3 (22·3)	751	83 (18·7)	NS
Pulse rate (bpm)	35	106 (41)	751	94 (27)	NS
Respiratory rate (breaths/min)	35	23 (8)	746	20 (2)	<·0001
Co-morbidities	19	54·3	227	3·2	·004
Hypertension	11	31·4	97	12·9	·003
Diabetes mellitus	7	2·0	70	9·3	·04
COPD/Asthma	5	14·3	46	6·1	NS
Heart disease	2	5·7	28	3·7	NS
Chronic kidney disease	1	2·9	7	0·9	NS
Chronic liver disease	1	2·9	1	0·1	·03
Laboratory diagnosis of dengue fever					
NS1	31	88·6	612	81·5	NS
IgM	19	54·3	348	46·3	NS
IgG	13	37·1	236	31·4	NS
Types of severe dengue					
Compensated shock	14	4·0	339	45·1	NS
Decompensated shock	15	42·9	174	23·2	·01
Severe bleeding	15	42·9	46	6·1	<·0001
Severe hepatitis	25	71·4	106	14·1	<·0001
Acute kidney impairment	28	8·0	185	24·6	<·0001

Encephalitis/encephalopathy	5	14·3	20	2·7	·0007
Cardiac complications	19	54·3	147	19·6	<·0001
Respiratory compromise	21	6·0	79	1·5	<·0001
Warning signs at diagnosis of severe dengue					
Persistent vomiting	23	65·7	369	49·1	NS
Persistent diarrhoea	23	65·7	280	37·3	·001
Abdominal pain	22	62·9	307	4·9	·01
Abdominal tenderness	12	34·3	141	18·8	·03
Third space fluid accumulation	16	45·7	125	16·6	<·0001
Spontaneous bleeding tendencies	13	37·1	98	13·0	·0002
Raised hematocrit with rapid drop of platelet	19	57·6	325	44·2	NS
Associated bloodstream infection	7	20	28	3·7	·0001
Timing of events					
Length of illness	35	6·9 (3·4)	750	7·6 (2·6)	NS
Length of stay	35	3·1(3·6)	750	4·0 (2·8)	NS
Day of onset of severe dengue	35	3·7 (1·7)	750	4·0 (2·0)	NS
Length of febrile phase	27	4·3 (2·3)	747	4·8 (2·1)	NS
Febrile severe dengue	14	51·9	561	75·1	·009
Febrile phase at admission	18	66·7	621	83·2	·03
Presentation as severe dengue	3	8·6	184	24·5	·04

Abbreviations: IQR, interquartile range; BMI, body mass index; BP, blood pressure; COPD, chronic obstructive pulmonary disease; NS1, non-structural protein 1 antigen; IgM, immunoglobulin M; IgG, immunoglobulin G; NS, not significant.

*a*Univariate logistic regression was used to compare differences between groups. Analysis was made on complete case basis.

Table 2. Laboratory variables.

	Died		Survived		
	n	Median (IQR) / %	n	Median (IQR) / %	*p*a value
At diagnosis of severe dengue					
WBC (×103/μL)	33	7·5 (5·7)	738	4·0 (2·9)	<·0001
Hb (g/dL)	33	14·8 (4·4)	738	14·6 (3·2)	NS
Platelet (×103/μL)	33	23 (35)	737	67 (99)	·0001
Hct (%)	34	43·8 (13)	742	43·5(8·6)	NS
AST (U/L)	34	813 (1565)	695	105 (202)	<·0001
ALT (U/L)	32	322 (480)	674	65·5 (123·5)	·0001
Albumin (g/dL)	33	32 (10)	678	36 (8)	<·0001
INR	21	1·41 (·46)	395	1·1 (1·43)	NS
Serum Bicarbonate (mmol/L)	34	17·9 (5·925)	685	23 (3·9)	<·0001
Serum Lactate (mmol/L)	34	3·92 (5·55)	626	1·6 (1·06)	<·0001
Serum Creatinine (μmol/L)	29	115 (46·4)	665	85 (41)	·003
At 24 hours after diagnosis of severe dengue					
WBC (×103/μL)	26	8·05 (9·9)	683	4·1 (3·95)	<·0001
Hb (g/dL)	26	12·55 (4·3)	683	13·8 (3)	·008
Platelet (×103/μL)	26	24·5 (33)	682	49 (73)	·007
Hct (%)	26	38·7 (12·2)	680	4·55 (7·8)	·02
AST (U/L)	21	3628 (9560)	446	153 (282)	<·0001
ALT (U/L)	23	1384 (1746·5)	444	89 (159)	<·0001
Albumin (g/dL)	21	30 (12)	435	31 (8)	NS
INR	20	1·67 (·76)	178	1·08 (·19)	<·0001
Serum Bicarbonate (mmol/L)	21	15·4 (1·4)	375	23 (3·2)	<·0001
Serum Lactate (mmol/L)	20	6·1 (12·3)	338	1·3 (·8)	<·0001
Serum Creatinine (μmol/L)	22	147·5 (77·5)	413	69 (34·9)	<·0001
Nadir and Peak Values					

Baseline Hct (%)	35	41 (7·05)	749	41·2 (7·6)	NS
Peak Hct (%)	35	49·9 (7·05)	749	46 (8·4)	·005
Day of Peak Hct	34	4·5 (2·3)	706	4·4 (2·2)	NS
Nadir Platelet (×103/μL)	35	8 (11)	749	25 (48)	·0007
Day of Nadir Platelet	34	4·1 (1·9)	738	5·4 (1·8)	NS
Peak AST (U/L)	35	4202 (12307·5)	750	165·5 (288)	<·0001
Day of Peak AST	35	5 (2·2)	744	5·0 (2·2)	NS
Peak ALT (U/L)	35	1510 (2550)	750	99 (166)	<·0001
Day of Peak ALT	34	5·2 (2·3)	743	5·2 (2·4)	NS
Peak Creatinine (μmol/L)	34	203·5 (13·5)	747	89 (41)	<·0001
Day of Peak Creatinine	35	5·3 (3·0)	739	4·1 (2·2)	NS

Abbreviations: IQR, interquartile range; WBC, white blood cell count; Hb, haemoglobin; Hct, haematocrit; AST, aspartate transaminase; ALT, alanine transaminase; INR, international normalised ratio; NS, not significant.

*a*Univariate logistic regression was used to compare differences between groups. Analysis was made on complete case basis.

Table 3. Univariate and multivariable analysis of selected variables for risk of death in severe dengue.

	Unadjusted ORa (95% CI)	*p*a value	Adjusted OR*b* (95% CI)	*pb* value
Clinical parameters at diagnosis of SD				
Age (years)	1·03 (1·01-1·06)	·001	-	-
Gender (male)	0·98 (·50-1·97)	NS	-	-
BMI (kg/m2)	1·07 (1·03-1·12)	·001	1·07 (1·02-1·12)	·006
Systolic BP (mmHg)	0·98 (·97-1·00)	NS	0·98 (·96-·99)	·008
Diastolic BP (mmHg)	0·99 (·97-1·01)	NS	0·98 (·96-1·00)	NS
Mean Arterial Pressure (mmHg)	0·98 (·96-1·01)	NS	0·98 (·95-1·00)	·03
Pulse rate (bpm)	1·02 (1·00-1·04)	NS	1·03 (1·01-1·05)	·01
Respiratory rate (breaths/min)	1·13 (1·07-1·19)	<·0001	1·12 (1·05-1·20)	·0003
Co-morbidities	2·74 (1·39-5·49)	·004	1·60 (·71-3·59)	NS
Associated bloodstream infection	6·46 (2·43-15·39)	·0001	3·58 (1·21-9·54)	·01
Laboratory confirmation of dengue fever				
NS1	1·76 (·68-5·99)	NS	1·61 (·59-5·68)	NS
IgM	1·38 (·70-2·75)	NS	1·45 (·71-2·99)	NS
IgG	1·29 (·62-2·57)	NS	1·37 (·63-2·90)	NS

Warning signs at diagnosis of SD				
Persistent vomiting	1·98 (·99-4·18)	NS	1·42 (·63-3·26)	NS
Persistent diarrhoea	3·22 (1·61-6·79)	·001	2·54 (1·23-5·49)	·01
Abdominal pain	2·45 (1·23-5·06)	·01	2·44 (1·18-5·26)	·02
Abdominal tenderness	2·26 (1·06-4·57)	·03	2·25 (1·01-4·79)	·04
Third space fluid accumulation	4·22 (2·09-8·43)	<·0001	3·15 (1·43-6·85)	·004
Spontaneous bleeding tendencies	3·94 (1·88-7·98)	·0002	2·7 (1·21-5·82)	·01
Raised hematocrit with rapid drop of platelet	1·72 (·85-3·54)	NS	1·27 (·60-2·73)	NS
Timing of events				
Length of illness	1·01 (·93-1·07)	NS	0·96 (·87-1·03)	NS
Length of stay	1·00 (·90-1·06)	NS	0·93 (·83-1·01)	NS
Febrile SD	0·36 (·16 - ·78)	·009	0·58 (·24-1·42)	NS
Febrile phase at admission	0·40 (·18 - ·96)	·03	0·46 (·18-1·20)	NS
Presentation as SD	0·29 (·07 - ·82)	·04	0·42 (·10-1·25)	NS
At diagnosis of severe dengue				
WBC (×103/µL)	1·16 (1·08 – 1·25)	<·0001	1·13 (1·04-1·22)	·004

Hb (g/dL)	1·08 (·94 -1·25)	NS	1·05 (·90- 1·23)	NS
Platelet (×103/μL)	0·98 (·97- ·99)	·0001	0·98 (·97-·99)	·004
Hct (%)	1·00 (·96 – 1·05)	NS	0·98 (·94- 1·04)	NS
AST (U/L)	1·00 (1·00 -1·00)	<·0001	1·00 (1·00- 1·00)	·0001
ALT (U/L)	1·00 (1·00 -1·00)	·0001	1·00 (1·00- 1·00)	·01
Albumin (g/dL)	0·85 (·80 – ·91)	<·0001	0·86 (·80-·93)	·0001
INR	1·43 (1·00 -2·78)	NS	1·15 (·79- 1·95)	NS
Serum Bicarbonate (mmol/L)	0·70 (·63 – ·77)	<·0001	0·72 (·63-·80)	<·0001
Serum Lactate (mmol/L)	1·50 (1·31 – 1·75)	<·0001	1·39 (1·23- 1·61)	<·0001
Serum Creatinine (μmol/L)	1·01 (1·00 -1·01)	·003	1·00 (1·00- 1·01)	NS

Abbreviations: CI, confidence interval; WBC, white blood cell count; Hb, haemoglobin; Hct, haematocrit; AST, aspartate transaminase; ALT, alanine transaminase; NS, not significant. Age, gender, centre and day of onset of SD were used as controlling variables.

*a*Univariate logistic regression was used to compare differences between groups. Analysis was made on complete case basis.

*b*Multivariable logistic regression was used to compare differences between groups; age, gender, centre and day of onset of severe dengue were used as controlling variables. Analysis was made on complete case basis.

Table 4. Performance measures of the final model, penalised logistic regression with elastic net penalty model·

Performance	Mean (95% CI)
AUROC	89·6 (89·5-89·8)
Sensitivity	99·62 (99·61-99·63)
Specificity	23·56 (23·13-24·00)
Positive predictive value	96·56 (96·53-96·58)
Negative predictive value	70·94 (70·05-71·83)
Positive likelihood ratio	1·45 (1·44-1·46)
Negative likelihood ratio	0·012 (·011-·012)

Abbreviations: CI, confidence interval; AUROC, area under receiver operating characteristic curve.

Figures

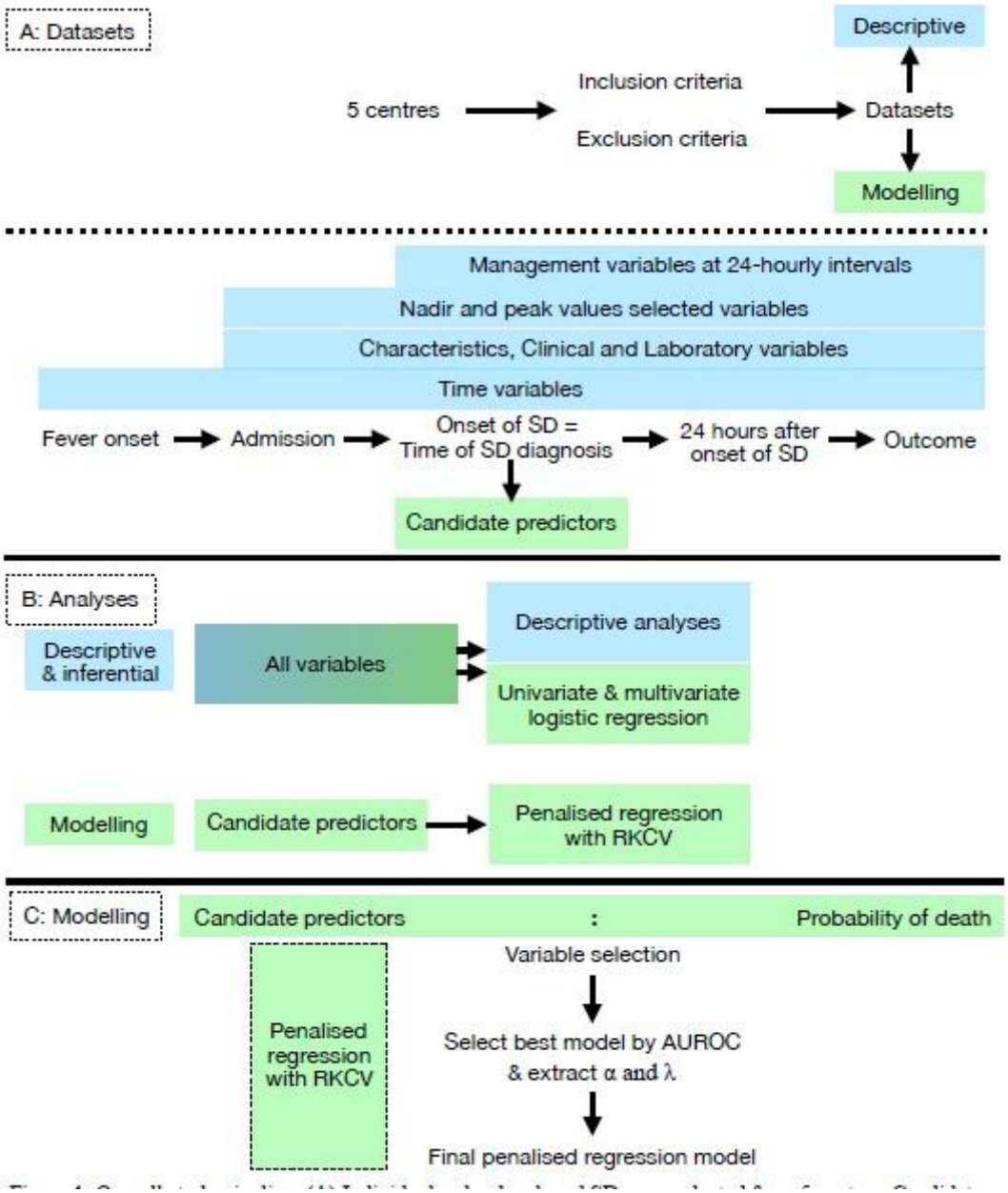

Figure 1

Overall study pipeline.

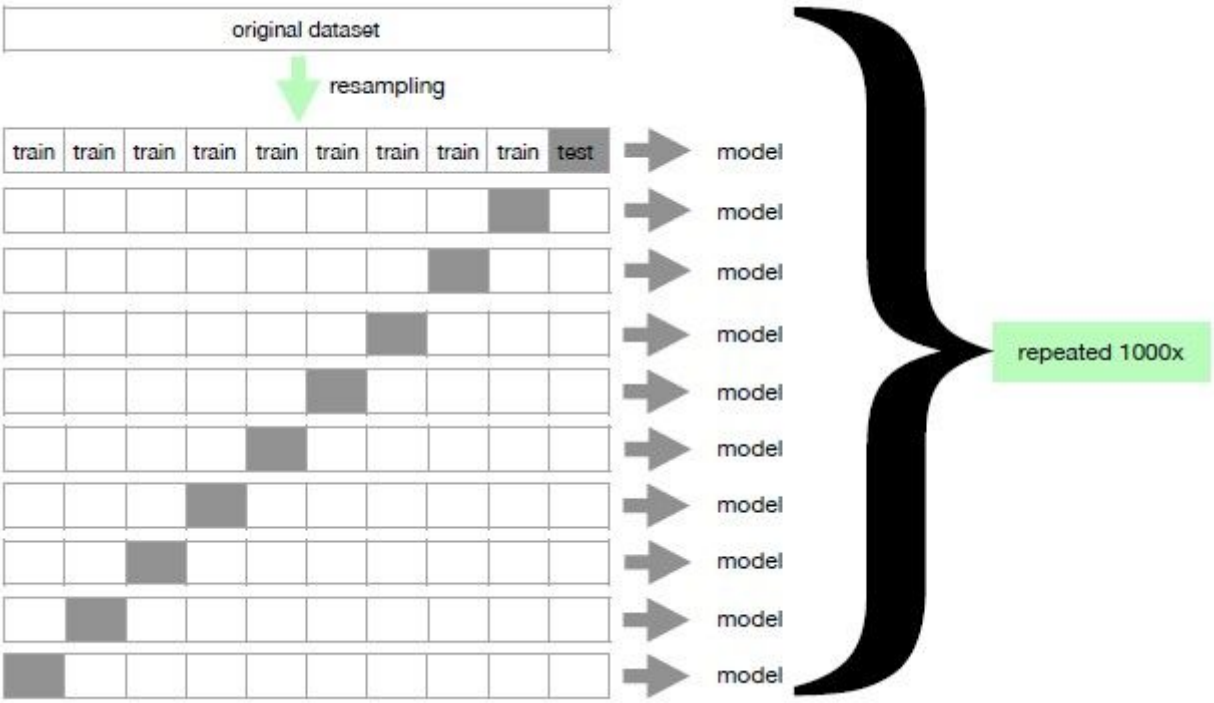

Figure 2

Internal validation scheme.

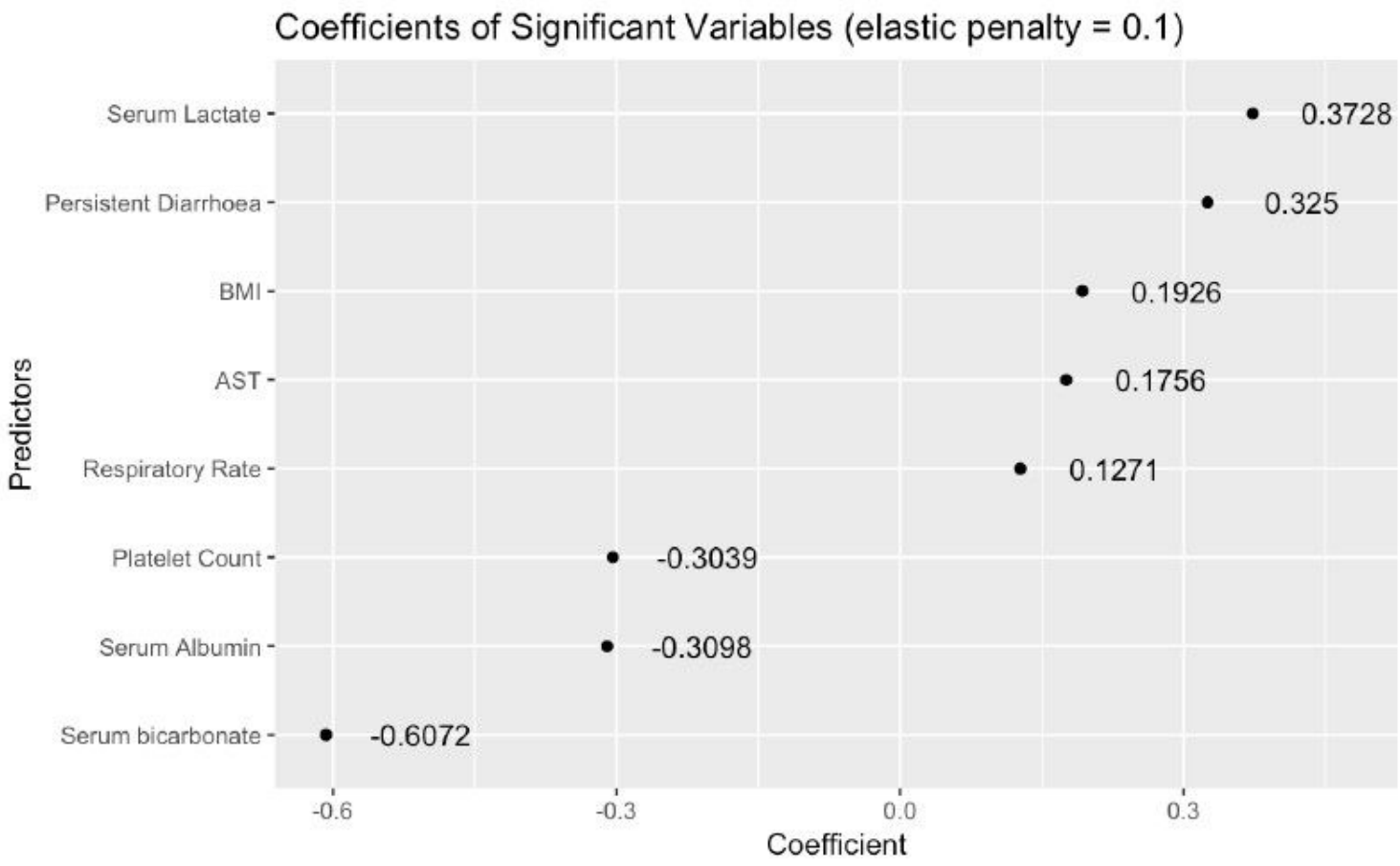

Figure 3

Coefficients of the final model, penalised logistic regression with elastic net penalty model.

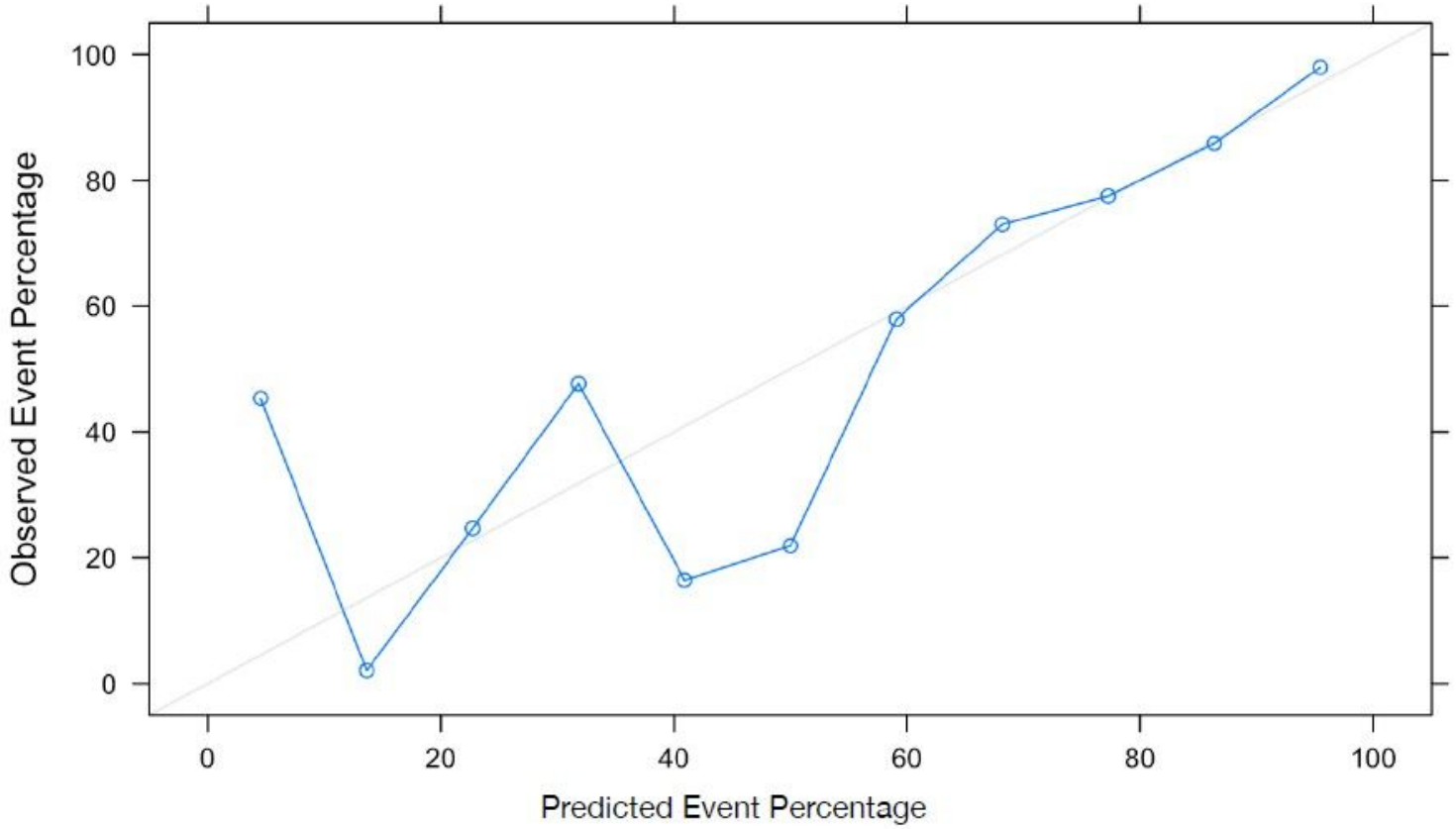

Figure 4

Calibration curve of the final model, penalised logistic regression with elastic net penalty model.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- supplement1.pdf
- supplement2.pdf
- supplement2.pdf
- supplement4.csv
- supplement5.csv