
Clinical risk prediction of progression to severe dengue illness during the febrile phase in primary healthcare settings

Sorawat Sangkaew

Imperial College London

Department of Infectious Disease

**Thesis submitted for the degree of Doctor of Philosophy in
Clinical Medicine Research (Infectious Disease)**

2021

Declaration of originality

I declare this work to be my own, conducted under the supervision of Prof. Alison Holmes, Dr Ilaria Dorigatti, and Dr Sophie Yacoub. Any participants involved in collaborative work have been acknowledged in the acknowledgement section of this thesis.

The copyright of this thesis rests with the author and is made available under the Creative Commons Attribution. Non-Commercial No Derivatives licence. Researchers are free to copy, distribute or transmit the thesis on the condition that they attribute it, that they do not use it for commercial purposes and that they do not alter, transform, or build upon it. For any reuse or redistribution, researchers must make clear to others the licence terms of this work.

Abstract

To date, no specific antiviral treatment is available for dengue. The recognition of clinical progression to severe disease during the early phases of the illness and the decisions made at triage are key for the clinical management and often determine the outcome. Previous studies have been developed to explore the relationship between risk factors and disease severity and have identified the risk predictors of severe dengue illness. However, the results of these studies have been inconsistent due to studies with small sample sizes and the significant variation in predictors on different days of illness. In addition, previous risk prediction models for early progression to severe disease were based on hospitalised patient data, thus missing to investigate early predictors potentially capable to inform triage and decision making in outpatient settings, where infections are first seen. This thesis aims to fill this knowledge gap and investigates how early risk predictors, collected within the first four days since symptom onset, can be used in conjunction with clinical prediction models to estimate individualised risks of progression to severe disease in outpatient settings.

This thesis provides evidence that platelet count, liver function test, and serum albumin can be used as risk predictors of progression to severe disease, backed both by a systematic review and meta-analysis of the existing literature, and by the analysis of clinical data from two large cohort studies conducted in Thailand and Vietnam. To investigate the development of the first signs of severe diseases and compensate for the infrequency of dengue shock syndrome, we investigate two clinical endpoints (i.e. dengue shock syndrome and a combined endpoint of dengue shock syndrome and/or moderate plasma leakage) using a variety of statistical models, including logistic regression and machine learning techniques, such as extreme gradient boosted tree and artificial neural network. The risk prediction models were developed following best practices, were optimised through internal and external validation, and using independent datasets from Thailand and Vietnam, two countries in dengue-endemic areas. We also characterised the temporal dynamics of clinical and laboratory parameters throughout disease progression and investigated the use and predictive power of sequential measurements for prediction, using recurrent neural networks.

The use of statistical predictive models to inform triage and clinical management of dengue in the outpatient setting is promising and has the potential to support decision making in the early phases

of disease in clinical practice. While further and extensive validation across healthcare contexts and populations is needed, this work lays the foundations for integrating evidence-based and data-driven methods into a decision support system in clinical practice, which in turn can contribute to informing decision making and optimising healthcare allocation and the optimal use of resources.

Acknowledgements

First and foremost, I would like to express my gratitude to Prof Alison Holmes, my main supervisor for the opportunity and freedom to pursue this research. I am very thankful for your support and positivity throughout the PhD supervision. Very special gratitude goes out to Dr Ilaria Dorigatti, my second supervisor, for incredible support, guidance, and encouragement. I have been very delighted to have you as a supervisor for both MSc and PhD programmes. Without your help and support, it would not be possible to get to this point. I am also very thankful to Dr Sophie Yacoub, my third supervisor for all support and guidance. I am also very pleased to join and to be a part of the Dengue Group at the Oxford University Clinical Research Unit (OUCRU) in Vietnam.

I would like to thank Prof Siripen Kalayanaroj, Dr Anon Srikiatkachorn, research and administration teams at Armed Forces Research Institute of Medical Sciences and Queen Sirikit National Institute of Child Health for all help and sharing data, mainly used in my PhD projects. In addition, I am very grateful to Prof Bridget Wills, Prof Cammeron Simmon, Dr Minh Tuan Nguyen, for sharing data, which was also used mainly in this thesis. Also, I am very thankful to Dr Nguyen Lam Vuong, Phung Khanh Lam, and all member of the Dengue group for sharing your valuable experiences on statistics and dengue case management. I also thank you for a warm welcome and for taking good care of me during the visit.

To colleagues and friends in the 7th-floor and 8th-floor offices, I am very thankful for all help and support during working here, such a very nice and supportive working environment. Also, very special thanks go to Dr Damien Ming, I have been very enjoyed working with you and very glad to have you as a friend in the office. I cannot thank you enough for all your help, support, and your contributions to my PhD projects.

Special gratitude also goes to the Collaborative Project to Increase Rural Doctor, Ministry of Public Health, Thailand, providing funding for this academic programme and to my dengue infected patients who inspired me to be interested in dengue. And finally, I would like to thank my family for believing in me and allowing me to explore and do whatever I made decisions. Five years away from home was not easy for me and my family, I do understand and now it is time to get back home, work and help people in Thailand. I would like to dedicate this achievement to my family.

Contents

Declaration of originality.....	1
Abstract	2
Acknowledgements.....	4
Contents	5
List of tables	9
List of figures	11
Abbreviations	13
Chapter 1 Introduction.....	15
1.1 Overview of dengue infection	15
1.1.1 Epidemiological data	15
1.1.2 Dengue virus.....	15
1.1.3 Mechanism of severe disease	16
1.1.4 Clinical manifestations	18
1.1.5 Diagnosis	19
1.1.6 Dengue disease classification.....	21
1.1.7 Clinical management.....	23
1.2 Factors associated with progression to severe dengue	24
1.2.1 Demographic factors	24
1.2.2 Clinical signs and symptoms.....	24
1.2.3 Conventional laboratory parameters.....	25
1.2.4 Imaging	25
1.2.5 Biomarkers	25
1.2.6 Novel vascular function techniques	28
1.3 Prognostic models for progression to severe dengue illness.....	28

1.4	Knowledge gaps and areas of interest	30
1.5	Aims and outline of the thesis	36
Chapter 2	Risk predictors for progression to severe dengue illness: a systematic review and meta-analysis	37
2.1	Rationale	37
2.2	Methodology	38
2.2.1	Search strategy	38
2.2.2	Study selection	38
2.2.3	Data extraction and quality assessment	40
2.2.4	Meta-analysis	40
2.2.5	Heterogeneity assessment	41
2.2.6	Analysis for special relationship	42
2.3	Results	42
2.3.1	Demographics	46
2.3.2	Clinical signs and symptoms	50
2.3.3	Immune response (primary vs secondary infection)	55
2.3.4	Virological parameters	55
2.3.5	Subgroup analyses of prognostic factors in children and adults	56
2.3.6	Risk of bias and quality assessment	56
2.4	Discussion	57
Chapter 3	Materials and analytical methods	65
3.1	Materials	65
3.1.1	Data of a prospective cohort study of Vietnamese children with dengue	65
3.1.2	Data of a prospective cohort study of Thai children with dengue	66
3.1.3	Laboratory-confirmed dengue diagnosis	67
3.2	Common analytical methods	68

3.2.1	Descriptive analysis	68
3.2.2	Treatment of missing data	68
3.2.3	Definition of prediction outcomes	69
3.2.4	Metrics of predictive performance	70
3.3	Methods for the development of prediction models using baseline clinical parameters measured within the first 72 hours of illness.....	72
3.3.1	Predictor selection	72
3.3.2	Statistical and machine learning models	76
3.3.3	Hyperparameter optimisation	78
3.3.4	Model development.....	81
3.3.5	Model validation	81
3.4	Sensitivity analysis	84
3.5	Methods for the analyses of the temporal dynamics of laboratory parameters	84
3.5.1	Analytical methods for describing temporal dynamics of clinical laboratory parameters	84
3.5.2	Predictor selection for risk prediction using daily clinical parameters.....	87
3.5.3	Analytical methods used for risk prediction using temporal dynamics of clinical parameters.....	91
3.5.4	Combination of multiple imputation and internal validation.....	95
Chapter 4	Risk prediction for progression to severe dengue illness using clinical parameters measured during the first 72 hours of illness	98
4.1	Rationale.....	98
4.2	Methods.....	99
4.3	Results.....	99
4.3.1	Characteristics and clinical baseline of patients	99
4.3.2	Selected predictors.....	105

4.3.3	Association of predictors and the clinical endpoints	111
4.3.4	Internal predictive performances	113
4.4	External validation	122
4.5	Sensitivity analysis	124
4.6	Discussion	124
Chapter 5	Risk prediction of progression to severe dengue illness from the observed dynamics of clinical parameters collected routinely during the febrile phase	128
5.1	Rationale	128
5.2	Method	129
5.3	Result	130
5.3.1	Baseline characteristics of patients in the Thai dataset	130
5.3.2	Clinical laboratory parameters of patients by the difference groups of clinical endpoints.....	132
5.3.3	Comparison of clinical parameters by day of illness.....	134
5.3.4	Temporal dynamics of clinical laboratory parameters	135
5.3.5	Selected dynamic predictors.....	138
5.3.6	Internal predictive performance	138
5.4	Discussion	141
Chapter 6	Conclusions.....	148
6.1	Summary and contributions	148
6.1.1	Risk predictors of the progression to severe dengue disease	148
6.1.2	Individualised risk prediction of the progression to severe dengue disease.....	150
6.2	Future Research Direction	153
References		156
Appendix		166

List of tables

Table 1.1. Summary of advantages and disadvantages of diagnostics	21
Table 1.2. Diagnostic criteria for severe dengue and a list of warning signs adapted from the 2009 WHO Classification	22
Table 1.3. Summary of 17 studies on the development risk prediction models for dengue severity included in this review	31
Table 2.1. Template used to assess the eligibility criteria	39
Table 2.2. Definitions of factors included in the systematic review and meta-analysis	44
Table 2.3. Summary of factors measured within the first 4 days of illness and their association with progression to severe dengue illness.....	58
Table 3.1. Clinical endpoints and definition	70
Table 3.2. Definition and type of data variables available in the Vietnamese dataset.....	72
Table 3.3. Table of the lower and upper bounds of hyperparameters used in each machine learning algorithms.....	79
Table 3.5. Definition and type of data variables available in Thai dataset.....	87
Table 4.1. Characteristics and clinical variables at enrolment of the Vietnamese patients in complete dataset with laboratory confirmed dengue infection developing Dengue Shock Syndrome and the combined endpoint of moderate plasma leakage and/or DSS.....	101
Table 4.2. Characteristics and clinical variables at enrolment of the Thai patients in complete dataset with laboratory confirmed dengue infection developing Dengue Shock Syndrome and the combined endpoint of moderate plasma leakage and/or DSS.	103
Table 4.3. Considered candidate predictors of Dengue Shock Syndrome (DSS) and the combined endpoint of moderate plasma leakage and/or DSS in the Vietnamese dataset.	106
Table 4.4. Global model, model selected by lasso selection with the minimum lambda and bootstraps-derived quantities for assessing the uncertainty of model for DSS.....	108
Table 4.5. Top 10 combinations of predictors selected using multivariable logistic regression with lasso selection to predict the risk of developing DSS.	109
Table 4.6. Global model, model selected by lasso selection and bootstraps-derived quantities for assessing the uncertainty of model for the combined endpoint.....	110

Table 4.7. Top 10 combinations of predictors selected using multivariable logistic regression with lasso selection to predict the risk of developing the combined endpoint (DSS and/or plasma leakage) using 1,000-bootstrap samples of the data.	111
Table 4.8. Associations of candidate predictors with dengue shock syndrome (DSS) and the combined endpoint of DSS and/or moderate plasma leakage derived from univariate logistic regression.....	112
Table 4.9. Table of optimal hyperparameters selected using Bayesian Global Optimisation with Gaussian Processes of models for DSS and the combined endpoint.....	113
Table 4.10. Summary table of the models with multivariable logistic regression with Lasso selection for DSS and the combined endpoint	115
Table 4.11. Predictive performance of the risk prediction models trained on the Vietnamese dataset on internal validation using 10-fold cross validation.	117
Table 4.12. Predictive performance of the five models fitted on the Vietnamese (training) dataset and validated on the Thai (external validation) dataset.	123
Table 4.13. Predictive performance of the risk prediction models for the combined endpoint of moderate plasma leakage and/or severe dengue trained on hospitalised patients in the Vietnamese dataset on internal validation using 10-fold cross validation.....	126
Table 5.1. Eight laboratory parameters included in the analysis of temporal dynamics	129
Table 5.2. Baseline characteristics of patients by groups of clinical endpoints	133
Table 5.3. Standardised QICs of models estimating the dynamics of clinical parameters during febrile phase of dengue.....	135
Table 5.4. Coefficients and p-value of the interaction terms between the outcome group and day of illness.....	136

List of figures

Figure 1.1. The viral genome and structure of dengue virus adapted	16
Figure 1.2. The clinical progression of dengue illness classified into the febrile phase, leakage phase and recovery phase. Reprinted from 2009 WHO Dengue Guidelines.	19
Figure 2.1. Flow diagram of study selection	43
Figure 2.2. Pooled SMD of factors potentially associated with progression to severe dengue.....	47
Figure 2.3. Pooled SMD of factors potentially associated with progression to severe dengue by age of study population.	48
Figure 2.4 Pooled ORs of progression to severe dengue given the presence of potential prognostic factors.....	51
Figure 2.5. Pooled OR of progression to severe dengue given the presence of potential prognostic factors by age of study population.	52
Figure 2.6. Risk of bias in the included studies assessed by the Quality in Prognostic Studies.	57
Figure 3.1. Description of the workflow used for model development and internal validation.....	82
Figure 3.2. Description of the workflow used for external validation.....	84
Figure 3.3. Different types of RNN by input and output structure.....	92
Figure 3.4. Forward propagation for a basic RNN model (A) and RNN unit (B)	93
Figure 3.5. LSTM basic structure, adapted	94
Figure 3.6. GRU basic structure, adapted	95
Figure 3.7. Combination of multiple imputation and internal validation using bootstrapping.	96
Figure 4.1. Flow chart describing patients' outcomes in the Vietnamese training and Thai external validation datasets	100
Figure 4.2. Calibration plots for the predicted risk of DSS (triangles) and the combined endpoint of moderate plasma leakage and/or DSS (points).	120
Figure 4.3. Decision curves, showing the net benefit against the threshold probability for DSS and the combined endpoint of moderate plasma leakage and/or DSS.....	121
Figure 4.4. Results of difference in means with 95% CI and relative difference with 95% CI in standard deviation of predicted logarithmic odds on the training and validation sets.....	122
Figure 5.1. Flow chart describing patients' clinical endpoints of the Thai dataset	131

Figure 5.2. Percentage of dengue infections by groups of the combined endpoint with laboratory predictors by day of illness and day of leakage.	132
Figure 5.3. Percentage of dengue infections by groups of DSS diagnosis with laboratory predictors by day of illness and day of leakage.....	132
Figure 5.4. P-value of the coefficient of groups by clinical endpoints (A for DSS and B for the combined endpoints of moderate plasma leakage and/or DSS) on the specified day of illness obtained from the multivariable models with adjustment of age and sex	140
Figure 5.5. Individual trajectories of clinical parameters for all patients who developed and did not develop the combined endpoint between day 1 and day7 of illness	143
Figure 5.6. Individual trajectories of clinical parameters for patients who developed and did not develop DSS between day 1 and day 7 of illness	144

Abbreviations

ADE	antibody dependence enhancement
AFRIMS	Armed Forces Research Institute of Medical Sciences
ALT	alanine aminotransferase
Ang	angiopoietin
ANN	Artificial neural network
AST	aspartate aminotransferase
AUROC	area under receiver operating characteristics curves
C protein	capsid protein
CART	classification and regression tree
CD	cluster of differentiation
CI	confidence interval
CITL	calibration in the large
CNS	central nervous system
CRI	compensatory reserve index
CRP	C-reactive protein
DCA	decision curve analysis
DENV	dengue virus
DF	dengue fever
DHF	dengue haemorrhagic fever
DSS	dengue shock syndrome
E protein	envelope protein
ELISA	enzyme-linked immunosorbent assay
GAM	generalised additive model
GEE	generalised estimating equations
GRU	gate recurrent unit
HCT	haematocrit
ICAM	intercellular adhesion molecule-1
ICU	intensive care unit
IFN	interferon
Ig	Immunoglobulin
IL	interleukin
IPD	inpatient department
Lasso	Least Absolute Shrinkage and Selection Operation
LSTM	long-short term memory
M protein	membrane protein

MICE	multivariable imputation by chained equations
MIF	macrophage migration inhibitory factor
NS	non-structural protein
OOB	out-of-bag
OPD	outpatient department
OPS	orthogonal polarisation spectral imaging
OR	odd ratio
OUCRU	Oxford University Clinical Research Unit
PAT	peripheral artery tonometry
PLT	platelet count
QIC	quasi-information criteria
QSNICH	Queen Sirikit National Institute of Child Health
RF	random forest
RNA	ribonucleic acid
RNN	recurrent neural network
RT-PCR	using reverse transcriptase-polymerase chain reaction
SD	severe dengue
SMD	standardised mean difference
SVM	support vector machine
TNF	tumour necrosis factor
VCAM	vascular cell adhesion molecule
WBC	white blood cell count
WHO	World Health Organisation
XGB	extreme gradient boosted tree

Chapter 1 Introduction

Dengue is a viral disease transmitted by mosquitos, which poses a high burden on public health systems worldwide. Although most infections result a self-resolving febrile illness, a minority of those infected develop a severe clinical syndrome that can be fatal without appropriate supportive treatment. Specific anti-viral treatment is not currently available, and the only available dengue vaccine is not suitable for use in seronegative individuals because of concerns regarding efficacy and safety^{1, 2}. To improve clinical management and reduce the healthcare burden, the early detection of progression to severe dengue illness is key. As there is no single prognostic factor which can effectively predict the risk of severe dengue in a timely manner, clinical prognostic models, which use a set of factors with a mathematical equation, can allow for improved prediction of severe disease. However, existing risk prediction models have been developed using mainly data from the hospitalised population and small data sets. Most of the models also only make use of prognostic factors at a single-time-point.

1.1 Overview of dengue infection

1.1.1 Epidemiological data

Dengue causes a high burden on public health systems globally; an estimated 390 million (95% credible interval 284-528 million) dengue infections occur per year³. Of those, 96 (67 to 137) million suffered from symptomatic infections and 10,000 to 20,000 people die from dengue each year³. The global economic burden in 2013 found 48.2% were treated in the outpatient settings; 18% were treated in inpatient settings; and the rest 33.8% remained outside of the healthcare system⁴. The aggregate costs of dengue care in hospitals were 46% (\$4,093-14 [95% UI \$1,572-\$9,098]) of the global total, while the costs of dengue care in outpatient settings were 34% (\$2,987 [95% UI \$1,174-\$6,547]). Approximately 70% of the infections occur in Asia, particularly in India -accounting for 30% of the global infections³. The rest of the infections are found in Latin America and Africa. However, under-reporting of dengue cases is likely to be an issue in many countries, particularly in Africa⁵.

1.1.2 Dengue virus

Dengue virus (DENV) is an enveloped virus with a single-stranded positive-sense RNA genome in the genus *Flavivirus* and family *Flaviviridae*. There are four serologically distinct serotypes (DENV-1,

DENV-2, DENV-3, and DENV-4). The four serotypes have a similar structure and share approximately 65% of their genome⁶. Figure 1.1 presents the RNA genome and structure of dengue virus. The RNA genome encodes three fundamental structural proteins (envelope protein [E], membrane protein [M], and capsid protein [C]) and seven non-structural (NS) proteins (NS1, NS2A, NS2B, NS3, NS4A, NS4B, and NS5). C protein encapsulates a viral genome in the core of viruses, known as the nucleocapsid. This is surrounded by a lipid bilayer membrane, called a viral envelope, which is acquired from the infected host cells. Embedded on the viral envelope, the E and M proteins, composing the outer layer of viruses, are responsible for viral entry into human cells and neutralisation by antibodies during viral infection. The NS form a complex that plays an important role in viral replication in the infected host cells.

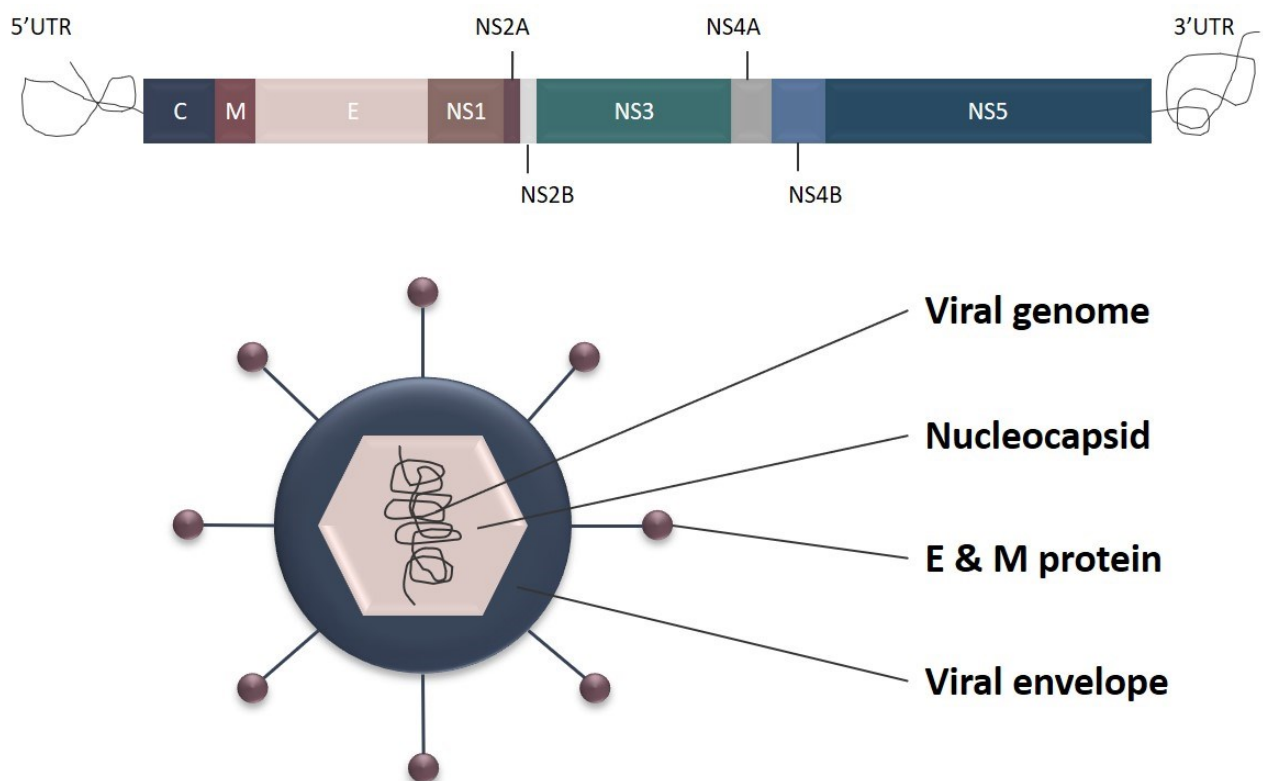


Figure 1.1. The viral genome and structure of dengue virus adapted from ⁶

1.1.3 Mechanism of severe disease

After the bite from an infectious mosquito, the virus disseminates through the blood stream and the lymphatic system. DENV infects mainly monocytes in blood and macrophages in the skin, liver, spleen, and thymus. Dengue viral particles or antigens have been found in neurons, microglia, and endothelial cells in the central nervous system (CNS) but no evidence of viral replication in the CNS.

DENV also infect hepatocytes (liver's cells) and cause cell apoptosis in the acute phase. In addition, previous *in vitro* studies found that DENV is capable to infect many types of cells including epithelial cells, endothelial cells, muscle cells, dendritic cells, bone marrow cells, and mast cells, but are not thought to play a role in replication⁷. After viral entry into cells, viral RNA is released and replicated in the cytoplasm. Synthesised virus is assembled in the endoplasmic reticulum and released via the trans-Golgi network.

An excessive immune response is considered to play a key part in the pathogenesis of severe disease. High levels of inflammatory cytokines have been found to be associated with severity of disease. Secondary infections with a heterotopic serotype have also been found to be associated with severe disease compared with primary infections. Three pathways have been suggested to explain the mechanism of severe disease: antibody dependence enhancement, original antigenic sin, and complement activation.

Antibody dependence enhancement (ADE) is a phenomenon whereby pre-existing heterologous antibodies increase viral uptake into cells and thus enhancing viral replication⁸. After primary dengue infection, heterologous antibodies and homologous antibodies are developed. While homologous antibody last life-long to protect from that particular dengue viral serotype, heterologous antibodies, which protect all dengue serotypes, only last up to a maximum of two years^{9, 10}. When heterologous antibodies wane to sub-neutralising levels, the antibodies fail to neutralise the current dengue infection instead binding the virus with Fc receptors expressed on target cells and facilitate viral entry into cells. As a result, ADE can result in a higher infected cell mass and consequently induce an excessive inflammatory response with downstream vascular and tissue damage. ADE risk appears to be limited to a particular range of antibody titres. A study showed that patients who had pre-existing antibodies within a narrow range of levels were at higher risk of ADE^{11, 12}. While heterologous antibodies with sub-neutralising levels are able to bind Fc receptors to enhance viral uptake, antibodies with high concentrations induce viral aggregates and bind to inhibitory Fc γ receptors, which reduce viral uptake¹³.

Original antigenic sin is another hypothesis for developing severe disease in secondary infections. After primary dengue infection, memory T cells are developed against both homotypic and heterotypic dengue serotypes. These pre-existing T cells rapidly respond to new dengue infections and take precedence over naïve T cells in secondary infection¹⁴. However, these pre-existing cells

have lower avidity for the current dengue serotype in secondary infection and therefore, promote excessive pro-inflammatory cytokines but with less viral killing¹⁵.

In addition to immune-mediated pathogenesis, viral factors also play a role with higher viral loads and NS1 protein, which is a product of viral replication complex, associated with disease severity. The secreted NS1 protein in blood stream is capable of protecting the viruses from complement and lectin-mediated neutralisations and inducing vascular leakage^{16, 17}. NS1 protein is also implicated in the vascular barrier disruption by binding to the endothelial glycocalyx layer and thus increasing vascular permeability and leakage^{18, 19}.

1.1.4 Clinical manifestations

Clinical symptoms of dengue infection broadly range from a relatively benign febrile illness to life-threatening manifestations. A minority of dengue-infected individuals develop severe clinical illness. The progression of clinical symptoms is typically characterised into three clinical phases: the febrile phase, the critical phase, and the recovery phase (Figure 1.2)⁵. After the incubation period of 3 to 7 days, approximately 25% of infected individuals develop symptoms³. The febrile phase begins with high-grade fever and is often accompanied by nausea, vomiting, headache, retrobulbar pain, myalgia, facial flushing, or transient macular rash. Minor haemorrhagic manifestations such as petechiae or bruising can occur, particularly at the sites of venepuncture or blood pressure cuff inflation. Mild to moderate thrombocytopenia and leukopenia are common laboratory findings. The febrile phase lasts 2 to 7 days. The transition period from the febrile phase to the critical phase typically occurs around day 4 of illness. In the critical phase, a small proportion of symptomatic individuals develop decompensated intravascular hypovolaemia from severe plasma leakage, which can become clinically apparent around the time of defervescence. The evidence of plasma leakage is found as increased haemoconcentration, pleural effusions or ascites. Initially patients with plasma leakage, through physiological compensation, are able to maintain adequate blood circulation to vital organs such as brain, heart, and kidneys. Therefore, patients may retain full consciousness but with relative tachycardia. Although systolic blood pressure remains normal or slightly high, pulse pressure becomes narrow (i.e. less than 20mmHg) accompanied by signs of vasoconstricted peripheral vascular networks, such as cold extremities, delayed capillary refill, or decreased urine output. Once hypotension develops, systolic blood pressure decreases rapidly, and patients can progress to irreversible shock with a significant risk of death. In addition, major haemorrhagic

manifestations such as gastrointestinal bleeding and mucosal bleeding requiring medical interventions can occur, but rarely, and usually occurs in patients following prolonged or profound shock. Bleeding manifestations are more common in adults than children²⁰. The critical phase lasts 24 to 48 hours. Then, the extravasated plasma returns to intravascular space over the next 48 to 72 hours, known as the recovery phase. Haemodynamic status becomes stable, and all symptoms gradually resolve.

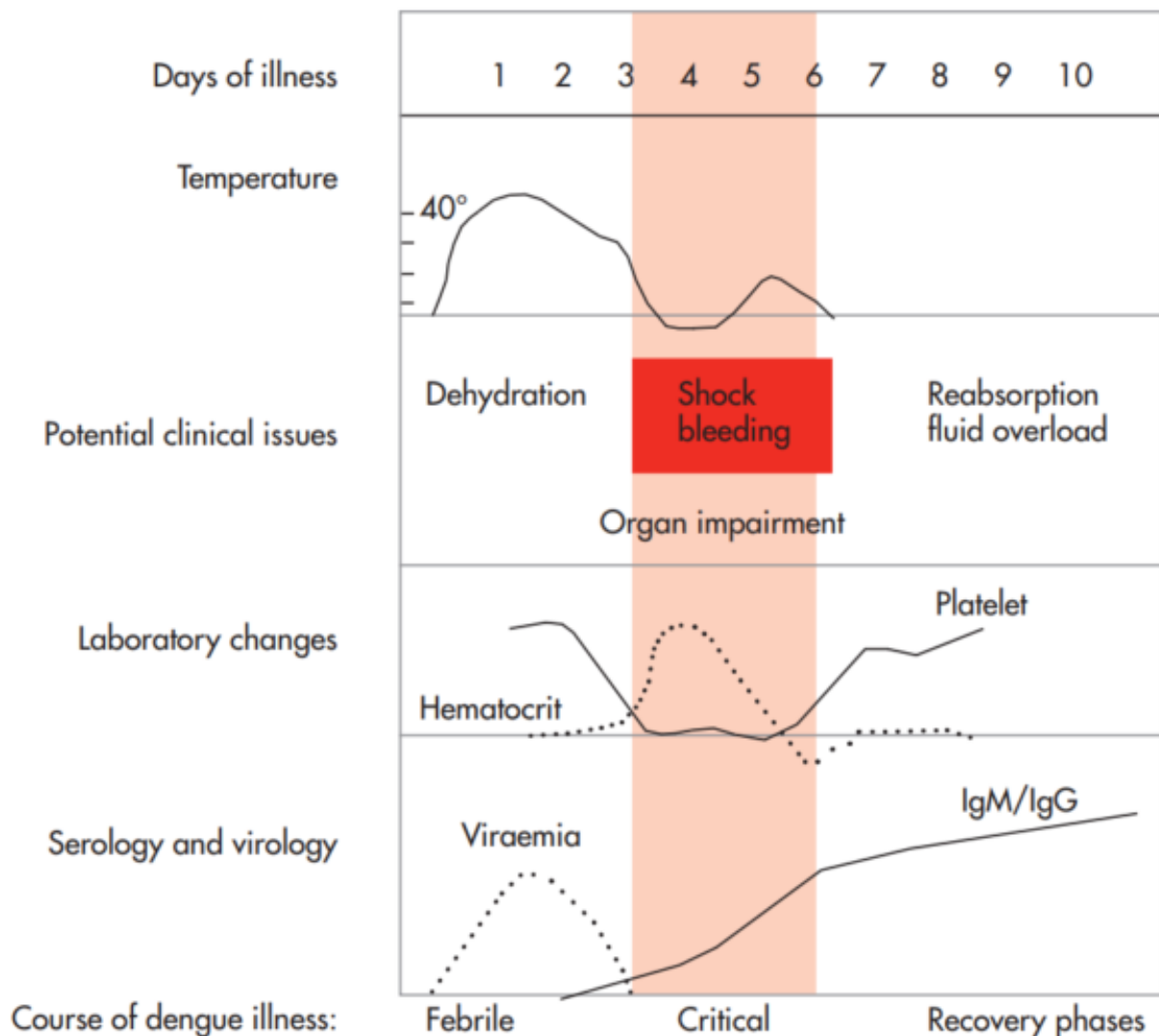


Figure 1.2. The clinical progression of dengue illness classified into the febrile phase, leakage phase and recovery phase. Reprinted from 2009 WHO Dengue Guidelines⁵.

1.1.5 Diagnosis

In the past and some resource limited areas, dengue can be diagnosed using just clinical diagnostic criteria. However, dengue is now diagnosed using a variety of laboratory tests. Each method must

be used in a specific context and time to achieve the highest accuracy and benefit to clinical management.

Clinical diagnosis was used as the main and primary method particularly in resource-limited areas without laboratory facilities. In 1997 WHO guidelines, dengue fever is clinically diagnosed as the presence of high-grade fever with two or more of the following findings: headache, retro-orbital pain, muscle pain, joint pain, maculopapular rash, haemorrhagic manifestations including positive tourniquet test, or leukopenia (white blood cell count less than $5,000 \text{ cells/ml}^3$)²¹. In the 2009 WHO guideline, the clinical diagnosis of dengue is defined as the presence of high-grade fever with two or more of the following: nausea, vomiting, maculopapular rash, muscle aches, leukopenia, or any warning signs⁵. Although clinical diagnosis does not require special laboratory facilities and is possible to be performed in all settings, clinical diagnosis is mainly dependent on the intuition and expertise of clinicians due to the lack of WHO guideline specificity.

Regarding laboratory methods (Table 1.1), the most common methods for dengue diagnosis involve the following: viral detection using viral isolation, nucleic acid detection using reverse transcriptase-polymerase chain reaction (RT-PCR), antigen detection with NS1 antigen, and serological assays via immunoglobulin (Ig) M and/or IgG enzyme-linked immunosorbent assay (ELISA)^{5, 6, 22}. For viral detection, cell culture and mosquito inoculation were the main techniques that have been used to detect dengue infection. A specimen can be collected from serum, plasma, or peripheral blood mononuclear cells. Virus can be also recovered from tissues at autopsy. Specimens should be collected during the viraemic phase, typically before day five of illness and kept at -70 degrees Celsius. These methods are mainly used in the research setting as they are time consuming and need special laboratory and expertise. RT-PCR has been used as the most common method for nucleic acid detection to quantify dengue virus and identify serotyping. Blood sampling for RT-PCR also should be also collected in the viraemic phase (days 1-5 of illness). The sensitivity and specificity of RT-PCR compared with viral isolation range from 77 to 96% and 99 to 100%, respectively²³. The turnaround time for the results using these methods can be as quick as two hours. Antigen detection is mainly performed using NS1 antigen, which is a glycoprotein produced from dengue infected cells. Overall, the sensitivity and specificity of current commercial diagnostic kits range from 77 to 87% for rapid diagnostic tests and 86 to 96% for an antibody-capture ELISA tests and varies according to serotype, immune status, and timing²⁴. The diagnostic performance is lower in secondary infections,

likely due to immune complex formation. For serological assays, the most widely used method is ELISA. This method is relatively easy to perform compared with viral isolation or nucleic acid detection, yet still needs special laboratory and experts. Commercial diagnostic kits are available at an affordable price e.g., DENV Detect IgM Capture ELISA (InBios International, Seattle, WA, USA), SD Bioline Dengue IgG ELISA (Standard Diagnostics, Suwon, Korea), or Panbio Dengue IgM Capture ELISA (Abbott, Macquarie Park, NSW, Australia). Sensitivity and specificity of commercial kits are 56 to 92% and 64 to 97%, respectively^{25, 26}. The limitation of this method is cross-reactivity with other flaviviruses which commonly circulate in the same endemic areas including Zika virus, yellow fever, and Japanese encephalitis virus. In addition, the presence of IgM and IgG is commonly detected only after day 5 of illness, and thus acute and convalescent samples are required. The use of rapid diagnostic tests using a combination of NS1 antigen, IgM and IgG improved the sensitivity of test to 87.4% (95% confidence interval (CI) 82.8-91.2) and reduced the number of false negatives to 14% (95%CI 11.4-18.6)²⁷. Although serological diagnostic kits can be used as a point of care testing for its affordable costs and ease of use, results should be interpreted with caution, particularly in the flavivirus endemic areas and when only one sample is available.

According to diagnosis for acute dengue infection, testing can be divided into two stages: the viraemic phase (which typically within day 5 of illness) and the early-post febrile phase (after day 5 of illness). Where suspected dengue cases present within the viraemic phase, clinical diagnosis, anti-NS1 antigen, and RT-PCR can be used in clinical practice depending on healthcare setting and laboratory facilities. In case of delayed presentation of patients with suspected dengue, the pairwise serological test can be adopted as the standard method for confirmatory laboratory tests.

1.1.6 Dengue disease classification

Dengue classification has been developed to assist clinicians in clinical decision-making in terms of appropriate case management for specific clinical phases and severity. Classifications widely used in standard practice is based on two systems established by the World Health Organisation (WHO) in 1997 and 2009^{5, 21}.

Table 1.1. Summary of advantages and disadvantages of diagnostics

Diagnostic tests	Advantages	Limitations
Viral isolation and identification	- Standard diagnosis (high specificity)	- Requires samples collected in viraemic phase (0-5 day of illness)

	- Identifies infecting serotypes	- Requires expertise and advance laboratory facilities - Time and labour consuming - Expensive - Not differentiate primary or post primary infection
RNA detection	- Standard diagnosis (high specificity) - Identifies infecting serotypes - Results in a day or two days	- Requires samples collected in viraemic phase (0-5 day of illness) - Requires expertise and advance laboratory facilities - Expensive - Not differentiate primary or post primary infection
NS1 antigen	- Easy to perform - Affordable price - Rapid turnaround time	- Moderate sensitivity
IgM or IgG seroconversion	- Easy to perform - Affordable price - Differentiates primary infection and post-primary infection	- Low to moderate sensitivity - Required paired samples testing
Single IgM detection	- Easy to perform - Inexpensive - Rapid results turnaround time	- Low sensitivity particularly in secondary infection and cross reactivity

The 1997 WHO classification categorises symptomatic infections into two main classes: dengue fever (DF), and dengue haemorrhagic fever (which comprises four severity grades: DHF I, II, III, and IV). The DHF grade III and IV is also called as dengue shock syndrome (DSS). Due to the complexity and low sensitivity of the 1997 WHO classification, a new classification system was established by a global dengue expert consensus meeting in 2009. This classification classifies symptomatic infection into two main classes: dengue (without or with warning signs) and severe dengue (SD). Diagnostic criteria for severe dengue and a list of the WHO warning signs are presented in Table 1.3. The 2009 WHO classification is more practical and shows higher sensitivity than the previous version, although it there have been issues regarding non-specificity of warning signs and the increase of unnecessary hospital admissions²⁸⁻³⁰.

Table 1.2. Diagnostic criteria for severe dengue and a list of warning signs adapted from the 2009 WHO Classification⁵

Severe dengue

Severe plasma leakage leading to

- ☐ Shock (Dengue Shock Syndrome)
- ☐ Fluid accumulation and respiratory distress

Severe bleeding (considered by clinicians)

Severe organ involvement

- ☐ Liver enzyme: AST or ALT > 1000 units/
 - ☐ Impaired consciousness
 - ☐ Heart and other organ involvement
-

Warning signs

Abdominal pain or tenderness

Persistent vomiting

Clinical fluid accumulation

Mucosal bleed

Lethargy or restlessness

Liver enlargement > 2 cm

Laboratory: increase in haematocrit concurrent with rapid decrease in platelet count

AST: aspartate aminotransferase; ALT: alanine aminotransferase

1.1.7 Clinical management

As there are no available specific treatments for dengue infection, symptomatic treatment and fluid therapy are key in the clinical management. Oral fluid intake and antipyretic drugs such as acetaminophen (but not non-steroidal anti-inflammatory drugs (NSAIDs) nor aspirin) are considered as the first choice of supportive treatment for patients with non-complicated dengue⁵. For patients with complicated dengue (i.e., unstable haemodynamic status or bleeding requiring intervention), balanced parenteral fluid is key to reduce mortality from dengue shock. The minimum amount of parenteral fluid should be given as necessary to maintain the cardiovascular system without causing fluid overload which can lead to respiratory distress and worse clinical outcomes. Regarding parenteral fluid, isotonic crystalloid solutions are the first-line treatment while isotonic colloid solutions are reserved for patients with re-shock and shock with signs of fluid overload^{5, 31}. Blood component transfusion is used in specific conditions, such as severe bleeding, to maintain the cardiovascular system and oxygenation^{5, 32}. Prophylactic platelet transfusion was demonstrated to have no benefit in reducing the risk of bleeding and can increase the risk of volume overload with patients with severe dengue^{33, 34}. Anti-viral drugs and host directed therapies (i.e., Chloroquine, Balapiravia, Celgosivir and Lovastatin) have shown no benefit in altering clinical outcome³⁵⁻³⁷. Corticosteroids, both high and low dose regimens in the early phase of infection, had no benefit in

reducing clinical severity but increased the risk of hyperglycaemia especially with high doses^{38, 39}. A recent randomised control trial also demonstrated that Ivermectin given in the febrile phase has no benefit on reducing dengue severity⁴⁰.

1.2 Factors associated with progression to severe dengue

As there is a wide range of clinical outcomes in dengue, from self-resolving illness to multi-organ failure and death, the early recognition of patients at high risk of severe illness is key, not only to reduce morbidity but also the demands placed on healthcare services. I reviewed the most relevant and recognised risk factors.

1.2.1 Demographic factors

Age: Two extremes of age – infant aged less than 9 months, young children, and elderly aged over 60 years, are most vulnerable to severe dengue illness⁴¹⁻⁴⁴. Whilst higher risk in young children can be explained by increased capillary permeability and less compensatory volume depletion from vascular leakage compared with adults⁴⁵, existing comorbidities impairing the cardiovascular system, such as diabetes mellitus, hypertension, or renal disease, and impaired immune system could increase the risk of severe illness in the elderly^{46, 47}.

Sex: Male sex has been associated with a higher prevalence of dengue infection, but female sex has a higher risk of severe illness and mortality⁴⁸. This could be explained from the difference in healthcare seeking behaviour between males and females⁴¹. Additionally, female sex may also be more prone to an increased inflammatory response to infection compared with males⁴⁹.

1.2.2 Clinical signs and symptoms

Abdominal pain or tenderness, persistent vomiting, clinical fluid accumulation, mucosal bleeding, lethargy or restlessness, and liver enlargement are in the list of WHO warning signs⁵. These factors show high sensitivity, but low specificity, for progression to severe dengue. As a result, using the warning signs to guide triage of patients during the febrile phase can increase the number of hospitalised patients by up to approximately 30%²⁸.

1.2.3 Conventional laboratory parameters

A rise in haematocrit accompanied by a rapid decrease in platelet count are also in the list of WHO warning signs⁵. A rise in haematocrit representing haemoconcentration is the most used surrogate measure of plasma leakage. An increase of greater than 20% from baseline haematocrit is regarded as significant leakage⁵⁰. However, this change occurs relatively late within the critical period, so cannot be used for early risk prediction. Platelet count is another haematological parameter considered as a potential prognostic factor. Platelet count starts decreasing in the febrile phase with the nadir usually around day 4-6 of illness. Although lower values are associated with severe dengue⁴⁸, no specific threshold has been defined yet. Additionally, the correlation between platelet concentration and haemorrhagic manifestations is not clear, in fact lower platelet count correlates more strongly with plasma leakage than bleeding. Elevated liver enzymes and hypoalbuminemia are also associated with progression to severe dengue^{48, 51}. Hypoalbuminemia can be explained by extravasation of albumin along with plasma during vascular leakage. The liver involvement could be due to a number of factors, including hepatocyte apoptosis from direct invasion of dengue virus, hypoxic damage from hypoperfusion, and excessive inflammatory response and hypercytokinaemia⁵²⁻⁵⁴.

1.2.4 Imaging

X-Ray and Ultrasound has been used to detect clinical and subclinical fluid accumulation from plasma leakage. Serial ultrasonography studies have demonstrated sub-clinical fluid accumulation is very common in dengue, already apparent in the febrile phase, in the form of small pleural effusions and ascites⁵⁵. Gall bladder wall thickening on ultrasound has been found to be one of the earlier markers of plasma leakage^{55, 56}. The main limitation of ultrasound is the need for equipment and expertise which can be lacking in LMICs.

1.2.5 Biomarkers

Viral related biomarkers

Higher viraemia levels were found to be associated with severe dengue implying a greater infected cell mass⁵⁷. However, viraemia levels differ depending on the infecting dengue serotype and existing immune status⁵⁸. During secondary infections with DENV-1 and DENV-2, the peak of viraemia was

significantly higher for patients with dengue shock syndrome compared with patients without dengue shock. The duration of viraemia was not significantly associated with disease severity^{57, 59}.

NS1 antigen is a protein product derived from viral replication and therefore levels correlate with viraemia levels. Although the levels of NS1 antigen vary by dengue serotype and immune status (primary and post-primary infections), previous studies also found that NS1 antigen was associated with dengue severity by involving in release of proinflammatory cytokines and directly disrupting endothelium thus causing vascular leakage^{16, 18, 19, 60}.

Immune activation biomarkers

Host immune response plays a significant role in the pathogenesis of progression to severe disease. Thus, the components of immune response including inflammatory cytokines, chemokines, and immune cell activation markers have been investigated as biomarkers of severe disease in dengue infection. TNF (tumour necrosis factor)- α , IFN- γ and matrix metalloproteases-2, 3 and 9 were higher amongst patients with progression to severe disease compared to patients without. These biomarkers are produced by dendritic cells and macrophages, which are the main target cells for dengue infection, and can cause plasma leakage⁶¹⁻⁶³. The number of CD4 and CD8 T cells, and NK (natural killer) cells also decreased during the early phase of illness⁶⁴. Soluble CD95, which is a mediator of cell apoptosis, was found higher in the early phase of illness. Pro-inflammatory cytokines including interferon (IFN)- γ , tumour necrosis factor (TNF)- α , interleukin (IL)-1, IL-6, IL-8, IL-10, CCL2 (MCP-1) and CCL5 (RANTES) were also excessive produced amongst patients with severe disease^{65, 66}. Mast cells are also activated in dengue and produce high levels of tryptase and chymase⁶⁷. These two proteinases were also found to be significantly higher in the serum of severe dengue patients.

The concentrations of cytokines and chemokines from severe dengue patients were found to be higher than in non-severe dengue patients, and this phenomenon has been termed a 'cytokine storm'. Significantly elevated levels of IFN- γ , TNF- α , IL-1 β , IL-4, IL-6, IL-8, IL-10, macrophage migration inhibitory factor (MIF) and chemokines CCL2, CCL4, CCL5, and CXCL10 (IP-10) were demonstrated amongst patients with severe disease during the febrile phase compared with patients with non-severe dengue. IL-6, IL-10, IFN- γ , MIF, and CCL-4 were also considered as predictors of severe dengue^{68, 69}. The levels of complement (C)-3, C3a and C5a were also found

higher in serums of patients with dengue haemorrhagic fever compared with dengue fever in acute phase⁶⁸.

Endothelial activation biomarkers

Dengue affects endothelial cells directly and indirectly resulting in increased vascular permeability. Significantly elevated levels of intercellular adhesion molecule-1 (ICAM-1) and vascular cell adhesion molecule-1 (VCAM-1) were found in dengue shock syndrome patients compared with dengue fever patients⁷⁰⁻⁷². Another non-specific biomarker of endothelial activation is thrombomodulin, which is released from endothelium and responsible for anticoagulation. Increased levels of thrombomodulin were found on 2 or 3 days before prior to defervescence and associated with progression to severe disease⁷³. The ratio of angiopoietin (Ang)-2 to Ang1 were found significantly higher amongst children with severe dengue⁷⁴. Elevated levels of chondroitin sulfate and syndecan-1, which are main components of endothelial glycocalyx layer maintaining vascular integrity, were significantly associated severe plasma leakage⁷⁵.

Biochemical biomarkers

C-reactive protein (CRP), which is an inflammatory biomarker, has been found to be higher amongst patients who progressed to severe dengue compared with patients who did not, when measured in the first three days of illness⁷⁶. In addition, elevated levels of CRP were associated with hospitalisation and a longer fever clearance time. Nitric oxide is well established as an immunoregulator and play a role in adjusting vascular diameters, preventing leukocyte adhesion and platelet aggregation. Significantly decreased levels of nitric oxide were found amongst patients with dengue haemorrhagic fever compared with patients with dengue fever⁷⁷⁻⁷⁹.

In summary there are many biomarker studies investigating the association with dengue severity, however the results of the studies are often inconsistent as they were conducted using differences in study designs, population, or laboratory techniques. Transparent and harmonized methodology and validation are urgently required to provide robust evidence of novel biomarkers. Particularly, these biomarker studies should explicitly report the day of illness of sample collection, viral serotype, immune status, and age.

1.2.6 Novel vascular function techniques

The main pathophysiology of severe dengue is increased vascular permeability which in turn results in plasma leakage and depleted intravascular volume. As a result, the assessment of vascular function, microcirculation, and intravascular volume status can be used as potential markers of disease progression. To evaluate clinical endothelial function, peripheral artery tonometry (PAT) has been used, which measures the reactive hyperaemia index (RHI). The RHI was associated with progression to severe dengue disease^{79, 80}. To monitor microvascular flow, bedside real-time imaging of the microcirculation has become possible with video-microscopy using orthogonal polarisation spectral imaging (OPS) and side-stream darkfield imaging⁸¹. Although microcirculation perfusion indices decreased in patients with plasma leakage compared with patients without, a significant difference was only detected in the critical phase. Microcirculatory parameters, therefore, did not change earlier enough for prediction⁸¹. In terms of assessment of intravascular volume, the compensatory reserve index (CRI) is a new parameter derived from machine learning algorithms and feature extraction of arterial pulse waveform. CRI has been shown to reflect central blood volume and has shown potential utility of the CRI to determine intravascular volume in a pilot study in dengue patients⁸². Another non-invasive monitoring technique for intravascular volume assessment is bedside echocardiography⁸³. This can assess both myocardial impairment as well as intravascular volume status.

1.3 Prognostic models for progression to severe dengue illness

As no standalone prognostic factor is sufficient to effectively predict progression to SD illness, clinical prognostic models are an alternative tool that uses mathematical functions incorporated with multiple prognostic factors to enhance individual risk prediction of severe dengue illness during the febrile phase. We have found 17 studies of clinical prediction models on a search of the Embase database up to August 2020⁸⁴⁻¹⁰⁰. The detail of each study is presented in Table 1.4. These selected studies were published between 2008 and 2020. They were all cohort studies. Half of them were prospective studies^{87, 88, 90, 92, 95, 96, 101}, and the rest were retrospective. Only six studies recruited participants at outpatient settings^{85, 87, 92, 93, 97, 98}, and the others recruited hospitalised patients. Five studies specified that the included participants had to have a fever lasting less than 72 hours, emphasising the clinical prediction nature during the early phase of the disease^{85, 87, 93, 97, 98}. Half of the studies recruited children as participants but the cut-off points for being defined as children

varied from those below 12 to 18 years old. The eligibility of subjects recruited in each study and the development of models were all distinct. Some studies distinguished the events of interest amongst either participants with laboratory-confirmed dengue infection or participants with acute febrile illness. Most studies were conducted in Southeast Asia (i.e., Cambodia, Thailand, Singapore, and Vietnam). The sample size of these studies ranged broadly from 20 to 2,500 patients. The outcomes of interest were also heterogeneous across the studies. Most studies applied either the 1997 (i.e., predicting the risk of DHF or DSS) or 2009 WHO classifications (i.e., predicting the risk of SD). One study constructed a prognostic model for profound shock and re-shock amongst patients in intensive care units and another study used moderate plasma leakage defined as 15 % haematocrit change as the primary outcome¹⁰¹.

In terms of the model development and validation, predictors included in the prognostic models were, demographic data, clinical signs and symptoms, and conventional laboratory parameters. However, two studies included novel biomarker genes and metabolomics parameters into the prognostic models and had a small sample size,^{87,96}. While most of the studies used baseline clinical parameters as predictors, three studies used dynamic changes of clinical parameters into their models^{84, 85, 101}. A prognostic model by Lam et al. included the change in the number of platelets as compared with the previous day¹⁰¹. The model outperformed a model with the number of platelets in a specific day. An alternative model in this study included the change in the number of platelets as compared with the previous day. This model outperformed a model consisting of the number of platelets in a specific day. Although a model by Park et al.⁸⁵ was constructed with a small sample size it consisted of clinical parameters from different illness days and achieved a high predictive performance (area under receiver operating characteristics curves (AUROC) 0.67 to 0.93 based on the day of illness). More than half of the studies employed models constructed with statistical modelling, including logistic regression, Cox regression, generalised additive modelling, and multivariate adaptive regression splines. The others applied novel machine learning in model development, specifically Bayesian feature reduction, classification and regression tree, and support vector machine^{86, 90, 95, 96, 98, 99}. Only half of the studies reported the internal validation of the model and one study performed external validation. Internal validation was performed with geographical validation, temporal validation, or simple data splitting validation. The most common metric of validation was AUROC, showing discriminative performance. Only five studies reported

calibration performance: three reported Hosmer-Lemeshow C-statistics^{88, 89, 96} and the other reported calibration in the large, calibration slope, or calibration curves^{90, 101}.

1.4 Knowledge gaps and areas of interest

The clinical manifestations of dengue infection range from a self-resolving illness to multi-organ failure and death. Most symptomatic infections do not progress to severe illness and are safe for home monitoring and care, whereas a minority will develop severe manifestations and need careful monitoring and prompt treatment in hospital. Therefore, effective triage of patients with dengue during the febrile phase is key to ensure appropriate patient management while also preventing overload of healthcare systems. Several studies have identified potential prognostic factors for severe dengue illness. Prognostics models have been also developed and implemented to assist clinicians to detect patients early with severe clinical illness. However, many gaps in our knowledge remain.

Most studies investigating clinical prognostic factors had small sample sizes and the results were inconsistent. To address the inconsistencies of the results from the existing studies, a systematic review and meta-analysis was conducted to identify the prognostic factors of severe dengue^{48, 102, 103}. However, the studies found high heterogeneity in clinical parameters due to significant variation on different days of illness. The clinical parameters presented early in the disease course that are associated with subsequent progression to severe illness still need to be identified.

Another area of need is to identify the optimal model for the risk prediction of dengue severity. Conflicting evidence to date has resulted from difference in populations used in model development, availability of predictors, inconsistent definition of both predictors and outcomes, or limited and unclear analytical methods of model development and validation. Additionally, most existing models used only baseline characteristics and clinical parameters. Models were mainly developed with regression models which do not account for non-linearity and can contribute to bias in prediction. Prognostic models incorporating both baseline and dynamic clinical parameters with a variety of modelling approaches would allow improved individual risk prediction. In addition, the comprehensive and reproducible analytical methods used in model development and validation will enhance the acceptance of model use in general practice.

Table 1.3. Summary of 17 studies on the development risk prediction models for dengue severity included in this review

No.	Study	Country	Study design	Population	Healthcare settings	Age groups	Outcomes of interest	Candidate predictor	Sample size	Number of events	Modelling approaches	Validation
1	Tan et al. (2020)	Singapore	retrospective	patients with dengue infection before the development of DHF	IPD	adults: aged over 14 years	DHF	A, B, and C	5,216	DHF: 606	logistic regression	internal: temporal validation
2	Park et al. (2018)	Thailand	retrospective	patients with dengue infection within first 72 hours of clinical onset	OPD	children: aged 6 months to 15 years old	1.Dengue 2.DHF 3.DSS	A, B, and C	257	1.Dengue: 105 2.DHF: 42 3.DSS: 9	structural equation model	internal
3	Low et al. (2018)	Malaysia	prospective	patients with dengue infection within first 72 hours of clinical onset	OPD	adults: aged 15 years or older	SD	A, B, C and D	82	SD: 29	logistic regression	no
4	Phakhounthong et al. (2018)	Cambodia	retrospective	patients with dengue infection	IPD	adults: aged under 16 years old	SD	C	160	SD: 38	CART and logistic regression	no

No.	Study	Country	Study design	Population	Healthcare settings	Age groups	Outcomes of interest	Candidate predictor	Sample size	Number of events	Modelling approaches	Validation
5	Sreenivasan et al. (2018)	India	Prospective	patients with dengue infection before the development of DHF	IPD	aged under 12 years	SD	B and C	359	SD: 93	logistic regression	no
6	Fernandez et al. (2017)	Honduras	retrospective	patients with dengue infection before the development of DHF	IPD	non-specific	moderate plasma leakage defined as 15% HCT rising from baseline	B and C	320	Moderate leakage: 34	logistic regression	no
7	Lam et al. (2017)	Vietnam	prospective	patients with dengue infection within first 96 hours of illness	IPD	children: aged 5 to 15 years	DSS	A, B, and C	2,301	DSS:143	logistic regression	no
8	Nguyen et al. (2017)	Vietnam	prospective	patients with acute febrile illness within first 96 hours of illness	OPD	children: aged 1 to 15 years	SD	B and C	7,563	SD: 117	logistic regression	internal: temporal validation

No.	Study	Country	Study design	Population	Healthcare settings	Age groups	Outcomes of interest	Candidate predictor	Sample size	Number of events	Modelling approaches	Validation
9	Pang et al. (2016)	Singapore	retrospective	patient with dengue infection within first 72 hours of clinical onset	OPD	Adults: aged under 18 years old	Dengue with warning signs	D	92	DwWS:47	logistic regression	internal: temporal validation
10	Lam et al. (2015)	Vietnam	prospective	patients with dengue shock syndrome	ICU	Children: aged under 15 years old	1. Recurrent shock 2. Profound shock	A, B, and C	1,207	recurrent shock: 222	logistic regression; logistic regression with lasso; CART; gradient boosted tree, GAM	internal: temporal validation
11	Pongpan et al. (2013)	Thailand	retrospective	patients with dengue infection	IPD	children: aged 1 to 15 years	1.DHF 2.DSS	A, B, and C	777	1.DHF: 296 2.DSS: 90	ordinal multinomial regression	no

No.	Study	Country	Study design	Population	Healthcare settings	Age groups	Outcomes of interest	Candidate predictor	Sample size	Number of events	Modelling approaches	Validation
12	Huy et al. (2013)	Vietnam	prospective	patients with dengue shock syndrome	IPD	children: age 6 months to 15 years old	DSS	A, B, and C	444	re-shock: 126	Logistic regression and additional ML (ANN, kNN, J48, LGBost, LRT, Logistic, NaiveB, RF, RSSpace, SMO and Vote))	no
13	Brasier et al. (2012)	Venezuela	prospective	prior to the development of DHF	OPD (unclear)	unspecific	DHF		51	DHF: 13	logistic regression, GAM, CART, and RF	no
14	Potts et al. (2010)	Thailand	retrospective	patients with dengue infection within first 72 hours of clinical onset	OPD and IPD	children: aged 6 months to 15 years old	DHF	A, B, and C	1,230	1.DF:374 2.DHF: 208	logistic regression	internal: geographical validation

No.	Study	Country	Study design	Population	Healthcare settings	Age groups	Outcomes of interest	Candidate predictor	Sample size	Number of events	Modelling approaches	Validation
15	Potts et al. (2010)	Thailand	prospective	patients with acute febrile illness within first 72 hours of clinical onset	OPD and IPD	children: 6 months to 15 years old	1. DSS 2. DSS and dengue fever with significant pleural effusion	C	1,227	1.DF: 614 2.DHF: 228	CART	no
16	Lee et al. (2009)	Singapore	retrospective	patients with dengue infection before the development of DHF	IPD	adults	DHF	B and C	1,855	DHF: 82	CART	internal: simple data splitting
17	Lee et al. (2008)	Singapore	retrospective	patients with dengue infection before the development of DHF	IPD	adults	DHF	B and C	1,973	DHF: 118	logistic regression	no

OPD: Outpatient departments; IPD: Inpatient department; DHF: Dengue haemorrhagic fever; SD: Severe dengue; CART: Classification and regression tree; RF: Random Forest; GAM: Generalised additive model; The categories of predictors: A = demographic data, B= Clinical signs and symptoms, C= Basic laboratory parameters, D= Novel parameters/techniques

1.5 Aims and outline of the thesis

In the context of heterogeneous clinical phenotypes from mild to severe illness and no single predictor sufficiently predicting progression to severe illness, this project aims to develop and determine robust clinical prediction models for individual risk of severe dengue illness. The prediction models will utilise both baseline and temporal dynamics of clinical parameters incorporating both statistical and machine learning approaches. My three specific aims are as follows:

1. To identify prognostic factors from the literature associated with progression to severe dengue illness during the febrile phase
2. To develop and determine clinical prediction models for individual risk of severe dengue illness using baseline clinical factors detected in the first 72 hours of fever onset
3. To describe the temporal dynamics of clinical laboratory parameter and to implement the dynamics in the development of individual risk prediction of severe dengue disease accompanied by the novel machine learning approached

This thesis is divided into six chapters including this introductory chapter:

- 1) Chapter-2: Systematic review and meta-analysis of prognostic factors in the febrile phase associated with progression to severe dengue illness
- 2) Chapter-3: Description of the clinical data sets and analytical methodology for the development and validation of prognostic models
- 3) Chapter-4: Clinical risk prediction models of static prognostic factors within 72 hours of illness onset using both statistical and machine learning approaches
- 4) Chapter-5: Temporal dynamics of clinical parameters and a case study of clinical risk prediction models with temporal dynamics of the clinical parameters using recurrent neural networks
- 5) Chapter-6: Discussion and suggestions for future research

Chapter 2 Risk predictors for progression to severe dengue illness: a systematic review and meta-analysis

2.1 Rationale

Dengue infection results in a broad range of clinical manifestations. Most infected individuals recover without hospitalisation, while a minority progress to severe illness that requires inpatient monitoring and meticulous fluid management. The early recognition of progression to severe illness, accompanied by prompt and appropriate fluid therapy are key to improving outcomes in dengue⁵. To assist in the early recognition of patients at high risk of disease progression, the WHO have recommended using warning signs as prognostic indicators²⁸. The use of the warning signs checklist provides high sensitivity but low specificity in identifying patients with a high risk of SD^{28, 29}. It therefore can increase the number of unnecessary hospital admissions and thus increasing demand on health care systems particularly in endemic areas during the rainy season when dengue cases peak. In addition, these warning signs occur relatively late in the disease course.

Several predictors of progression to severe dengue have been identified from previous studies. However, most of these studies had small sample sizes and the results are often conflicting. Systematic review and meta-analysis studies have also been conducted to increase the statistical power with aggregated data. Previous systematic reviews and meta-analyses considered the whole time-course of illness and did not differentiate the febrile phase, which occurs during the first 1 to 4 days from symptoms onset, from the later clinical phases when severe illness occur^{48, 102, 103}. These previous reviews and meta-analyses may have failed to identify significant associations between early risk predictors and progression to severe dengue due to the dynamic nature of the disease. Clinical signs and symptoms, patient laboratory parameters, and viral-related biomarkers vary markedly across the different phases of the disease.

In this chapter, we conducted a systematic review and meta-analysis to identify risk predictors associated with progression to severe dengue illness that are identifiable during the febrile

phase. The identified risk predictors could improve the risk assessment in the early febrile phase. Improving the prediction of progression to severe dengue in the early phase of disease would have major implications for the management of patients, and health care resource allocation.

2.2 Methodology

We conducted a systematic review and meta-analysis following the Preferred Reporting Items for Systematic Reviews and Meta-analyses for Protocols (PRISMA-P) guidelines¹⁰⁴ and the steps recommended in the guide for systematic review and meta-analysis of prognostic factor studies¹⁰⁵. The protocol was registered and updated with PROSPERO (CRD42018093363) (Appendix 1)¹⁰⁶.

2.2.1 Search strategy

We searched for both published and grey literature on eight main medical databases (EMBASE, MEDLINE, Web of Science, Scopus, CINAHL, Global Health, Open Grey, and ProQuest Dissertations and Theses Global) to ensure covering almost all studies. We searched for studies conducted between January 1997 and January 2020. Original studies written in English were included; reviews, conference abstracts, comments, notes, and letters were excluded. MeSH terms and explode function were applied where applicable to comprehensively extend a set of terms in hierarchical list (Appendix 2-3). A medical reference librarian also verified the search terms and strategy to ensure that the searching syntax was appropriate and complete.

2.2.2 Study selection

After all studies were identified from the search, we removed duplicate studies with EndNote X8 software (Clarivate Analytics, Philadelphia, PA, USA) and subsequently checked manually. Both abstract and full-text screening were performed by the first and second reviewers in accordance with the eligibility criteria Table 2.1 on a Covidence systematic review software (Veritas Health Innovation, Melbourne, Australia). Discrepancies between the two reviewers resulted in a third reviewer to make the final decision.

80 **Table 2.1.** Template used to assess the eligibility criteria

Reviewer name:	Date:
Author name/Study ID:	Year:
Title	Journal:
Study designs	Include Original papers Research in human Exclude Descriptive studies (case reports, case studies), epidemiological studies, ecological or mathematical modelling studies
Population	Include Symptomatic infected individuals in the febrile phase AND Laboratory confirmed dengue diagnosis according to WHO Dengue Guideline in 1997 and 2009* Exclude Individuals who had developed clinical manifestation of severe dengue (severe plasma leakage, hemodynamic instability, and organ impairment)
Exposure(s)	Include Demographic features, clinical manifestations (signs and symptoms), laboratory parameters, or imaging techniques AND Parameters collected during the febrile phase Exclude Genetic factors or gene expression
Outcome(s)	Include Association between predictors and risk of development severe clinical manifestations of dengue infection according to WHO guideline (dengue haemorrhagic fever grade I-IV, dengue shock syndrome, and severe dengue) compared with dengue fever
Overall:	Include Exclude

2.2.3 Data extraction and quality assessment

Briefly, a data extraction form included the first author's surname, year of publication, participant recruitment, study design, data collection (prospective or retrospective), dengue classification, population characteristics (age groups or specific features), sample size, associated factors, and associations between disease severity and clinical factors; the full details of the data extraction form is in Appendix 4. Both crude associations and adjusted associations with a set of adjustment variables were extracted if applicable. The data extract form was modified to comprehensively cover the central and disperse estimates of associations because the association between factors and progression to clinical outcomes were reported in different statistical parameters. Studies of the same factors that were conducted in the same place and time were considered as duplicates. In this case, the study with the larger sample size was selected and the other was discarded. Ten percent of the data extraction was performed by the first reviewer and verified by either the second or third reviewer before all extraction was performed.

We assessed the quality and risk of bias of the included studies using the quality in prognostic factor studies (QUIPS) tool¹⁰⁷. The first reviewer and either second or third reviewers assessed the quality independently. Discrepancies were solved by consensus amongst the three reviewers. The data extraction and assessment of quality and risk of bias were performed on Microsoft Excel.

2.2.4 Meta-analysis

We performed meta-analysis to aggregate the association of factors for progression to severe dengue illness using R package Meta and Metafor^{108, 109}. We adopted clinical outcomes based on the definitions of severe dengue illness according to the 1997 and 2009 classifications^{5, 21}, which have been used worldwide as a standard. To increase the number of included studies and generalisability, we combined the two classifications. As a result, dengue progression was classified into two main groups: patients who did not progress to severe dengue illness (DF and non-SD) and patients who progress to severe dengue illness (DHF and SD). We performed

a meta-analysis for factors reported in at least four studies and each factor was analysed using univariable meta-analysis separately. For continuous variables, since the included studies did not report a statistical estimation of factors in a consistent format, we transformed the reported parameters to standardised sample means and standard deviations according to methods proposed by Luo et al.¹¹⁰. The crude associations from univariable analyses and adjusted associations from the multivariable models were aggregated separately. Aggregated associations were estimated using random effects models since we assumed that the populations were different. The variance between studies in the random effects models were estimated using DerSimonian-Laird method.

We also performed subgroup analyses for children and adults since the risk of progression to severe disease differs with age. Studies where the patient population was 18 years or older were defined as adult studies, and those which included patients below 18 years were defined as studies with children. Furthermore, age cut-offs in some studies were inconsistent and not always explicitly defined. For instance, patients aged between 14 and 18 years old have been classified as both adults in some studies and as children in others. Given the results may be presented jointly with other age categories, it was not always possible to separate the estimates for adults and children, according to our definition. Therefore, we assigned studies as paediatric or adults according to the definition given in the original studies. In addition, sensitivity analysis was also performed accounting for the outcome based on the two classifications where factors were reported in at least four studies in each subgroup.

2.2.5 Heterogeneity assessment

Heterogeneity between studies was assessed using Cochran Q statistic and I^2 . Heterogeneity was considered where p-value of Cochran Q statistic was > 0.1 . The levels of heterogeneity were interpreted following the suggestion by Higgings et al.- as low ($I^2 < 25\%$), low to moderate (25% to $< 50\%$), moderate to substantial (50% to $< 75\%$), or substantial ($\geq 75\%$)^{111, 112}. If heterogeneity was suspected, subgroup analysis was performed in cases with ≥ 4 studies in each subgroup. Sensitivity analysis was performed by removing studies with extreme effect size and heterogeneity to investigate the robustness of pooled effect sizes. The sensitivity

analysis of platelet count was additionally performed by omitting the two studies reporting odds ratios (ORs) instead of standardised mean differences (SMDs). Finally, the small-study effect was assessed to detect any publication bias using funnel plot and Egger's test. Publication bias was suspected when the p-value of the Egger's test < 0.05 ¹¹³. The actual effect size was then estimated using Duval and Tweedie's trim-and-fill procedure¹¹⁴.

2.2.6 Analysis for special relationship

As risk factors can be expressed as ordinal and nominal variables, in addition to continuous variables, dose-response meta-analysis were performed to combine the effect sizes. For age factors reported as ordinal variables, dose-response meta-analysis was performed. The middle value of categorical prognostic factors in each category was chosen. When the upper and lower limits were not available, 1.5 times the closest lower category and 1 were used, respectively. We then used both linear and non-linear models to fit the potential age trend and risk of developing severe manifestations. The Akaike information criterion and the Bayesian information criterion were used to select the optimal model. For categorical covariates, we used dose-response meta-analysis with package Doseresmeta in R version 3.4.4¹¹⁵.

2.3 Results

We identified 6,643 articles in our initial searches, of which 3,796 were excluded due to duplication (Figure 2.1). 2,847 articles were screened on titles and abstracts and 457 articles were passed to full-text assessment for eligibility; 307 articles were excluded as they did not meet the eligibility criteria. 150 studies were included in the systematic review, 122 of which were included in the meta-analysis – these studies all reported predictors measured during the febrile phase of dengue or the first 4 days of illness. In the study selection, the agreement between the two reviewers was moderate, Kappa index 0.70 (95% CI 0.63-0.77) and percentage agreement 87% (95% CI 82-90%).

More than 80% of the included studies were conducted in Asia and Latin Americas and published between 2008 and 2020 (121 of 150). These studies were carried out using a cohort

design (60%, 90/150), a case-control design (29%, 44/150), and a cross-section design (11%, 16/150). Approximately 60% (87/150) of the included studies defined severity outcomes according to the 1997 WHO Classification²¹ and the remaining 40% used the 2009 WHO Classification⁵.

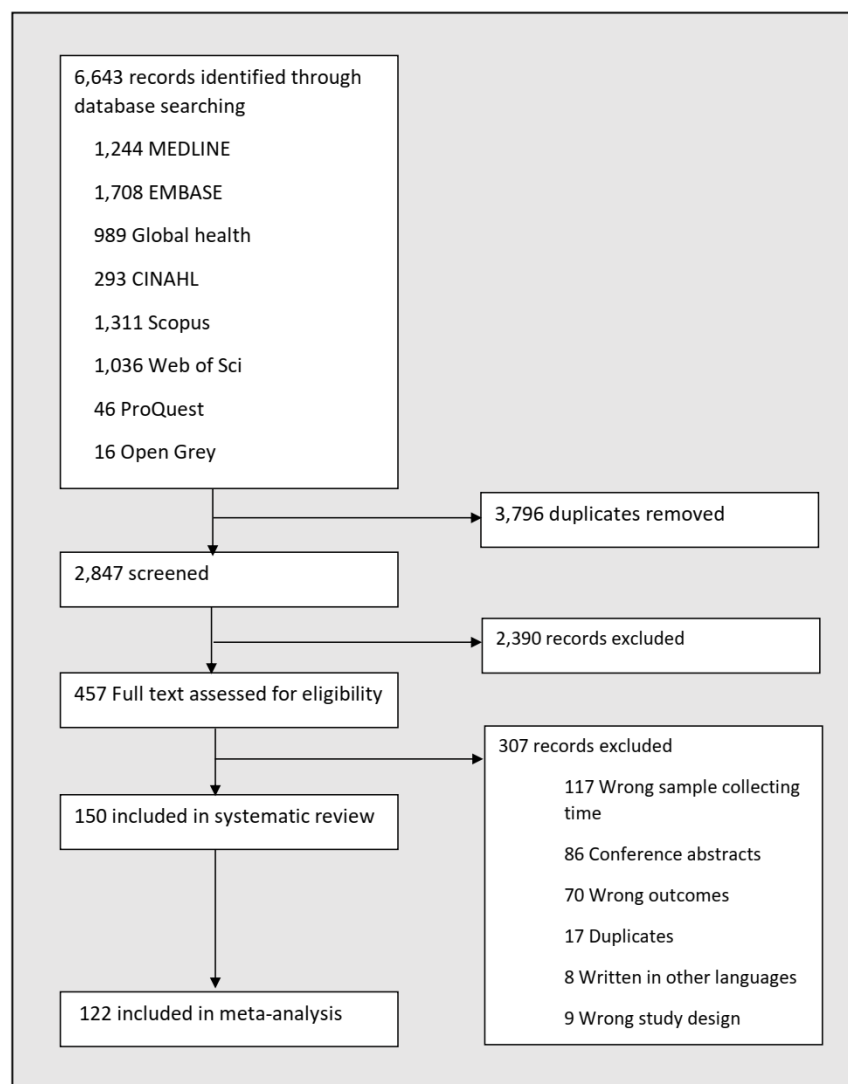


Figure 2.1. Flow diagram of study selection

Amongst the 150 studies included in the systematic review, we identified 203 factors that were investigated with dengue severity outcomes, 25 factors of which were reported in at least four studies and therefore were included in the meta-analysis. These factors were classified into four domains: nine factors in demographics and comorbidities, seven factors in clinical signs and symptoms, seven factors in patient laboratory parameters, and two factors

in viral related biomarkers. The definition of these factors included in this meta-analysis are presented in Table 2.2. The pooled crude associations for all 25 factors and the pooled adjusted associations for age, diabetes, and vomiting were reported.

Table 2.2. Definitions of factors included in the systematic review and meta-analysis

Associated factors	Definition
Age	Age of participants in year or month units recorded in medical records or history taking
Sex	sex recorded in medical records
Nutritional status	For children (≤ 18 years old), weight for age based on local standardized guidelines or BMI-for-age >2 Z score according to WHO sex-specific growth reference for children. For adults (>18 years old) the BMI criterion was used, malnourished BMI < 18 and obese BMI > 30 kg/m ²
Weight	Weight of patients measured in the hospitals or recorded in medical records
Mixed comorbidity	Unspecific pre-existing comorbidities e.g., cancer, asthma, chronic obstructive pulmonary disease, etc. according to medical records or history taking
Hypertension	Pre-existing hypertension recorded in medical records or history taking
Diabetes mellitus	Pre-existing diabetes mellitus recorded in medical records or history taking
Renal disease	Pre-existing renal disease excluding acute kidney injury related to dengue episode
Cardiovascular disease	Pre-existing cardiovascular disease e.g., chronic heart failure, myocardial infarction according to medical records
Rash	Presence of unspecific rash during the first four days of the illness according to history taking or physical examinations

Associated factors	Definition
Vomiting	Presence of any episodes of vomiting during the first four days of the illness according to history taking
Abdominal pain and tenderness	Presence of abdominal pain and tenderness of any intensity during the first four days of the illness according to history taking or physical examinations
Headache	Presence of headache during the first four days of illness according to history taking
Minor bleeding	Presence of either mucosal or other spontaneous bleeding that did not require any intervention e.g., blood transfusion and present during the first four days of illness. This includes skin bleeding (petechiae, ecchymosis, or purpura), mucosal bleeding (epistaxis, gum bleeding, or other sites), haematemesis or melena.
Positive tourniquet test	Tourniquet test with the presence of ≥ 20 petechiae per square inch.
Immune status	Secondary infection determined with the use of serological tests (i.e. either IgM or IgG ELISA during the first four days of illness or the convalescence phase) based on criteria in each study
Clinical fluid accumulation	Presence of fluid accumulation (i.e. either pleural effusion or ascites) during the first four days of illness by either chest X-ray or ultrasonography, or physical examination
Serotype	Dengue serotype measured by qualitative RT-PCR
Viraemia levels	Viral RNA levels quantified using quantitative RT-PCR assays from patients' plasma collected during the first four days of illness
Platelet count	Platelet count during the first four days of illness. If there were more than one result, the platelet count on day 3 of illness was chosen.
Leukocyte cell count	The number of leukocytes during the first four days of illness. If there were leukocyte count results for more

Associated factors	Definition
	than one day, leukocyte count on day 3 of illness was chosen.
Haematocrit	Haematocrit measured during the first four days of illness. If there are reports of haematocrit for more than one day, haematocrit on day 3 of illness was chosen.
Aspartate transaminase (AST)	AST levels measured during the first four days of illness. If there are reports of AST on more than one day, AST on day 3 of illness was chosen.
Alanine transaminase (ALT)	ALT levels measured during the first four days of illness. If there are reports of ALT on more than one day, ALT on day 3 of illness was chosen.
Serum albumin	Serum albumin levels measured during the first four days of illness. If there are reports of serum albumin on more than one day, albumin levels on day 3 of illness was chosen.

2.3.1 Demographics

Age

In the main analysis, the age profiles of patients who progressed to severe disease were not significantly different from the age profile of the patients who did not progress to severe disease (Figure 2.2). Heterogeneity in the pooled results was substantial in both the main meta-analysis and in the sensitivity analysis (Figure 2.2). In the subgroup analysis based on studies by age, adults who progressed to severe disease were significantly older than individuals who did not progress (SMD = 0.1 and 95% CI = 0.02 to 0.18) (Figure 2.3).

For the adjusted associations of age taken from studies reporting multivariable models, all the studies were conducted in children and used the 1997 WHO classification. We found that younger children had a higher risk of progression to severe disease than older children. In the children aged under 15 years, the odd of developing severe disease decreases by 8% for every year of increase in age (Appendix 5). In a dose-response meta-analysis between age and

disease severity, a linear model provided the best goodness of fit according to Akaike Information Criterion and Bayesian Information Criterion. The odds ratio of progression to severe disease was 1.03 (95% CI 0.96-1.11) but this association was not statistically significant (Appendix 6-7).

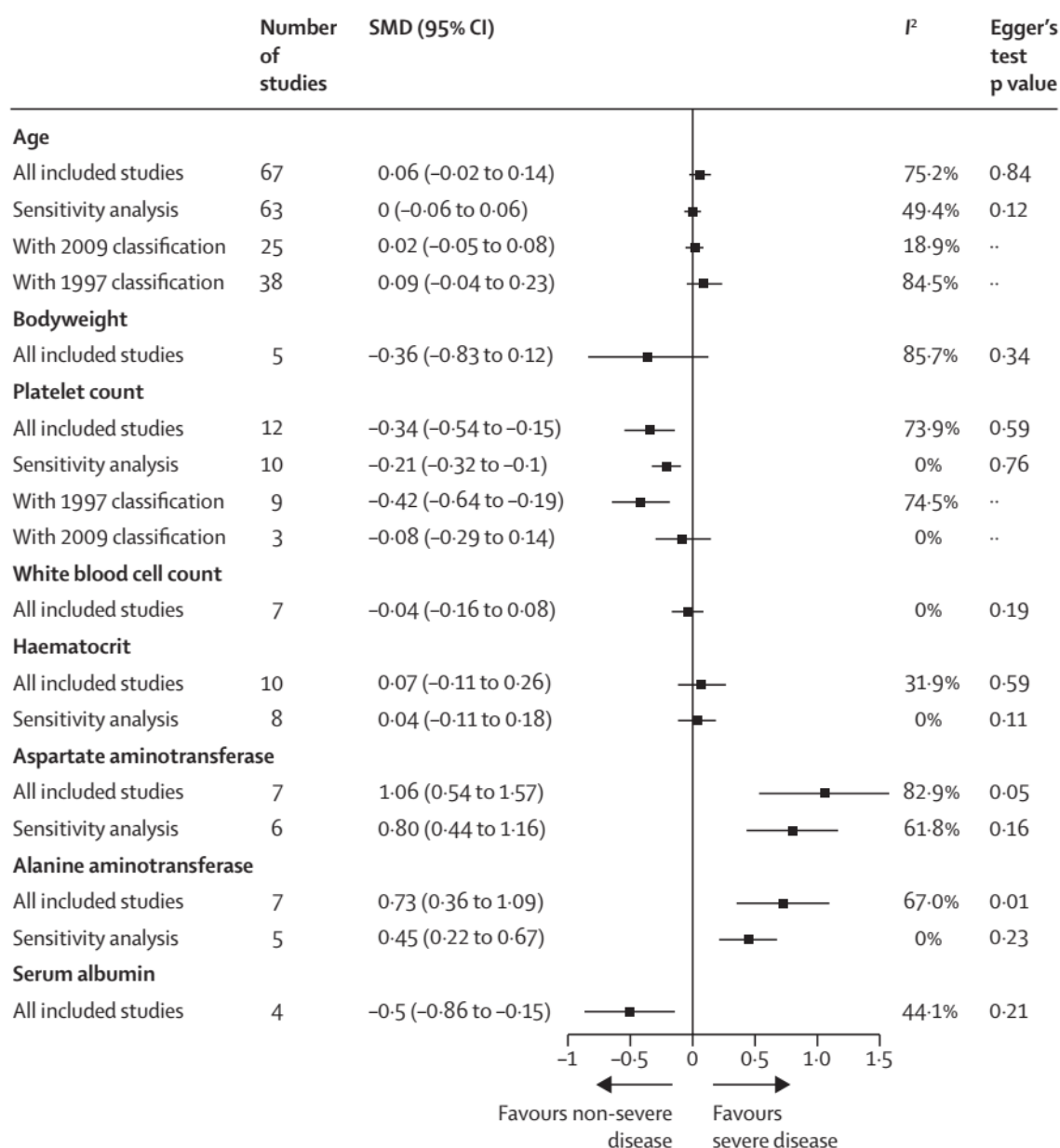


Figure 2.2. Pooled SMD of factors potentially associated with progression to severe dengue. In sensitivity analyses of all factors except platelet count, studies were excluded due to extreme effect sizes and

heterogeneity. For platelet count, the sensitivity analysis was done excluding studies that reported odds ratios instead of SMDs; SMD=standardised mean difference. Reprinted from¹¹⁶.

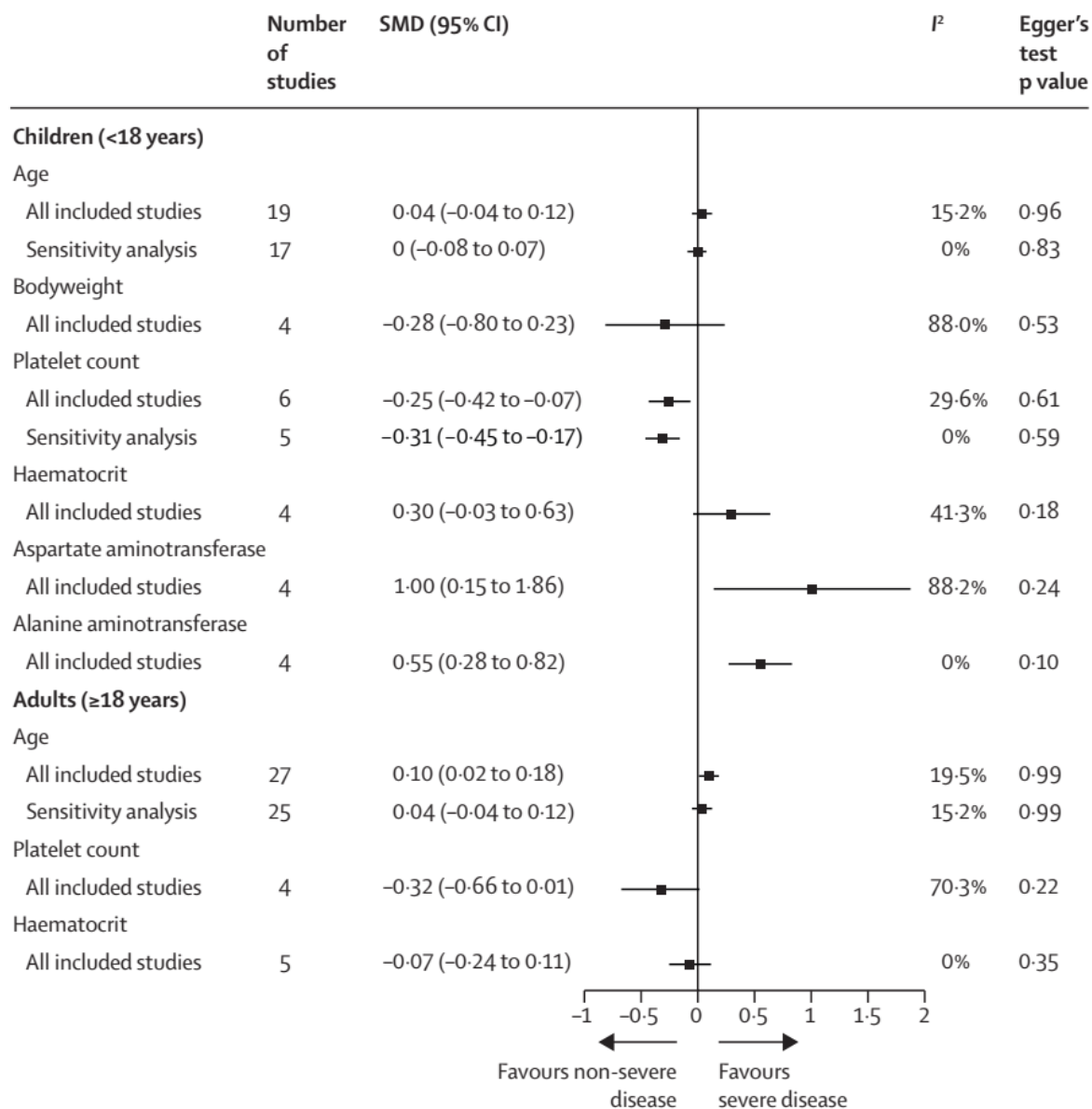


Figure 2.3. Pooled SMD of factors potentially associated with progression to severe dengue by age of study population. In sensitivity analyses, studies were excluded due to extreme effect sizes and heterogeneity. SMD=standardised mean difference. Reprinted from¹¹⁶.

Sex

Female sex significantly increased the risk of progression to severe dengue illness in the main analysis, OR 1.13 (95% CI 1.01-1.26) (Figure 2.4). However, this association was not significant

in the sensitivity analysis removing two studies with the extreme effect size and heterogeneity (Figure 2.4). We found evidence of publication bias in the pooled results included in the meta-analysis, Egger's test p-value 0.021 and thus performed Duval and Tweedie's trim and fill procedure to adjust the pooled OR. The adjusted OR was 0.95 with 95% CI 0.84 to 1.08, which was not statistically significant (Appendix 8). In the subgroup analysis by age groups, we found that female sex was associated with a higher risk of progression to severe disease in adults but not in children with an OR of 1.31 (95% CI 1.03-1.66) and 1.01 (95% CI 0.88-1.17), respectively (Figure 2.5).

Comorbidity

We conducted four meta-analyses for comorbidities: one with mixed comorbidities referring to non-specific underlying diseases such as respiratory diseases, cancers, or gastrointestinal diseases; and one for each diabetes mellites, hypertension, renal disease, or cardiovascular disease performed separately. The presence of mixed comorbidities was not associated with progression to severe disease in either the main analysis or sensitivity analysis (Figure 2.4). Hypertension, diabetes mellites, renal disease, and cardiovascular disease were positively associated with progression to severe disease (Figure 2.4). Diabetes mellitus and renal disease were explicitly associated with progression to severe illness with pooled ORs 4.38 (95% CI 2.58-7.43) and 4.67 (95% CI 2.21-9.88), respectively (Figure 2.4). Hypertension and cardiovascular disease were also positively associated with progression to severe illness with pooled ORs 2.19 (95% CI 1.36-3.53) and 2.79 (95% CI 1.04-7.5), respectively (Figure 2.4). We found heterogeneities in the pooled ORs of the four comorbidities in either the main studies or sensitivity analyses were at moderate to substantial levels. In addition, we performed a meta-analysis for the adjusted associations of diabetes mellitus and progression to severe illness obtained from multivariable models in the five included studies. The presence of pre-existing diabetes mellitus increased the odds of developing severe illness by 89% compared with no diabetes mellitus, OR 1.89 (95% CI 1.30-2.73) (Appendix 9).

Nutritional status

We included five studies reporting associations of nutritional status and progression to severe illness and four of which were conducted only in children. The nutritional status in studies with children was defined using the weight for age according to local standardised guidelines while the BMI criterion was used in the studies with adults. Where BMI is less than 18 kg/m² or more than 30 kg/m², malnourished or obese status was assigned, respectively. We found that malnourished individuals were at a lower risk of progression to severe dengue than the obese individuals, OR 0.76 (95% CI 0.58-0.99) but this was not significant when compared with normal weight individuals, OR at 0.80 (95% CI 0.63-1.01) (Figure 2.4). Obese individuals were likely to have higher risk of progression to severe illness than individuals with normal nutritional status, but the risk was not statistically significant, OR at 1.06 (95% CI 0.90-1.24) (Figure 2.4).

2.3.2 Clinical signs and symptoms

The associations of the seven clinical signs and symptoms are presented in Figure 2.4. We found that the presence of vomiting, abdominal pain or tenderness, bleeding, or clinical fluid accumulation increased the risk of progression to severe illness. The details of vomiting definition were not reported; only two of eight studies defined the presence of vomiting as at least two episodes during the febrile phase. The pooled crude association and adjusted association of vomiting and progression to severe illness were at ORs 2.25 (95% CI 1.87 to 2.71) and 2.00 (95% CI 1.69 to 2.49) respectively (Figure 2.4 and Appendix 10). No evidence of heterogeneity or publication bias were found in the pooled results of both parameters.

We also found that the presence of abdominal pain and tenderness increased the odds of progression to severe illness by 92% compared with individuals without this symptom, OR 1.92 (95% CI 1.35-2.74) (Figure 2.4). However, the precise definition was not given in the majority of the studies included and the heterogeneity was found to be high ($I^2 = 75.1\%$) (Figure 2.4). In the main analysis, we defined the presence of bleeding as the mucosal or other spontaneous bleeding. We found that individuals with bleeding had a higher risk of progression to severe illness with OR at 1.57% (95% CI 1.13-2.19) (Figure 2.4). The sensitivity analysis was also performed with the two definitions of bleeding (i.e., spontaneous, and

262 mucosal bleeding) and no evidence of a difference was found (Figure 2.4).

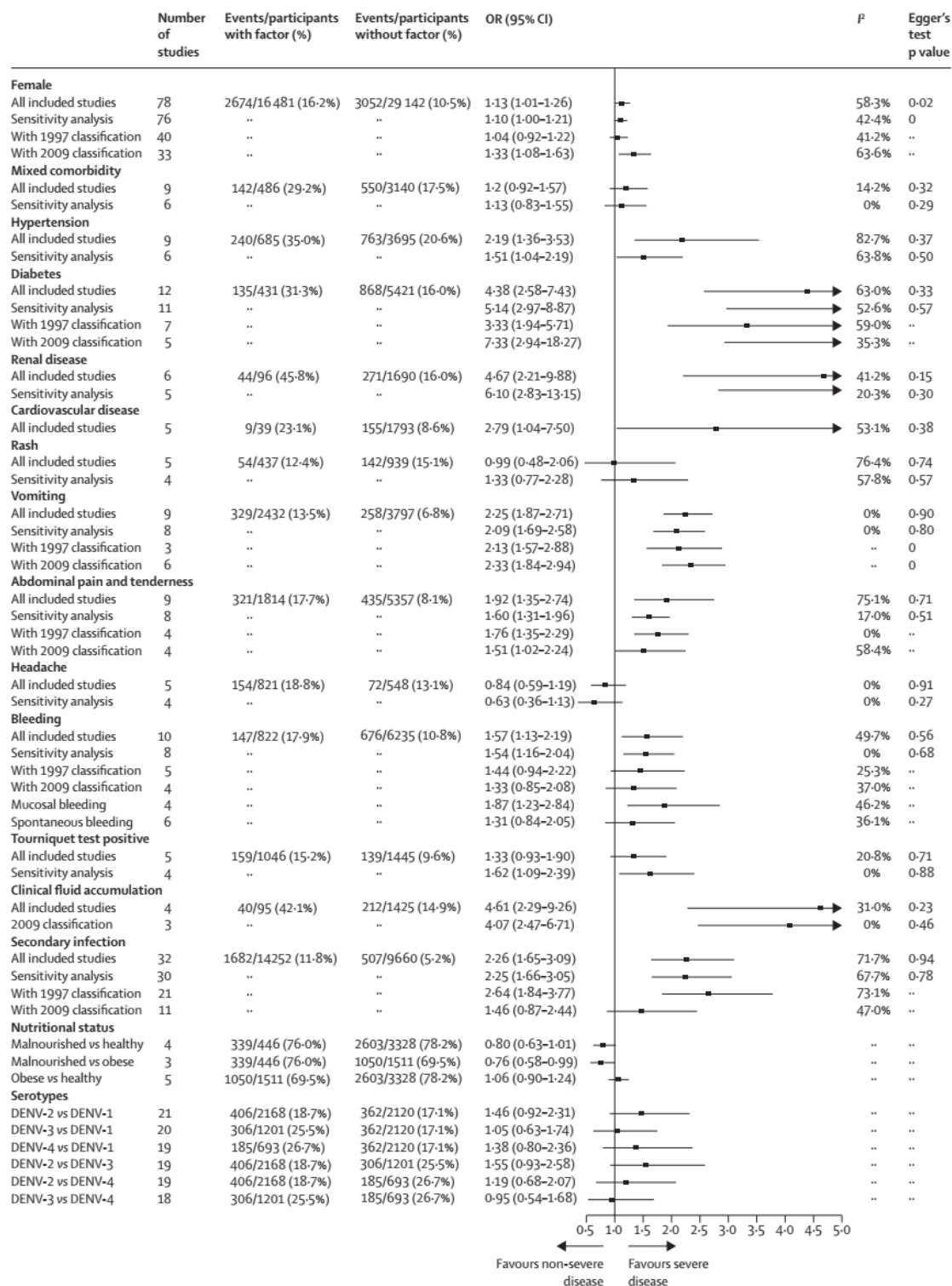


Figure 2.4 Pooled ORs of progression to severe dengue given the presence of potential prognostic factors. In the sensitivity analyses, studies were excluded due to extreme effect sizes and heterogeneity. DENV=dengue virus serotype. OR=odds ratio. Reprinted from¹¹⁶.

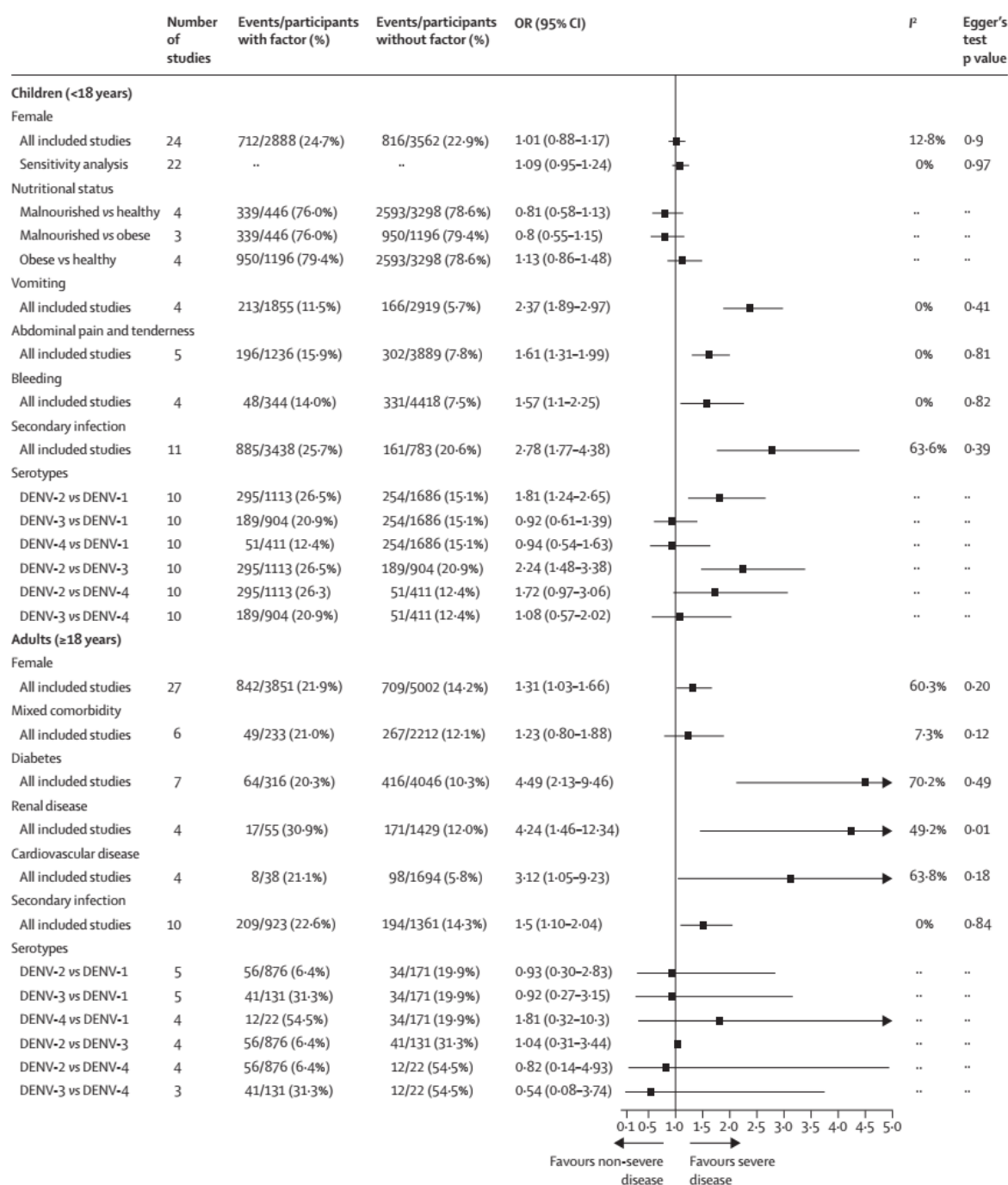


Figure 2.5. Pooled OR of progression to severe dengue given the presence of potential prognostic factors by age of study population. In sensitivity analyses, studies were excluded due to extreme effect sizes and heterogeneity. DENV=dengue virus serotype. OR=odds ratio. Reprinted from¹¹⁶.

The association of bleeding and progression to severe illness obtained from the sensitivity analysis (i.e., having removed two studies) remained statistically significant, OR 1.54 (95% CI 1.16-2.04) and no heterogeneity was demonstrated in the pooled results (Figure 2.4).

Regarding clinical fluid accumulation (detected during the first four days of illness), four studies were included; three of which referred to both ascites and pleural effusions and one referred to pleural effusions only. Even with the limited number of studies included in this meta-analysis analysis and the limited information on methodologies used for detecting pleural effusions, a significant association between the presence of clinical fluid accumulation during the first four days of illness and progression to severe dengue illness was found, with an OR of 4.61 (95% CI 2.29 to 9.26) with low-level heterogeneity ($I^2 = 0.31$) (Figure 2.4).

No association was found with a positive tourniquet test, rash, or headache and progression to severe illness. We found no heterogeneity in the pooled data of headache and positive tourniquet test, but there was moderate heterogeneity in the pooled results of rash ($I^2 = 0.74$) (Figure 2.4).

Laboratory parameters

Haematological parameters

Three haematological laboratory parameters were included in the meta-analysis- haematocrit (HCT), white blood cell count (WBC), and platelet count (PLT). Only lower PLT was significantly associated with progression to severe illness (measured during the febrile phase). In the main analysis including 12 studies, we found that patients with lower platelet count were associated with progression to severe illness. This finding was consistent with the pooled results from the sensitivity analysis which omitted the two studies originally reported ORs. Neither heterogeneity in the pooled results nor publication bias was not found in the results of the sensitivity analysis. Our meta-analyses found no association between dengue severity and either HCT or WBC measured during the febrile phase. In a subgroup analysis of HCT by age, children who progressed to severe illness had higher HCTs than children who did not progress,

but the association was not significant, with a pooled SMD of 0.30 (95% CI -0.03 to 0.63) (Figure 2.3).

Aspartate and alanine aminotransferase (AST and ALT)

We included seven studies investigating the association of the liver enzymes, AST and ALT with progression to severe illness. We found that either AST or ALT measured in the febrile phase was significantly higher amongst individuals who progressed to severe illness than individuals who did not progress, with a SMD of 1.06 (95% CI 0.54 -1.57) and 0.73 (95% CI 0.36-1.09), respectively (Figure 2.2). The positive associations of both AST and ALT, and progression to severe illness were consistent within the results of sensitivity analysis and subgroup analysis. In the sensitivity analysis omitting one study for AST and two studies for ALT, the pooled SMDs of AST and ALT of individuals with progression to severe illness remained significantly higher than for individuals who did not progress, with a pooled SMD of 1.06 (95% CI 0.54-1.57) and 0.73 (95% CI 0.36-1.09), respectively (Figure 2.2). Although the heterogeneity in the pooled SMDs of AST and ALT were found to be moderate to high, the heterogeneity in the pooled associations of AST and ALT decreased to zero and moderate, respectively. We also found evidence of publication bias in both the pooled associations of AST and ALT. However, the pooled associations of AST and ALT with disease progression after Duval and Tweedie's trim and fill procedure remained statistically significant, SMD 0.81 (95% CI 0.20–1.42) and 0.73 (95% CI 0.36–1.09), respectively (Appendix 11 and 12).

Serum albumin

We included four studies that investigated the association of serum albumin and progression to severe dengue. All the studies used dengue haemorrhagic fever according to the definition by the 1997 WHO classification as the main outcome (none used the 2009 WHO classification). Individuals who progressed to severe illness had lower serum albumin levels compared with individuals who did not progress, the pooled SMD of -0.5 (95% CI -0.86- -0.15) (Figure 2.2). The heterogeneity in the pooled results was low to moderate and no evidence of publication bias was found.

2.3.3 Immune response (primary vs secondary infection)

Thirty-two studies investigated the association of primary/secondary dengue infection and progression to severe illness. Individuals with secondary infections had over twice the odds of progression to severe illness compared to individuals with primary infections, OR 2.26 (95% CI 1.65-3.09) (Figure 2.4). This finding was consistent with the pooled associations of both sensitivity analysis and subgroup analysis. In the sensitivity analysis two studies were omitted, and the pooled OR was 2.25 (95% CI 1.66-3.05) (Figure 2.4). In the subgroup analysis by age, secondary infections amongst children (OR 2.78; 95% CI 1.77-4.38) had a stronger association with progression to severe illness compared with adults (OR 1.5 (95% CI 1.10-2.04) (Figure 2.5). However, the heterogeneity in the pooled ORs of the subgroup analysis in children was moderate to substantial while no heterogeneity was found in the pooled ORs in studies with adults. We found no evidence of publication bias in the subgroup analysis for both children and adults.

2.3.4 Virological parameters

Viraemia

Of the four studies included in the meta-analysis, five associations between viraemia and progression to severe illness were reported and included in the meta-analysis. We found substantial heterogeneity in the pooled association ($I^2 = 92$), so we did not report this association in the meta-analysis (Appendix 13). Descriptively, two associations were not statistically significant, one reported that higher viraemia was associated with progression to severe illness and the other reported the opposite association. One study investigated the association of viraemia and severity in two different population; they found a positive association between viraemia and progression to severe illness amongst adults with secondary infections, with DENV-1, 2, and 3 but found a negative association in children with secondary infection with DENV-1.

Viral serotypes

The main meta-analysis consisted of 21 studies reporting dengue viral serotype and progression to severe illness. DENV-2 tended to be associated with progression to severe illness compared with the other serotypes, but no statistically significant difference found in the main analysis. However, we found that DENV-2 increased the odds of progression to severe illness by 80% (OR 1.81; 95% CI 1.24-2.65) amongst studies with children and two times (OR 2.24; 95% CI 1.48-3.38) when compared with DENV-1 and DENV-3, respectively (Figure 2.4).

2.3.5 Subgroup analyses of prognostic factors in children and adults

In a subgroup analysis by age, 12 risk factors were identified for adults and 10 for children which were then included in the meta-analyses (Appendix 14). While secondary infection was associated with an increased risk of severe dengue in studies including both children and adults, haematocrit during the first four days of illness was not significantly associated with disease severity in either age group (Figures 2.3). Lower platelet counts and serotype DENV-2 was associated with a significant risk of developing severe dengue in children but not amongst adults (Figures 2.3 and 2.5). Age was not significantly associated with disease severity amongst children, whereas increasing age in adults was associated with severe dengue (SMD = 0.1; 95% CI = 0.02-0.18) (Figure 2.3). Female sex was also significantly associated with developing severe dengue amongst adults (pooled OR = 1.31; 95% CI = 1.03-1.66) but not amongst children (pooled OR = 1.01; 95% CI 0.88-1.17) (Figure 2.5).

2.3.6 Risk of bias and quality assessment

Using QUIPS tool to assess risk of bias in six domains (Figure 2.6), we found high risk of bias where potential confounders were not addressed and adjusted appropriately. There was also a significant proportion of moderate to high risk of bias in patient selection, as some studies recruited hospitalised patients, which could have missed patients presenting with mild symptoms. The risk of bias in study attrition was low because illness duration was relatively short, generally around seven days. Although the risk of bias in the outcome and prognostic

factors were low as the all the studies included used definitions based on the WHO Guidelines, 25% of included studies were considered to have a moderate risk of bias for outcome and prognostic factors.



Figure 2.6. Risk of bias in the included studies assessed by the Quality in Prognostic Studies.

2.4 Discussion

This systematic review and meta-analysis revealed that younger children, older adults, and female sex were demographical risk factors for progression to severe dengue illness. Individuals with pre-existing diabetes mellitus, hypertension, renal disease, or cardiovascular disease were at higher risk of disease progression. Patients with the following symptoms/signs of vomiting, abdominal pain and/or tenderness, bleeding, or clinical fluid accumulation during the febrile phase were at higher risk of progression to severe illness. Regarding laboratory parameters, patients with lower PLT, serum albumin level, and higher AST or ALT levels measured during the febrile phase were associated with progression to severe illness. Of the

different serotypes, DENV-2 infection amongst children and secondary dengue infections also increased a risk of progression to severe illness (Table 2.3).

Table 2.3. Summary of factors measured within the first 4 days of illness and their association with progression to severe dengue illness

Domain	Associated with progression	Not associated with progression
Demographics and comorbidities	Older age (in adults), younger age (in children), female sex, diabetes mellitus, hypertension, renal disease, cardiovascular disease	Nutritional status, bodyweight, mixed comorbidity
Signs and symptoms	Vomiting, abdominal pain or tenderness, bleeding, and pleural effusion or ascites	Rash, headache, and positive tourniquet test
Laboratory parameters	Platelet count, aspartate aminotransferase, alanine aminotransferase, serum albumin, and secondary infection	White blood cell count, haematocrit
Virological profile	Dengue viral serotype 2 (in children)	Viral load

Although we found no significant association between age and progression to severe illness in the main analysis, older adults were at a higher risk of progression to severe illness in a subgroup analysis by age groups. The tendency of older adults to progress to severe illness amongst adults could be due to co-existing (and often undiagnosed) diabetes mellitus, hypertension, renal disease, or cardiovascular disease, which are commonly found amongst older adults and associated with progression to severe illness. In addition, a recent study found immunosenescence, a dysfunctional immune system due to ageing, could enhance viral infection and thus increasing a risk of severe disease amongst elderly¹¹⁷. We also found that younger children were at a higher risk of progression to severe illness observed in the pooled adjusted association taken from multivariable models. Biologically this could be explained by children having increased vascular filtration capacity of their developing capillaries and less capacity to compensate for plasma leakage than adults⁴⁵.

Female gender was associated with a higher risk of progression to severe illness in both the main analysis and in subgroup analysis with adult studies. The increased tendency of progression to severe illness in women could be explained by the differences in the immune

response observed in female individuals. The study⁴⁹ found that the antibody titres of female patients dengue was significantly higher than male patients in secondary infections. This could be implied that female sex prone to develop more inflammatory response and more severe plasma leakage than male patients. In addition, this finding could be related to the difference of health seeking behaviours between males and female individuals⁴¹.

Diabetes mellitus and hypertension have been highlighted as potential risk factors in both the 1997 and 2009 WHO guidelines⁵. These risk factors are confirmed in our systematic review and meta-analysis⁴⁸. However, the actual pathophysiology of diabetes mellitus contributing to progression to severe illness is still not fully known. In a study by Lee et al¹¹⁸, patients with poor glycaemic control defined as A haemoglobin A1c (HbA1C) > 7% were at the higher risk of progression to DHF and DSS compared with diabetic patients with adequate glycaemic control. One possible explanation is that pre-existing vascular dysfunction and associated endothelial activation in diabetes patients predisposes them to more severe dengue-associated vasculopathy. We also demonstrated the association between renal disease and cardiovascular disease and progression to severe illness. In patients with renal disease, a pre-existing pro-inflammatory state has been observed, along with the associated endothelial dysfunction which could also increase the risk of vascular injury during dengue illness¹¹⁹. Uraemia, which is commonly found in chronic kidney disease, also increases the risk of endothelial dysfunction,^{79, 120}. The association between either renal disease or cardiovascular disease should be interpreted with caution since this finding was not adjusted for potential confounders such as diabetes mellitus or hypertension, which are often present in the same individuals.

Patients with the signs/symptoms including vomiting, abdominal pain and tenderness, bleeding, and clinical fluid accumulation detected during the febrile phase were also associated with progression to severe illness. Heterogeneity of these predictors were moderate to substantial. This could be due to the subjective nature of these signs and symptoms which depend on patient reports as well as clinicians who performed the physical examination. In addition, the measurement and definition of these signs were not uniform³⁰. Bleeding manifestations in dengue patients range broadly from minor bleeding to severe

haemorrhage which needs an intervention to control the bleeding and/or blood component transfusion. Minor bleeding manifestations during the febrile phase including either mucosal or spontaneous bleeding were associated with progression to severe illness.

The detection of clinical fluid accumulation during the febrile phase was associated with progression to severe illness. Although all dengue patients develop some evidence of plasma leakage usually manifesting in the critical phase, early fluid accumulation detected in the febrile phase suggests plasma leakage is already more pronounced and these patients are at risk of intravascular volume depletion and decompensating later in the critical phase. Portable ultrasonography can facilitate early detection of fluid accumulation and as this technology becomes more available and affordable in low-to-middle-income settings, this is likely to assist patient triage in the future.

The presence of either clinical fluid accumulation or bleeding manifestations is part of the diagnostic criteria of DHF. As a result, this will influence the association between these factors and severe illness. However, these factors could still have prognostic utility as these parameters were measured specifically during the febrile phase or within the first 96 hours of clinical onset (i.e., prior to the development of severe dengue). In addition, the results of subgroup analysis by WHO classification found no statistically significant pooled effect sizes between the two subgroups.

Haematological laboratory parameters have been used in both dengue diagnosis and clinical classification. We found only PLT was useful as a prognostic factor; WBC and HCT measured during the febrile phase did not predict progression to severe illness. This finding supports the use of the warning sign rapid fall in PLT (accompanied by a rise in haematocrit) in the 2009 WHO guideline. The utility of PLT count as a predictor of progression to severe illness is also consistent with the findings in a study by Lam et al. that showed incorporating the dynamics of platelet count increased the predictive performance of progression to dengue shock syndrome¹⁰¹. No study has been able to provide robust evidence of a clinically useful cut-off value for platelets which can identify individuals at risk of progression to severe illness. We observed that the average platelet count in the group of patients who progressed to severe

dengue was less 100,000 cells/dl, which is consistent with recommendations for hospitalisation of the 1997 WHO dengue classification and 2011 SEARO dengue guideline^{21, 51}.

The liver enzymes AST and ALT are not included in the warning signs in the 1997 and 2009 WHO Guidelines^{5, 21}. However, this study demonstrated that higher levels of AST and ALT in the febrile phase were significantly associated with progression to severe illness. We would suggest monitoring AST or ALT during the febrile phase is useful to identify high risk patients. This study could not provide the cut-off value of AST or ALT for predicting progression to severe illness. However, we would suggest that levels above three times upper limit of normal AST and ALT could be used as a clinically relevant cut-off value for progression to severe illness (derived from the average AST and ALT levels in the included studies).

Patients with lower serum albumin levels measured during the febrile phase were also found to be at a higher risk of disease progression. The lower level of serum albumin can be explained by the extravasation of albumin with plasma during early vascular leakage. The use of albumin as an indicator of the degree of early plasma leakage could therefore also be considered. A limitation of measuring serial albumin levels is cost, and levels can be influenced by pre-existing nutritional status. Our data support the 2011 SEARO Dengue guideline recommendation; a serum albumin level of lower than 3.5, or reduction by 0.5 mg/dl during the febrile phases compared with baseline levels is associated with progression to severe illness⁵¹.

In the main analysis, certain dengue serotypes were not associated with progression to severe illness. However, infection with DENV-2 increased the risk of progression to severe dengue compared to DENV-1 or DENV-3 amongst children, but not adults. The different associations of serotypes and progression to severe disease between adults and children could be due to potential confounders such as immune status, the sequence of infecting serotypes, and time interval between infections.

The increase in risk to severe illness during a secondary dengue infection has been well documented in previous studies. This is thought to be due to the phenomenon of antibody-

dependent enhancement whereby non-neutralising antibodies elicited from a primary infection facilitates viral uptake into cells and increases viral replication during secondary infections. The association of secondary dengue infection and progression to severe illness in adults was not as strong as the association amongst studies with children. This was in line with the effect of age on secondary infection and clinical progression¹²¹. This finding may be related to the high force of infection in some endemic settings where individuals experience their second dengue infection as children and their tertiary and quaternary infections as adults which are usually asymptomatic^{122, 123}. There are no widely used assays to distinguish tertiary and quaternary from the secondary infection in general practice^{122, 124}. The assay (i.e. Hemagglutination inhibition assay and Plaque reduction neutralization test) was merely used in a research setting with prospective design studies. A reliable assay for immune status at a single time-point during the febrile phase is still challenging but could be clinically useful in terms of risk prediction and vaccination. Secondary infections are clearly more associated with progression to severe illness but until more robust point-of-care tests are available in general practice, immune status cannot be currently included in any triage system.

Malnourished children were found to be at a lower risk of progression to severe illness than children with obesity⁶. This could be related to a less robust immune response with less inflammation in malnourished children causing less tissue injury during the acute illness. We found weak evidence of association between obesity and progression to severe illness whereas a previous meta-analysis reported obesity was positively associated with dengue shock syndrome⁴⁸. The inconsistency may be due to the different outcomes used in the studies.

We found no evidence that a positive tourniquet test is useful to predict a more severe clinical outcome. However, we did not assess the use of tourniquet test as a diagnostic tool in dengue as suggested by the 1997 and 2009 WHO guidelines^{5, 21}. Although white blood cell count and haematocrit levels measured within the first 96 hours of illness are not able to predict progression to severe illness later in the critical phase, monitoring haematocrit has a key role for detecting and monitoring plasma leakage in the critical phase.

The pooled results of viraemia had substantial heterogeneity (I^2 92%). This could be due to large individual variability including infecting serotype, immune status, and age groups. In addition, samples were collected at different time points and viraemia was measured with different laboratory methods. Future studies on viraemia and dengue severity with large sample size and specific clinical time points are needed to address this gap. The most recent studies found the higher peak of viraemia was detected in the febrile phase, the more likely to develop severe disease regardless serotype and immune status⁵⁹. Using viraemia as a prognostic factor is still challenging as peak viraemia is commonly missed by the time patients present to hospitals. In addition, quantification of peripheral blood viraemia might not reflect the total body viral load.

This systematic review and meta-analysis focused on parameters detectable during the febrile phase, thus specifically identifying the factors that can predict the progression to severe illness early upon the onset of symptoms before the critical phase. We found lower heterogeneity in pooled effect sizes compared to previous studies^{24, 36, 102} because of the shorter and more specific time-course of illness considered and relatively lower variation in the clinical parameters.

This study has limitations. Since there were limited number of studies reporting adjusted associations derived from multivariable models, we reported pooled crude associations, which could have been overestimated, as these were not adjusted for confounding effects. In the few instances where meta-analyses were performed with a limited number of studies, pooled effect sizes should be interpreted with caution. The definition of some associated factors was not clearly provided and may thus present inconsistencies among studies. The quality and method that were used to assign the final diagnosis including quality and evidence supporting the final diagnoses were not clearly mentioned in the included studies. This could increase the heterogeneity of pooled associations because final outcomes were not assigned. We did not include genetic factors in this study because genetic tests have not been routinely analysed along with clinical outcomes. The study design and methodologies adopted for genetic studies differ to those used to analyse other clinical parameters. We limited the selection of the studies to those written in English which might have reduced the precision of

the estimates. However, over 80% of included studies were conducted in Southeast Asia and Latin America and thus the findings demonstrated in this study are therefore likely representative of general practice in dengue endemic areas. In addition, our meta-analysis has not been able to capture all pre-existing comorbidities or co-infections with other tropical infectious disease (e.g. sickle cell disease or malaria infection^{125, 126}), the use of non-invasive clinical monitoring (e.g. ultrasonography, arterial pulse waveform, or side-stream darkfield imaging) or the role of novel biomarkers in dengue which are of great research interest and represent a promising avenue to better understand dengue pathogenesis and identify new biomarkers^{76, 81, 127, 128}. Unfortunately, the number of studies involving these biomarkers is limited to date and significant heterogeneity in study designs hindered our ability to define roles for every clinical factor or novel biomarker.

In this study we have demonstrated that monitoring serum albumin, AST, and ALT during the febrile phase of illness, in addition to the warning signs detailed in the 2009 WHO guidelines, can enhance ability to predict the risk of developing severe illness. In addition, identifying the infecting viral serotype and the immunological status (e.g., primary versus secondary infections) during the febrile phase could further improve the accuracy of the risk predictions. Further prospective studies focusing on the febrile phase of illness will allow investigation of the role of novel biomarkers and non-invasive methods such as ultrasonography in predicting the early onset of severe dengue. This study adds to the overriding objective of improving patient triage, optimal allocation of resources, reduction in morbidity and mortality with the aim of curbing the economic and public health impact of dengue worldwide.

Chapter 3 Materials and analytical methods

In this chapter, I provide a general description of the materials and analytical methods used to address research objectives 2 and 3, i.e., the development and validation of risk prediction models for severe dengue illness using static (at a single time point) clinical observations and the daily monitoring of clinical parameters in time, respectively. The first part of this chapter describes the two data resources used in the development of clinical prediction models, including the study designs, study populations, study sites and the data collecting protocol. In the second part of this chapter, common analytical methods used in the subsequent chapters are specified including the protocols used for the descriptive analyses, variable selection, and dealing with missing data. In addition, the details of the analytical methods used to develop the clinical risk prediction models described in Chapter 4 and 5 (with both static and dynamical clinical observations) are presented in this chapter.

3.1 Materials

In the thesis, I developed risk prediction models using two independent datasets obtained from prospective cohort studies of children with dengue conducted in Vietnam and Thailand. The Vietnamese dataset was used as the development set in the development of prognostic models with clinical baseline parameters. The Thai dataset was used as the external validation set for the models with clinical baseline parameters (Chapter 4). The Thai dataset was also used to quantitatively describe and develop risk prediction models with the temporal dynamics of clinical parameters (Chapter 5).

3.1.1 Data of a prospective cohort study of Vietnamese children with dengue

This dataset was obtained from a prospective cohort study of children with dengue conducted at the Oxford University Clinical Research Unit (OUCRU) in Ho Chi Minh City with seven collaborator hospitals in Southern Vietnam (Hospital for Tropical Diseases, Children's Hospital Number 1, Children's Hospital Number 2, Tien Giang Provincial Hospital, District 8 Hospital) between 2010 and 2013. Children aged between 1 year and 15 years who presented at outpatient departments within the first 72 hours of fever and with suspected dengue infection

were recruited into the prospective study. Children with an identified source of infection or who could not be contacted via mobile phone during a follow-up period were excluded.

Trained study nurses discussed the purpose, the risks and benefits of being a part of the study, and the process of the study with family members or guardians. The family members or guardians were asked for written informed consent for the children to participate in the study. Demographic details, clinical history and the results of physical examinations were collected using an electronic case record form at enrolment. First blood samples for haematology, biochemistry, diagnostic laboratory (i.e., non-structural protein 1 (NS1) rapid test (NS1 Ag STRIP, Bio-Rad), and quantitative RT-PCR) were collected also at the time of patient enrolment. All ambulatory patients were followed up by daily phone calls to determine whether they were at home or had been hospitalised since the previous call. Daily phone calls were made until the fever had resolved and the patient had returned to the normal daily activities. Ten percent of ambulatory cases were drawn a second blood sample for serological diagnostic purposes during the recovery phase. For patients who were hospitalised at any time during the six days after enrolment, a case report form was completed to capture data on whether the patient evolved to severe dengue according to the World Health Organization (WHO) 2009 classification scheme and a second blood sample was collected at the time of discharge from hospitals. During the period of study, participants were managed following the standard treatment guidelines of the hospitals.

3.1.2 Data of a prospective cohort study of Thai children with dengue

This dataset was obtained from an existing database collected in a prospective cohort study conducted by the Armed Forces Research Institute of Medical Sciences (AFRIMS) at the Queen Sirikit National Institute of Child Health (QSNICH), in Bangkok, Thailand between 1994 and 2008. The fully anonymous data were stored at AFRIMS as electronic files. The study recruited children aged between 6 months and 14 years and presented at the outpatient department of the QSNICH within 72 hours of fever. Children with other specific sources of infection, anaemia, malnutrition, or history of chronic illness were excluded.

Children and family members or guardians were informed about the purpose, possible risks and benefits of participating in the study and then asked for a written consent form to participate in the study. Demographic details and the results of the clinical evaluation were recorded using case record forms at enrolment. All participants were admitted to the QSNICH hospital and monitored until the convalescence/recovery phase, defined as 24 hours after two consecutive temperatures below 38 °C (defervescence). The first blood samples were collected within the first 24 hours of admission and daily until discharge or a maximum of five consecutive days. Case management was conducted according to the 1997 and 2009 WHO Dengue Guidelines^{5, 21}. In addition to the standard guidelines, all participants in the defervescence phase (i.e., 1 to 2 days since defervescence) were tested with a finger-stick haematocrit test for measuring haemoconcentration and a right lateral decubitus chest x-ray was conducted for capturing pleural effusion. The effusion was measured as pleural effusion index defined to $100 \times (\text{maximum width of right pleural effusion}) / (\text{maximum width of right hemithorax})$.

3.1.3 Laboratory-confirmed dengue diagnosis

Laboratory confirmation was conducted according to a serological or viral molecular detection assay. All enrolment blood samples were tested with (1) RT-PCR assay; (2) NS1 ELISA (Platelia Dengue NS1 Ag ELISA, Bio-Rad); or (3) IgM (Panbio, Brisbane, Australia). The convalescence blood samples (i.e., typically after day 6 of illness) were tested with the IgM assay only. Laboratory-confirmed dengue was defined as a positive result to either of the three diagnostic tests (RT-PCR, NS1 ELISA, or IgM seroconversion). Since enrolment occurred early during the febrile phase of disease (≤ 72 hrs after fever onset), RT-PCR was positive for the near totality of all laboratory-confirmed dengue infections. As a result, the serological tests of paired blood sample confirmed only the status of RT-PCR negative patients as either true dengue or not dengue cases.

In the Vietnamese dataset, convalescent blood samples were not collected from all participants. Instead, a second blood sample was determined by a randomisation code and represented 10% of the ambulatory population. A convalescent blood sample was also

collected from all hospitalised patients, irrespective of clinical diagnosis, at the time of discharge from the hospital. In the Thai dataset, a convalescent blood sample was collected from all participants.

3.2 Common analytical methods

3.2.1 Descriptive analysis

Continuous data were summarised using the mean and standard deviation or the median and interquartile range, according to whether data normally distributed or not. The Shapiro-Wilk test was used for testing the normality of data. The data were considered to be normally distributed when the p-value of the Shapiro-Wilk test was more than 0.05. Categorical data were summarised in terms of their frequency and/or percentage. In comparing variables between two groups, we used the independent t-test and the Mann Whitney U test for normally and non-normally distributed variables, respectively. For categorical data, Fisher Exact test was performed if possible and otherwise a Chi-squared test was used. The temporal dynamics of daily clinical parameters over a day of illness was visualised using trajectory plots.

3.2.2 Treatment of missing data

Missing data could impact model precision. I investigated patterns of missingness and proportion of missingness. The data frames in this analysis were created to long format data, in which each row is one time point per subject. The proportion of missingness was defined as the number of rows with any missing values divided by the total rows of data. Missing completely at random was assumed for missingness in this thesis. Where the percentage of missingness was less than 5%, 5% to 15%, and more than 15%, available (complete) data analysis, single data imputation, and multiple imputation were performed, respectively, as recommended by Harrell et al¹²⁹.

In the Vietnamese dataset, there were only 51 (2.17%) missing values in either outcome or candidate predictors in the whole dataset (Appendix 17). I, therefore, performed complete

data analysis, analysing the data excluding the cases with missing information, without performing data imputation.

In the Thai dataset, the percentage of missingness was 25% and, thus, multiple imputation was performed. This was performed using multivariable imputation by chained equations (MICE), which is implemented in the R package mice¹³⁰. The mice package was developed to impute missing values using all available information in the dataset. Specifically, the algorithm starts by replacing the missing values of each variable with temporary central values derived from non-missing values of the variable. Then the temporary replacements of one variable are set back to missing and regression is performed to predict the individual-level missing values using all other (complete) covariates. These steps are iteratively repeated for each variable with missing data. For this dataset, data imputation was performed using predictive mean matching models and logistic regression for continuous and binary variables, respectively¹³¹. The clinical endpoints (i.e., DSS and the combined clinical endpoints of moderate plasma leakage and/or DSS) were excluded from the imputation models as recommended¹³². We generated ten imputed datasets, which were used for the model development and internal validation following recommendation by Wahl ¹³²(see Section 3.5.4).

3.2.3 Definition of prediction outcomes

In this study, we considered two clinical endpoints: DSS as the first endpoint, and a secondary combined endpoint including DSS and/or moderate plasma leakage (Table 3.2). DSS was defined by the presence of either narrow pulse pressure (≤ 20 mmHg) or hypotension according to normal range by age with signs of poor peripheral perfusion e.g. cold-clammy skin, increased capillary refill time (> 2 seconds), peripheral cyanosis, or skin mottling. For the secondary combined endpoint, moderate plasma leakage was defined as evidence of plasma leakage with either: greater than 15% rise in haematocrit during illness from baseline and/or evidence of new pleural effusion and/or ascites on ultrasound or chest x-ray in accordance with recommendations from dengue experts⁹. The haematocrit baseline was measured at enrolment within the first 72 hours of illness onset in the absence of any clinical signs of plasma leakage nor haemodynamic instability. The secondary combined endpoint aims to

identify the subset of patients who have less severe disease but require close monitoring and prompt medical intervention in hospitals, given DSS alone was a rare endpoint and represents an extreme in physiological derangement.

Table 3.1. Clinical endpoints and definition

Domain	Clinical outcome	Definition
Dengue shock syndrome	Haemodynamic instability	Plasma leakage resulting in shock defined as: - Narrow pulse pressure less than 20 mmHg OR - Systolic blood pressure less than 90 mmHg (or hypotension in accordance with the sex and age specific reference of population data) AND - Poor tissue perfusion such as clammy skin, urine output less than 0.5 ml/kg/hr
Moderate plasma leakage	15% rise of haematocrit levels from haematocrit baseline	Haematocrit rising 15% from baseline haematocrit (haematocrit for age or measured during the first 72 hours of illness with no signs of plasma leakage and haemodynamic instability)
	Evidence of clinical fluid accumulation	Fluid accumulation (i.e., pleural effusion or ascites) by chest x-ray, ultrasonography, or physical examinations (depending on health facility and completeness of medical records)

3.2.4 Metrics of predictive performance

The predictive performance metrics used in this thesis evaluate the overall prediction performance (the Brier Score), the discrimination and calibration of risk prediction models,

and decision curve analysis (DCA)^{133, 134}. Brier Score assesses a composite of discrimination and calibration of prediction models, with BS equal to 0 identifying the optimal model. With a binary outcome, this metric measures a variation in the observed probability across predicted risks. This metric may be inaccurate where the outcome is rare (severe class imbalance in a prediction outcome). Model discrimination identifies the ability of a model to correctly assign individuals into a certain group according to a true outcome, which is also denoted model accuracy. Discrimination performance was evaluated using the AUROC where 0.5 indicates predictions occurring by chance and 1.0 indicates the perfect discrimination performance. AUROC determines the performance of prediction models to all binary outcome thresholds. Model calibration describes the agreement of the predicted probability to the observed probability in the study population. Calibration was assessed and presented using calibration plots, calibration in the large (CITL, aka calibration intercept), and calibration slope. Calibration plots visualise the agreement between predicted risks and actual risk for observation (this was only used for the static models developed in Chapter 4)¹³⁵. In this plot, all individuals are stratified into equal-size groups, commonly 5 to 10 groups, then ranked based on the predicted probability. The mean predicted probability of each group is plotted on the x-axis while the observed proportion of subjects in the group with the relative outcome is plotted on the y-axis. The perfect prediction is on the diagonal line, where the x-axis and y-axis are the predicted and observed risks, respectively. The calibration intercept and calibration slope were used to characterise the calibration performance. The calibration intercept indicates systemic deviations in the predictions, whether over- or under-estimated, compared with the observed probability. The calibration slope assesses the spread of the predicted risks, where the perfect calibration slope is 1, while on the contrary, a calibration slope less than 1 indicates that the predicted risk is too extreme. In this case, the predicted risk is too high for a patient with high risk and too low for patients with low risk. On the contrary, a calibration slope larger than 1 indicates that the predicted risk is too moderate. Decision curve analysis evaluates the overall net-benefit of prediction models across threshold probabilities and compared to default strategies of treating all individuals as positive outcomes or negative outcomes across threshold probabilities (this was applied only to the static models in Chapter 4). Net-benefit is defined as $\text{sensitivity} \times \text{prevalence} - (1 - \text{specificity}) \times (1 - \text{prevalence}) \times w$, where w is the odds

at the threshold probability, defined as the accepted number of false positive for each true positive of the prediction outcome^{136, 137}. In addition, I also assessed accuracy, sensitivity, and specificity of prediction models, which are the most common metrics for classification models. The threshold probability for classification was determined using the minimum distance from the left upper corner, which aims at minimising false positive and false negative proportions. The threshold probability, accuracy, sensitivity, and specificity were implemented using the pROC package¹³⁸ and BS was calculated using the DescTools package¹³⁹ in R.

3.3 Methods for the development of prediction models using baseline clinical parameters measured within the first 72 hours of illness

This section describes the details of the methods used in the development of risk prediction models for progression to both clinical endpoints (DSS and a combined endpoint of DSS and moderate plasma leakage) using baseline clinical parameters measured during the first 72 hours of illness (Chapter 4).

3.3.1 Predictor selection

Candidate predictors were selected based on the systematic review and meta-analysis¹¹⁶ presented in Chapter 2; discussion with three dengue experts in Thailand and Vietnam (Prof Kalayanarooj, Dr Suntarattiwong, and Dr Yacoub); and data availability (Table 3.2). The number of candidate predictors was considered in the range of 10 to 25 events per predictive variable. Since there were 110 cases with DSS and 295 cases with a combined endpoint of moderate plasma leakage and/or DSS in the Vietnamese dataset, the number of candidate predictors considered ranged from 10 to 20 predictors.

Table 3.2. Definition and type of data variables available in the Vietnamese dataset

Domain	Predictor	Definition	Type of data
Demographics	Age	Age of participant at enrolment	Continuous
	Sex	Physical sex of patients	Binary

Domain			Predictor	Definition	Type of data
			Weight	Body weight at enrolment	Continuous
			Height	Height at enrolment	Continuous
			Nutritional status	Overweight is defined where BMI for age is higher than +2SD (in the Vietnamese data set)	Binary
Clinical signs and symptoms			Presence of vomiting	“Yes” if patients had a history of at least one episode of vomiting during the febrile phase; otherwise “No”	Binary
			Presence of abdominal pain or tenderness	“Yes” if patients had a history of abdominal pain or abdominal tenderness from physical examination during the febrile phase; otherwise “No”	Binary
			Presence of mucosal bleeding	“Yes” if patients had a history of mucosal bleeding (gum, epistaxis, gastrointestinal or vaginal bleeding) which does not need invasive interventions or blood component transfusion; otherwise “No”	Binary
			Presence of headache	“Yes” if patients had a history headache during the febrile phase; otherwise “No”	Binary
			Presence of anorexia	“Yes” if patients had a history anorexia during the febrile phase; otherwise “No”	Binary
			Presence of nausea	“Yes” if patients had a history nausea during the	Binary

Domain	Predictor	Definition	Type of data
		febrile phase; otherwise "No"	
	Presence of rash	"Yes" if patients had a history rash or maculopapular rash from physical examination during the febrile phase; otherwise "No"	Binary
Laboratory parameters	White blood cell count	Total white blood cell count (cells/dl) measured during the first 72 hours of illness	Continuous
	Neutrophil count	Total neutrocyte count (cells/dl) during measured the first 72 hours of illness	Continuous
	Lymphocyte count	Total lymphocyte count (cells/dl) during measured the first 72 hours of illness	Continuous
	Platelet count	Platelet count (cell/dl) measured during the first 72 hours of illness	Continuous
	Aspartate aminotransferase	Aspartate aminotransferase measured during the first 72 hours of illness	Continuous
	Alanine aminotransferase	Alanine aminotransferase measured during the first 72 hours of illness	Continuous
	Serum albumin	Serum albumin (g/dl) measured during the first 72 hours of illness	Continuous
	Dengue serotypes	Dengue serotypes detected using rt-PCR	Categorical

Domain		Predictor	Definition	Type of data
			during the first 72 hours of illness	
		Immune status	Serological definition relied upon IgM and IgG ELISA test in the febrile phase and convalescent phase; Primary infection: two negative dengue-specific IgG on febrile and convalescence samples; secondary infection: at least one positive dengue-specific IgG on febrile and convalescence samples; Indeterminate: single specimen (a febrile sample or a convalescent sample) available	Categorical
Clinical parameters	progression	Developing haemoconcentration	“Yes” if a patient develops 15% haematocrit rising from baseline haematocrit (defined as haematocrit levels measured during the first 72 hours of illness or average haematocrit levels based on population data); otherwise “No”	categorical
		Presence of vital sign instability (shock)	“Yes” if systolic blood pressure falls below the normal range by age, pulse pressure less than 20 mmHg, or undetected blood pressure, otherwise “No”	categorical
Final outcomes		Developing DSS	“Yes” if a hospitalised patient develops DSS	categorical

Domain	Predictor	Definition	Type of data
	Developing a combined endpoint of moderate plasma leakage and/or DSS	“Yes” if a hospitalised patient develops a combined endpoint of moderate plasma leakage or dengue shock, otherwise “No”	categorical

For logistic regression, the candidate predictors to be included in the risk prediction models were selected using lasso regression (described in Section 3.3.2). The set of predictors with the highest model accuracy and model stability was included in the final model¹⁴⁰. Model accuracy was evaluated using the regularisation parameter (λ) that provides the minimal misclassification error. I determined the minimal λ using stratified 10-fold cross-validation. Model stability and variable selection were performed using 1,000 bootstrap samples. The set of predictors with the most frequent selection across the 1,000 bootstrap samples was selected. For the risk prediction models developed using machine learning techniques, all candidate predictors were included.

3.3.2 Statistical and machine learning models

Logistic regression with lasso selection

Logistic regression is a generalised linear model where the outcome variable is binary, and the covariates can be either continuous or categorical. The Least Absolute Shrinkage and Selection Operation (lasso) is a method that performs regularisation and variable selection. With a negative log-likelihood, lasso seeks the model coefficients by adding a shrinkage penalty, which is the sum of the absolute values of the model parameters. The tuning parameter (λ) controls the relative impact of these two terms of the regression coefficient estimates. When λ is equal to 0 the penalty term has no effect, and lasso is equivalent to ordinary logistic regression. When λ is sufficiently large, the impact of the shrinkage penalty grows, and the coefficient becomes 0. To select the predictors to be included in multivariable logistic regression, I used lasso selection with 10-fold cross-validation, which is available in the glmnet package¹⁴¹ in R.

788 **Random forest**

789 Random forest is a tree-based model that combines the concept of bootstrapping and the
790 concept of subsampling predictors. This approach ensembles the results of the set of decision
791 trees that are constructed using bootstrapped training sets with a randomly selected set of
792 predictors. Firstly, a bootstrap sample of the original dataset (random resample of the data
793 with replacement) is created, and a random subset of predictors selected. A decision tree is
794 then built on the bootstrap sample. These steps are repeatedly performed until termination.
795 A single prediction for each subject is obtained using the most frequent prediction for
796 classification problems (predicting a class output); for regression problem (predicting a
797 continuous quantity output), the average predicted output is estimated. The main advantage
798 of random forest is achieving a substantial reduction in variance by using only a subset of the
799 predictors at each split in which decision trees are built. On the contrary, decision trees built
800 using all predictors can be very similar when using all predictors since they are built with
801 strong predictors in the first split. Averaging highly correlation tree-based results dose not
802 lead to large reduction in variance.

803 **Extreme gradient boosted tree**

804 Extreme gradient boost (XGB) is an extension of tree-based decision that builds decision trees
805 sequentially intending to reduce the remaining error from previous trees. Thus, each
806 subsequent tree is updated based on the estimated errors of the previous one. This algorithm
807 first builds a simple tree that typically provides low predictive performance, called weak
808 learners. However, XGB identifies an optimal split differently to a simple tree-based
809 classification with regularity parameter, lambda (λ). The optimal split is defined as the one
810 that achieves the highest improvement compared with the previous split. In addition, XGB
811 uses hyperparameter, gamma (γ), in pruning trees. Trees are pruned when the improvement
812 score is less than γ . As a result, the higher the γ , the more tree is pruned, giving a tree with
813 more bias and less variance.

814 **Support vector machine**

815 Support vector machine (SVM) is a machine learning algorithm that uses hyperplanes to
816 classify data points, minimising classification errors and maximising the distance between
817 classes. It is also improved to allow non-linear classifiers using a Kernel function to find support
818 vector classifiers in high dimensions. A kernel function is used to calculate the relationship
819 between each pair of observed data points as if they are in high dimensions without data
820 transformation to high dimensional values (reducing computational power). SVM then finds
821 the optimal support vector classifiers in high dimensions. In this thesis, I used the most
822 common SVM, the one with radial basis function kernel, also known as the Gaussian kernel.
823 The output value of the radial kernel is a relationship between two observations in infinite
824 dimensions and SVM then identifies the optimal hyperplane in the infinite dimensions.

825 **Artificial neural networks**

826 Artificial neural network (ANN) is a subset of machine learning algorithms inspired by the
827 biological neural network that transfers and processes information. ANN comprises several
828 interconnected artificial neurons, known as nodes or process elements, which are structured
829 in at least three layers: input, hidden and output layers. Each neuron in one layer connects to
830 all neurons in another layer. The strength of connections between neurons is weighted based
831 on the performance in the training process. Artificial neural networks are fitted using
832 feedforward with the backpropagation algorithms. In this thesis, because there was a limited
833 number of predictors and the size of the dataset was not extensive (with approximately 2,000
834 samples in the Vietnamese dataset and 500 samples in the Thai datasets), I considered
835 constructing up to three hidden layers.

836 3.3.3 Hyperparameter optimisation

837 Hyperparameter optimisation is essential in the development of prediction models with
838 machine learning approaches to increase predictive power and avoid overfitting thus
839 increasing the generalisability of models. Three main approaches of hyperparameters
840 optimisations have been commonly used: Cartesian grid search, random grid search and

purposive grid search. In this study, the purposive grid search was used with Bayesian Optimisation¹⁴². This method is an algorithm that employs a probabilistic model to optimise hyperparameters used in machine learning algorithms to improve the predictive performance of prediction models. There are two main steps applied in Bayesian Optimisation: (1) determining values of the objective function (i.e., logarithmic loss function for binary classification) with a Bayesian statistical model and (2) identifying the hyperparameters that need to be sampled with an acquisition function (i.e., Gaussian Upper Confidence bound). To determine the objective function, a Gaussian process is commonly used due to flexibility and tractability. The Gaussian process is started with prior distributions drawn at random from five prior probability distribution and provides the posterior distributions of the selected hyperparameters. The acquisition function, which guides the exploration of the parameter space during Bayesian optimisation, was used to identify the set of hyperparameters; their lower and upper bounds are presented in in Table 3.4.

Table 3.3. Table of the lower and upper bounds of hyperparameters used in each machine learning algorithms

Algorithm	Hyperparameters	Range
Random forest	the number of predictors	2 to the maximum number of predictors
	the least number of samples in each leaf	1 to 50
XGB	number of trees	10 to 1000
	the maximum number of splits in each tree	1 to 20
	learning rate	0.01 to 0.1
	gamma	0.0001 to 0.1
	the least number of samples in each leaf	1 to 50
	proportion of used predictors	0.5 to 1
	proportion of used predictors	0.5 to 1
SVM	cost	0.1 to 10
	gamma	0.01 to 1

Algorithm	Hyperparameters	Range
ANN-1	epochs	100 to 1000
	hidden neuron_1	1 to the max number of predictors
	l1	0.000001 to 0.1
	Rho	0.9 to 0.99
	epsilon	0.000001 to 0.0001
	hidden drop	0 to 0.5
ANN-2	epochs	100 to 1000
	hidden neuron_1	1 to the max number of predictors
	hidden neuron_2	1 to the max number of predictors
	l1	0.000001 to 0.1
	Rho	0.9 to 0.99
	epsilon	0.000001 to 0.0001
ANN-3	epochs	100 to 1000
	hidden neuron_1	1 to the max number of predictors
	hidden neuron_2	1 to the max number of predictors
	hidden neuron_3	1 to the max number of predictors
	l1	0.000001 to 0.1
	Rho	0.9 to 0.99
	epsilon	0.000001 to 0.0001
	hidden drop	0 to 0.5

855 Logistic regression with lasso selection (LR), random forest (RF), Extreme gradience boosted tree (XGB), support
856 vector machine (SVM), and artificial neural network with 2 and 3 hidden layers (ANN-2 and ANN-3, respectively)

3.3.4 Model development

I developed six different models for predicting the risk of developing DSS (the primary endpoint) and the combined endpoint of moderate plasma leakage and/or DSS (the secondary endpoint): a logistic regression model using the warning signs identified by WHO as predictors (this was taken as the reference model), a multivariable logistic regression model with lasso selection, and four machine learning-based models. The reference model included only the WHO warning signs (i.e., presence of vomiting, mucosal bleeding, abdominal pain or tenderness, clinical fluid accumulation, lethargy or restlessness, and platelet count). The model with multivariable logistic regression included all predictors selected by lasso regression. Machine learning-based models were developed using RF, XGB, SVM, and ANN using all candidate predictors.

The conceptual framework for the data analysis is outlined in Figure 3.1. Continuous predictors were standardised for the SVM and ANN algorithms. For the logistic regression models, RF, and XGB we used continuous predictors with the real units. PLT and AST were log-transformed (with natural base) due to the right long tail distribution. To avoid overfitting, I used Bayesian Global Optimisation with Gaussian Processes in the `rBayesianOptimization` package¹⁴² to optimise the hyperparameters of the machine learning models (See Section 3.3.3). In the Bayesian optimisation, logarithmic loss and Gaussian process upper confidence bound were applied as a metric and acquisition function, respectively. The lower and upper bounds of each hyperparameter in this study are presented in Table 3.5. Machine learning models were calibrated using Platt's (sigmoid) calibration with 10-fold stratification by the outcome.

3.3.5 Model validation

Internal validation

I assessed the predictive performance using overall measurement (i.e., Brier scores), discrimination measurement (area under receiver operation curves, AUROC), calibration measurements (i.e., calibration curves, CITL, and calibration slope) and DCA (See Section

3.2.4). To robustly assess model predictive performances (i.e., Brier score, AUROC, CITL, and calibration slope), 10-fold cross-validation with 45-Block Jackknife estimation was used to estimate the average and confidence intervals (Figure 3.1).

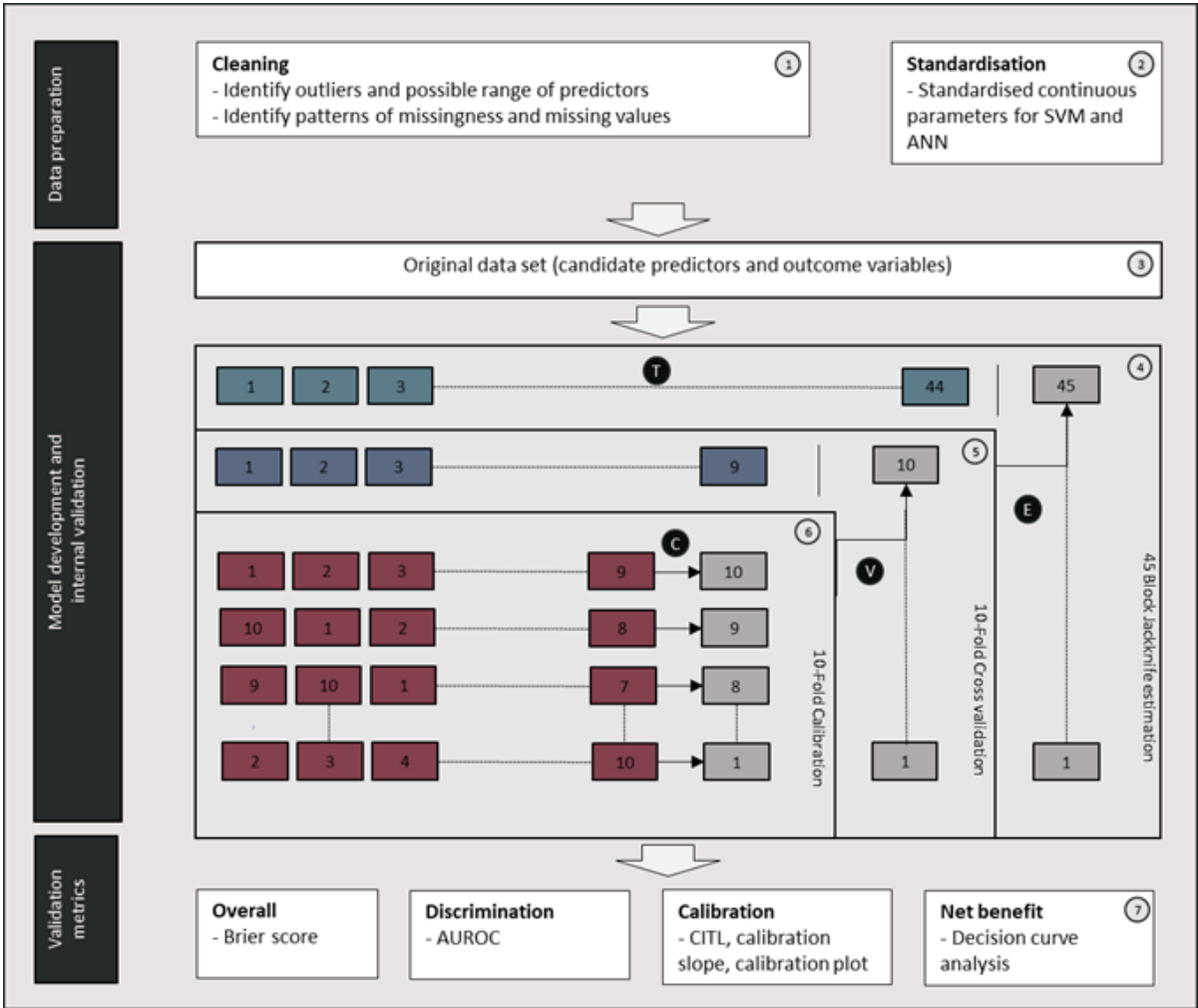


Figure 3.1. Description of the workflow used for model development and internal validation. After data preparation (Step 1 and 2), the data are used for the development of risk prediction models (Step 3). The dataset is split into 45 random blocks stratified by the outcome (Step 4). Each block (green boxes) is then used for hyperparameter optimisation (Process T). The model is trained and validated on 44 of the 45 blocks with 10-fold cross-validation (Process V), leaving one block out in turn (Step 5). The training set (blue boxes) is divided into ten folds (in Calibration loop): nine folds (red boxes) are used to develop the model and the other fold (grey boxes) is used for Platt model calibration (Process C). Outputs from the calibrated model

Specifically, the original data set was split into 45 blocks with stratification based on outcomes to ensure that all splits had the same proportions of complicated and non-complicated dengue

cases. I performed 10-fold cross-validation on a Block Jackknife data set, having left out each block in turn (i.e., repeat model training 45 times, leaving a different block out each time). In 10-fold cross-validation, the validation set was one-fold-out, in turn, using it as out-of-sample validation. The training sets (the other nine folds) were used to run 10-fold calibration to train and calibrate the models. The outputs of the Calibration loop were the combined calibration set of raw probability outputs that were used for training the calibrator models. These calibrator models were then fitted on the raw probability in a validation loop to generate calibrated probability, which was used to validate with the validation set. Calibration curves and decision curve analysis was performed on out-of-fold prediction.

External validation

The developed models were externally validated on the independent cohort data set of Thai children with a laboratory-confirmed diagnosis of dengue infection. The external validation was performed following the framework proposed by Debray et al.¹⁴³ to enhance the interpretability of results obtained from external validation (Figure 3.2). I performed the following steps: (1) I investigated the relatedness between the training and the validation datasets, (2) I assessed the model performance on the external validation dataset, and (3) we interpreted model transportability and reproducibility according to the results from steps one and two. I quantified the degree of relatedness between the training dataset and validation data set using AUROC of the membership. I also estimated a relative standard deviation of linear predictors and mean difference of linear predictors obtained between a training dataset and a validation dataset. Predictive performance on the external validation was assessed using overall measurement (i.e., Brier scores), discrimination measurement (area under receiver operation curves), and calibration measurements (i.e., calibration curves, calibration in the large, and calibration slope). Given the results of data relatedness and model performance on external validation dataset, I interpreted how well models reproduce to similar case-mix or how well models transport to a different and related different case-mix.

3.4 Sensitivity analysis

In a sensitivity analysis, I applied the risk prediction models to hospitalised cases with a combined endpoint of moderate plasma leakage and/or DSS. I used the same candidate predictors, models and performance metrics adopted in the main analysis.

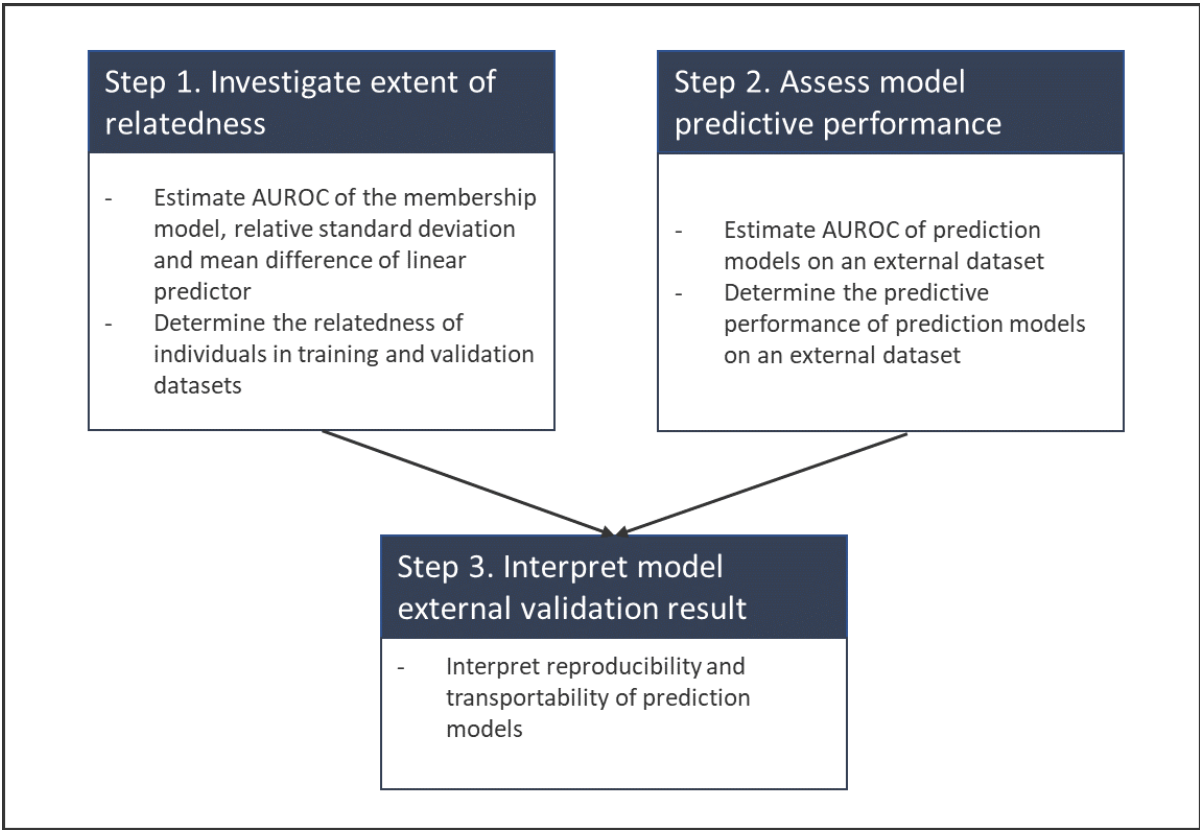


Figure 3.2. Description of the workflow used for external validation, adapted from Debray et al.¹⁴³

3.5 Methods for the analyses of the temporal dynamics of laboratory parameters

3.5.1 Analytical methods for describing temporal dynamics of clinical laboratory parameters

Longitudinal data are commonly found in biomedical studies where an outcome variable (dependent covariate) is repeatedly measured across different time points on the same subjects. Longitudinal data analysis is, therefore, used to estimate a change in the dependent variable across different time points and to determine the association of the change with independent variables or predictors. Outcome variables collected over time are not

independent of each other. This violates the assumption of regression models. Longitudinal analysis accounts for the correlated variance of the outcome variables in the same subject across different time points as seen the following matrix¹⁴⁴:

$$\begin{bmatrix} Var(Y_{i1}) & Cov(Y_{i1}, Y_{i2}) & \cdots & Cov(Y_{i1}, Y_{in}) \\ Cov(Y_{i2}, Y_{i1}) & Var(Y_{i2}) & \cdots & Cov(Y_{i2}, Y_{in}) \\ \vdots & \vdots & \ddots & \vdots \\ Cov(Y_{in}, Y_{i1}) & Cov(Y_{in}, Y_{i2}) & \cdots & Var(Y_{in}) \end{bmatrix}$$

where Y denotes an outcome variable of the ith patient across n measurements. In general, the structure of the covariance is assumed to be: exchangeable (i.e., the correlation between any two measurements is assumed to be the same), stationary (i.e., the correlation between two measurements relies on the distance between the two time points e.g., measurement at time 1 vs 3, 2 vs 4), and autoregressive (i.e., the correlation between two measurements changes with the distances). In this study, the following first-order autoregressive correlation structure was used:

$$Cov = \sigma^2 \begin{bmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{n-1} \\ \rho & 1 & \rho & \cdots & \rho^{n-2} \\ \rho^2 & \rho & 1 & \cdots & \rho^{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{n-1} & \rho^{n-2} & \rho^{n-3} & \cdots & 1 \end{bmatrix}$$

where σ^2 is the variance, and $\rho^{|j-k|}$ is an autoregressive correlation between the measurement at time j and k. This correlation structure was used according to the following reasons. Firstly, the correlation of clinical parameters tends to decrease as the lag between observations increases. Secondly, some individuals had only two measurements on each clinical parameter, thus a correlation structure depended on multiple measurements could not be estimated. Lastly, the time interval of measurements was approximately equal at 24 hours since laboratory were routinely collected every morning.

To investigate the difference in the temporal dynamics of the clinical parameters between the two groups of patients developing and not developing the endpoints (i.e., DSS as the primary clinical endpoint and the combined endpoint of moderate plasma leakage and/or DSS as the

secondary clinical endpoint), I used marginal models with generalised estimating equations (GEE). I assessed the dynamics of the average daily clinical laboratory parameters (E(PARAM)) across each group on each day of illness (DAY_ILLNESS) using four different regression models (Equation 1-4). The four models are formulated in equation 3.1 to 3.4:

Linear model unadjusted for age and sex

$$E(PARAM) = \beta_0 + \beta_1(GROUP) + \beta_2(DAY_ILLNESS) + \beta_3(GROUP)(DAY_ILLNESS) \quad (3.1)$$

Quadratic model unadjusted for age and sex

$$E(PARAM) = \beta_0 + \beta_1(GROUP) + \beta_2(DAY_ILLNESS) + \beta_3(DAY_ILLNESS)^2 + \beta_4(GROUP)(DAY_ILLNESS) \quad (3.2)$$

Linear model adjusted for age and sex

$$E(PARAM) = \beta_0 + \beta_1(AGE) + \beta_2(SEX) + \beta_3(GROUP) + \beta_4(DAY_ILLNESS) + \beta_5(GROUP)(DAY_ILLNESS) \quad (3.3)$$

Quadratic model adjusted for age and sex

$$E(PARAM) = \beta_0 + \beta_1(AGE) + \beta_2(SEX) + \beta_3(GROUP) + \beta_4(DAY_ILLNESS) + \beta_5(DAY_ILLNESS)^2 + \beta_6(GROUP)(DAY_ILLNESS) \quad (3.4)$$

Interaction terms between the group and the day of illness were used to determine the differences in the change of the clinical parameter in time between the two groups. Where the p-value of the interaction terms are statistically significant, we interpreted that the dynamics of the clinical parameter are significantly different between individuals with and without the clinical endpoint. Marginal models with GEE were implemented using geepack package¹⁴⁵ in R.

3.5.2 Predictor selection for risk prediction using daily clinical parameters

The predictors included in the model were decided according to the previous systematic review and meta-analysis (presented in Chapter 2); discussion with the three dengue experts in Thailand and Vietnam; the findings derived from the temporal dynamics of clinical parameters; and data availability (Table 3.6).

Table 3.4. Definition and type of data variables available in Thai dataset

Domain	Parameter	Definition	Type
Patient identification	patient id	Patients' identification number	categorical
	year of recruitment	Year when patients were enrolled	continuous
	date of patient enrolment	Date of patient enrolment	date
	date of data collection	Date when data was collected	date
	date of onset	Date of fever onset obtained from patients' history taking	date
	date of defervescence	Date on which 2 consecutive temperatures below 38°C	date
	day of illness	Number of days since fever onset, where the day of fever onset was counted as day 1	continuous
	day of defervescence	Number of days since defervescence, where the day of defervescence was counted as day 0	continuous
Demographics	age	Age of participant at enrolment	continuous
	sex	Physical sex	categorical
	weight	Body weight at enrolment	continuous
	Z score of weight for age at enrolment	Z score of weight for age according to the Thai national standardised weight for age	continuous

Domain	Parameter	Definition	Type
Clinical parameters at enrolment	presence of headache	“Yes” if patients had a history of headache during the febrile phase; otherwise “No”	categorical
	presence of anorexia	“Yes” if patients had a history of anorexia during the febrile phase; otherwise “No”	categorical
	presence of nausea	“Yes” if patients had a history of nausea during the febrile phase; otherwise “No”	categorical
	presence of vomiting	“Yes” if patients had a history of vomiting during the febrile phase; otherwise “No”	categorical
	presence of abdominal pain	“Yes” if patients had a history of abdominal pain during the febrile phase; otherwise “No”	categorical
	presence of drowsiness	“Yes” if patients had a history of drowsiness during the febrile phase; otherwise “No”	categorical
	presence of rash	“Yes” if patients had a history of rash during the febrile phase; otherwise “No”	categorical
	presence of itching	“Yes” if patients had a history of itching during the febrile phase; otherwise “No”	categorical
	presence of mucosal bleeding	“Yes” if patients had a history of mucosal bleeding (gastrointestinal or vaginal bleeding) which does not need invasive interventions or blood component transfusion; otherwise “No”	categorical
Daily clinical parameters	maximum body temperature	The maximum body temperature in Celsius degree	continuous
	Lowest systolic blood pressure	The lowest systolic blood pressure in mmHg	continuous

Domain	Parameter	Definition	Type
	Minimum pulse pressure	The minimum pulse pressure in mmHg	continuous
	haematocrit*	The fractional contribution of erythrocytes (red blood cells) in the blood volume daily measured in per cent unit	continuous
	platelet count*	Platelets (thrombocytes) are nonnucleated cell fragments that trigger substances for the formation of blood clots daily measured in cells/dl	continuous
	white blood cell count*	White Blood Cells (leukocytes) are nucleated cells responsible for producing a wide range of immune responses daily measured in cells/dl	continuous
	neutrophil count*	Leukocytes important in phagocytosis daily measured in cells/dl	continuous
	lymphocyte count*	Leukocytes made up of B and T lymphocytes and null cells that are important for direct and memory immune responses daily measured in cells/dl	continuous
	AST levels*	Aspartate aminotransferase, liver enzyme that catalyses the transfer of the amino group of aspartates to α -keto-glutamate to produce oxaloacetate and glutamate daily measured in units/l	continuous
	ALT levels*	Alanine aminotransferase, liver enzyme that catalyses the transfer of the amino group of alanine to α -keto-glutamate to produce oxaloacetate and glutamate daily measured in units/l	continuous
	serum albumin levels*	The main protein component in blood and is has a major role in carrying of free fatty acids in blood daily measured in g/l	continuous

Domain	Parameter	Definition	Type
Diagnostic laboratory parameters	Infecting dengue serotypes	Dengue serotypes detected using rt-PCR testing on blood specimen collected during the first 72 hours of illness	categorical
	Immune status	Serological definition relied upon IgM and IgG ELISA test in the febrile phase and convalescent phase; Primary infection: two negative dengue-specific IgG on febrile and convalescence samples; secondary infection: at least one positive dengue-specific IgG on febrile and convalescence samples; Indeterminate: single specimen (a febrile sample or a convalescent sample) available	categorical
Clinical progression parameters	Evidence of fluid accumulation	"Yes" if a patient develops sign of either pleural effusion or ascites during illness; otherwise "No"	categorical
	Developing haemoconcentration	"Yes" if a patient develops 15% haematocrit rising from baseline haematocrit (defined as haematocrit levels measured during the first 72 hours of illness or average haematocrit levels based on population data); otherwise "No"	categorical
	Presence of vital sign instability (shock)	"Yes" if systolic blood pressure falls below the normal range by age, pulse pressure less than 20 mmHg, or undetected blood pressure: otherwise "No"	categorical
Final outcomes	Developing a combined endpoint of moderate plasma leakage and/or dengue shock syndrome	"Yes" if a hospitalised patient develops a combined endpoint of moderate plasma leakage or dengue shock; otherwise "No"	categorical
	Developing dengue shock syndrome	"Yes" if a hospitalised patient develops dengue shock syndrome	categorical

*Clinical parameters were included in statistical tests for comparison temporal dynamics between the complicated and non-complicated group

3.5.3 Analytical methods used for risk prediction using temporal dynamics of clinical parameters

In Chapter 5, I developed risk prediction models for DSS and the combined endpoint of moderate plasma leakage and/or DSS using the temporal dynamics of laboratory parameters and a variation of recurrent neural networks (RNNs). I applied three recurrent neural network-based approaches: simple (vanilla) RNN, gate recurrent unit (GRU), and long-short term memory (LSTM). The performance of these models was compared with the performance of the models developed on a single clinical parameter measurement at enrolment (described in Section 3.3.2). The details of three RNN based approaches are described below.

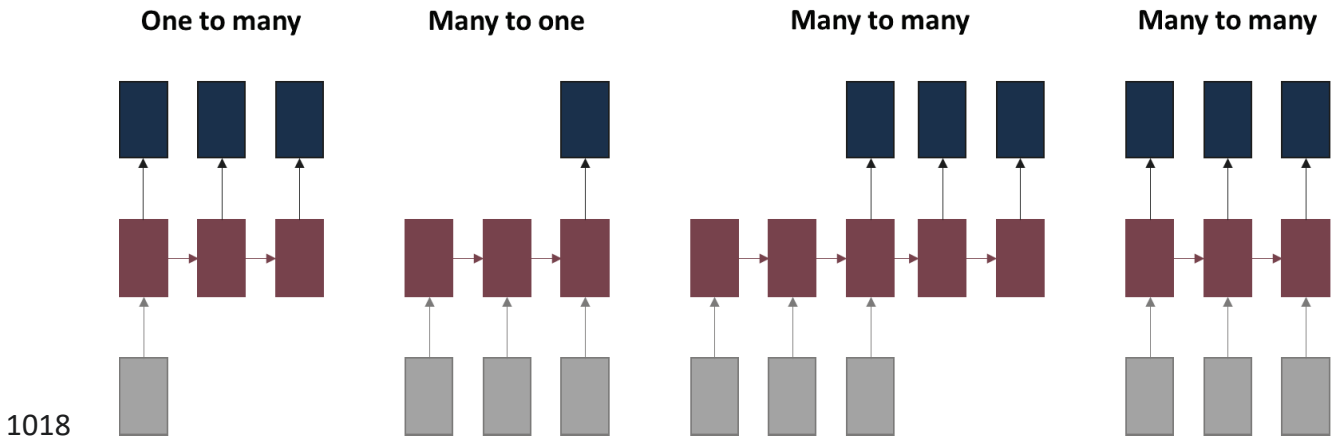
Recurrent neural networks

A recurrent neural network (RNN) is a type of neural network that allows employing either sequential measured input or output data. This method can be classified based on input and output structure: a sequence output, a sequence input, a sequence output and input, and synched sequence input and output RNNs (Figure 3.3). In this thesis, sequence input RNNs were used since daily measured clinical parameters were used as predictors and the risk of an endpoint over an illness was the single prediction outcome. With sequence input RNNs, previous input (h^{tn-1} , aka a hidden activation vector) and current input (x^{tn} , aka an embedding vector) in the sequence are aggregated over time to neural networks and estimate an output (y^{tn} , aka a context vector), as seen in Figure 3.4 and in equation 3.5¹⁴⁶.

$$h^{tn} = \tanh (W^{hh}h^{tn-1} + W^{xh}x^{tn}) \quad (3.5)$$

where W is a weight parameter updated with gradient in back propagation through time. When the final output is compared with an observed outcome, the error is calculated and used in back-propagation through time to adjust the weight of neural networks at each stage. As a result, sequence input RNNs give a single output (i.e., a predicted risk of the combined endpoint on a day of defervescence) based on a sequence of input data (i.e., daily clinical parameters). However, RNNs have a major drawback called vanishing or exploding gradients. During the backpropagation through time, the weights of the network are adjusted using

1015 multiplication. As a result, the further backpropagate moves backwards, the larger or smaller
 1016 weight is adjusted. This means a simple RNN estimates an output mainly on the most recent
 1017 sequence data.



1019 **Figure 3.3.** Different types of RNN by input and output structure adapted from Karpathy¹⁴⁷. Each squares and
 1020 arrow depict a vector and function respectively: input vectors (grey), RNN's state (red), and output vectors (blue).
 1021 From the left to right, (1) a sequence output, (2) a sequence input, (3) a sequence output and input, and (4)
 1022 synced sequence input and output RNNs.

1023 Long short-term memory

1024 Long short-term memory network (LSTM) is a special type of RNN designed to employ long-
 1025 term dependencies, solving vanishing and exploding gradience introduced by Hochreiter and
 1026 Schmidhumber¹⁴⁸. In addition to simple RNNs, LSTMs include three types of gate units to
 1027 manage information passed to neural networks: a forget-gate unit (f^{tn}), an input-gate unit
 1028 (i^{tn}), and an output gate unit (o^{tn}). A simplified flow chart of LSTMs is presented in Figure 3.5.

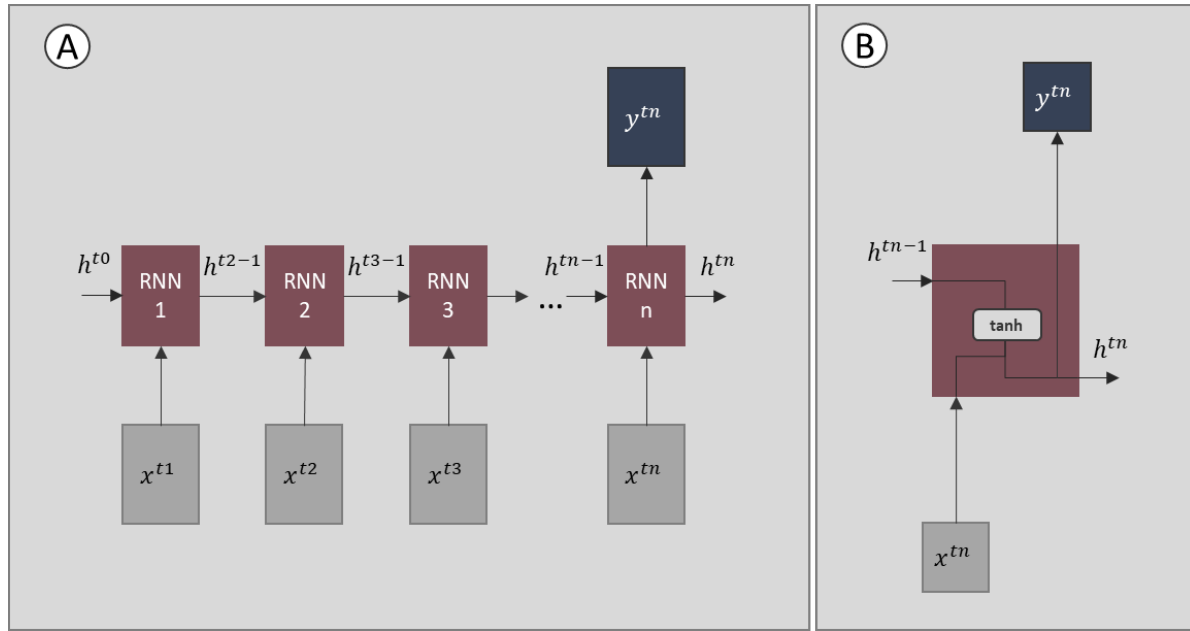


Figure 3.4. Forward propagation for a basic RNN model (A) and RNN unit (B), adapted from Olah¹⁴⁹

LSTMs use both context vectors (c^{tn-1}, c^{tn}) and hidden activation vectors (h^{tn-1}, h^{tn}) in algorithms. LSTMs first decide which information obtained from a previous cell will be ignored using forget-gate units. Then input-gate units are used to select which information obtained from the previous cell will be used in a current cell and passed to the next cell. Lastly, update gate units are used to estimate output of the current cell. The three types of gate units are formulated in equation 3.6 to 3.11.

$$f^{tn} = \sigma(W^f \cdot [h^{tn-1}, x^{tn}] + b^f) \quad (3.6)$$

$$i^{tn} = \sigma(W^i \cdot [h^{tn-1}, x^{tn}] + b^i) \quad (3.7)$$

$$\check{c}^{tn} = \tanh(W^c \cdot [h^{tn-1}, x^{tn}] + b^c) \quad (3.8)$$

$$o^{tn} = \sigma(W^o \cdot [h^{tn-1}, x^{tn}] + b^o) \quad (3.9)$$

$$h^{tn} = o^{tn} \cdot \tanh(c^{tn}) \quad (3.10)$$

$$c^{tn} = f^{tn} \cdot c^{tn-1} + i^{tn} \cdot \check{c}^{tn} \quad (3.11)$$

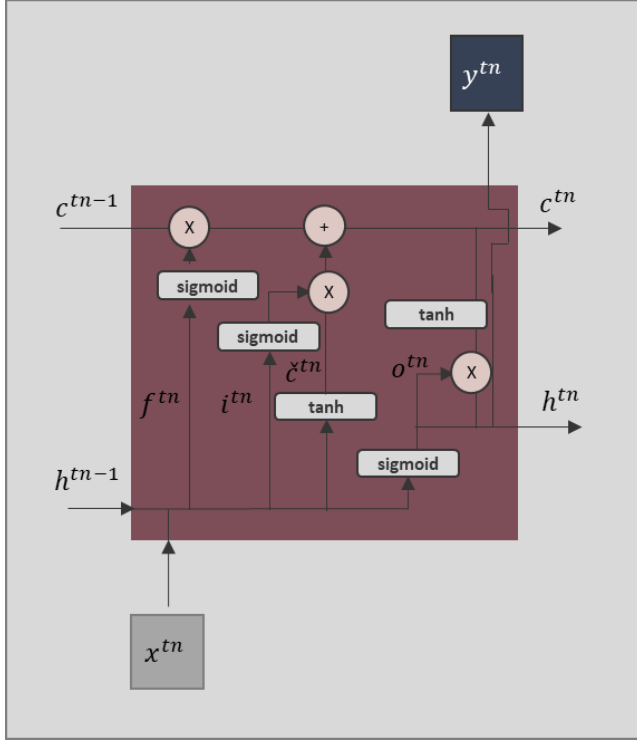


Figure 3.5. LSTM basic structure, adapted from Olah¹⁴⁹

Gated Recurrent Units

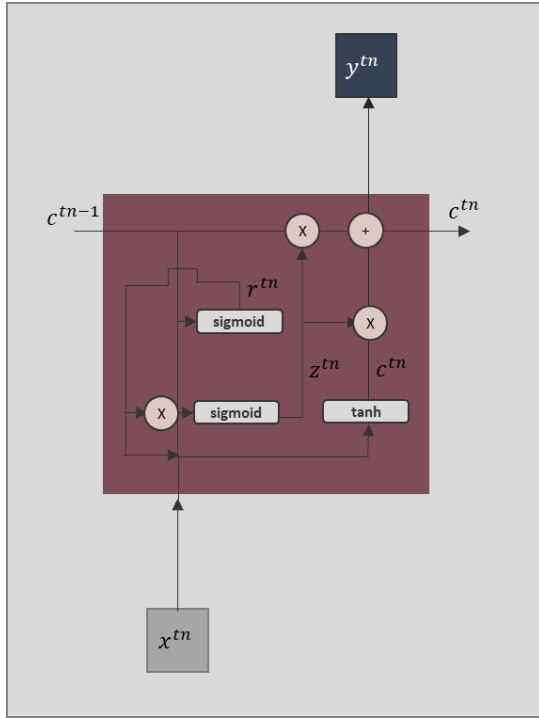
Gated Recurrent Units (GRUs) are a variety of long short-term memory networks that combines forget and input gate units to a single update gate unit. GRUs also merges context vectors (c^{tn-1}, c^{tn}) and hidden activation vectors (h^{tn-1}, h^{tn}) together and run through update gate and reset gate functions resulting model is simpler than LSTMs. These functions are added to prevent a difficulty of vanishing gradient or exploded gradient¹⁵⁰. They are instructed to control whether previous input data (ht-1, aka hidden activation vectors) should be included or excluded in RNNs (Figure 3.6). An update gate (z_{tn}) determines how much input data from previous sequence should be passed to RNNs. A reset gate (r_{tn}) determines how much input data from the previous sequence should be diminished. In other words, RNNs determine memory content, which is information obtained from previous sequence and which will be used in neural networks c_{tn} . With an update gate, RNNs then determine information that will be passed to the next RNNs as an output of a current RNN (y_{tn})

$$z_{tn} = \sigma(W^z \cdot [c^{tn-1}, x^t]) \quad (3.12)$$

1060
$$r_{tn} = \sigma(W^r \cdot [c^{tn-1}, x^t]) \quad (3.13)$$

1061
$$c_{tn} = (1 - z_{tn}) \cdot (W^r \cdot c^{tn-1} + z_{tn} \cdot c_{tn-1} \quad (3.14)$$

1062
$$y_{tn} = \tanh(W \cdot [r^{tn} \cdot c^{tn-1}, x^{tn}]) \quad (3.15)$$



1063
1064 **Figure 3.6.** GRU basic structure, adapted from Olah¹⁴⁹

1065 3.5.4 Combination of multiple imputation and internal validation

1066 In the chapter 5, I combined multiple imputation and internal validation following the
 1067 methods proposed by Wahl et al¹³². Multiple imputation was performed on the predictors
 1068 (excluding the outcome variables), which was followed by bootstrapping to estimate
 1069 confidence intervals. I imputed ten datasets (without outcome variables) using the mice
 1070 framework in R¹³⁰ (Figure 3.7). Each imputed dataset was then bootstrapped to ten
 1071 bootstrapped datasets with out-of-bag labels (i.e., samples were not included in each
 1072 bootstrapped dataset). Each bootstrapped dataset was used for training each risk prediction
 1073 model. Hyperparameter optimisation was performed using Bayesian optimisation (see Section
 1074 3.3.3). The developed models were validated with overall (Brier Score), discrimination

(AUROC), and calibration performances (CITL and calibration slope). Predictive performances were calculated using out-of-bag (OOB) samples (a set of samples which are not included in bootstrap datasets) and a weighted average of apparent and OOB performance with 95% CI calculated using bootstrapping empirical confidence interval.

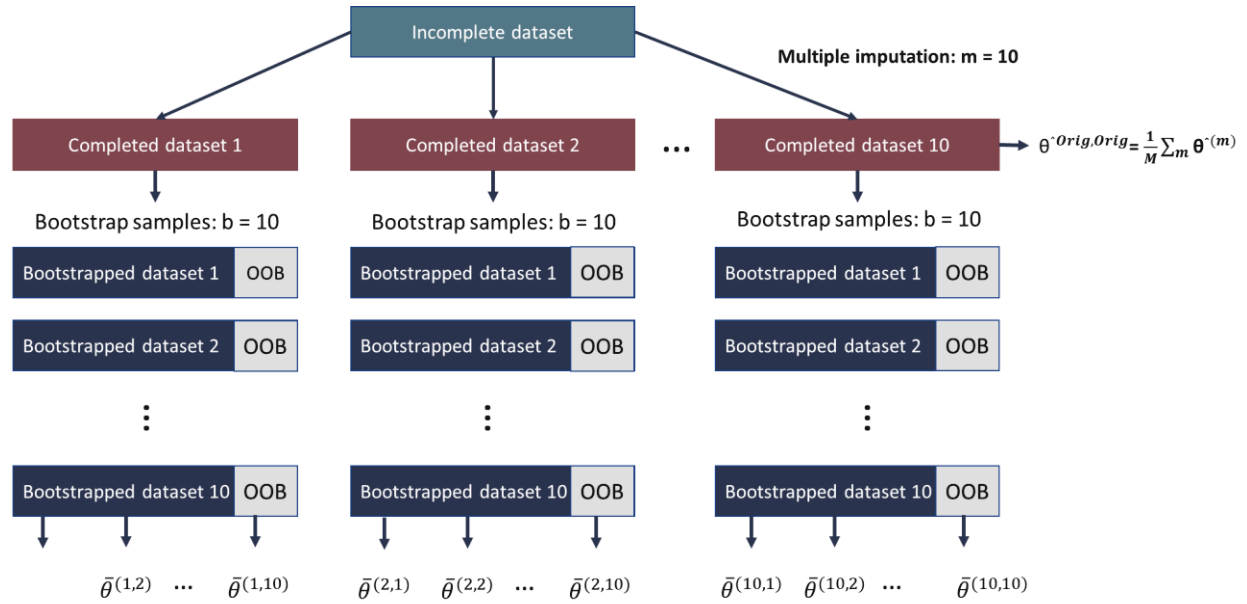


Figure 3.7. Combination of multiple imputation (m times) and internal validation using bootstrapping (b times).

I estimated the average performance of the models, denoted $\hat{\theta}^{0.632+}$. The central estimate of $\hat{\theta}^{0.632+}$ was estimated as equation 3.16 to 3.18:

$$\hat{\theta}^{0.632+} = (\hat{\theta}^{Orig,Orig} + W \cdot \hat{\theta}^{OOB}) \quad (3.16)$$

$$W = 0.632 / (1 - 0.368 \cdot R) \quad (3.17)$$

$$R = (\hat{\theta}^{OOB} - \hat{\theta}^{Orig,Orig}) / (\hat{\theta}^{non-info} - \hat{\theta}^{Orig,Orig}) \quad (3.18)$$

Where $\hat{\theta}^{Orig,Orig}$ is calculated as the average over the 10 complete imputed datasets; $\hat{\theta}^{OOB}$ is calculated as the average over B·M $\hat{\theta}^{BS(b,m),OOB}$ (performance on the OOB samples having trained the model on each $BS^{(b,m)}$ set); and $\hat{\theta}^{non-info}$ is calculated as the average performance having randomly generated the prediction outcome. The 95% bootstrapping empirical confidence interval was estimated as equation 3.19.

$$1091 \quad (\hat{\theta}^{0.632+} - \xi_{1-\alpha/2}), (\hat{\theta}^{0.632+} - \xi_{1-\alpha}) \quad (3.19)$$

1092 Where $\xi_{1-\alpha}$ and $\xi_{1-\alpha/2}$ was obtained from the 2.5% and 97.5% percentiles of $W_{b,m}$ (over all b
1093 and m, i.e. B x M values). Matrix $W_{b,m}$ was calculated as equation 3.20.

$$1094 \quad W_{b,m} = \hat{\theta}^{BS(b,m),BS(b,m)} - \hat{\theta}^{Orig,Orig} \quad (3.20)$$

Chapter 4 Risk prediction for progression to severe dengue illness using clinical parameters measured during the first 72 hours of illness

4.1 Rationale

Dengue causes a high burden on public health systems worldwide. A study in 2013 estimated that, globally, approximately 50% of patients with dengue are evaluated and treated in outpatient settings, 30% are not reported or do not have access to healthcare systems, and the rest 20% are admitted and treated in inpatient settings⁴. However, approximately 50% of the total costs of global dengue care were spent in inpatient settings and 30% were used in outpatient settings⁴.

Since the majority of infections are mild or moderate illness which are self-resolving, while only a minority develop severe clinical illness (typically on day four to six of illness). Most patients are first evaluated and receive care mainly in outpatient settings, particularly during the febrile phase^{5, 21}. In such settings, there is also an opportunity in the first three to four days of illness to diagnose and recognise patients at a high risk of developing severe dengue illness. The World Health Organisation (WHO), in collaboration with the Special Programme for Research and Training in Tropical Diseases (TDR), promotes the use of warning signs to assist the recognition of patients in the febrile phase who have a higher risk of developing severe dengue⁵. The warning sign checklist shows high sensitivity and moderate specificity, potentially causing a large number of unnecessary admissions^{28, 29}. Notably, the use of this checklist does not provide an individualised prediction for the risk of developing severe dengue and these criteria are often used inconsistently in different settings³⁰.

To improve patient triage during the early febrile phase of dengue, clinical prediction models can play a key role in estimating individualised risk prediction of developing severe dengue illness. Although several prediction models for dengue severity have been developed, only a few of them have been implemented into clinical practice. This is likely due to the limited

specificity of these models, heterogeneity in the surveyed populations, as well as differences in the availability of predictors and variability of the definitions. In addition, most models were developed using data from hospitalised patients while most patients during the early phase of illness are evaluated in outpatient settings. The aims of this chapter are: (1) to identify predictors of severe dengue illness detected during the febrile phase and (2) to develop risk prediction models using a variety of statistical and machine learning approaches, to provide robust early risk prediction of progression to moderate plasma leakage or severe dengue during the first 72 hours of illness for individuals in an outpatient context.

4.2 Methods

This study was approved by the scientific and ethical committees of collaborating hospitals and the Oxford University Tropical Research Ethical Committee (NCT01421732) (Appendix 16) and The Research Ethics Review Committee of QSNICH (REC.082/2562) (Appendix 17). This chapter reports the development of prediction models following the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD)¹⁵¹.

This chapter presents the characteristics of patients included in the development and validation of models based on clinical parameters measured at enrolment. I investigated and identified predictors associated with the two clinical endpoints (i.e., DSS and the combined endpoint of moderate plasma leakage and/or DSS). In addition, I developed and evaluated the performance of the clinical risk prediction models using the Vietnamese and Thai datasets. The details of the datasets and analytical methods are described in Chapter 3.

4.3 Results

4.3.1 Characteristics and clinical baseline of patients

In the Vietnamese dataset, 8,100 patients were enrolled, of which 2,245 were laboratory-confirmed dengue (Figure 4.1A); 1,226 patients were managed as outpatients (non-hospitalised patients); and 1,019 patients were hospitalised. Of the outpatients, 1,185 (96.66%) completed follow-up until recovery and 41 (3.34%) did not complete follow up.

Amongst the hospitalised patients, 110 developed DSS and 185 developed moderate plasma leakage (Figure 4.1A). In the Thai dataset, 1,210 patients were enrolled in the study and 524 were laboratory confirmed dengue; amongst 218 hospitalised cases, 182 children developed moderate plasma leakage and 36 DSS alone (Figure 4.1B).

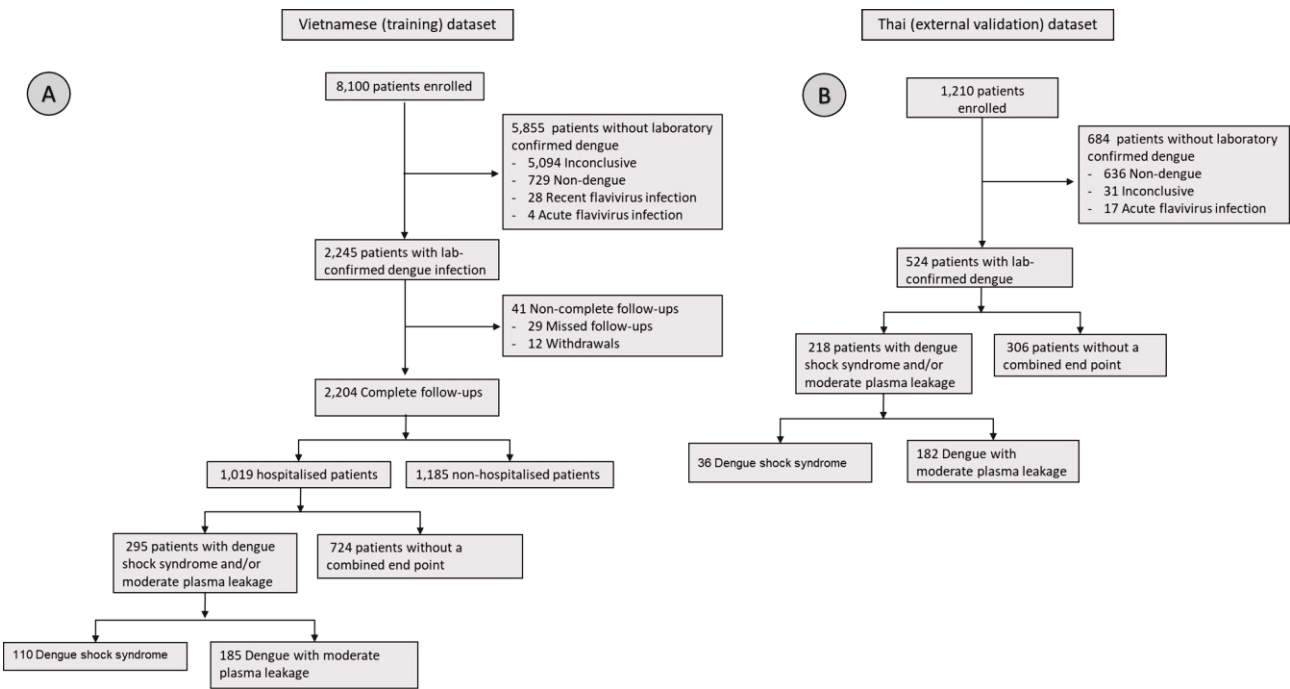


Figure 4.1. Flow chart describing patients' outcomes in the Vietnamese training (A) and Thai external validation datasets (B)

The characteristics of the patients with complete data on the day of enrolment is shown in Table 4.1 for the Vietnamese dataset and in Table 4.2 for the Thai dataset. In both datasets, the mean age was 8 to 9 years, and we observed the same proportion of male patients (approximately 55%). Patients were enrolled from day 2 to 4 of illness in both datasets. In the Vietnamese dataset, history of vomiting (p-value= 0.012 and <0.001) and abdominal pain or tenderness (p-value= 0.009 and <0.001) were clinical signs and symptoms associated with both DSS and the combined endpoint of moderate plasma leakage and/or DSS, while mucosal bleeding was associated with the combined endpoint (p-value= 0.040) but not with DSS (p-value= 0.427) (Table 4.1). In the Thai dataset, we observed no statistically significant difference in the clinical signs and symptoms of children with DSS and with the combined outcome (Table 4.2). Higher AST and lower platelet counts were more commonly found in

children with DSS and the combined endpoint in both datasets. In the Vietnamese dataset, significantly higher haematocrit levels (p-value= <0.001 and <0.001) and lymphocyte count (p-value= 0.017 and <0.001), and lower serum albumin levels (p-value= 0.003 and <0.001) were associated with DSS and the combined endpoint of moderate plasma leakage and/or DSS (Table 4.1). In the Thai dataset, we did not find significant association between haematocrit or serum albumin measured in the first 72 after fever onset and DSS or the combined endpoint (Table 4.2). Secondary infection was associated with the development of DSS and the combined endpoint in both the Vietnamese (p-value= <0.001 and <0.001) and Thai datasets (p-value= 0.015 and <0.001). In the Vietnamese dataset, we found significant differences in the frequency of the infecting dengue serotypes among children with and without the combined endpoint (p-value= <0.001) but not amongst children with and without DSS (p-value= 0.078) (Table 4.1). We did not observe significant associations between the serotype of infection and both endpoints in the Thai dataset (p-value= 0.878 and 0.330) (Table 4.2).

Table 4.1. Characteristics and clinical variables at enrolment of the Vietnamese patients in complete dataset with laboratory confirmed dengue infection developing (= Yes) Dengue Shock Syndrome and the combined endpoint of moderate plasma leakage and/or DSS.

Parameters	Overall	Dengue Shock Syndrome (DSS)			Combined endpoint: moderate plasma leakage and/or DSS		
		No	Yes	p-value	No	Yes	p-value
Subjects	2191	2084	107		1900	291	
Day of illness: mean (SD)	2.85 (0.79)	2.84 (0.79)	3.07 (0.75)	0.002	2.84 (0.79)	2.92 (0.78)	0.09
Age (years): mean (SD)	8.89 (3.41)	8.89 (3.44)	8.92 (2.79)	0.91	8.85 (3.44)	9.12 (3.25)	0.22
Sex = Male (%)	1230 (56.1)	1174 (56.3)	56 (52.3)	0.476	1064 (56.0)	166 (57.0)	0.75
Overweight (%)	281 (12.8)	262 (12.6)	19 (17.8)	0.157	246 (12.9)	35 (12.0)	0.71
Vomiting Hx (%)	902 (41.2)	845 (40.5)	57 (53.3)	0.012	750 (39.5)	152 (52.2)	<0.001

Parameters	Overall	Dengue Shock Syndrome (DSS)			Combined endpoint: moderate plasma leakage and/or DSS		
		No	Yes	p-value	No	Yes	p-value
Mucosal bleeding PE (%)	99 (4.5)	92 (4.4)	7 (6.5)	0.427	77 (4.1)	22 (7.6)	0.04
Abdominal tenderness PE (%)	63 (2.9)	55 (2.6)	8 (7.5)	0.009	46 (2.4)	17 (5.8)	<0.001
HCT (%): mean (SD)	38.77 (3.39)	38.69 (3.35)	40.25 (3.71)	<0.001	38.86 (3.32)	38.15 (3.76)	<0.001
PLT (1000 cell/dL): mean (SD)	186.16 (69.50)	188.81 (69.08)	134.69 (56.48)	<0.001	191.14 (69.01)	153.64 (63.71)	<0.001
WBC (1000 cell/dL): mean (SD)	5.64 (3.08)	5.66 (3.09)	5.20 (2.73)	0.128	5.69 (3.11)	5.33 (2.81)	0.06
NC (1000 cell/dL): mean (SD)	3.73 (2.64)	3.74 (2.64)	3.57 (2.53)	0.507	3.73 (2.66)	3.77 (2.50)	0.82
LC (1000 cell/dL): mean (SD)	1.30 (0.91)	1.31 (0.91)	1.10 (0.66)	0.017	1.35 (0.94)	1.02 (0.57)	<0.001
AST (units/dL): mean (SD)	62.73 (49.02)	60.31 (44.12)	109.79 (95.07)	<0.001	59.37 (42.81)	84.64 (74.75)	<0.001
ALT (units/dL): mean (SD)	32.14 (37.31)	30.74 (34.37)	59.31 (68.91)	<0.001	30.31 (34.40)	44.04 (50.97)	<0.001
ALB (g/L): mean (SD)	43.73 (3.12)	43.77 (3.08)	42.87 (3.74)	0.003	43.87 (3.03)	42.78 (3.51)	<0.001
Immune status (%)				<0.001			<0.001
indeterminate	1193 (53.1)	1144 (54.9)	4 (3.7)		1138 (59.6)	16 (5.4)	
primary infection	302 (13.5)	292 (14.0)	7 (6.5)		265 (13.9)	35 (11.9)	
secondary infection	750 (33.4)	648 (31.1)	96 (89.7)		505 (26.5)	244 (82.7)	
Serotype (%)				0.078			<0.001
1	835 (40.4)	779 (37.4)	36 (33.6)		720 (41.2)	99 (35.1)	

Parameters	Overall	Dengue Shock Syndrome (DSS)			Combined endpoint: moderate plasma leakage and/or DSS		
		No	Yes	p-value	No	Yes	p-value
2	456 (22.1)	412 (19.8)	31 (29.0)		363 (20.8)	82 (29.1)	
3	197 (9.5)	185 (8.9)	6 (5.6)		173 (9.9)	19 (6.7)	
4	579 (28.0)	542 (26.0)	30 (28.0)		490 (28.1)	82 (29.1)	
negative	170 (7.8)	166 (8.0)	4 (3.7)		157 (8.3)	13 (4.5)	

1181

1182 **Table 4.2.** Characteristics and clinical variables at enrolment of the Thai patients in complete dataset with
1183 laboratory confirmed dengue infection developing (= Yes) Dengue Shock Syndrome and the combined endpoint
1184 of moderate plasma leakage and/or DSS.

Parameters	Overall	Dengue Shock Syndrome (DSS)			Combined endpoint: moderate plasma leakage and/or DSS		
		No	Yes	p-value	No	Yes	p-value
Subjects	446	417	28	-	262	184	-
Day of illness: mean (SD)	3.13 (0.83)	3.11 (0.83)	3.46 (0.69)	0.03	3.11 (0.83)	3.16 (0.84)	0.51
Age (years): mean (SD)	8.59 (2.99)	8.62 (3.02)	8.14 (2.52)	0.412	8.60 (2.89)	8.57 (3.13)	0.94
Sex = Male (%)	240 (53.8)	228 (54.7)	11 (39.3)	0.166	140 (53.4)	100 (54.3)	0.93
Overweight (%)	74 (16.6)	69 (16.5)	4 (14.3)	0.961	46 (17.6)	28 (15.2)	0.60
Vomiting Hx (%)	305 (68.4)	286 (68.6)	19 (67.9)	1	180 (68.7)	125 (67.9)	0.95
Mucosal bleeding PE (%)	48 (10.8)	44 (10.6)	4 (14.3)	0.763	26 (9.9)	22 (12.0)	0.60

Parameters	Overall	Dengue Shock Syndrome (DSS)			Combined endpoint: moderate plasma leakage and/or DSS		
		No	Yes	p-value	No	Yes	p-value
Abdominal tenderness PE (%)	142 (31.8)	131 (31.4)	10 (35.7)	0.792	79 (30.2)	63 (34.2)	0.42
HCT (%): mean (SD)	38.14 (3.58)	38.20 (3.57)	37.11 (3.68)	0.117	38.45 (3.05)	37.69 (4.19)	0.03
PLT (1000 cell/dL): mean (SD)	193.14 (78.01)	195.70 (77.46)	151.23 (74.24)	0.003	204.51 (77.38)	176.94 (76.23)	<0.001
WBC (1000 cell/dL): mean (SD)	5.14 (3.55)	5.16 (3.63)	4.61 (2.00)	0.43	5.21 (3.41)	5.03 (3.76)	0.61
NC (1000 cell/dL): mean (SD)	3.67 (2.67)	3.69 (2.72)	3.31 (1.82)	0.475	3.70 (2.86)	3.63 (2.38)	0.77
LC (1000 cell/dL): mean (SD)	1.04 (1.38)	1.04 (1.40)	0.97 (1.06)	0.789	1.04 (0.96)	1.03 (1.83)	0.91
AST (units/dL): mean (SD)	61.95 (64.51)	59.81 (63.20)	94.93 (76.30)	0.005	52.76 (33.90)	75.03 (90.49)	<0.001
ALT (units/dL): mean (SD)	33.59 (54.54)	32.58 (54.95)	48.79 (47.42)	0.129	28.85 (26.67)	40.34 (78.37)	0.03
ALB (g/L): mean (SD)	4.61 (0.52)	4.61 (0.53)	4.51 (0.44)	0.321	4.63 (0.52)	4.57 (0.53)	0.27
Immune status (%)				<0.001			<0.001
indeterminate	4 (0.9)	3(0.7	0(0.0)	)	3 (1.1)	1 (0.5)	
primary infection	95 (21.3)	95 (21.3)	0(0.0)		70 (26.7)	25 (13.6)	
secondary infection	347 (77.8)	319(76.5)	28(100)		189 (72.1)	158 (85.9)	
Serotype (%)				0.878			0.33
1	141 (36.1)	133(36.1)	7(31.8)		91 (38.7)	50 (32.1)	

Parameters	Overall	Dengue Shock Syndrome (DSS)			Combined endpoint: moderate plasma leakage and/or DSS		
		No	Yes	p-value	No	Yes	p-value
	2	77 (19.7)	72(19.6)	5(22.7)	43 (18.3)	34 (21.8)	
	3	108 (27.6)	101(27.4)	7(31.8)	58 (24.7)	50 (32.1)	
	4	51 (13.0)	48(13)	3(13.6)	34 (14.5)	17 (10.9)	
	negative	14 (3.6)	14(3.8)	0(0.0)	9 (3.8)	5 (3.2)	

1185 4.3.2 Selected predictors

1186 Of the total 22 clinical parameters, 21 clinical parameters were collected and measured during
1187 the first 72 hours of illness. The other parameter, representing the immune status, was
1188 defined using either laboratory parameters collected at enrolment or in the convalescent
1189 phase (See section 3.3.1). Based on our previous systematic review and meta-analysis (See
1190 Chapter 2)¹¹⁶, and discussion with dengue experts, 11 candidate predictors were selected,
1191 comprising demographic information (i.e., age and overweight), signs and symptoms (i.e.,
1192 vomiting, abdominal pain, skin haemorrhage, and mucosal bleeding), and laboratory
1193 parameters (i.e., PLT, WBC, LC, AST, and serum albumin) at recruitment (detected within the
1194 first 72 hours of clinical onset) (Table 4.3). HCT and haemoglobin were excluded from the
1195 predictors of the combined endpoint because they are used as a part of the outcome
1196 classification criteria but not for DSS. Neutrophile count and ALT were also excluded due to
1197 high collinearity with white blood cell count and AST, respectively. The infecting serotype and
1198 immune status were excluded since a test for either parameter is not currently available in
1199 clinical practice. I performed variable selection following the recommended¹³³ number of
1200 events per variable amounting to 8 (107 events/12 variables) and 26 (291 events/11 variables)
1201 for the models predicting the combined endpoint (moderate plasma leakage and/or DSS) and
1202 DSS, respectively.

1203 **Table 4.3.** Considered candidate predictors of Dengue Shock Syndrome (DSS) and the combined endpoint of
 1204 moderate plasma leakage and/or DSS in the Vietnamese dataset. Predictors included in each analysis are noted
 1205 by +. Numbers denote the variables included in each analysis.

No.	Variable	Meta-analysis ¹¹⁶	Expert opinion*	Candidate predictors	Lasso selection for DSS	Lasso selection for the combine endpoint
1	Age	+	+	+ (1)	-	-
2	Sex = male	-	-	-	-	-
3	Overweight	-	+	+ (2)	-	-
4	Headache	-	-	-	-	-
5	Anorexia	-	-	-	-	-
6	Nausea	-	-	-	-	-
7	Vomiting	+	+	+ (3)	-	+(1)
8	Abdominal pain	+	+	+ (4)	-	+(2)
9	Skin haemorrhage	-	+	+ (5)	-	-
10	Presence of Rash	-	-	-	-	-
11	Mucosal bleeding	+	+	+ (6)	-	+(3)
12	Haemoglobin	-	+	- (High correlation with haematocrit)	-	-
13	Haematocrit	-	+	+ (7) *	+(1)	-*
14	Platelet count	+	+	+ (8)	+(2)	+(4)
15	White blood cell count	-	+	+ (9)	-	+(5)
16	Neutrophil count	-	+	- (High correlation with white	-	-

No.	Variable	Meta-analysis ¹¹⁶	Expert opinion*	Candidate predictors	Lasso selection for DSS	Lasso selection for the combine endpoint
				blood cell count)		
17	Lymphocyte counts	-	+	+ (10)	+(3)	+(6)
18	ALT	+	+	- (High correlation with AST)	-	-
19	AST	+	+	+ (11)	+(4)	+(7)
20	Serum albumin	+	+	+ (12)	-	+(8)
21	Serotypes	+ (DENV2)	-	- (no POCT availability)	-	-
22	Immune status	+	+	- (no POCT availability)	-	-

*Haematocrit was not considered among the candidate predictors of the combine endpoints because it was part of the outcome definition; AST: aspartate aminotransferase; ALT alanine aminotransferase; DENV: dengue virus; POCT: point of care tests.

With lasso selection and bootstrapping, I quantified the inclusion frequency for each candidate predictors (Table 4.4). Using DSS as a prediction outcome, I found that platelet count and AST were always included as predictors, followed by lymphocyte count (68.9%) and haematocrit (60.8%). The four predictors were jointly selected with a frequency of 18% (Table 4.5). For the models predicting the combined endpoint, platelet count, lymphocyte count, and AST were selected, in combination with other variables, with a frequency of 100% (Table 4.7); each of serum albumin, white blood cell count, vomiting, mucosal bleeding, ad abdominal pain or tenderness were selected with a frequency of over 60%. Each of age, skin bleeding and obese status were selected with a frequency of less than 30% (Table 4.6). Overall, the set of predictors including platelet count, lymphocyte count, AST, serum albumin, white blood cell count, vomiting, mucosal bleeding, and abdominal pain or tenderness was jointly selected

with a frequency of 23.2% (Table 4.7). With multivariable regression and lasso selection, the final predictions for the risk of developing DSS were generated using four predictors (haematocrit, platelet count, lymphocyte count, and AST). Eight predictors (vomiting, abdominal pain or tenderness, mucosal bleeding, platelet count, white blood cell count, AST, and serum albumin) were used to generate the final predictions of the risk of developing the combined endpoint using multivariable regression with lasso selection (Table 4.3).

Table 4.4. Global model, model selected by lasso selection with the minimum lambda and bootstraps-derived quantities for assessing the uncertainty of model for DSS

Predictors	Global model (including all predictors)			Selected models by lasso on bootstrap datasets						
	Coefficients	SE	Bootstrap inclusion frequency (%)	Coefficients	SE	RMSD ratio	Relative conditional bias (%)	Bootstrap median	Bootstrap 2.5 th percentile	Bootstrap 97.5 th percentile
Intercept	-0.764	2.41	100	-1.54	2.207	1.108	210.672	-2.383	-6.937	1.648
PLT	-1.665	0.293	100	-1.678	0.278	2.615	-43.568	-0.933	-1.428	-0.466
AST	1.011	0.19	100	1.044	0.186	1.285	-14.923	0.866	0.473	1.211
LC	-0.93	0.211	68.9	-0.896	0.204	4.125	-90.042	-0.044	-0.242	0
HCT	0.077	0.034	60.8	0.075	0.033	1.913	-65.937	0.008	0	0.066
Vomiting	0.066	0.218	18.6	0	0	0.36	86.905	0	0	0.237
Obese	0.377	0.29	16.6	0	0	1.242	-60.88	0	0	0.272
WBC	0.948	0.265	12.9	0.902	0.259	3.516	-83.649	0	0	0.244
Abdominal tenderness	-0.292	0.565	12.3	0	0	0.608	-188.498	0	0	0.395
Mucosal bleeding	0.362	0.479	10.8	0	0	0.732	-41.012	0	0	0.328
Skin bleeding	0.054	0.321	9	0	0	0.21	139.384	0	0	0.161

Predictors	Global model (including all predictors)			Selected models by lasso on bootstrap datasets						
	Coefficients	SE	Bootstrap inclusion frequency (%)	Coefficients	SE	RMSD ratio	Relative conditional bias (%)	Bootstrap median	Bootstrap 2.5 th percentile	Bootstrap 97.5 th percentile
ALB	-0.022	0.036	4.3	0	0	0.615	-40.528	0	-0.008	0
Age	-0.074	0.04	0.1	-0.077	0.039	1.861	-98.123	0	0	0

SE: standard error; PLT: platelet count; LC: lymphocyte count; AST: aspartate aminotransaminase; WBC; white blood cell count; ALB: serum albumin

Table 4.5. Top 10 combinations of predictors selected using multivariable logistic regression with lasso selection to predict the risk of developing DSS using 1,000-bootstrap samples of the data.

Rank	Predictors	Frequency
1	PLT, AST, LC, HCT	18%
2	PLT, AST, LC	10.5%
3	PLT, AST, HCT	7.9%
4	PLT, AST	4.5%
5	PLT, AST, LC, HCT, Vomiting	3.5%
6	PLT, AST, LC, HCT, Obese	3.2%
7	PLT, AST, LC, HCT, WBC	3.2%
8	PLT, AST, LC, Obese	3%
9	PLT, AST, HCT, Vomiting	2.5%
10	PLT, AST, LC, HCT, Abdominal pain or tenderness	2.3%

PLT: platelet count; LC: lymphocyte count; AST: aspartate aminotransaminase; WBC; white blood cell count; ALB: serum albumin

Table 4.6. Global model, model selected by lasso selection with the minimum lambda and bootstraps-derived quantities for assessing the uncertainty of model for the combined endpoint

Predictors	Global model (including all predictors)			Selected models by lasso on bootstrap datasets						
	Coefficients	SE	Bootstrap inclusion frequency (%)	Coefficients	SE	RMSD ratio	Relative conditional bias (%)	Bootstrap median	Bootstrap 2.5 th percentile	Bootstrap 97.5 th percentile
Intercept	3.791	0.985	100	3.824	0.984	2.424	-57.522	1.627	-0.235	3.512
PLT	-0.01	0.001	100	-0.01	0.001	2.854	-36.763	-0.006	-0.009	-0.004
LC	-0.902	0.13	100	-0.898	0.129	3.287	-46.177	-0.483	-0.682	-0.315
AST	0.004	0.001	100	0.004	0.001	1.218	-11.34	0.004	0.002	0.006
ALB	-0.09	0.022	99.2	-0.091	0.022	1.834	-36.819	-0.056	-0.099	-0.011
WBC	0.168	0.028	98.4	0.168	0.028	3.872	-61.361	0.064	0.008	0.118
Vomiting	0.249	0.137	87.8	0	0	1.076	-24.548	0.157	0	0.413
Mucosal bleeding	0.625	0.305	73.3	0	0	1.464	-46.444	0.204	0	0.78
Abdominal pain or tenderness	0.415	0.352	61.8	0	0	0.917	-18.552	0.108	0	0.809
Age	-0.038	0.023	29.7	-0.039	0.023	1.484	-61.618	0	-0.032	0
Skin bleeding	-0.112	0.219	12.3	0	0	0.557	-123.046	0	-0.057	0.104
Obese	0.038	0.209	11.7	0	0	0.252	-46.719	0	-0.05	0.08

SE: standard error; PLT: platelet count; LC: lymphocyte count; AST: aspartate aminotransaminase; WBC; white blood cell count; ALB: serum albumin

Table 4.7. Top 10 combinations of predictors selected using multivariable logistic regression with lasso selection to predict the risk of developing the combined endpoint (DSS and/or plasma leakage) using 1,000-bootstrap samples of the data.

Rank	Predictors	Frequency
1	PLT, LC, AST, ALB, WBC, Vomiting, Mucosal bleeding, and Abdominal pain or tenderness	23.2
2	PLT, LC, AST, ALB, WBC, Vomiting, and Mucosal bleeding	13.5
3	PLT, LC, AST, ALB, WBC, Vomiting, and Abdominal pain or tenderness, Age	8.4
4	PLT, LC, AST, ALB, WBC, Vomiting, Mucosal bleeding, Abdominal pain or tenderness, Age	7.2
5	PLT, LC, AST, ALB, WBC, Vomiting, Abdominal pain or tenderness	4.8
6	PLT, LC, AST, ALB, WBC, Vomiting, Mucosal bleeding, Age	4.3
7	PLT, LC, AST, ALB, WBC, Vomiting	3.6
8	PLT, LC, AST, ALB, WBC, Vomiting, Abdominal pain or tenderness, Age	2.6
9	PLT, LC, AST, ALB, WBC, Mucosal bleeding, Abdominal pain or tenderness	23.2
10	PLT, LC, AST, ALB, WBC, Vomiting, Mucosal bleeding, Abdominal pain or tenderness, Obese	13.5

PLT: platelet count; LC: lymphocyte count; AST: aspartate aminotransaminase; WBC; white blood cell count; ALB: serum albumin

4.3.3 Association of predictors and the clinical endpoints

Table 4.8 shows the association of the candidate predictors with the prediction outcomes. We find that vomiting (OR = 1.67 [95%CI 1.13, 2.47] and OR = 1.68 [1.31, 2.15]) and abdominal pain (OR = 2.98 [1.28, 6.09] and OR = 2.50 [1.38, 4.34]) are associated with both DSS and the combined clinical endpoint, respectively. Skin bleeding was associated with DSS (OR = 1.89 [1.14, 3.02]) and mucosal bleeding showed an OR = 1.94 [1.16; 3.11]) of developing the combined endpoint. During the first 72 hour of illness, we also found a significant association between higher haematocrit (OR = 1.14 [1.08, 1.20]), higher AST (OR = 1.01 [1.01, 1.01] and OR= 1.01 [1.01, 1.01]), lower PLT (OR = 0.98 [0.98, 0.99] and OR= 0.99 [0.99, 0.99]), lower

lymphocyte count (OR = 0.69 [0.50, 0.92] and OR = 0.52 [0.42, 0.65]), and lower serum albumin (OR = 0.91 [0.86, 0.97] and OR = 0.89 [0.86, 0.93]) for both DSS and the combined clinical endpoint, respectively.

Table 4.8. Associations of candidate predictors with dengue shock syndrome (DSS) and the combined endpoint of DSS and/or moderate plasma leakage derived from univariate logistic regression.

Parameters	DSS		Combined endpoint	
	Odds ratio (95%CI)	p-value	Odds ratio	p-value
Age (years)	1.00 (0.95, 1.06)	0.910	1.02 (0.99, 1.06)	0.187
Nutritional status: Overweight	1.50 (0.87, 2.45)	0.120	0.92 (0.62, 1.32)	0.662
Vomiting: Yes	1.67 (1.13, 2.47)	0.010	1.68 (1.31, 2.15)	<0.001
Skin bleeding: Yes	1.89 (1.14, 3.02)	0.010	1.38 (0.97, 1.94)	0.064
Mucosal bleeding: Yes	1.52 (0.62, 3.13)	0.305	1.94 (1.16, 3.11)	0.008
Abdominal pain or tenderness: Yes	2.98 (1.28, 6.09)	0.005	2.50 (1.38, 4.34)	0.002
Haematocrit (%)	1.14 (1.08, 1.20)	<0.001	-	-
White blood cell count (1,000 cells/dl)	0.94 (0.88, 1.01)	0.127	0.96 (0.92, 1.00)	0.061
Platelet count (1,000 cells/dl)	0.985 (0.981, 0.988)	<0.001	0.991 (0.988, 0.993)	<0.001
Lymphocyte counts (1,000 cells/dl)	0.69 (0.50, 0.92)	0.017	0.52 (0.42, 0.65)	<0.001
Serum albumin (g/L)	0.91 (0.86, 0.97)	0.003	0.89 (0.86, 0.93)	<0.001
AST (units/dl)	1.009 (1.007, 1.0113)	<0.001	1.007 (1.005, 1.009)	<0.001

4.3.4 Internal predictive performances

The optimal hyperparameters and the coefficients of the multivariable logistic regression models with Lasso selection are presented in Table 4.9 and 4.10. The predictive performance of the models developed on the Vietnamese dataset evaluated internally (on out-of-sample Vietnamese data) and externally (on the Thai dataset) using 10-fold cross-validation are presented in Table 4.11 and 4.12, respectively. Among the prediction models for DSS, the reference model (logistic regression with the WHO warning signs) achieved the best overall and discrimination performance with an average Brier score of 0.041 [95% CI 0.04, 0.041] and AUROC of 0.789 [95% CI 0.787, 0.791] (Table 4.11). However, logistic regression with lasso selection obtained better calibration performance (CITL= -0.001 [95% CI -0.002, 0.001] and calibration slope 0.942 [95% CI 0.933, 0.951]) with comparable discrimination performance and higher mean sensitivity of 0.789 [95% CI 0.787, 0.791]) (Table 4.11). Amongst the machine learning models, XGB performed best both in terms of discrimination and calibration and was followed by ANN with 2 hidden layers (Table 4.11).

Table 4.9. Table of optimal hyperparameters selected using Bayesian Global Optimisation with Gaussian Processes of models for DSS and the combined endpoint

Algorithm	Hyperparameters	Range	Selected values
Models for dengue shock syndrome			
Random forest	the number of predictors considered in each data split	2 to the max number of predictors	2
	the minimal number of members in each classification group	1 to 50	50
XGB	the number of trees	10 to 1000	326
	the maximum number of splits	1 to 20	1
	learning rate	0.0001 to 0.1	0.079
	gamma	0.0001 to 0.1	0.033
	cost	0.1 to 10	5.904
SVM	cost	0.1 to 10	5.904

Algorithm	Hyperparameters	Range	Selected values
	gamma	0.1 to 1	0.104
ANN-2	hidden layer-1 drop out proportion	0 to 0.5	0.3
	hidden layer-2 drop out proportion	0 to 0.5	0.3
	hidden neuron in layer 1	1 to the max number of predictors	7
	hidden neuron in layer 2	1 to the max number of predictors	9
	learning rate	0.0001 to 0.01	0.0009
Models for a combined endpoints of moderate plasma leakage or severe dengue			
Random forest	the number of predictors considered in each data split	2 to the max number of predictors	2
	the minimal number of members in each classification group	1 to 50	49
XGB	number of trees	10 to 1000	648
	the maximum number of splits	1 to 20	1
	learning rate	0.0001 to 0.1	0.0366
	gamma	0.0001 to 0.1	0.0983
SVM	cost	0.1 to 10	3.8284
	gamma	0.1 to 1	0.2229
ANN-2	hidden layer-1 drop out proportion	0 to 0.5	0.3
	hidden layer-2 drop out proportion	0 to 0.5	0.5
	hidden neuron in layer 1	1 to the max number of predictors	6
	hidden neuron in layer 2	1 to the max number of predictors	11
	learning rate	0.0001 to 0.01	0.0057

1273 **Table 4.10.** Summary table of the models with multivariable logistic regression with Lasso selection for DSS and
1274 the combined endpoint

Parameter	Coefficient	SE	z value	p-value
Outcome: DSS				
Intercept	-2.787	2.173	-1.283	0.200
Haematocrit (%)	0.055	0.030	1.856	0.063
log (Platelet count (1,000 cells/dl))	-1.235	0.254	-4.871	0.000
Lymphocyte counts (1,000 cells/dl)	-0.474	0.166	-2.852	0.004
log (AST (units/dl))	1.053	0.182	5.781	0.000
Outcome: combined endpoint of moderate plasma leakage and/or severe dengue				
Intercept	3.364	0.949	3.544	0.000
Vomiting: Yes	0.276	0.136	2.039	0.041
Skin bleeding: Yes	0.570	0.277	2.057	0.040
Mucosal bleeding: Yes	0.406	0.347	1.170	0.242
White blood cell count (1,000 cells/dl)	0.173	0.028	6.279	0.000
log (Platelet count (1,000 cells/dl))	-0.010	0.001	-7.173	0.000
Lymphocyte counts (1,000 cells/dl)	-0.840	0.122	-6.886	0.000
Serum albumin (g/L)	-0.092	0.022	-4.198	0.000
log (AST (units/dl))	0.004	0.001	3.975	0.000

1275

Among the models predicting the combined endpoint of moderate plasma leakage and/or DSS, the reference model (logistic regression using WHO warning signs) obtained the best BS and calibration performance. However, the discriminative performance in terms of AUROC was relatively low, with an average score of 0.697 [95% CI 0.696, 0.698] compared with the other models (Table 4.11). Multivariable logistic regression models using predictors selected by Lasso selection increased the discriminative performance by 6% on average. Multivariable logistic regression tended to overestimate the risk of developing the combined endpoint in patients with a low risk, and to underestimate the risk in patients with a high risk, as shown by the calibration slope: 0.954 [95% CI 0.948 to 0.960] (Table 4.11) and a calibration plot (Figure 4.2). XGB achieved the highest discrimination ability (AUROC 0.745 [95% CI 0.744, 0.747] and accuracy 0.721 [95% CI 0.717, 0.724]) but the calibration performance (CITL -0.018 [95% CI -0.022, -0.015] and calibration slope 1.053 [95% CI 1.037, 1.07]) was lower than that obtained with logistic regression using Lasso selection (CITL = -0.001 [95% CI -0.002, 0.001] and Calibration slope = 0.942 [95% CI 0.933, 0.951]). RF and SVM underperformed in terms of overall performance, discrimination, and calibration ability compared to the other models (Table 4.11).

Table 4.11. Predictive performance of the risk prediction models trained on the Vietnamese dataset on internal validation using 10-fold cross validation.

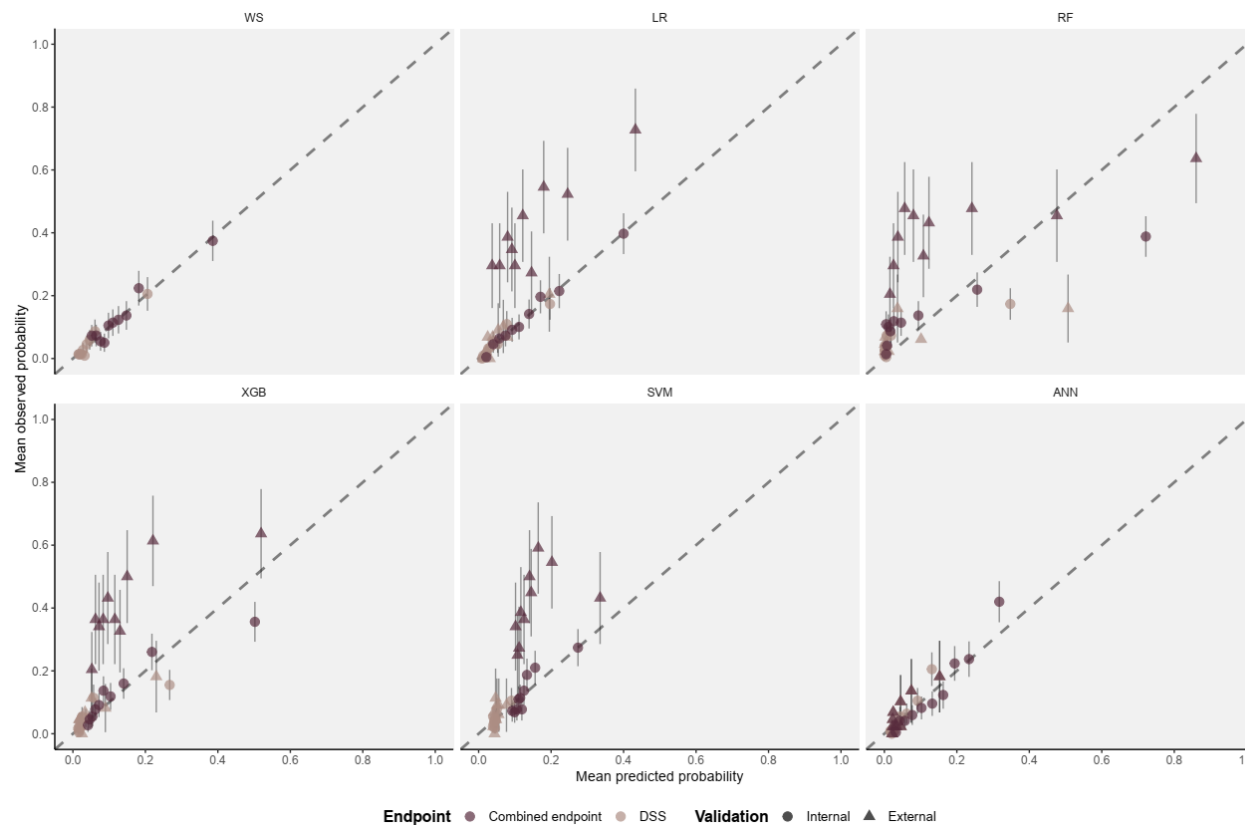
Model	Brier score (95% CI) (<i>the lower, the better</i>)	AUROC (95% CI)	Calibration intercept (95% CI) (<i>the closer to 0, the better</i>)	Calibration slope (95% CI) (<i>the closer to 1, the better</i>)	Accuracy (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)
Outcome: dengue shock syndrome									
Logistic regression using WHO warning signs	0.041 (0.040, 0.041)	0.789 (0.787, 0.791)	-0.003 (-0.005, - 0.001)	1.027 (1.011, 1.043)	0.756 (0.752, 0.761)	0.779 (0.773, 0.785)	0.755 (0.75, 0.76)	0.154 (0.15, 0.158)	0.985 (0.985, 0.986)
Logistic regression with variables selected by Lasso selection	0.044 (0.043, 0.044)	0.789 (0.787, 0.791)	-0.001 (-0.002, 0.001)	0.942 (0.933, 0.951)	0.731 (0.725, 0.736)	0.78 (0.774, 0.787)	0.728 (0.722, 0.734)	0.138 (0.135, 0.142)	0.985 (0.984, 0.985)
Random forest	0.044 (0.044, 0.044)	0.755 (0.753, 0.757)	-0.007 (-0.009, - 0.004)	1.026 (1.013, 1.038)	0.715 (0.71, 0.72)	0.771 (0.764, 0.779)	0.712 (0.706, 0.717)	0.128 (0.127, 0.13)	0.984 (0.983, 0.984)
Extreme gradient boosted tree	0.044 (0.044, 0.044)	0.785 (0.783, 0.787)	-0.018 (-0.022, - 0.015)	1.053 (1.037, 1.07)	0.749 (0.743, 0.755)	0.751 (0.744, 0.758)	0.749 (0.742, 0.755)	0.146 (0.143, 0.149)	0.983 (0.983, 0.984)

Model	Brier score (95% CI) (<i>the lower, the better</i>)	AUROC (95% CI)	Calibration intercept (95% CI) (<i>the closer to 0, the better</i>)	Calibration slope (95% CI) (<i>the closer to 1, the better</i>)	Accuracy (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)
Support vector machine	0.047 (0.047, 0.047)	0.613 (0.607, 0.619)	-0.016 (-0.02, -0.012)	0.404 (0.238, 0.57)	0.756 (0.752, 0.761)	0.779 (0.773, 0.785)	0.755 (0.75, 0.76)	0.154 (0.15, 0.158)	0.985 (0.985, 0.986)
Artificial neural network:2 hidden layers	0.045 (0.044, 0.045)	0.775 (0.771, 0.779)	-0.047 (-0.056, -0.039)	1.742 (1.598, 1.886)	0.731 (0.725, 0.736)	0.78 (0.774, 0.787)	0.728 (0.722, 0.734)	0.138 (0.135, 0.142)	0.985 (0.984, 0.985)
Outcome: combined endpoint of moderate plasma leakage and/or DSS									
Logistic regression using WHO warning signs	0.101 (0.101, 0.101)	0.697 (0.696, 0.698)	-0.002 (-0.003, -0.001)	0.965 (0.957, 0.973)	0.684 (0.68, 0.689)	0.661 (0.656, 0.667)	0.688 (0.682, 0.693)	0.254 (0.251, 0.257)	0.93 (0.929, 0.931)
Logistic regression using variables selected by Lasso selection	0.104 (0.104, 0.105)	0.741 (0.74, 0.742)	0.001 (0, 0.001)	0.968 (0.962, 0.974)	0.703 (0.7, 0.707)	0.706 (0.701, 0.711)	0.703 (0.698, 0.707)	0.273 (0.27, 0.275)	0.94 (0.939, 0.941)
Random forest (RF)	0.104 (0.104, 0.105)	0.735 (0.734, 0.736)	-0.004 (-0.005, -0.003)	1.035 (1.027, 1.043)	0.716 (0.712, 0.72)	0.693 (0.687, 0.699)	0.72 (0.714, 0.725)	0.281 (0.278, 0.285)	0.939 (0.938, 0.94)

Model	Brier score (95% CI) (<i>the lower, the better</i>)	AUROC (95% CI)	Calibration intercept (95% CI) (<i>the closer to 0, the better</i>)	Calibration slope (95% CI) (<i>the closer to 1, the better</i>)	Accuracy (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)
Extreme gradient boosted tree (XGB)	0.104 (0.104, 0.104)	0.745 (0.744, 0.747)	-0.005 (-0.007, -0.002)	1.04 (1.031, 1.05)	0.721 (0.717, 0.724)	0.696 (0.692, 0.701)	0.724 (0.72, 0.729)	0.289 (0.285, 0.292)	0.94 (0.939, 0.94)
Support vector machine (SVM)	0.111 (0.11, 0.111)	0.673 (0.671, 0.676)	-0.009 (-0.011, -0.007)	1.078 (1.056, 1.099)	0.678 (0.673, 0.684)	0.637 (0.631, 0.644)	0.685 (0.678, 0.692)	0.245 (0.241, 0.249)	0.925 (0.924, 0.926)
Artificial neural network:2 hidden layers (ANN2)	0.107 (0.107, 0.108)	0.724 (0.722, 0.727)	-0.01 (-0.013, -0.007)	1.31 (1.264, 1.355)	0.691 (0.684, 0.697)	0.689 (0.683, 0.696)	0.691 (0.683, 0.699)	0.26 (0.256, 0.264)	0.935 (0.934, 0.936)

1293

AUROC: area under receiver operating characteristics curves; CI: Confidence interval



1294

1295

1296

1297

1298

Figure 4.2. Calibration plots for the predicted risk of DSS (triangles) and the combined endpoint of moderate plasma leakage and/or DSS (points) using: logistic regression with the WHO warning signs as predictors (Ref) (validated only internally), logistic regression with variables selected by Lasso selection (LR), random forest (RF), extreme gradient boosted tree (XGB), support vector machine (SVM), and artificial neural network with 2 hidden layers (ANN). In each panel, the diagonal dotted line represents the perfect agreement between predicted and observed risk.

The decision curve analysis, shown in Figure 4.3, visualises the probability of predicting the outcome (y-axis) as a function of the threshold probability (x-axis) and can be interpreted as the number of false positives (i.e. incorrectly identified DSS cases or, in the secondary analysis, case developing moderate plasma leakage and/or DSS) for each true positive. Given the low threshold probability (less than 0.1), ANN provided the higher net benefit for predicting DSS. For the combined endpoint of moderate plasma leakage and/or DSS, we found that the multivariable logistic regression model with variables selected by Lasso selection, XGB, and ANN-2 achieved higher net benefits overall compared with the reference model. However, where the thresholds probability exceeded 25%, the reference model achieved a higher net benefit. RF and SVM achieved a lower net benefit throughout the whole threshold probability spectrum on both outcomes (Figure 4.3).

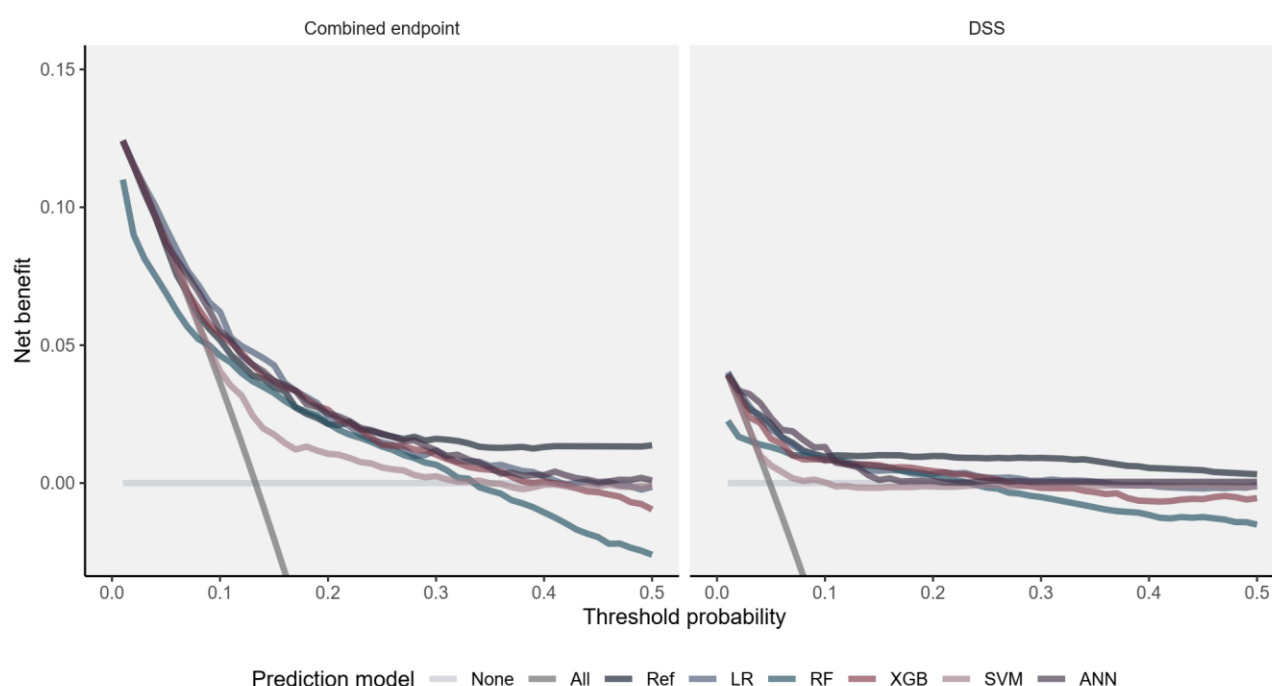


Figure 4.3. Decision curves, showing the net benefit (y-axis) against the threshold probability (x-axis) for DSS (right) and the combined endpoint of moderate plasma leakage and/or DSS (left) obtained with logistic regression using the WHO warning signs (Ref), logistic regression with variables selected by Lasso selection (LR), random forest (RF), Extreme gradient boosted tree (XGB), support vector machine (SVM), and artificial neural network with 2 hidden layers (ANN).

4.4 External validation

The assessment of the relatedness between the Vietnamese (training) and Thai (external validation) datasets showed that the case-mix of the two datasets were different. The concordance statistic of the membership model was on average 0.84 [95% CI 0.81, 0.86]. Comparison of the predicted risk in the training and validation datasets, including the mean difference and relative difference in standard deviation of linear predictor between training and external validation datasets further supports the observed differences between the two datasets (Figure 4.4). While the XGB model achieved the highest AUROC at 0.679 and 0.650 for predicting DSS and the combined endpoint, respectively, the ANN model accomplished the best calibration performance for DSS (CITL = 0.331 and calibration slope = 0.912) and the combined endpoint (CITL = 1.345 and calibration slope = 0.617).

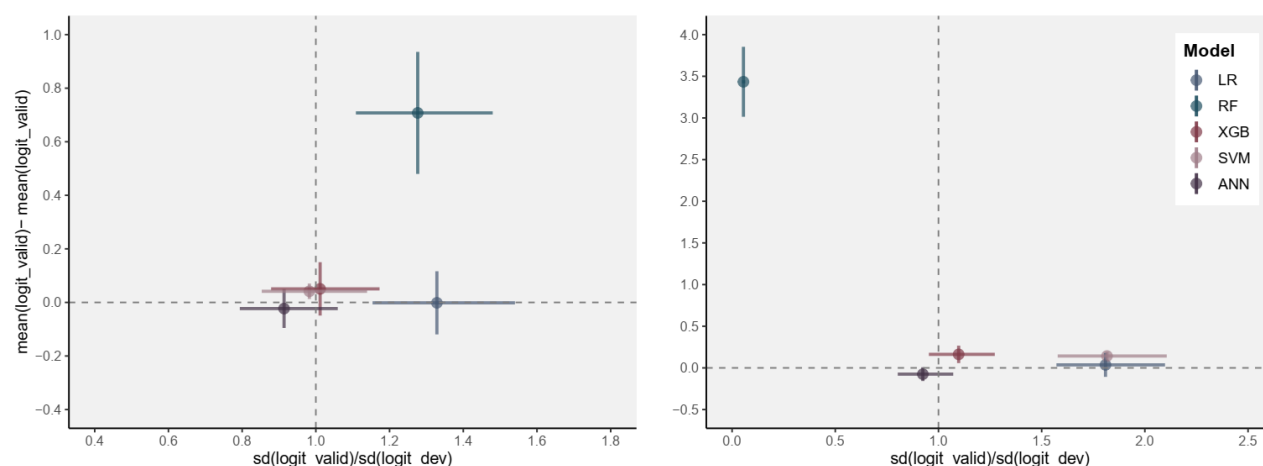


Figure 4.4. Results of difference in means with 95% CI (the x axis) and relative difference with 95% CI in standard deviation (the y axis) of predicted logarithmic odds on the training and validation sets. The vertical and horizontal dotted lines reflect no relative difference in standard deviation and no difference in means of predicted logarithmic odds between the training and validation sets. LR: models with logistic regression and Lasso selection; RF: models with random forest; XGB: models with extreme gradient boosted tree; SVM: models with support vector machine; ANN: models with artificial neural networks with 2 hidden layers.

The predictive performance obtained on the external validation dataset was lower than that observed on the out-of-sample validation sets both in terms of calibration and discrimination. The AUROC obtained on the external validation dataset were on average 11% and 17% lower than the AUROC obtained on the internal validation datasets using DSS and the combined endpoint,

respectively (Table 4.12). The calibration metrics suggest that the prediction of both endpoints is generally underestimated (i.e., predicted risks were generally lower than observed risks as seen CITL higher than 0) and relatively overly extreme (i.e., predicted risks were likely close to 0 or 1 as seen calibration slope lower than 1) (Table 4.12 and Figure 4.2).

Table 4.12. Predictive performance of the five models fitted on the Vietnamese (training) dataset and validated on the Thai (external validation) dataset.

Model	Fitted to Vietnamese dataset and externally validated on the Thai dataset			
	Brier score (<i>the lower, the better</i>)	AUROC	Calibration intercept (<i>the closer to 0, the better</i>)	Calibration slope (<i>the closer to 1, the better</i>)
Outcome: dengue shock syndrome				
Logistic regression with Lasso selection	0.059	0.677	0.188	0.531
Random forest	0.085	0.665	-0.234	0.223
Extreme gradient boosted tree	0.060	0.679	0.248	0.544
Support vector machine	0.057	0.648	0.234	1.626
Artificial neural network:2 hidden layers	0.057	0.690	0.331	0.912
Outcome: combined endpoint of moderate plasma leakage and/or severe dengue				
Logistic regression with Lasso selection	0.2985	0.6243	1.6362	0.2874
Random forest	0.3117	0.6339	1.9028	0.2136
Extreme gradient boosted tree	0.3002	0.6500	1.6130	0.7310
Support vector machine	0.3034	0.5715	1.3515	0.3909
Artificial neural network:2 hidden layers	0.2891	0.6148	1.3447	0.6172

AUROC: area under receiver operating characteristics curves; CI: Confidence interval

4.5 Sensitivity analysis

In a sensitivity analysis, we restricted the analysis to hospitalised cases with a combined endpoint of moderate plasma leakage and/or DSS. In this subset of hospitalised patients, we found that models with logistic regression and Lasso selection achieved the highest discrimination and calibration performance (AUROC 0.648 [95% CI 0.647, 0.65], CITL 0 [95% CI -0.001, 0.001], and calibration slope 0.903 [95% CI 0.892, 0.914]) (Table 4.13). Its average sensitivity (0.640 [95% CI 0.634, 0.646]) and specificity (0.646 [95% CI 0.641, 0.652]) were relatively high compared to the other models. The performance obtained in the main analysis exceeded that obtained in the sensitivity analysis for all metrics.

4.6 Discussion

In outpatient settings, where patients with dengue infection are first seen and evaluated, effective triage is key for identifying the patients at risk of developing severe disease who require hospital care. In this work we developed risk prediction models using data obtained from outpatient departments in Vietnam and Thailand to identify clinical and laboratory predictors widely available during the early (less than 72 hours from fever onset) phases of illness linked with an increased risk of progression to severe disease. We found that multivariable logistic regression with predictors selected using Lasso selection achieved the highest predictive performance, followed by XGB and ANN.

In comparison with models developed in previous studies^{84, 86, 90, 92, 99}, the models in this thesis use predictors collected in outpatient settings less than 72 hours from fever onset and thus provide earlier estimates of the risk of developing DSS and a combined endpoint of plasma leakage and/or DSS. While DSS is well defined and uniform in clinical practice, it is a relative rare outcome at the extreme end of the disease spectrum. For this reason, and in order to evaluate the risk of developing severe symptoms and disease which require hospitalisation, we also developed models for predicting a combined endpoint including moderate plasma leakage too. We included eleven candidate predictors in the machine learning models. Using the multivariable logistic regression

model with Lasso selection, we found that respectively four and eight predictors were selected for predicting DSS and the combined endpoint of moderate plasma leakage and/or DSS, respectively.

For DSS, the predictive variables were haematocrit levels, platelet count, lymphocyte count, and AST. For the combined endpoint of moderate plasma leakage and/or severe dengue the predictive variables were vomiting, abdominal pain, mucosal bleeding, platelet count, white blood cell count, lymphocyte count, AST and serum albumin levels. Platelet count and AST were consistently selected as predictors in all models, suggesting that these are strong predictors of the development of severe dengue illness. This finding is consistent with the results of our literature review and meta-analysis¹¹⁶ showing a significant association of these predictors with progression to severe dengue disease and reinforce the usefulness of testing and monitoring platelet count and AST during the febrile phase for the early detection of progression to severe dengue illness. In addition, our results confirm that the presence of vomiting, mucosal bleeding, and abdominal pain, which are included as warning signs in the 2009 WHO Dengue guideline^{5, 116}, are established risk factors for developing severe dengue illness. Notably, all predictors used in this study are conventional laboratory parameters which are generally collected in outpatient setting. Our finding that serum albumin levels are significantly lower among patients developing severe illness¹¹⁶ is not surprising and can be a consequence of the extravasation of serum albumin during plasma leakage. A key strength of this study is the use of an external validation dataset obtained independently and in a different country. We found good predictive performance of the models on the external validation dataset, despite the differences between the training and validation set. In addition, we demonstrate that models with logistic regression and Lasso selection performed relatively better than the other models, including the reference models. This finding was consistent in all analyses, including in the sensitivity analysis.

This study, however, has limitations and there are some important caveats. Given the combined endpoint defined as evidence of moderate plasma leakage in the Vietnamese dataset, the ambulatory patients receiving home care were followed up with daily phone calls and we assumed that, unless moderate plasma leakage was detected during the outpatient hospital visits, they did not develop moderate plasma leakage. It is possible that a few of the ambulatory patients who recovered without hospitalisation developed moderate plasma leakage but were not reported.

Table 4.13. Predictive performance of the risk prediction models for the combined endpoint of moderate plasma leakage and/or severe dengue trained on hospitalised patients in the Vietnamese dataset on internal validation using 10-fold cross validation.

Model	Brier score (95% CI) (<i>the lower, the better</i>)	AUROC (95% CI)	Calibration intercept (95% CI) (<i>the closer to 0, the better</i>)	Calibration slope (95% CI) (<i>the closer to 1, the better</i>)	Accuracy (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)
Logistic regression using WHO warning signs	0.189 (0.188, 0.189)	0.61 (0.608, 0.613)	-0.002 (-0.003, - 0.001)	0.893 (0.852, 0.934)	0.62 (0.617, 0.623)	0.605 (0.599, 0.612)	0.626 (0.62, 0.632)	0.403 (0.4, 0.406)	0.798 (0.796, 0.8)
Logistic regression with variables selected by Lasso selection	0.194 (0.194, 0.194)	0.648 (0.647, 0.65)	0 (-0.001, 0.001)	0.903 (0.892, 0.914)	0.644 (0.642, 0.647)	0.64 (0.634, 0.646)	0.646 (0.641, 0.652)	0.428 (0.425, 0.431)	0.817 (0.815, 0.819)
Random forest	0.197 (0.197, 0.197)	0.624 (0.621, 0.626)	0 (-0.002, 0.001)	1.12 (1.098, 1.142)	0.624 (0.62, 0.628)	0.616 (0.609, 0.623)	0.628 (0.62, 0.635)	0.406 (0.403, 0.41)	0.803 (0.801, 0.805)
Extreme gradient boosted tree	0.197 (0.196, 0.197)	0.633 (0.631, 0.636)	0 (-0.002, 0.001)	1.127 (1.104, 1.15)	0.646 (0.643, 0.65)	0.611 (0.604, 0.618)	0.66 (0.654, 0.667)	0.428 (0.424, 0.431)	0.808 (0.807, 0.81)
Support vector machine	0.201 (0.201, 0.202)	0.586 (0.581, 0.591)	-0.002 (-0.005, 0)	1.032 (0.936, 1.128)	0.601 (0.597, 0.606)	0.605 (0.598, 0.613)	0.6 (0.592, 0.607)	0.384 (0.38, 0.388)	0.791 (0.788, 0.794)
Artificial neural network:2 hidden layers	0.197 (0.197, 0.198)	0.633 (0.63, 0.636)	-0.008 (-0.011, - 0.006)	1.384 (1.32, 1.447)	0.631 (0.626, 0.635)	0.627 (0.621, 0.634)	0.632 (0.624, 0.64)	0.412 (0.408, 0.416)	0.808 (0.806, 0.809)

1 This assumption can also affect the relatedness between the Vietnamese (training) and Thai
2 (external validation) datasets, where all participants enrolled at the outpatient clinics were
3 subsequently admitted to the hospital for daily monitoring purposes.

4 The use of predictive models to inform triage and clinical practice is promising and can provide
5 further support to decision making in the treatment of dengue as well as other infectious
6 diseases. However, its use will need to be validated across settings and populations. With
7 further testing and validation, this work lays the foundations for integrating evidence-based
8 and data-driven methods into a decision support system in clinical practice, which in turn can
9 contribute to optimising healthcare allocation and the optimal use of resources.

Chapter 5 Risk prediction of progression to severe dengue illness from the observed dynamics of clinical parameters collected routinely during the febrile phase

5.1 Rationale

Most patients with symptomatic dengue infection are first seen and evaluated in outpatient settings during the early febrile phase. Prediction as to whether patients will progress to severe disease is challenging in this phase. Various clinical features and laboratory parameters of severe disease may unfold at a time relatively close to the critical phase. In fact, the WHO guidelines recommend outpatient daily assessments to monitor disease progression, focusing on the early detection of signs of leakage, which typically manifest on day 4 to 6 of illness⁵. This is particularly important not only to provide adequate care, but also to optimise resources. In fact, while patients with warning signs need hospitalisation both for close monitoring and prompt appropriate interventions, patients without warning signs can continue daily outpatient evaluation while recovering⁵.

Since the clinical progression of dengue is dynamic, daily clinical evaluation is necessary for the early detection of clinical deterioration. To understand the dynamics of clinical progression, I analysed a longitudinal description of clinical parameters prior to progression to severe dengue illness collected during the febrile phase. Although several studies have been performed and reported clinical indicators of severe dengue illness, a limited number of studies investigated the dynamics of clinical parameters in dengue patients. In addition, the availability of clinical risk prediction models based on sequential rather than single measurements may be more precise and appealing^{84, 101}. In Chapter 4, I developed risk prediction models using single time-point measurements collected at Enrolment within the first 72 hours of fever onset. In this chapter, RNNs, a set of neural network algorithms employing sequential input or output information, were used to develop risk prediction models, using repeatedly measured clinical and laboratory data as a case study. Instead of using only baseline or single-time-point measured data, RNNs use sequential data to make

clinical predictions. This technique has been adopted to predict clinical outcomes for several diseases, but not for dengue.

This chapter aims to quantitatively describe the dynamics of clinical laboratory parameters and to present the results of applying RNN to sequential clinical and laboratory data routinely collected during the early febrile phase of dengue illness.

5.2 Method

This study was approved by The Research Ethics Review Committee of QSNICH (REC.082/2562). The development and validation of the risk prediction models followed the guidelines in Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD)¹⁵¹.

This chapter uses data from the prospective cohort study of Thai children with laboratory-confirmed dengue infection. The study enrolled children presenting at outpatient clinics within 72 hours of fever onset. Demographic details and the results of the clinical evaluation were recorded at enrolment. The first blood samples were collected within the first 24 hours after enrolment and daily until discharge or a maximum of five consecutive days. A more detailed description of the dataset is provided in Chapter 3. Marginal models with GEE were used to describe and investigate the temporal dynamics of eight parameters (Table 5.1) between two groups of patients (developing and not developing DSS and the combined endpoint). Details of the analytical methods, prediction outcome, and available data are described in Chapter 3.

Table 5.1. Eight laboratory parameters included in the analysis of temporal dynamics

Predictor	Definition	Type of data
White blood cell count (WBC)	Total white blood cell count (cells/dl) measured during the first 72 hours of illness	Continuous

Predictor	Definition	Type of data
Neutrophil count (NC)	Total neutrocyte count (cells/dl) during measured the first 72 hours of illness	Continuous
Lymphocyte count (LC)	Total lymphocyte count (cells/dl) during measured the first 72 hours of illness	Continuous
Platelet count (PLT)	Platelet count (cell/dl) measured during the first 72 hours of illness	Continuous
Aspartate aminotransferase (AST)	Aspartate aminotransferase measured during the first 72 hours of illness	Continuous
Alanine aminotransferase (ALT)	Alanine aminotransferase measured during the first 72 hours of illness	Continuous
Serum albumin (ALB)	Serum albumin (g/dl) measured during the first 72 hours of illness	Continuous

5.3 Result

5.3.1 Baseline characteristics of patients in the Thai dataset

Of 1,210 children with suspected dengue infection recruited in the prospective cohort study, 524 children were laboratory confirmed dengue infections; 682 children were excluded due to a lack of laboratory confirmation of dengue infection, and one child was excluded due to a single day admission record (Figure 5.1). As a result, the data of 523 children with laboratory confirmed dengue were used in this chapter. Of those with laboratory confirmed dengue, 218 children subsequently developed a combined endpoint of moderate plasma leakage and/or DSS (specifically, 36 developed DSS and 182 developed moderate plasma leakages alone). As many as 305 children did not develop either moderate plasma leakage or DSS.

Table 5.2 summarises the baseline characteristics of the patients included in this analysis. The average age of the children overall and in the different outcome groups were eight years with the similar proportion (p-value = 0.434 for DSS outcome and 0.609 for the combined outcome) of

males and females. The proportion of overweight children, defined as over the 75th percentile of weight for age¹⁵², was approximately 20% and 15% in a group of patients with DSS and the combined endpoint, respectively.

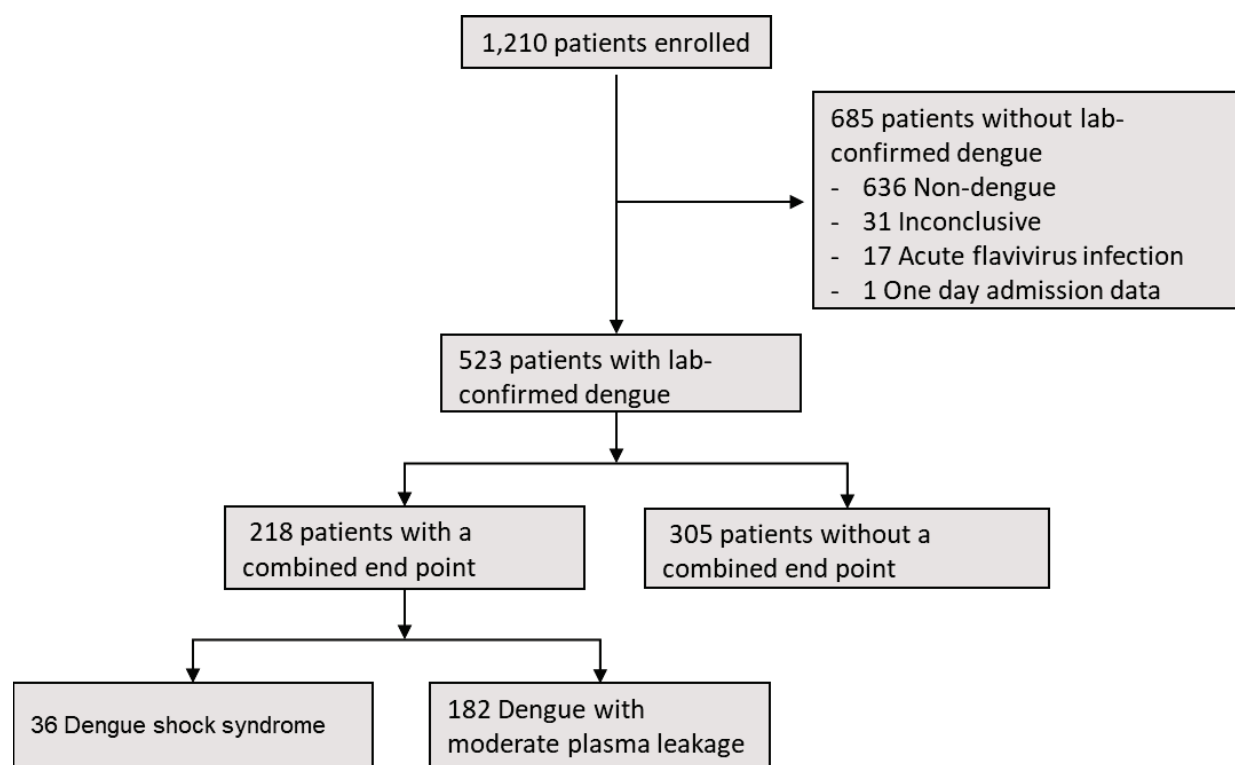


Figure 5.1. Flow chart describing patients' clinical endpoints of the Thai dataset

I observed no statistically significant differences in the proportion of children overweight between the groups with and without clinical endpoints (p-value = 0.299 for DSS outcome and 0.053 for the combined outcome). Patients were enrolled approximately on day 3 or 4 of illness. Three quarters (75%) of patients were secondary infections and the proportion of secondary infections was significantly higher among the patients with DSS and with the combined endpoint of moderate plasma leakage and/or DSS (p-value 0.016 and 0.002, respectively). Half of the patients were infected by each DENV1 and DENV3 and the rest were infected by DENV2 or DENV4. There was no statistically significant difference in the infecting serotype by outcome group with both endpoints. Data availability on each day of illness and day of leakage are presented in Figure 5.2 and 5.3.

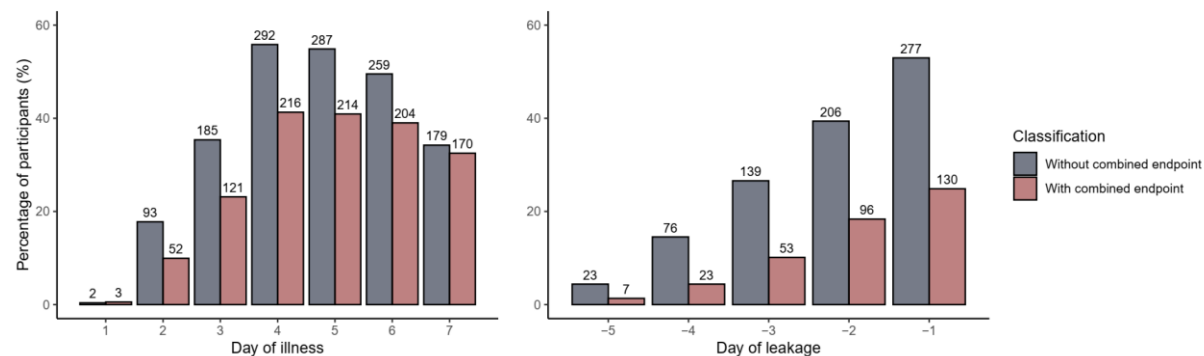


Figure 5.2. Percentage of dengue infections by groups of the combined endpoint with laboratory predictors by day of illness and day of leakage. The numbers at the top of each bar represent the number of observed participants with clinical and laboratory information which was available for analysis by day of leakage.

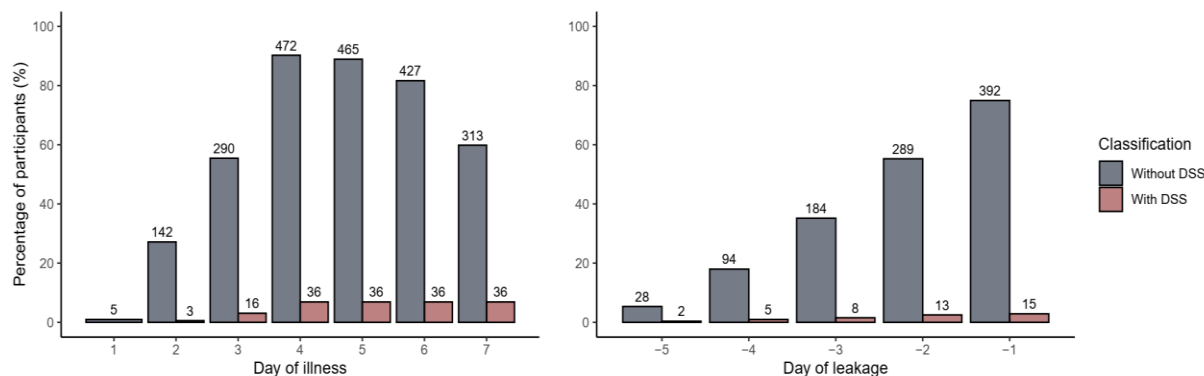


Figure 5.3. Percentage of dengue infections by groups of DSS diagnosis with laboratory predictors by day of illness and day of leakage. The numbers at the top of each bar represent the number of observed participants with clinical and laboratory information by day of leakage.

87 5.3.2 Clinical laboratory parameters of patients by the difference groups of clinical
88 endpoints

89 To investigate the difference in clinical parameters between the two groups of patients
90 developing and not developing the two clinical endpoints, I performed two analyses:
91 multivariable linear regression and marginal models with GEE.

92 **Table 5.2.** Baseline characteristics of patients by groups of clinical endpoints

Parameters	Overall	DSS			p-value	Combined endpoint		p-value
		Without	With			Without	With	
Number of subjects	523	487	36			305	218	
Age (years): median (IQR)	8.83 (6.33, 10.96)	8.83 (6.33, 11.04)	8.91 (6.80, 10.14)	0.618		8.83 (6.25, 10.92)	8.96 (6.50, 11.08)	0.487
Sex: male	287 (54.9)	270 (55.4)	17 (47.2)	0.434		164 (53.8)	123 (56.4)	0.609
Weight (kg): median (IQR)	25.80 (18.75, 34.85)	25.80 (18.90, 34.92)	21.50 (18.45, 33.10)	0.432		25.30 (18.90, 34.00)	26.00 (18.50, 35.38)	0.642
Nutritional status				0.299				0.053
normal	434 (83.0)	406 (83.4)	28 (77.8)			253 (83.0)	181 (83.0)	
overweight	85 (16.3)	78 (16.0)	7 (19.4)			52 (17.0)	33 (15.1)	
NA	4 (0.8)	3 (0.6)	1 (2.8)			0 (0.0)	4 (1.8)	
Day of illness at enrolment	3.00 (2.00, 4.00)	3.00 (2.00, 4.00)	4.00 (3.00, 4.00)	0.011		3.00 (2.00, 4.00)	3.00 (3.00, 4.00)	0.091
Immune status				0.016				0.002
indeterminate	5 (1.0)	5 (1.0)	0 (0.0)			4 (1.3)	1 (0.5)	
primary infection	109 (20.8)	108 (22.2)	1 (2.8)			79 (25.9)	30 (13.8)	
secondary infection	409 (78.2)	374 (76.8)	35 (97.2)			222 (72.8)	187 (85.8)	

Parameters	Overall	DSS		p-value	Combined endpoint		p-value
		Without	With		Without	With	
Infecting serotypes				0.811			0.462
DENV-1	172 (32.9)	160 (32.9)	12 (33.3)		107 (35.1)	65 (29.8)	
DENV-2	91 (17.4)	85 (17.5)	6 (16.7)		52 (17.0)	39 (17.9)	
DENV-3	125 (23.9)	116 (23.8)	9 (25.0)		69 (22.6)	56 (25.7)	
DENV-4	59 (11.3)	56 (11.5)	3 (8.3)		38 (12.5)	21 (9.6)	
Indeterminate	15 (2.9)	15 (3.1)	0 (0.0)		9 (3.0)	6 (2.8)	
Not known	61 (11.7)	55 (11.3)	6 (16.7)		30 (9.8)	31 (14.2)	

5.3.3 Comparison of clinical parameters by day of illness

Multivariable linear regression was used to assess differences in the individual-level average daily clinical parameter in patients developing and not developing each clinical endpoint, adjusted for the effect of age and sex. Figure 5.4 and Appendix 18 shows the p-value of the group coefficients obtained from the multivariable linear model in quadratic model with adjustment for age and sex for all laboratory predictors, excepted for serum albumin which obtained from models in linear model without adjustment for age and sex.

In the comparison of laboratory parameters by day of illness between patients with and without DSS, there was no difference in the daily average laboratory parameters between groups until day 4 of illness. Liver function enzymes (AST and ALT) were the earliest laboratory parameters showing significant statistical differences, followed by platelet count and lymphocyte count, from day 5 of illness.

In the comparison of laboratory parameters between the groups of patients with and without the combined endpoint, most laboratory parameters were significantly different on day 5 of illness onwards. On the other hand, I found no difference in the daily average laboratory parameters measured on day 1 or 2 of illness between patients developing and not developing the combined endpoint. Platelet count, AST and ALT were found to be significantly different (p -value < 0.05) between the two groups from day 3 until day 7 of illness. The levels of haematocrit, neutrophil count, lymphocyte count and serum albumin were significantly different on day 5 and 6 of illness (p -value < 0.001). These findings suggest that platelet count and liver function test (AST and ALT) are the earliest clinical parameters showing statistically significant differences between the patients developing and not developing the combined endpoint of moderate leakage and/or DSS.

5.3.4 Temporal dynamics of clinical laboratory parameters

I investigated the temporal dynamics of the clinical parameters between the two groups of patients developing and not developing DSS and the combined endpoint using marginal models with GEE (described in Chapter 3). I considered both linear and quadratic models and selected optimal models using the quasi-information criteria (QIC). The QIC of models estimating the dynamics of clinical parameters is presented in Table 5.3 and Appendix 19. Models with quadratic terms and adjustment for age and sex achieved the best goodness of fit according to the QIC metric for all laboratory parameters except serum albumin levels, for which a linear model without age and sex adjustment performed best.

Table 5.3. Standardised QICs of models estimating the dynamics of clinical parameters during febrile phase of dengue

Parameters	Combined endpoint				DSS			
	Non-adjusted for age and sex		Adjusted		Non-adjusted for age and sex		Adjusted	
	Linear	Quadratic	Linear	Quadratic	Linear	Quadratic	Linear	Quadratic
Haematocrit	0.78	0.76	-1.04	-1.06	1.09	1.07	-0.79	-0.81
Platelet count	-0.09	-0.14	-1.27	-1.30	1.38	1.32	0.07	0.03

Parameters	Combined endpoint				DSS			
	Non-adjusted for age and sex		Adjusted		Non-adjusted for age and sex		Adjusted	
	Linear	Quadratic	Linear	Quadratic	Linear	Quadratic	Linear	Quadratic
White blood cell count	1.29	-0.48	0.47	-1.22	1.22	-0.49	0.42	-1.21
Neutrophil count	1.02	-0.88	0.91	-0.97	0.96	-0.89	0.85	-0.99
Lymphocyte count	0.92	0.91	-0.94	-0.96	0.96	0.95	-0.91	-0.93
AST	-0.42	-0.60	-1.10	-1.26	1.35	1.23	0.45	0.35
ALT	0.18	0.08	-1.11	-1.21	1.44	1.38	-0.34	-0.42
Serum albumin	-1.63	-0.58	-0.65	0.38	-0.32	0.93	0.33	1.53

Colour scale ranges from low (in blue) to high values (in red)

Table 5.4 shows the estimated coefficients and 95% CI of the interaction terms between the day of illness and the group obtained from the selected models for each laboratory parameter (i.e., haematocrit, white blood cell count, neutrophil count, lymphocyte count, platelet count, AST, ALT, and serum albumin). The significance of the coefficient of the interaction term identifies significant differences in clinical parameters across the days of illness between subjects developing and not developing the endpoint.

Table 5.4. Coefficients and p-value of the interaction terms between the outcome group and day of illness, as described in Equation 3.1 and 3.4.

Parameters	Models		Model with adjustment for age and sex	
	Coefficient (95%CI)	p value	Coefficient (95%CI)	p value
Endpoint: Dengue shock syndrome				
HCT	0.4 (-0.38; 1.18)	0.321	0.31 (-0.52; 1.14)	0.464
PLT	-14034.07 (-23248.58; -4819.55)	0.003	-13381.71 (-22638.07; -4125.36)	0.005

Parameters	Models		Model with adjustment for age and sex	
	Coefficient (95%CI)	p value	Coefficient (95%CI)	p value
WBC	233.32 (-240.79; 707.44)	0.335	280.95 (-170.77; 732.66)	0.223
NEU	163.84 (-174.1; 501.78)	0.342	165.07 (-173.37; 503.5)	0.339
LYM	64.54 (-101.45; 230.53)	0.446	104.12 (-52.95; 261.18)	0.194
AST	23.27 (-4.99; 51.53)	0.107	23.01 (-5.09; 51.1)	0.109
ALT	8.53 (-3.35; 20.41)	0.159	8.23 (-3.7; 20.15)	0.176
ALB	-0.26 (-0.43; -0.09)	0.002	-0.26 (-0.43; -0.1)	0.002
Endpoint: Combined endpoint of moderate plasma leakage and/or dengue shock syndrome				
HCT*	-	-	-	-
PLT	-9690.33 (-14524.11; -4856.56)	<0.001	-9247.28 (-13997.2; -4497.35)	<0.001
WBC	152.79 (-66.75; 372.32)	0.173	160.46 (-55.86; 376.77)	0.146
NEU	58.28 (-118.11; 234.67)	0.517	59.38 (-117.15; 235.91)	0.510
LYM	26.59 (-61.95; 115.13)	0.556	32.71 (-49.08; 114.5)	0.433
AST	17.66 (5.91; 29.42)	0.003	17.48 (5.77; 29.18)	0.003
ALT	6.4 (0.87; 11.93)	0.023	6.33 (0.82; 11.85)	0.024
ALB	-0.1 (-0.14; -0.05)	<0.001	-0.1 (-0.14; -0.05)	<0.001

*HCT is included in the diagnostic criteria for moderate plasma leakage and was thus excluded from the analysis;
HCT: haematocrit; PLT: platelet count; WBC: white blood cell count; NC: neutrophil count; LC: lymphocyte count;
AST: aspartate transaminase; ALT: alanine transaminase; ALB; serum albumin

In comparing the laboratory measurements in the patients developing the combined endpoint and those who did not, there were significant differences in the dynamics of platelet count, AST, ALT and serum albumin levels between subjects developing and not developing the combined endpoint of moderate plasma leakage and/or DSS. The platelet count and serum albumin of patients with the combined endpoint were also significantly lower than the ones

of the patients who did not develop the endpoints (Table 5.4). The AST and ALT of patients with the combined endpoint were significantly (p -value = 0.003 and 0.024, respectively) higher compared to the group of patients not developing moderate plasma leakage and/or DSS. The increases in haematocrit, AST and ALT were significantly higher amongst patients with the combined endpoint (p -value <0.001, 0.003, and 0.024) as diseases progressed; however, the same trends were not found when comparing patients with and without DSS. The median of the laboratory variables estimated from the GEE model are presented in Figure 5.5 and 5.6.

5.3.5 Selected dynamic predictors

The analysis in the previous section suggests that the dynamics of platelet count, AST, ALT and serum albumin measured during the febrile phase were statistically different between patients with and without the combined endpoint of moderate plasma leakage and/or DSS. When considering patients developing DSS and not developing DSS, we found that only the dynamics of platelet count and serum albumin were significantly different between the two groups. We also found that PLT and liver enzymes (i.e., AST and ALT) were the earliest predictors of the combined endpoints. Since the levels of AST are highly correlated with the levels of ALT throughout dengue illness, in the following analyses only AST was selected as predictor to avoid multicollinearity. Given these results and the results of the systematic review and meta-analysis (Chapter 2 to 4), the dynamics of four laboratory parameters (platelet count, lymphocyte count, AST and serum albumin) were used in model development in the next section.

5.3.6 Internal predictive performance

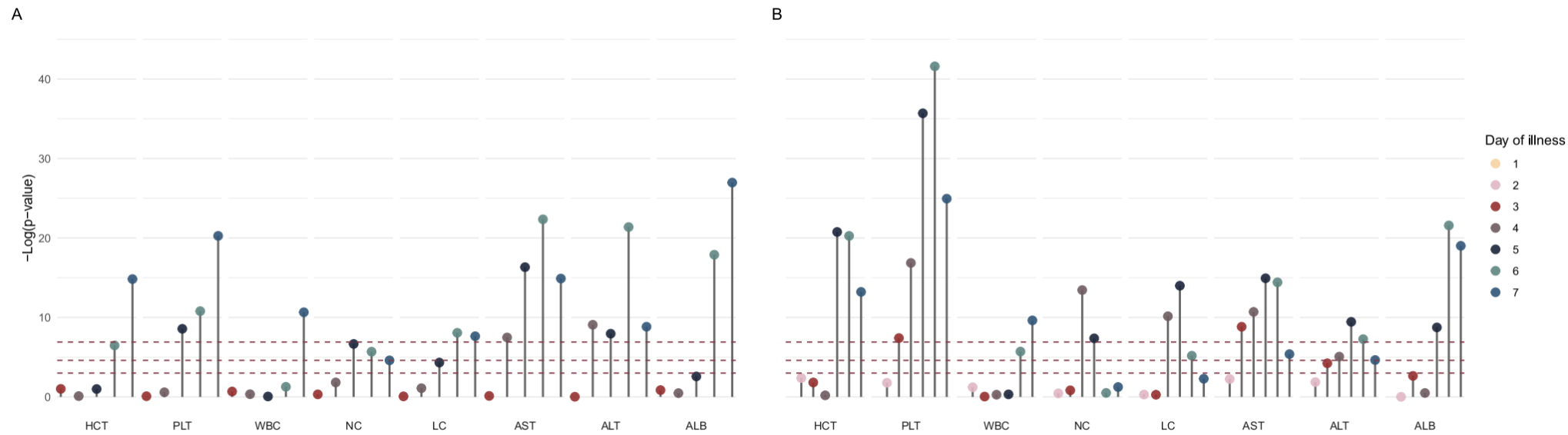
The predictive performance of the three RNN-based models (i.e., simple RNN, GRU and LSTM) and eight models using logistic regression and ANN with predictors measured on day of enrolment and on each day of illness (i.e., day 2 to 4 of illness) are presented in Figure 5.7 and Appendix 20. The models were developed to predict two separate clinical endpoints: DSS and the combined endpoints of moderate plasma leakage and/or DSS. The predictive performance of the models was estimated using a combination of multiple imputation and

171 internal validation (described in Section 3.5.4). The predictive performance of risk prediction
172 models was assessed using the Brier Score (BS), AUROC, CITL and calibration slope. Sensitivity,
173 specificity, accuracy, PPV and NPV were also estimated with the classification threshold
174 defined by the minimal distance to the left corner of AUROC curves.

175 As expected, we found that the BS of the models predicting DSS was better than the BS of the
176 models predicting the combined endpoints (approximately 0.05 for DSS and 0.20 for the
177 combined endpoint). While we obtained similar BS from the models developed for DSS, the
178 BS of the models for the combined endpoint varied across different types of models. The BS
179 of RNN, GRU and LSTM models for both two outcomes were not different and were
180 approximately at the average of BS obtained from other models.

181 The AUROC of the models for DSS was generally higher than those of the models developed
182 for the combined endpoint, but with wider confidence intervals. Both logistic regression and
183 ANN using laboratory predictors measured on later day of illness (i.e., day 4 of illness) achieved
184 higher AUROC compared with the models with predictors measured at early of illness (either
185 day 2 or 3 of illness). The AUROC of the models with predictors measured at enrolment was
186 comparable with models developed using laboratory predictors measured at day 4 of illness.

187



188

189

190

191

192

Figure 5.4. P-value of the coefficient of groups by clinical endpoints (A for DSS and B for the combined endpoints of moderate plasma leakage and/or DSS) on the specified day of illness obtained from the multivariable models with adjustment of age and sex. The red dotted horizontal lines at 2.995, 4.605 and 6.908 represent the p-value at 0.05, 0.01, and 0.001 significance threshold on the log scale. HCT: haematocrit; PLT: platelet count; WBC: white blood cell count; NC: neutrophil count; LC: lymphocyte count; AST: aspartate transaminase; ALT: alanine transaminase; ALB; serum albumin

Although, the AUROC of models with RNNs, GRU and LSTM were similar to the AUROC of the models with predictors measured at enrolment and on day 4 of illness, the RNN-based models had wider confidence intervals. There were no differences in AUROC across models with RNNs, GRU and LSTM.

In terms of calibration performance, the central estimates of both CITL and calibration slope were close to 0 and 1, respectively. However, models with logistic regression provided more consistent performance compared with the calibration performance obtained from machine learning-based models. Both the CITL and calibration slope of logistic regression models were considerably closer to 0 and 1, respectively, compared with the machine learning models. The confidence interval of calibration performance of ANN was significantly wider than for the other models, including RNN, GRU and LSTM.

The accuracy of RNN, GRU and LSTM was higher than those obtained with the logistic regression models and ANN, with predictors measured on each day of illness. The models for DSS achieved higher accuracy, but with wider confidence interval compared with the accuracy of the models for the combined endpoint. The sensitivity and specificity of RNN, GRU and LSTM were similar to the performance of logistic regression models using laboratory predictors measured on day 4 of illness or enrolment. Models with laboratory predictors measured on later day of illness achieved better sensitivity and specificity while models with laboratory predictors measured on enrolment and on day 4 of illness also obtained the better performance.

5.4 Discussion

In this chapter, I described the temporal dynamics of laboratory data routinely collected during the febrile phase of dengue illness. The dynamics of platelet count, AST, ALT, and serum albumin were statistically different between the group of patients with and without the combined endpoint of moderate plasma leakage and/or DSS. Only platelet count and serum albumin were statistically different between the groups of patients with and without DSS. I also found that platelet count and liver function test (AST and ALT) were the earliest indicators

of developing DSS and the combined endpoint of moderate plasma leakage and/or DSS. The other laboratory parameters, including haematocrit, showed statistical difference between the two groups during the late phase of illness, which is close to the critical phase.

Platelet count is a well-known predictor of progression to severe dengue disease and this parameter is also well-established as a warning sign in the WHO classifications. This finding agrees with our previous findings, from the systematic review and meta-analysis described in Chapter 2, which showed that platelet count in the febrile phase was associated with progression to severe disease (severe dengue and/or dengue haemorrhagic fever). Platelet count was significantly different between groups with and without DSS and/or the combined endpoint on day 3 of illness for the combined endpoint and day 4 of illness for DSS, which is one or two days before the critical phase, which typically occurs in day 4 to 6 of illness. This finding is also in line with previous analyses of the dynamics of platelet in dengue^{96, 153, 154}.

Liver function enzymes (AST and ALT) were also significantly different between the groups of patients with and without the combined endpoints, but not for DSS. Higher AST and ALT was found more commonly amongst patients with the combined endpoints from day 3 of illness. This finding is also supported by our systematic review and meta-analysis as well as by previous studies^{116, 155, 156}. Elevated liver enzymes can be caused by multiple factors, including indirect hepatocellular injury from dengue virus, hypoxic damage due to impaired hepatic perfusion from plasma leakage, or immune-mediated injury^{52, 116}. Serum albumin was significantly different between groups of patients with and without DSS and the combined endpoints. Serum albumin starts being significantly different on day 5 of illness in both clinical endpoints in multiple logistic regression analysis, which agrees with a previous study¹⁵⁵. Amongst patients with DSS and/or the combined endpoint, serum albumin decreased more rapidly than in the group of patients without DSS and/or the combined endpoint. Also this finding agrees with the result obtained from our systematic review and meta-analysis¹¹⁶. Therefore, monitoring liver function enzymes (AST and ALT) and serum albumin, in addition to haematocrit and platelet count during the febrile phase, could assist clinicians to identify patient at risk of progression to severe disease (i.e., DSS and the combined endpoints of moderate plasma leakage and/or DSS).

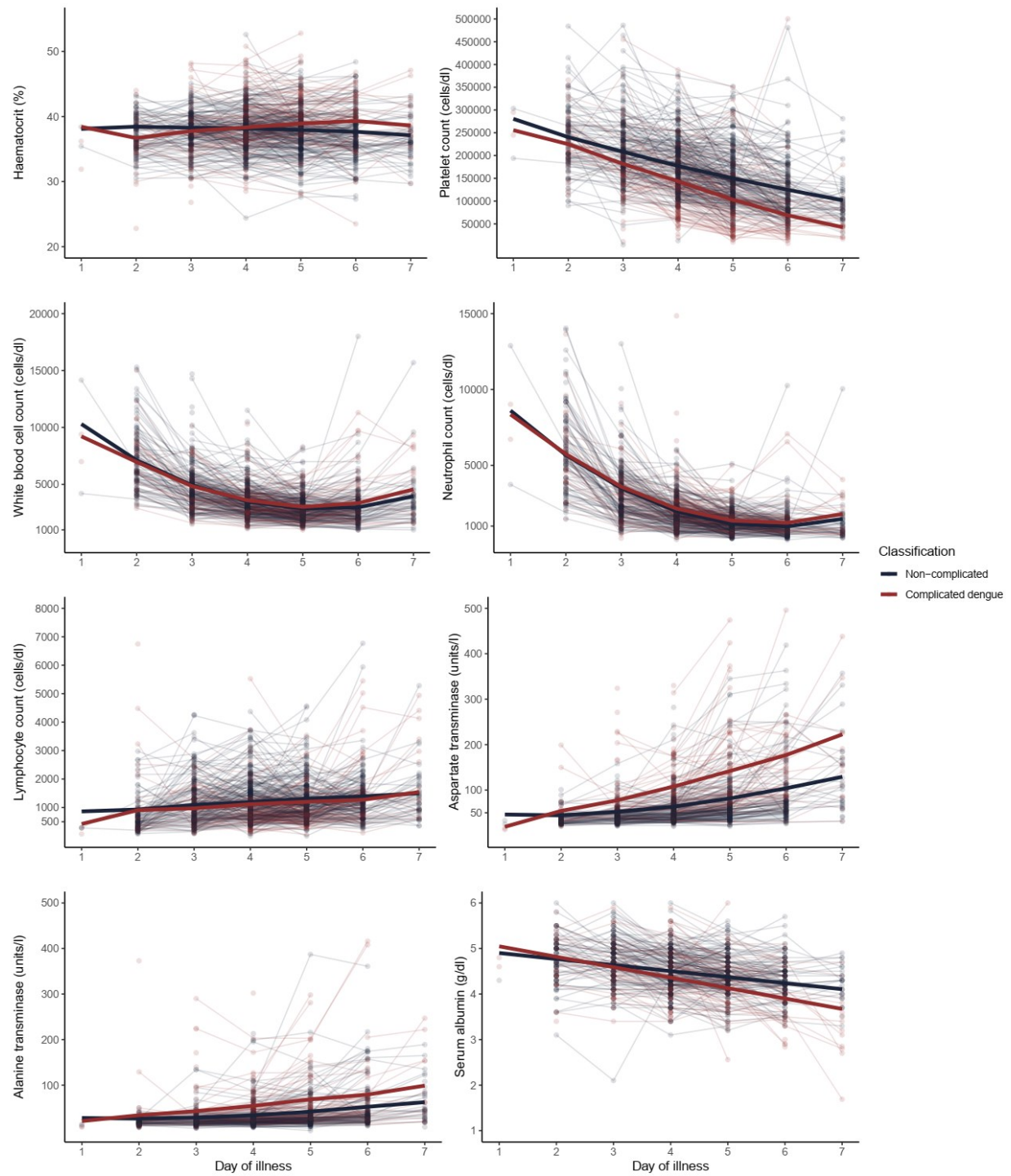


Figure 5.5. Individual trajectories of clinical parameters for all patients who developed (pale red) and did not develop (pale navy) the combined endpoint between day 1 and day 7 of illness. Bold lines display the estimated mean of each clinical parameter from the selected marginal models with Equation 3.4 for the other variables.

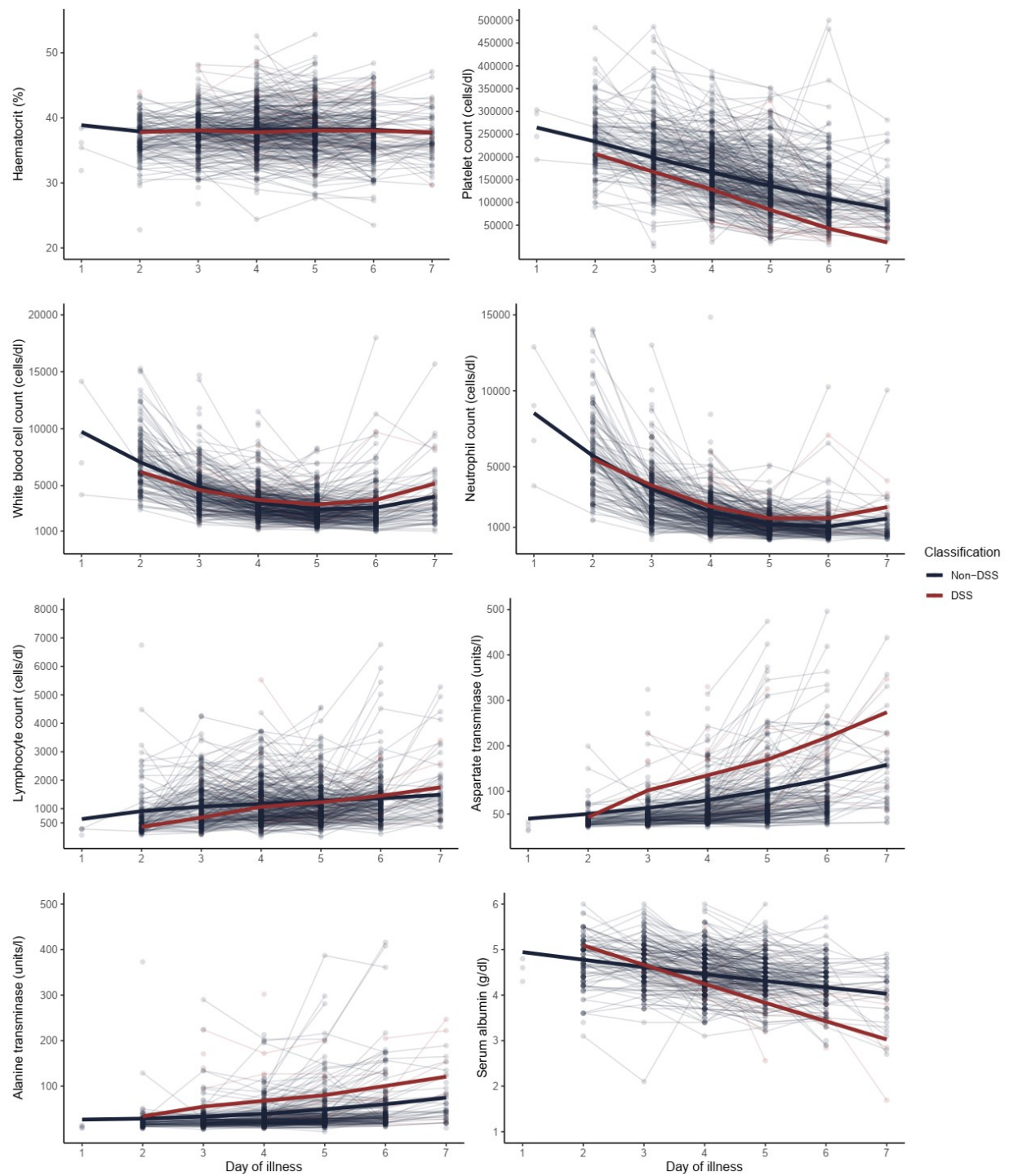


Figure 5.6. Individual trajectories of clinical parameters for patients who developed (pale red) and did not develop (pale navy) DSS between day 1 and day 7 of illness. Bold lines display the estimated mean of each clinical parameter from the selected marginal models with Equation 3.1 for serum albumin and with Equation 3.4 for the other variables.

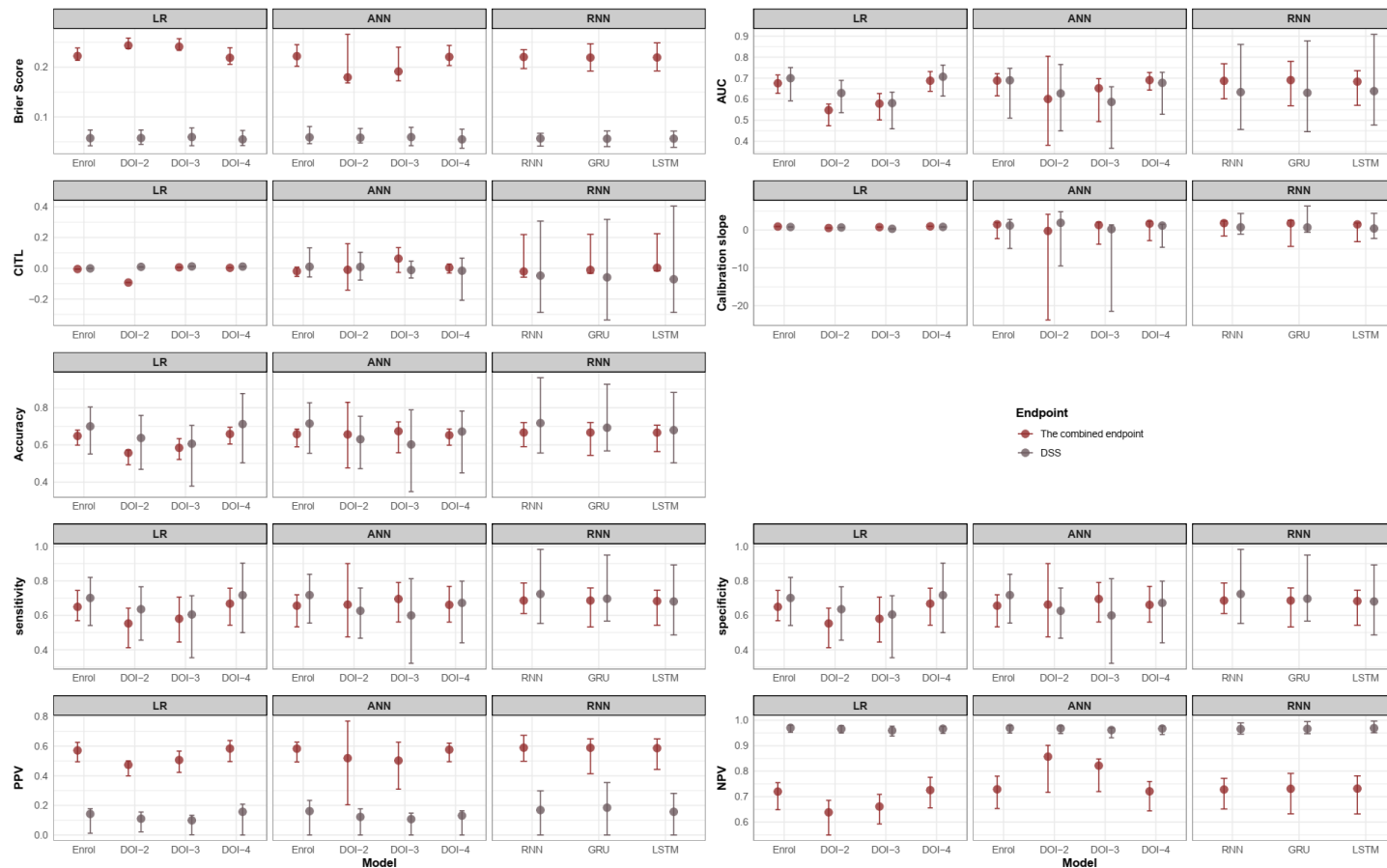


Figure 5.7. Predictive performance (Brier Score, AUROC, CITL, Calibration slope, accuracy, sensitivity, specificity, PPV, and NPV) of the risk prediction models for DSS and the combined endpoints of moderate plasma leakage and/or DSS using logistic regression (LR), artificial neural network (ANN), and recurrent neural network (RNN)-based models (RNN, gate recurrent unit (GRU), and long-short term memory (LSTM))

Regarding the performance of the models, we found that BS was significant higher for models targeting DSS. This result could be affected by class imbalance, which is why other metrics need to be used, including discrimination and calibration performance. The AUROC, accuracy, sensitivity, and specificity of models gained better performance where predictors measured close to the critical phase were used. This suggests that laboratory parameters change and become more different between the groups as severe disease develops, especially closely to the critical phase. Despite this, the sequential measurement of laboratory data is still useful and essential in triage of patients with dengue infection during the febrile phase. The predictive performance of RNN-based models was lower than that of logistic regression with laboratory predictors measured at enrolment or on each day of illness from day 2 to 4. Given the limited number of laboratory measurements and the relatively small sample size, complex and sophisticated models such RNN not always improve the performance compared to simpler models. However, RNN-based models represent a new promising avenue which will need to be further explored with continuous physiology monitoring data, which are expected to become widely available in the close future¹⁵⁷.

The analysis presented in this chapter has limitations. The main one is the relatively small sample size of the data and therefore the limited number of DSS cases. Another limitation is the relatively limited frequency of the measurements. In addition, because most of the patients were enrolled on day 3 to 4 of illness, a limited number of laboratory variables measured on day 1 to 3 of illness were available. Reduced sample sizes limit the powers to identify relationships and to evaluate the utility of RNN-based models. While this chapter focused on RNN, simpler statistical models such as partly conditional models or joint models could be explored in the future. In this chapter, model validation was performed using only internal validation and no external validation due to the lack of available prospective data in other settings. Lastly, real world outpatient setting data might be different from the cohort study data used in the chapter. Specifically, the proportion of patients with the combined endpoint analysed in this study might represent an overestimate of the real frequency of the combined endpoint and the availability and monitoring frequency of the laboratory measurements could be different from the one observed in the data analysed in this chapter.

In conclusion, from the analysis conducted in this chapter shows that the laboratory measurements monitored sequentially and widely available in dengue patients changed in time throughout the illness. We found evidence that most changes in the laboratory data unfold relatively close to the critical phase. However, monitoring platelet count, AST, ALT and serum albumin can help to identify patient at high risk of severe disease. Given the limited time window of observation, RNN-based models provided inferior performance than simpler models such as multiple logistic regression or ANN. In the future, once that more data obtained from larger scale studies or data of continuous monitoring become available, it will be critical to re-evaluate the use of complex and sophisticated RNN-based models for predicting the progression to severe dengue from the early phases of the illness.

Chapter 6 Conclusions

6.1 Summary and contributions

6.1.1 Risk predictors of the progression to severe dengue disease

Dengue poses an enormous disease burden across the tropics and subtropics. As no effective dengue vaccine or specific anti-viral drug currently exists, the early prediction of the progression to severe disease plays a vital role in patient triage, recognising patients who are at a higher risk of progression to severe disease and need close monitoring in hospitals. Although the prognostic factors of dengue severity have already been studied, there is limited consistency in the results due to the small sample sizes of the studies. To overcome the issue, systematic review and meta-analysis studies were conducted. Previous systematic reviews and meta-analyses considered clinical parameters collected during the whole-time course of the illness, did not differentiate the febrile phase, which occurs during the first 1 to 4 days since fever onset, from the later clinical phases when severe manifestations occur. Since clinical signs and symptoms, patient laboratory parameters, and viral-related biomarkers vary markedly across the time-course of the dengue disease, previous reviews and meta-analyses may have failed to identify significant associations between early prognostic factors and progression to a severe form of the disease.

This thesis identifies early predictors of progression to severe dengue disease based on both the existing literature and comprehensive datasets of two perspective cohort studies conducted in Thailand and Vietnam. The systematic review and meta-analysis described in Chapter 2 reveals the risk factors associated with severe dengue and dengue haemorrhagic fever from clinical data measured within 96 hours after fever onset. This specification intends to identify the risk factors observable in the early phase of the illness and the use of these risk factors in outpatient settings, where patients are first seen and evaluated. In our systematic review and meta-analysis¹¹⁶, we found that female patients had a higher risk of developing severe dengue than male patients, but this association was not observed among children. Pre-existing comorbidities associated with the severe form of the disease included diabetes,

hypertension, renal disease, and cardiovascular disease. I also found that there were clinical features during the febrile phase associated with progression to the severe phase of the disease including vomiting, abdominal pain and tenderness, spontaneous or mucosal bleeding, and the presence of clinical fluid accumulation. During the first 4 days after fever onset, we found that platelet count and serum albumin were lower, and aminotransferase concentrations (AST and ALT) were higher among individuals who progressed to severe disease. Dengue virus serotype 2 was found to be associated with severe disease in children but not in adults. Secondary infections were also associated with more frequent progression to severe disease compared with primary infections.

Chapter 4 presents the results obtained from the analysis of two prospective cohort studies conducted in Vietnam and Thailand and identifies the predictors associated with two clinical endpoints, DSS and the combined endpoint of moderate plasma leakage and/or DSS. In this analysis, the predictors were selected based on (i) the results of the systematic review and meta-analysis (summarised in Chapter 2), (ii) discussion with dengue experts, and (iii) lasso selection. Haematocrit, platelet count, lymphocyte count, and AST were found to be early predictors associated with the development of DSS; and the presence of abdominal pain or tenderness, vomiting, mucosal bleeding, white blood cell count, lymphocyte count, platelet count, aspartate aminotransferase, and serum albumin were found to be associated with the combined endpoint. The risk predictors identified using the Vietnamese datasets in Chapter 4 are in line with the results of the systematic review and meta-analysis conducted in Chapter 2.

Since the clinical progression of dengue is dynamic and adverse events typically unfold at a time close to the critical period, I analysed the temporal dynamics of the laboratory data collected prior to the development of the two clinical endpoints (DSS and the combined endpoints of moderate plasma leakage and/or DSS) in Chapter 5. Chapter 5 provides a better understanding of the dynamics of clinical progression, and laboratory variables associated with progression to the severe disease during the febrile phase. Platelet count and serum albumin amongst patients who subsequently developed either DSS or the combined endpoint decreased significantly faster compared to amongst patients who did not develop the

endpoints. In addition, liver enzymes (AST and ALT) of the group of patients who subsequently developed the combined endpoint increased significantly faster compared with those of the patients who did not. Platelet count and liver enzymes (AST and ALT) were also found to be one of the earliest indicators of the progression to severe disease, both for DSS and the combined endpoint.

This thesis provides solid evidence that early predictors associated with the progression to severe dengue disease exist, and these are backed by the existing literature and the observations collected in the cohort studies. The findings of this thesis support the monitoring of the warning signs described in the 2009 WHO guidelines for dengue. In addition to the warning signs, the results of this thesis reinforce the usefulness of testing and monitoring platelet count, AST, ALT, and serum albumin during the febrile phase, to aid the early detection of progression to severe forms of dengue. These findings provide strong evidence for clinicians who work in primary care settings. Because the predictors identified in this thesis are widely available in the context of clinical practice in outpatient settings, the findings could be useful for frontline clinicians who evaluate and triage patients during the early phase of disease. Some parameters such as secondary infections or infecting dengue serotype are associated with a higher risk of progression to severe disease, but until more robust point-of-care diagnostics are developed, incorporating immune status into early prediction models is difficult. The thesis also makes a strong call for point of care testing of the infecting serotype and immune status during the febrile phase of the illness, which can inform the early identification of patients at high risk of severe dengue.

6.1.2 Individualised risk prediction of the progression to severe dengue disease

This thesis uses data collected in two prospective cohort studies conducted independently in two different countries for the development and validation of risk prediction models. Even though the two datasets were found to differ in terms of data relatedness, due to different patient follow-up protocols and study design to detect clinical outcomes, both studies recruited children with laboratory-confirmed dengue infection presenting at outpatient settings within the first 72 hours of fever onset. Most clinical parameters and disease

progression measures recorded in both datasets followed the WHO guidelines^{5, 21}. This allowed us to use the Vietnamese dataset, which provides a larger sample size, as training dataset and the Thai data as an external validation dataset to assess the reproducibility and validity of the developed models. The risk prediction models were developed using a variety of regression and machine learning techniques with outpatient data collected within the first 3 days of illness. The optimal models were selected according to various predictive performance metrics including the Brier Score, discrimination and calibration performance, and decision curve analysis in Chapter 4. This thesis also provides a case study on the model development using RNNs involving sequential laboratory data measured during the febrile phase in Chapter 5.

Clinical and laboratory data measured on the day of enrolment (within 72 hours from fever onset) from two prospective cohort studies independently conducted in Vietnam and Thailand were used to develop and assess the performance of 6 predictive models: logistic regression, multivariable logistic regression with lasso selection, random forest, extreme gradient boosted classification, a support vector machine and an artificial neural network. The predictive performance was assessed using the Brier score, the area under the receiver operating characteristics curves (AUROC), the calibration in the large (CITL), the calibration slope, the calibration plots, and decision curve analysis. The sensitivity, specificity, positive predictive values, and negative predictive values of the models were also estimated. The model using logistic regression with lasso selection turned out to be the optimal model according to all metrics. Amongst the machine learning models, XGB and ANN provided the optimal performance.

This thesis also provides a case study on the development of models using RNN and sequential laboratory data collected during the febrile phase. The performance of these models was compared with logistic regression and ANN using laboratory data measured at enrolment as well as on each day of illness, taken separately, during the febrile phase. The performance of the models using laboratory data collected on the close day to the critical phase was higher compared with the models using earlier data for both clinical endpoints. The performance of the models developed using the data collected mainly on day 3 and 4 of illness was

comparable to the performance of the models using laboratory parameters measured on day 4 of the illness, where the critical phase typically occurs on day 5. The uncertainty of RNN-based models was considerably larger than the uncertainty observed with ANN and logistic regression. The use of GRU and LSTM models did not improve the model's predictive performance, although it did increase the uncertainty of the predictive performance.

The main contribution of this thesis is the development of risk prediction models for the two clinical endpoints using outpatient data inclusive of a variety of models. These models were validated on internal and external datasets, in the latter case using independent data from different countries in dengue endemic areas. The results of this thesis suggest that clinical risk prediction models produce more accurate predictions than using the warning signs check list because the former not only predict the classification outputs but also the individualised risk of developing the clinical endpoints. Using the individualised risk prediction models in clinical practice could be promising to assist clinicians to triage patients in the early phase of disease. In addition, since the clinical parameters vary as the illness progresses, the use of risk prediction models applied to sequential data indicates trends of progression to severe disease. Monitoring the trends could be also useful for patients who are evaluated daily in the outpatient settings.

Although standard recommendations and guidance for the development of clinical risk prediction models are already available and are well established, the quality of the model development and validation varies in the literature. This thesis follows the standard guidance and recommendations for model development^{133, 134, 151, 158}. To enhance transparency and reproducibility of the development and validation of risk prediction models, this thesis illustrates all steps of model development both for regression and machine learning techniques. This includes selecting the candidate and final predictors included in the prediction models, defining the prediction outcome, pre-processing the data, and dealing with missing data, hyperparameter optimisation, model calibration (for machine learning models) and by internal and external validation.

Finally, this thesis explores the development of models using the dynamic changes of laboratory data measured prior to the critical phase of illness. The finding that RNN models produce similar predictive performance to the models developed using data collected at enrolment is likely due to the small sample size of the dataset and data granularity. This could suggest that sophisticated and complex models do not always achieve better performance given the small data set, even though they do increase the variability of the predictive model performance.

6.2 Future Research Direction

Platelet count, AST, ALT, and serum albumin were found to be the predictors of progression to severe disease in addition to being warning signs. The understanding of the pathophysiology of these parameters has the potential to reinforce the prognostic ability and use of these predictors in clinical practice. In addition to the conventional predictors that have been mentioned in this thesis, novel biomarkers and advances in technology that allow for point of care tests such as ultrasonography or echocardiography in primary healthcare settings can enable better triage and recognising patients at high risk of progression to severe disease, particularly in the early phase of the illness. The further development of technology used to detect and monitor continuously clinical and laboratory data will be also useful for triage patients with dengue throughout the febrile phase and during the transitional period between the febrile phase and the critical phase. In fact, progression to severe disease often unfolds at a time close to the critical period. Including novel biomarkers and continuous monitoring will strengthen the efficacy of triage, thus reducing the demand on the healthcare system and improving the healthcare resource allocation.

The RNN models involving sequential laboratory data developed in Chapter 5 performed as well as the models developed on data collected at enrolment. The datasets that we had available had a relatively small number of children developing the clinical endpoints and monitored a relatively small number of laboratory indicators. In future work, it will be interesting to further explore the use of these models especially when larger datasets, obtained from large-scale perspective cohort studies, become available.

Although this thesis used data from well-designed clinical studies conducted in multiple hospitals, which ensures representation over a wide range of dengue patients, the future use of models to inform triage and clinical practice will need to be validated in new settings and populations. The new populations can be different to the ones used in this thesis and involve different contexts of care and different endemic diseases. This thesis focused on patients with laboratory confirmed dengue infection. In future work, it will be useful to test the performance of the risk prediction models including in febrile patients without laboratory confirmation of dengue infection, which may thus include infections of co-circulating tropical diseases such as malaria, leptospirosis, or typhus. The further evaluation of the prediction models on real world data in clinical practice in different contexts is important not only for model validation, but also to measure the impact of the predictions on the patient outcomes and cost effectiveness of care. To assess the impact of the utilisation of prediction models to inform triage and hospitalisation, further studies are needed. Cluster randomised trials would be the most effective design to assess the impacts of prediction models, where centres (clinical practices) are assigned to the intervention (in this context, the use of predictive models) or not, and the benefits of utilising risk prediction models to predict the risk of developing DSS or plasma leakage can be measured in terms of bed use and the number of unplanned or unnecessary admission compared with the control group¹⁵⁹, for instance. As an alternative design, which would be less costly and time consuming, a prospective before-after study, comparing the individual outcomes between those managed without the prediction models in an earlier period and those managed in a later period after the introduction of the prediction model, could be conducted in order to assess the impact of the models in real clinical practice.

In future work, together with my team in Thailand, and the team at Imperial College London, I am planning to implement and test the use of the risk prediction models developed in this thesis in real clinical practice. I am planning to continue the research developed in this thesis in 2 phases: (i) validate the models using real, in-house, already available data collected from the clinical practice conducted in the Hatyai Regional hospital and Prince of Songkha University Hospital in Thailand, and (ii) assess the impact of the model implementation in

clinical practice in primary healthcare hospitals in the south of Thailand. The aims are (i) to evaluate the performance of the models in new contexts and (ii) assess whether the prediction models can assist clinicians in complex decision-making to increase the precision of disease management and patient safety. The model implementation will have to be designed jointly by developers and end users, i.e., healthcare professionals in primary care settings, and will have to be affordable and meet the users' computer literacy. The implementation of the models will also aim at developing trust and acceptance by its users, specifically clinicians in primary healthcare settings who engage with the patient's primary evaluation while also triaging patients during the febrile phase. Ultimately, I hope that the statistical models and results developed in this thesis will inform triage and the optimal use of resources, to alleviate the burden of dengue hospitalisation and offer the best possible care to the patients.

References

1. Ferguson NM, Rodriguez-Barraquer I, Dorigatti I, Mier YT-RL, Laydon DJ, Cummings DA. Benefits and risks of the Sanofi-Pasteur dengue vaccine: Modeling optimal deployment. *Science*. 2016;353(6303):1033-6.
2. Hadinegoro SR, Arredondo-Garcia JL, Capeding MR, Deseda C, Chotpitayasunondh T, Dietze R, et al. Efficacy and Long-Term Safety of a Dengue Vaccine in Regions of Endemic Disease. *N Engl J Med*. 2015;373(13):1195-206.
3. Bhatt S, Gething PW, Brady OJ, Messina JP, Farlow AW, Moyes CL, et al. The global distribution and burden of dengue. *Nature*. 2013;496(7446):504-7.
4. Shepard DS, Undurraga EA, Halasa YA, Stanaway JD. The global economic burden of dengue: a systematic analysis. *Lancet Infect Dis*. 2016;16(8):935-41.
5. World Health Organization SPfRaTiTD. Dengue guidelines for diagnosis, treatment, prevention and control. new edition ed. Geneva: World Health Organization; 2009.
6. Guzman MG, Halstead SB, Artsob H, Buchy P, Farrar J, Gubler DJ, et al. Dengue: a continuing global threat. *Nat Rev Microbiol*. 2010;8(12 Suppl):S7-16.
7. Rosen L, Drouet MT, Deubel V. Detection of dengue virus RNA by reverse transcription-polymerase chain reaction in the liver and lymphoid organs but not in the brain in fatal human infection. *Am J Trop Med Hyg*. 1999;61(5):720-4.
8. Halstead SB, Nimmannitya S, Cohen SN. Observations Related to Pathogenesis of Dengue Hemorrhagic Fever .4. Relation of Disease Severity to Antibody Response and Virus Recovered. *Yale J Biol Med*. 1970;42(5):311-+.
9. Anderson KB, Gibbons RV, Cummings DAT, Nisalak A, Green S, Libraty DH, et al. A Shorter Time Interval Between First and Second Dengue Infections Is Associated With Protection From Clinical Illness in a School-based Cohort in Thailand. *J Infect Dis*. 2014;209(3):360-8.
10. Montoya M, Gresh L, Mercado JC, Williams KL, Vargas MJ, Gutierrez G, et al. Symptomatic versus inapparent outcome in repeat dengue virus infections is influenced by the time interval between infections and study year. *PLoS Negl Trop Dis*. 2013;7(8):e2357.
11. Chan KR, Wang X, Saron WAA, Gan ES, Tan HC, Mok DZL, et al. Cross-reactive antibodies enhance live attenuated virus infection for increased immunogenicity. *Nat Microbiol*. 2016;1:16164.
12. 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(6365):929-32.
13. Boonnak K, Slike BM, Donofrio GC, Marovich MA. Human FcγRIII cytoplasmic domains differentially influence antibody-mediated dengue virus infection. *J Immunol*. 2013;190(11):5659-65.
14. Rothman AL, Ennis FA. Immunopathogenesis of Dengue hemorrhagic fever. *Virology*. 1999;257(1):1-6.
15. Mongkolsapaya J, Dejnirattisai W, Xu XN, Vasanawathana S, Tangthawornchaikul N, Chairunsri A, et al. Original antigenic sin and apoptosis in the pathogenesis of dengue hemorrhagic fever. *Nat Med*. 2003;9(7):921-7.

16. Thiemmecca S, Tamdet C, Punyadee N, Prommool T, Songjaeng A, Noisakran S, et al. Secreted NS1 Protects Dengue Virus from Mannose-Binding Lectin-Mediated Neutralization. *J Immunol*. 2016;197(10):4053-65.
17. Avirutnan P, Punyadee N, Noisakran S, Komoltri C, Thiemmecca S, Auethavornanan K, et al. Vascular leakage in severe dengue virus infections: a potential role for the nonstructural viral protein NS1 and complement. *J Infect Dis*. 2006;193(8):1078-88.
18. Glasner DR, Ratnasiri K, Puerta-Guardo H, Espinosa DA, Beatty PR, Harris E. Dengue virus NS1 cytokine-independent vascular leak is dependent on endothelial glycocalyx components. *PLoS Pathog*. 2017;13(11):e1006673.
19. Modhiran N, Watterson D, Muller DA, Panetta AK, Sester DP, Liu L, et al. Dengue virus NS1 protein activates cells via Toll-like receptor 4 and disrupts endothelial cell monolayer integrity. *Sci Transl Med*. 2015;7(304):304ra142.
20. Wichmann O, Hongsiriwon S, Bowonwatanuwong C, Chotivanich K, Sukthana Y, Pukrittayakamee S. Risk factors and clinical features associated with severe dengue infection in adults and children during the 2001 epidemic in Chonburi, Thailand. *Trop Med Int Health*. 2004;9(9):1022-9.
21. World Health Organization. Dengue hemorrhagic fever: diagnosis, treatment, prevention and control. 2nd Edition ed. Geneva: World Health Organization; 1997.
22. Guzman MG, Kouri G. Dengue diagnosis, advances and challenges. *Int J Infect Dis*. 2004;8(2):69-80.
23. Tang KF, Ooi EE. Diagnosis of dengue: an update. *Expert Rev Anti Infect Ther*. 2012;10(8):895-907.
24. Zhang H, Li W, Wang J, Peng H, Che X, Chen X, et al. NS1-based tests with diagnostic utility for confirming dengue infection: a meta-analysis. *Int J Infect Dis*. 2014;26:57-66.
25. Namekar M, Ellis EM, O'Connell M, Elm J, Gurary A, Park SY, et al. Evaluation of a new commercially available immunoglobulin M capture enzyme-linked immunosorbent assay for diagnosis of dengue virus infection. *J Clin Microbiol*. 2013;51(9):3102-6.
26. Blacksell SD, Jarman RG, Gibbons RV, Tanganuchitcharnchai A, Mammen MP, Jr., Nisalak A, et al. Comparison of seven commercial antigen and antibody enzyme-linked immunosorbent assays for detection of acute dengue infection. *Clin Vaccine Immunol*. 2012;19(5):804-10.
27. Chong ZL, Sekaran SD, Soe HJ, Peramalah D, Rampal S, Ng CW. Diagnostic accuracy and utility of three dengue diagnostic tests for the diagnosis of acute dengue infection in Malaysia. *BMC Infect Dis*. 2020;20(1):210.
28. Barniol J, Gaczkowski R, Barbato EV, da Cunha RV, Salgado D, Martinez E, et al. Usefulness and applicability of the revised dengue case classification by disease: multi-centre study in 18 countries. *BMC Infect Dis*. 2011;11:106.
29. Kalayanaroj S. Dengue classification: current WHO vs. the newly suggested classification for better clinical application? *J Med Assoc Thai*. 2011;94 Suppl 3:S74-84.
30. Morra ME, Altibi AMA, Iqtadar S, Minh LHN, Elawady SS, Hallab A, et al. Definitions for warning signs and signs of severe dengue according to the WHO 2009 classification: Systematic review of literature. *Rev Med Virol*. 2018;28(4):e1979.
31. Wills BA, Nguyen MD, Ha TL, Dong TH, Tran TN, Le TT, et al. Comparison of three fluid solutions for resuscitation in dengue shock syndrome. *N Engl J Med*. 2005;353(9):877-89.

32. Simmons CP, Farrar JJ, Nguyen v V, Wills B. Dengue. *N Engl J Med*. 2012;366(15):1423-32.
33. Thomas L, Kaidomar S, Kerob-Bauchet B, Moravie V, Brouste Y, King JP, et al. Prospective observational study of low thresholds for platelet transfusion in adult dengue patients. *Transfusion*. 2009;49(7):1400-11.
34. Lye DC, Lee VJ, Sun Y, Leo YS. Lack of efficacy of prophylactic platelet transfusion for severe thrombocytopenia in adults with acute uncomplicated dengue infection. *Clin Infect Dis*. 2009;48(9):1262-5.
35. Tricou V, Minh NN, Van TP, Lee SJ, Farrar J, Wills B, et al. A randomized controlled trial of chloroquine for the treatment of dengue in Vietnamese adults. *PLoS Negl Trop Dis*. 2010;4(8):e785.
36. Nguyen NM, Tran CN, Phung LK, Duong KT, Huynh Hle A, Farrar J, et al. A randomized, double-blind placebo controlled trial of balapiravir, a polymerase inhibitor, in adult dengue patients. *J Infect Dis*. 2013;207(9):1442-50.
37. Whitehorn J, Nguyen CVV, Khanh LP, Kien DTH, Quyen NTH, Tran NTT, et al. Lovastatin for the Treatment of Adult Patients With Dengue: A Randomized, Double-Blind, Placebo-Controlled Trial. *Clin Infect Dis*. 2016;62(4):468-76.
38. Murray CJ, Vos T, Lozano R, Naghavi M, Flaxman AD, Michaud C, et al. Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet*. 2012;380(9859):2197-223.
39. Zhang F, Kramer CV. Corticosteroids for dengue infection. *Cochrane Database Syst Rev*. 2014(7):CD003488.
40. Suputtamongkol Y, Avirutnan P, Mairiang D, Angkasekwinai N, Niwattayakul K, Yamasmith E, et al. Ivermectin Accelerates Circulating Nonstructural Protein 1 (NS1) Clearance in Adult Dengue Patients: A Combined Phase 2/3 Randomized Double-blinded Placebo Controlled Trial. *Clin Infect Dis*. 2021.
41. 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(1):127-34.
42. Lee CC, Hsu HC, Chang CM, Hong MY, Ko WC. Atypical presentations of dengue disease in the elderly visiting the ED. *Am J Emerg Med*. 2013;31(5):783-7.
43. Rowe EK, Leo YS, Wong JG, Thein TL, Gan VC, Lee LK, et al. Challenges in dengue fever in the elderly: atypical presentation and risk of severe dengue and hospital-acquired infection [corrected]. *PLoS Negl Trop Dis*. 2014;8(4):e2777.
44. Hammond SN, Balmaseda A, Perez L, Tellez Y, Saborio SI, Mercado JC, et al. Differences in dengue severity in infants, children, and adults in a 3-year hospital-based study in Nicaragua. *Am J Trop Med Hyg*. 2005;73(6):1063-70.
45. Gamble J, Bethell D, Day NP, Loc PP, Phu NH, Gartside IB, et al. Age-related changes in microvascular permeability: a significant factor in the susceptibility of children to shock? *Clin Sci (Lond)*. 2000;98(2):211-6.
46. Lee IK, Liu JW, Yang KD. Clinical and laboratory characteristics and risk factors for fatality in elderly patients with dengue hemorrhagic fever. *Am J Trop Med Hyg*. 2008;79(2):149-53.

47. Opal SM, Girard TD, Ely EW. The immunopathogenesis of sepsis in elderly patients. *Clin Infect Dis.* 2005;41 Suppl 7:S504-12.
48. Huy NT, Van Giang T, Thuy DH, Kikuchi M, Hien TT, Zamora J, et al. Factors associated with dengue shock syndrome: a systematic review and meta-analysis. *PLoS Negl Trop Dis.* 2013;7(9):e2412.
49. Halstead SB. Observations related to pathogenesis of dengue hemorrhagic fever. VI. Hypotheses and discussion. *Yale J Biol Med.* 1970;42(5):350-62.
50. Tomashek KM, Wills B, See Lum LC, Thomas L, Durbin A, Leo YS, et al. Development of standard clinical endpoints for use in dengue interventional trials. *PLoS Negl Trop Dis.* 2018;12(10):e0006497.
51. World Health Organization. Regional Office for South-East Asia. Comprehensive guidelines for prevention and control of dengue and dengue haemorrhagic fever. Rev. and expanded. ed. New Delhi, India: World Health Organization Regional Office for South-East Asia; 2011. xiv, 196 p. p.
52. Martina BEE, Koraka P, Osterhaus ADME. Dengue Virus Pathogenesis: an Integrated View. *Clin Microbiol Rev.* 2009;22(4):564-+.
53. Gil L, Martinez G, Tapanes R, Castro O, Gonzalez D, Bernardo L, et al. Oxidative stress in adult dengue patients. *Am J Trop Med Hyg.* 2004;71(5):652-7.
54. Diamond MS, Edgil D, Roberts TG, Lu B, Harris E. Infection of human cells by dengue virus is modulated by different cell types and viral strains. *J Virol.* 2000;74(17):7814-23.
55. Srikiatkachorn A, Krautrachue A, Ratanaprakarn W, Wongtapradit L, Nithipanya N, Kalayanarooj S, et al. Natural history of plasma leakage in dengue hemorrhagic fever: a serial ultrasonographic study. *Pediatr Infect Dis J.* 2007;26(4):283-90; discussion 91-2.
56. Michels M, Sumardi U, de Mast Q, Jusuf H, Puspita M, Dewi IM, et al. The predictive diagnostic value of serial daily bedside ultrasonography for severe dengue in Indonesian adults. *PLoS Negl Trop Dis.* 2013;7(6):e2277.
57. 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(1):2-9.
58. 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.
59. Vuong NL, Quyen NTH, Tien NTH, Tuan NM, Kien DTH, Lam PK, et al. Higher plasma viremia in the febrile phase is associated with adverse dengue outcomes irrespective of infecting serotype or host immune status: an analysis of 5642 Vietnamese cases. *Clin Infect Dis.* 2020.
60. Puerta-Guardo H, Glasner DR, Espinosa DA, Biering SB, Patana M, Ratnasiri K, et al. Flavivirus NS1 Triggers Tissue-Specific Vascular Endothelial Dysfunction Reflecting Disease Tropism. *Cell Rep.* 2019;26(6):1598-613 e8.
61. Pichyangkul S, Endy TP, Kalayanarooj S, Nisalak A, Yongvanitchit K, Green S, et al. A blunted blood plasmacytoid dendritic cell response to an acute systemic viral infection is associated with increased disease severity. *J Immunol.* 2003;171(10):5571-8.
62. Bittencourt MD, Martial J, Cabie A, Thomas L, Cesaire R. Decreased Peripheral Dendritic Cell Numbers in Dengue Virus Infection. *J Clin Immunol.* 2012;32(1):161-72.

63. Luplertlop N, Misse D, Bray D, Deleuze V, Gonzalez JP, Leardkamolkarn V, et al. Dengue-virus-infected dendritic cells trigger vascular leakage through metalloproteinase overproduction (vol 7, pg 1176, 2006). *Embo Rep*. 2006;7(12):1290-.
64. Green S, Pichyangkul S, Vaughn DW, Kalayanarooj S, Nimmannitya S, Nisalak A, et al. Early CD69 expression on peripheral blood lymphocytes from children with dengue hemorrhagic fever. *J Infect Dis*. 1999;180(5):1429-35.
65. Mongkolsapaya J, Duangchinda T, Dejnirattisai W, Vasanawathana S, Avirutnan P, Jairungsri A, et al. T cell responses in dengue hemorrhagic fever: are cross-reactive T cells suboptimal? *J Immunol*. 2006;176(6):3821-9.
66. Bozza FA, Cruz OG, Zagne SM, Azeredo EL, Nogueira RMR, Assis EF, et al. Multiplex cytokine profile from dengue patients: MIP-1beta and IFN-gamma as predictive factors for severity. *BMC Infect Dis*. 2008;8.
67. Furuta T, Murao LA, Lan NT, Huy NT, Huong VT, Thuy TT, et al. Association of mast cell-derived VEGF and proteases in Dengue shock syndrome. *PLoS Negl Trop Dis*. 2012;6(2):e1505.
68. Nascimento EJM, Silva AM, Cordeiro MT, Brito CA, Gil LHV, Braga-Neto U, et al. Alternative Complement Pathway Deregulation Is Correlated with Dengue Severity. *PLoS One*. 2009;4(8).
69. Bethell DB, Flobbe K, Phuong CXT, Day NPJ, Phuong PT, Buurman WA, et al. Pathophysiologic and prognostic role of cytokines in dengue hemorrhagic fever. *J Infect Dis*. 1998;177(3):778-82.
70. Koraka P, Murgue B, Deparis X, van Gorp ECM, Setiati TE, Osterhaus ADME, et al. Elevation of soluble VCAM-1 plasma levels in children with acute dengue virus infection of varying severity. *J Med Virol*. 2004;72(3):445-50.
71. Khongphatthanayothin A, Phumaphuti P, Thongchalprasit K, Poovorawan Y. Serum levels of sICAM-1 and sE-Selectin in patients with dengue virus infection. *Jpn J Infect Dis*. 2006;59(3):186-8.
72. Azeredo EL, Zagne SM, Alvarenga AR, Nogueira RMR, Kubelka CF, de Oliveira-Pinto LM. Activated peripheral lymphocytes with increased expression of cell adhesion molecules and cytotoxic markers are associated with dengue fever disease. *Mem Inst Oswaldo Cruz*. 2006;101(4):437-49.
73. Butthep P, Chunhakan S, Tangnararatchakit K, Yoksan S, Pattanapanyasat K, Chuansumrit A. Elevated soluble thrombomodulin in the febrile stage related to patients at risk for dengue shock syndrome. *Pediatr Infect Dis J*. 2006;25(10):894-7.
74. Michels M, van der Ven AJAM, Djamiatun K, Fijnheer R, de Groot PG, Griffioen AW, et al. Imbalance of Angiopoietin-1 and Angiopoietin-2 in Severe Dengue and Relationship with Thrombocytopenia, Endothelial Activation, and Vascular Stability. *Am J Trop Med Hyg*. 2012;87(5):943-6.
75. Suwanto S, Sasmono RT, Sinto R, Ibrahim E, Suryamin M. Association of Endothelial Glycocalyx and Tight and Adherens Junctions With Severity of Plasma Leakage in Dengue Infection. *J Infect Dis*. 2017;215(6):992-9.
76. Vuong NL, Le Duyen HT, Lam PK, Tam DTH, Vinh Chau NV, Van Kinh N, et al. C-reactive protein as a potential biomarker for disease progression in dengue: a multi-country observational study. *BMC Med*. 2020;18(1):35.

77. Valero N, Espina LM, Anez G, Torres E, Mosquera JA. Short report: Increased level of serum nitric oxide in patients with dengue. *Am J Trop Med Hyg.* 2002;66(6):762-4.
78. Trairatvorakul P, Chongsrisawat V, Ngamvasinont D, Asawarachun D, Nantasook J, Poovorawan Y. Serum nitric oxide in children with Dengue infection. *Asian Pac J Allergy Immunol.* 2005;23(2-3):115-9.
79. Yacoub S, Lam PK, Huynh TT, Nguyen Ho HH, Dong Thi HT, Van NT, et al. Endothelial Nitric Oxide Pathways in the Pathophysiology of Dengue: A Prospective Observational Study. *Clin Infect Dis.* 2017;65(9):1453-61.
80. Thein TL, Wong J, Leo YS, Ooi EE, Lye D, Yeo TW. Association Between Increased Vascular Nitric Oxide Bioavailability and Progression to Dengue Hemorrhagic Fever in Adults. *J Infect Dis.* 2015;212(5):711-4.
81. Yacoub S, Lam PK, Vu le HM, Le TL, Ha NT, Toan TT, et al. Association of Microvascular Function and Endothelial Biomarkers With Clinical Outcome in Dengue: An Observational Study. *J Infect Dis.* 2016;214(5):697-706.
82. Moulton SL, Mulligan J, Srikiatkachorn A, Kalayanarooj S, Grudic GZ, Green S, et al. State-of-the-art monitoring in treatment of dengue shock syndrome: a case series. *J Med Case Rep.* 2016;10(1):233.
83. Yacoub S, Trung TH, Lam PK, Thien VHN, Hai DHT, Phan TQ, et al. Cardio-haemodynamic assessment and venous lactate in severe dengue: Relationship with recurrent shock and respiratory distress. *PLoS Negl Trop Dis.* 2017;11(7):e0005740.
84. Tan KW, Tan B, Thein TL, Leo YS, Lye DC, Dickens BL, et al. Dynamic dengue haemorrhagic fever calculators as clinical decision support tools in adult dengue. *Trans R Soc Trop Med Hyg.* 2020;114(1):7-15.
85. Park S, Srikiatkachorn A, Kalayanarooj S, Macareo L, Green S, Friedman JF, et al. Use of structural equation models to predict dengue illness phenotype. *PLoS Negl Trop Dis.* 2018;12(10):e0006799.
86. Phakhounthong K, Chaovalit P, Jittamala P, Blacksell SD, Carter MJ, Turner P, et al. Predicting the severity of dengue fever in children on admission based on clinical features and laboratory indicators: application of classification tree analysis. *BMC Pediatr.* 2018;18(1):109.
87. Low GKK, Gan SC, Zainal N, Naidu KD, Amin-Nordin S, Khoo CS, et al. The predictive and diagnostic accuracy of vascular endothelial growth factor and pentraxin-3 in severe dengue. *Pathog Glob Health.* 2018;112(6):334-41.
88. Sreenivasan P, S G, K S. Development of a Prognostic Prediction Model to Determine Severe Dengue in Children. *Indian J Pediatr.* 2018;85(6):433-9.
89. Fernandez E, Smieja M, Walter SD, Loeb M. A retrospective cohort study to predict severe dengue in Honduran patients. *BMC Infect Dis.* 2017;17(1):676.
90. Lam PK, Tam DT, Dung NM, Tien NT, Kieu NT, Simmons C, et al. A Prognostic Model for Development of Profound Shock among Children Presenting with Dengue Shock Syndrome. *PLoS One.* 2015;10(5):e0126134.
91. Pongpan S, Wisitwong A, Tawichasri C, Patumanond J, Namwongprom S. Development of dengue infection severity score. *ISRN Pediatr.* 2013;2013:845876.
92. Nguyen MT, Ho TN, Nguyen VV, Nguyen TH, Ha MT, Ta VT, et al. An Evidence-Based Algorithm for Early Prognosis of Severe Dengue in the Outpatient Setting. *Clin Infect Dis.* 2017;64(5):656-63.

93. Pang J, Lindblom A, Tolfvenstam T, Thein TL, Naim AN, Ling L, et al. Discovery and Validation of Prognostic Biomarker Models to Guide Triage among Adult Dengue Patients at Early Infection. *PLoS One*. 2016;11(6):e0155993.
94. Keene DJ, Schluskel MM, Thompson J, Hagan DA, Williams MA, Byrne C, et al. Prognostic models for identifying risk of poor outcome in people with acute ankle sprains: the SPRAINED development and external validation study. *Health Technol Assess*. 2018;22(64):1-112.
95. Huy NT, Thao NT, Ha TT, Lan NT, Nga PT, Thuy TT, et al. Development of clinical decision rules to predict recurrent shock in dengue. *Crit Care*. 2013;17(6):R280.
96. Brasier AR, Ju H, Garcia J, Spratt HM, Victor SS, Forshey BM, et al. A three-component biomarker panel for prediction of dengue hemorrhagic fever. *Am J Trop Med Hyg*. 2012;86(2):341-8.
97. Potts JA, Gibbons RV, Rothman AL, Srikiatkachorn A, Thomas SJ, Supradish PO, et al. Prediction of dengue disease severity among pediatric Thai patients using early clinical laboratory indicators. *PLoS Negl Trop Dis*. 2010;4(8):e769.
98. Potts JA, Thomas SJ, Srikiatkachorn A, Supradish PO, Li W, Nisalak A, et al. Classification of dengue illness based on readily available laboratory data. *Am J Trop Med Hyg*. 2010;83(4):781-8.
99. Lee VJ, Lye DC, Sun Y, Leo YS. Decision tree algorithm in deciding hospitalization for adult patients with dengue haemorrhagic fever in Singapore. *Trop Med Int Health*. 2009;14(9):1154-9.
100. Lee VJ, Lye DC, Sun Y, Fernandez G, Ong A, Leo YS. Predictive value of simple clinical and laboratory variables for dengue hemorrhagic fever in adults. *J Clin Virol*. 2008;42(1):34-9.
101. Lam PK, Ngoc TV, Thu Thuy TT, Hong Van NT, Nhu Thuy TT, Hoai Tam DT, et al. The value of daily platelet counts for predicting dengue shock syndrome: Results from a prospective observational study of 2301 Vietnamese children with dengue. *PLoS Negl Trop Dis*. 2017;11(4):e0005498.
102. Soo KM, Khalid B, Ching SM, Tham CL, Basir R, Chee HY. Meta-analysis of biomarkers for severe dengue infections. *PeerJ*. 2017;5:e3589.
103. Zhang H, Zhou YP, Peng HJ, Zhang XH, Zhou FY, Liu ZH, et al. Predictive symptoms and signs of severe dengue disease for patients with dengue fever: a meta-analysis. *Biomed Res Int*. 2014;2014:359308.
104. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gotzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *BMJ*. 2009;339:b2700.
105. Riley RD, Moons KGM, Snell KIE, Ensor J, Hooft L, Altman DG, et al. A guide to systematic review and meta-analysis of prognostic factor studies. *BMJ*. 2019;364:k4597.
106. Sangkaew S, Ming D, Boonyasiri A, Honeyford K, Kalayanaroj S, Yacoub S, et al. Predictors of dengue severity: a systematic review and meta-analysis 2019. Available from: https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42018093363.
107. Hayden JA, van der Windt DA, Cartwright JL, Cote P, Bombardier C. Assessing bias in studies of prognostic factors. *Ann Intern Med*. 2013;158(4):280-6.
108. Balduzzi S, Rücker G, Schwarzer G. How to perform a meta-analysis with R: a practical tutorial. *Evidence-Based Mental Health*. 2019:153-60.

109. Viechtbauer W. Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*. 2010;36:1-48.
110. Luo D, Wan X, Liu J, Tong T. Optimally estimating the sample mean from the sample size, median, mid-range, and/or mid-quartile range. *Stat Methods Med Res*. 2018;27(6):1785-805.
111. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ*. 2003;327(7414):557-60.
112. Cumpston M, Li T, Page MJ, Chandler J, Welch VA, Higgins JP, et al. Updated guidance for trusted systematic reviews: a new edition of the Cochrane Handbook for Systematic Reviews of Interventions. *Cochrane Database Syst Rev*. 2019;10:ED000142.
113. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ*. 1997;315(7109):629-34.
114. Duval S, Tweedie R. Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*. 2000;56(2):455-63.
115. Crippa A, Orsini N. Multivariate Dose-Response Meta-Analysis: The dosresmeta R Package. *Journal of Statistical Software*. 2016;72(Cs1).
116. Sangkaew S, Ming D, Boonyasiri A, Honeyford K, Kalayanaroj S, Yacoub S, et al. Risk predictors of progression to severe disease during the febrile phase of dengue: a systematic review and meta-analysis. *Lancet Infect Dis*. 2021.
117. Hsieh TH, Tsai TT, Chen CL, Shen TJ, Jhan MK, Tseng PC, et al. Senescence in Monocytes Facilitates Dengue Virus Infection by Increasing Infectivity. *Front Cell Infect Microbiol*. 2020;10:375.
118. Lee IK, Hsieh CJ, Lee CT, Liu JW. Diabetic patients suffering dengue are at risk for development of dengue shock syndrome/severe dengue: Emphasizing the impacts of co-existing comorbidity(ies) and glycemic control on dengue severity. *J Microbiol Immunol Infect*. 2020;53(1):69-78.
119. Pecoits-Filho R, Heimbürger O, Barany P, Suliman M, Fehrman-Ekholm I, Lindholm B, et al. Associations between circulating inflammatory markers and residual renal function in CRF patients. *Am J Kidney Dis*. 2003;41(6):1212-8.
120. Aznar-Salatti J, Escolar G, Cases A, Gomez-Ortiz G, Anton P, Castillo R, et al. Uraemic medium causes endothelial cell dysfunction characterized by an alteration of the properties of its subendothelial matrix. *Nephrol Dial Transplant*. 1995;10(12):2199-204.
121. Guzman MG, Kouri G, Bravo J, Valdes L, Vazquez S, Halstead SB. Effect of age on outcome of secondary dengue 2 infections. *Int J Infect Dis*. 2002;6(2):118-24.
122. Olkowski S, Forshey BM, Morrison AC, Rocha C, Vilcarromero S, Halsey ES, et al. Reduced risk of disease during postsecondary dengue virus infections. *J Infect Dis*. 2013;208(6):1026-33.
123. Wikramaratna PS, Simmons CP, Gupta S, Recker M. The effects of tertiary and quaternary infections on the epidemiology of dengue. *PLoS One*. 2010;5(8):e12347.
124. Nisalak A, Clapham HE, Kalayanaroj S, Klungthong C, Thaisomboonsuk B, Fernandez S, et al. Forty Years of Dengue Surveillance at a Tertiary Pediatric Hospital in Bangkok, Thailand, 1973-2012. *Am J Trop Med Hyg*. 2016;94(6):1342-7.
125. Kotepui M, Kotepui KU. Prevalence and laboratory analysis of malaria and dengue co-infection: a systematic review and meta-analysis. *BMC Public Health*. 2019;19(1):1148.

126. Wilder-Smith A, Leong WY. Risk of severe dengue is higher in patients with sickle cell disease: a scoping review. *J Travel Med.* 2019;26(1).
127. Manamperi M, Jayamanne BDW, Somaratne T, Perera N, Fernando L. Predictive value of persistent NS1 antigen positivity beyond 3rd day for dengue haemorrhagic fever in Sri Lankan children. *BMC Res Notes.* 2019;12(1):214.
128. Lee YH, Leong WY, Wilder-Smith A. Markers of dengue severity: a systematic review of cytokines and chemokines. *J Gen Virol.* 2016;97(12):3103-19.
129. Harrell JFE. *Regression Modeling Strategies : With Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis.* Cham: Springer International Publishing : Imprint: Springer,; 2015.
130. Zhang Z. Multiple imputation with multivariate imputation by chained equation (MICE) package. *Ann Transl Med.* 2016;4(2):30.
131. Huque MH, Carlin JB, Simpson JA, Lee KJ. A comparison of multiple imputation methods for missing data in longitudinal studies. *BMC Med Res Methodol.* 2018;18(1):168.
132. Wahl S, Boulesteix AL, Zierer A, Thorand B, van de Wiel MA. Assessment of predictive performance in incomplete data by combining internal validation and multiple imputation. *BMC Med Res Methodol.* 2016;16(1):144.
133. Moons KG, Kengne AP, Woodward M, Royston P, Vergouwe Y, Altman DG, et al. Risk prediction models: I. Development, internal validation, and assessing the incremental value of a new (bio)marker. *Heart.* 2012;98(9):683-90.
134. Moons KG, Kengne AP, Grobbee DE, Royston P, Vergouwe Y, Altman DG, et al. Risk prediction models: II. External validation, model updating, and impact assessment. *Heart.* 2012;98(9):691-8.
135. Van Calster B, McLernon DJ, van Smeden M, Wynants L, Steyerberg EW, Topic Group 'Evaluating diagnostic t, et al. Calibration: the Achilles heel of predictive analytics. *BMC Med.* 2019;17(1):230.
136. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Making.* 2006;26(6):565-74.
137. Vickers AJ, Van Calster B, Steyerberg EW. Net benefit approaches to the evaluation of prediction models, molecular markers, and diagnostic tests. *BMJ.* 2016;352:i6.
138. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics.* 2011;12:77.
139. Andri Signorell et mult. al. DescTools: Tools for Descriptive Statistics. {2017},.
140. Heinze G, Wallisch C, Dunkler D. Variable selection - A review and recommendations for the practicing statistician. *Biom J.* 2018;60(3):431-49.
141. Friedman J, Hastie T, Tibshirani R. Regularization Paths for Generalized Linear Models via Coordinate Descent. *J Stat Softw.* 2010;33(1):1-22.
142. Yan Y. *rBayesianOptimization: Bayesian Optimization of Hyperparameters.* 2016.
143. Debray TP, Vergouwe Y, Koffijberg H, Nieboer D, Steyerberg EW, Moons KG. A new framework to enhance the interpretation of external validation studies of clinical prediction models. *J Clin Epidemiol.* 2015;68(3):279-89.
144. Fitzmaurice GM, Laird NM, Ware JH. *Applied longitudinal analysis.* 2nd ed. Hoboken, N.J.: Wiley; 2011. xxv, 701 p. p.

145. Halekoh U, Højsgaard S, Yan J. The R Package geepack for Generalized Estimating Equations. *Journal of Statistical Software*. 2006;15/2:1--11.
146. Ghatak A. Deep learning with R. New York, NY: Springer Berlin Heidelberg; 2019. pages cm p.
147. Karpathy A. The unreasonable effectiveness of recurrent neural networks [Available from: <http://karpathy.github.io/2015/05/21/rnn-effectiveness/>].
148. Hochreiter S, Schmidhuber J. Long Short-Term Memory. *Neural Comput*. 1997;9(8):1735-80.
149. Olah C. Understanding LSTM Networks 2015 [Available from: <https://colah.github.io/posts/2015-08-Understanding-LSTMs/>].
150. Cho K, Merrienboer Bv, Bahdanau D. On the Properties of Neural Machine Translation: Encoder–Decoder Approaches. 2014.
151. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): the TRIPOD statement. *Ann Intern Med*. 2015;162(1):55-63.
152. Gamliel A, Ziv-Baran T, Siegel RM, Fogelman Y, Dubnov-Raz G. Using weight-for-age percentiles to screen for overweight and obese children and adolescents. *Prev Med*. 2015;81:174-9.
153. Azin FR, Goncalves RP, Pitombeira MH, Lima DM, Branco IC. Dengue: profile of hematological and biochemical dynamics. *Rev Bras Hematol Hemoter*. 2012;34(1):36-41.
154. Srichaikul T, Nimmannitya S. Haematology in dengue and dengue haemorrhagic fever. *Baillieres Best Pract Res Clin Haematol*. 2000;13(2):261-76.
155. Fernando S, Wijewickrama A, Gomes L, Punchihewa CT, Madusanka SD, Dissanayake H, et al. Patterns and causes of liver involvement in acute dengue infection. *BMC Infect Dis*. 2016;16:319.
156. Trung DT, Thao le TT, Hien TT, Hung NT, Vinh NN, Hien PT, et al. Liver involvement associated with dengue infection in adults in Vietnam. *Am J Trop Med Hyg*. 2010;83(4):774-80.
157. Ming DK, Sangkaew S, Chanh HQ, Nhat PTH, Yacoub S, Georgiou P, et al. Continuous physiological monitoring using wearable technology to inform individual management of infectious diseases, public health and outbreak responses. *Int J Infect Dis*. 2020;96:648-54.
158. Steyerberg EW, Vergouwe Y. Towards better clinical prediction models: seven steps for development and an ABCD for validation. *Eur Heart J*. 2014;35(29):1925-+.
159. Kappen TH, van Klei WA, van Wolfswinkel L, Kalkman CJ, Vergouwe Y, Moons KGM. Evaluating the impact of prediction models: lessons learned, challenges, and recommendations. *Diagn Progn Res*. 2018;2:11.

Appendix

Relevant Publications

1. **Sangkaew S**, Ming D, Boonyasiri A, Honeyford K, Kalayanarooj S, Yacoub S, et al. Risk predictors of progression to severe disease during the febrile phase of dengue: a systematic review and meta-analysis. *Lancet Infect Dis*. 2021.
2. **Sangkaew S**, Ming D, Hernandez B, Herrero P, Suntarattiwong P, Kalayanarooj S, et al. Early individualised risk prediction using routine clinical data for children with acute dengue in outpatient settings in Vietnam and Thailand. (Being submitted)
3. Ming DK, **Sangkaew S**, Chanh HQ, Nhat PTH, Yacoub S, Georgiou P, et al. Continuous physiological monitoring using wearable technology to inform individual management of infectious diseases, public health and outbreak responses. *Int J Infect Dis*. 2020;96:648-54.

Presentations

1. Poster presentation: Clinical predictors of dengue severity in the febrile phase: a systematic review. (2019) European Congress of Clinical Microbiology & Infectious Diseases (ECCMID) 29th.
2. Poster presentation: Enhancing risk prediction of progression to severe disease during the febrile phase of dengue: A systematic review and meta-analysis. (2020) International Society for Infectious Diseases 19th (Cancellation due to the unfolding COVID-19 pandemic)

Awards

1. A travel grant for a research visit at Oxford University Clinical Research Unit (OUCRU) in Ho Chi Minh City, Vietnam, between November and December 2020 by London Centre of Neglected Infectious Disease Research

2. Funding for a research visit at Queen Sirikit National Institute of Child Health (QSNICH) in Bangkok, Thailand, in January 2021 by Faculty of Medicine Dean's Internship award

Appendix 1. Registered protocol of a systematic review on Prospero

Citation

Sorawat Sangkaew, Adhiratha Boonyasiri, Damien Ming, Kate Honeyford, Siripen Kalayanaroj, Sophie Yacoub, Ilaria Dorigatti, Alison Holmes. Predictors of dengue severity: a systematic review and meta-analysis. PROSPERO 2018 CRD42018093363 Available from: https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42018093363

Review question

Among individuals with dengue infection during the febrile phase, what are clinical and laboratory parameters associated with dengue severity?

Searches

- The following databases will be searched: PubMed/MEDLINE, EMBASE, Global Health, CINAHL, Web of Science, and Scopus. Reference lists of the included articles will be screened for relevant articles. Grey literature will be reviewed using OpenGrey. Dissertations and theses will be searched for in ProQuest Dissertations and Theses Global. Search terms will be tailored for each database, will use MeSH terms where applicable and will be checked by a medical reference librarian to ensure the syntax is correct and complete.

- No date limits will be imposed on the search

- Only studies written in English.

Types of study to be included

Inclusion criteria

- Original articles conducted with comparative or analytic study designs

- Studies written in English

Exclusion criteria

- Review articles

- Descriptive studies (case report, case series, ecological study)

- In vitro studies

- Seroepidemiological study

- Books (whole book, series, chapters), conference abstract, comments, notes and letters

Condition or domain being studied

Reducing dengue morbidity/mortality and its burden on healthcare systems requires the early detection of individuals at risk of dengue deterioration. Currently, the prediction of dengue deterioration during the early febrile phase is extremely challenging. Predictions are based on clinical features and measurements of serial laboratory data, which require clinical experience and often a subjective evaluation of both case management strategy prior to the onset of the critical phase. For individuals predicted to be at a higher risk of dengue deterioration, hospitalization for meticulous monitoring are required. Monitoring is invasive for patients and labour intensive for healthcare personnel.

Recent, technological advances allow the timely detection of pathogens and the continuous monitoring of clinical changes. A vast amount of literature related to the predictors of the risk of developing severe dengue

has been published. However, many studies were conducted with small sample sizes and the results are not always consistent. We aim to summarise the conventional and novel clinical and laboratory parameters associated with dengue deterioration including emerging promising biomarkers and imaging.

Participants/population

Inclusion criteria

- Dengue infection confirmed by laboratory diagnosis according to the WHO guideline in 1997 and 2009

Exclusion criteria

- Participants who have had the manifestations of severe dengue (evidence of severe plasma leakage, hemodynamic instability and organ impairment).

Intervention(s), exposure(s)

Inclusion criteria

- Clinical features (signs and symptoms), laboratory parameters, and imaging techniques which collected during the febrile phase

Exclusion criteria

- Genetic factors.

Comparator(s)/control

Not applicable.

Context

Main outcome(s)

- Clinical and laboratory parameters associated with dengue severity
- Association between those parameters and dengue severity

Measures of effect

Mean differences, odds ratio, relative risks and confidence interval

Additional outcome(s)

None.

Measures of effect

Not applicable.

Data extraction (selection and coding)

- De-duplicating the articles with Endnote software and manual removal. Two independent reviewers will screen titles and abstracts and select full-text articles yielded by the main searches against the screening and selection criteria. Disagreement will be resolved by the third reviewer

- A standardised data extraction form will be piloted on a selection of included studies. This form will then be used to extract data by the first reviewer then will be checked by one of the second and third reviewers. Disagreement in the extracted data will be resolved by discussion among the three reviewers.

- Authors, year of publication, year(s) of study, country, study design, number of participants (N), age range of participants, the clinical period of variable collection, associated factors, outcomes (DF, DHF, DSS, Dengue with warning signs, and severe dengue), the kinetics of parameters, how outcome was defined

(WHO guideline 1997, 2009, or both), unadjusted relative difference between dengue severity (risk ratio, odds ratio etc.), adjusted variables, adjusted relative difference between groups.

Risk of bias (quality) assessment

The first reviewer will assess the quality of the included full-text articles and double check with one of the two reviewers. The Quality in Prognostic Factor Studies (QUIPS) tool will be followed to assess the quality and risk of bias of the studies. Any discrepancies will be resolved by discussion or by the third reviewer if necessary.

Strategy for data synthesis

- If the data are heterogeneous, a descriptive synthesis based on the predictive parameters and clinical outcomes (dengue severity). In addition, potential biomarkers will be reported along with general study characteristics.

- If the data are relatively homogenous, meta-analyses will be performed by calculating the odds ratios or risk ratios of severe forms of dengue infection for each included study and presenting an overall pooled odds ratio with 95% confidence intervals.

Analysis of subgroups or subsets

Subgroup analysis will be performed for (1) different age groups (infancy, childhood, and adulthood), (2) high and low endemic regions, and (3) Discrepancy in dengue severity diagnosis according to WHO DENGUE GUIDELINE 1997 and 2009.

Contact details for further information

Sorawat Sangkaew
ss1116@ic.ac.uk

Organisational affiliation of the review

Health Protection Research Unit in Healthcare Associated Infection and Antimicrobial Resistance at Imperial College London, London, UK
<https://www.imperial.ac.uk/medicine/hpru-amr>

Review team members and their organisational affiliations

Dr Sorawat Sangkaew. Health Protection Research Unit in Healthcare Associated Infection and Antimicrobial Resistance at Imperial College London, London, UK
Assistant/Associate Professor Adhiratha Boonyasiri. Health Protection Research Unit in Healthcare Associated Infection and Antimicrobial Resistance at Imperial College London, London, UK
Dr Damien Ming. Health Protection Research Unit in Healthcare Associated Infection and Antimicrobial Resistance at Imperial College London, London, UK
Dr Kate Honeyford. Dr Foster Unit, Department of Primary Care and Public Health, Imperial College London, London, UK
Professor Siripen Kalayanaroj. Department of Pediatrics, Queen Sirikit National Institute of Child Health, Bangkok, Thailand
Dr Sophie Yacoub. Oxford University Clinical Research Unit, Wellcome Trust Major Overseas Programme, Ho Chi Minh City
Dr Ilaria Dorigatti. MRC Centre for Outbreak Analysis and Modelling, Department of Infectious Disease Epidemiology, Imperial College London
Professor Alison Holmes. Health Protection Research Unit in Healthcare Associated Infection and Antimicrobial Resistance at Imperial College London, London, UK

Type and method of review

Meta-analysis, Prognostic, Systematic review

Anticipated or actual start date

09 March 2018

Anticipated completion date

15 December 2018

Funding sources/sponsors

None

Conflicts of interest

None known

Language

English

Country

England

Stage of review

Review Completed not published

Subject index terms status

Subject indexing assigned by CRD

Subject index terms

Dengue; Humans; Severe Dengue

Date of registration in PROSPERO

27 April 2018

Date of first submission

19 April 2018

Stage of review at time of this submission

Stage	Started	Completed
Preliminary searches	Yes	Yes
Piloting of the study selection process	Yes	Yes
Formal screening of search results against eligibility criteria	Yes	Yes
Data extraction	Yes	Yes
Risk of bias (quality) assessment	Yes	Yes
Data analysis	Yes	Yes

Revision note

We added Prof. Siripen Kalayanaroj as a co-author because she has contributed and shared advices in terms of dengue in children. In addition, we have used the Quality in Prognostic Factor Studies (QUIPS) tool instead of JBI tools as the QUIPS was more specific to studies of prognostic factors.

The record owner confirms that the information they have supplied for this submission is accurate and complete and they understand that deliberate provision of inaccurate information or omission of data may be construed as scientific misconduct.

The record owner confirms that they will update the status of the review when it is completed and will add publication details in due course.

Versions

27 April 2018
30 October 2018
28 November 2019

Appendix 2. Search terms used in the literature review. We searched for all combinations of the terms in column 1 and 2 (e.g. “dengue shock syndrome AND risk factor(s)”, “dengue shock syndrome AND risk parameter(s)”, “dengue shock syndrome AND risk variable(s)”, “dengue shock AND risk factor(s)” etc).

Complicated dengue	Predictors
- dengue shock syndrome- dengue shock - dengue haemorrhagic fever/dengue - haemorrhagic fever - dengue with warning sign(s) - severe dengue - complicated dengue - dengue severity	- risk factor(s) - risk parameter(s) - risk variable(s) - predictive/predicting factor(s) - predictive/predicting parameter(s) - prognostic factor(s) - prognostic parameter(s) - associated factor(s) - associated parameter(s) - associated characteristic(s) - risk(s) - prognosis - biomarker(s) - marker(s)

Appendix 3. Search syntaxes applied in the Ovid platform for searching MEDLINE database. We searched for all combinations of the terms in column 1 and 2. The explode command (exp) was used to retrieve records that contained the specific term and its subheadings (e.g. Severe dengue, risk factors, and prognosis). Terms in the columns were search using multiple-purpose command (.mp), searching keywords in a set of fields usually including Title, Original Title, Abstract, and Subject Heading.

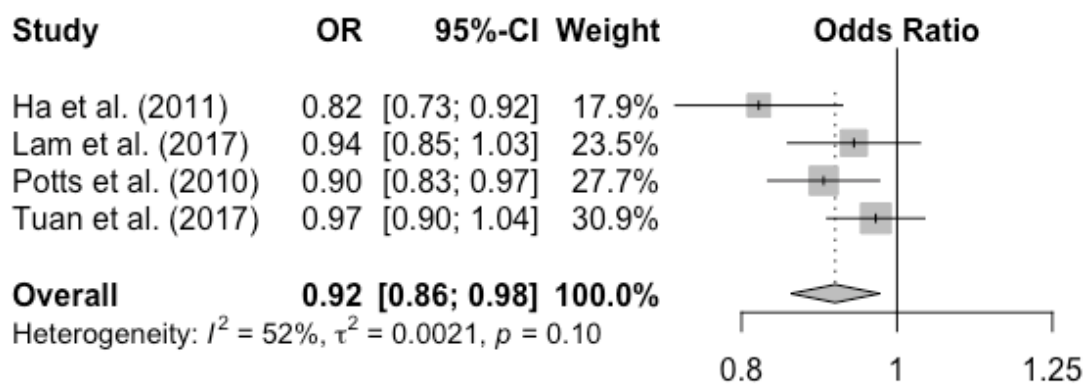
Terms for dengue with progression to severe disease	Terms for predictors
• [exp Severe Dengue/] • severe dengue.mp. • dengue shock syndrome*.mp. • dengue h?emorrhagic fever.mp • dengue with warning sign*.mp. • dengue severity.mp. • dengue death.mp. • dengue death.mp. • dengue shock.mp. • 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9	• [exp Risk Factors/] • risk factor*.mp. • [exp RISK/] • risk*.mp. • [exp PROGNOSIS/] • prognosis.mp. • [exp BIOMARKERS/] • biomarker*.mp. • marker*.mp. • ((risk* or predictive or prognostic or associated) adj (factor* or parameter* or characteristic* or variable*)).mp. [mp=title, abstract, full text, caption text] • predictor*.mp. • prediction.mp. • 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22

Terms for dengue with progression to severe disease	Terms for predictors
10 AND 23	

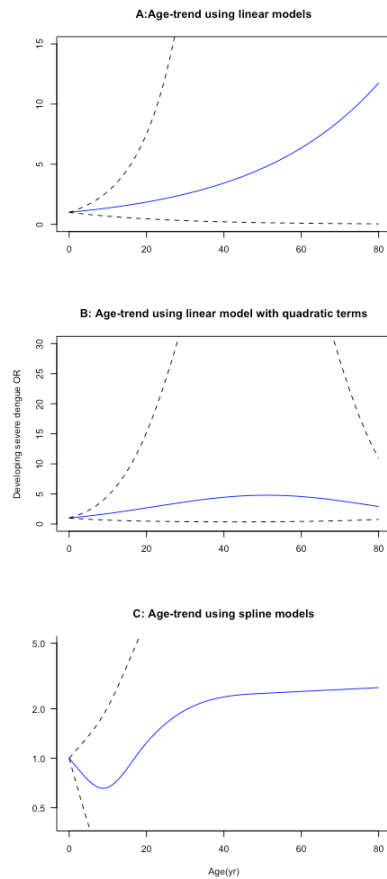
Appendix 4. Template used for extracting information from the studies included in the systematic review.

Paper ID
Title
Reviews
First author's surname
Year of publication
Study design
Data collection
Year of patient recruitment (Month/Year - Month/year)
Country where patients were recruited
Hospital/centre names where patient was recruited
Population inclusion and exclusion criteria
Dengue classification 1: WHO 1997 (DS, DHF, and DSS) 2: WHO 2009 (DWS-, DWS+ SD)
When were samples or information collected
Reference group (e.g. a DF group)
Comparative group (i.e. DHF, DSS, DHF/DSS, DWS+/SD, or SD)
Continuous factors_1 name (unit)
Description (if applicability)
How were samples or information measured (if applicable)
Sample size of a reference groups
Sample size of a comparative group
Ref: Mean, Median, Mean difference
Ref: SD, Upper 95% CI, Lower 95% CI, IQR, Upper IQR, Lower IQR
Ref: Odds ratio, Relative risk, Adjusted odds ratio, Adjusted relative risk
Ref: 95% CI of OR
Comp: Mean, Median, Mean difference
Comp: SD, Upper 95% CI, Lower 95% CI, IQR, Upper IQR, Lower IQR
p-value
Statistic test
Comp: Odds ratio, Relative risk, Adjusted odds ratio, Adjusted relative risk

Comp: 95% CI of OR
p-value
Statistic test
Adjusted variables
Categorical factors _1 name (unit):
Description (if necessary)
How were samples or information measured (if necessary)
Sample size of a reference groups
Sample size of a comparative group
Level 1: The number of non-events (in reference group)
Level 1: The number of events (in comparison group)
Level 2: The number of non-events (in reference group)
Level 2: The number of events (in comparison group)
Level 3: The number of non-events (in reference group)
Level 3: The number of events (in comparison group)
p-value
Statistic test
Comp: Odds ratio, Relative risk, Adjusted odds ratio, Adjusted relative risk
Comp: Upper 95% CI
Comp: Lower 95% CI
Comp: Odds ratio, Relative risk, Adjusted odds ratio, Adjusted relative risk
Comp: Upper 95% CI
Comp: Lower 95% CI
p-value
Statistic test



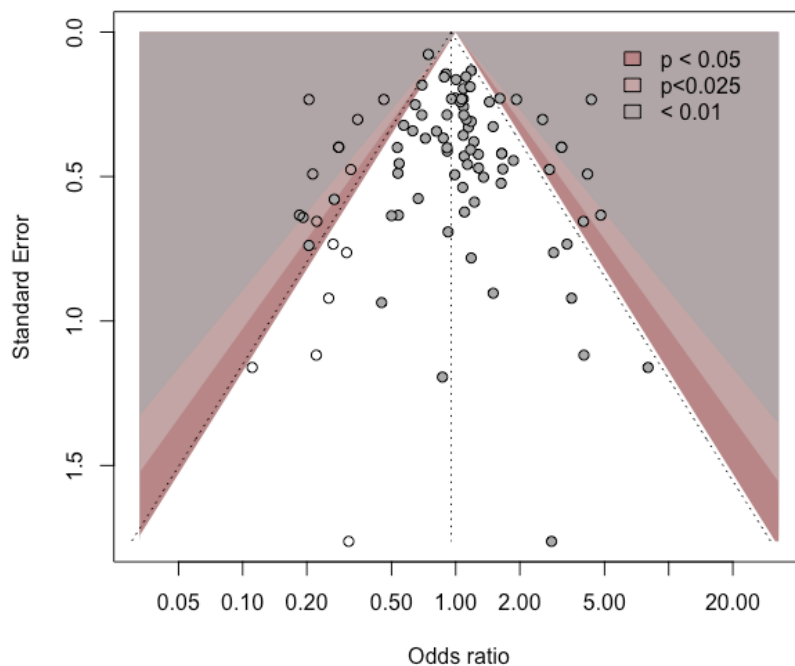
Appendix 5. Forest plot of associations from multivariable models between age and severe progression. Four studies conducted among children (≤ 18 years old) were included in the meta-analysis. The pooled OR was 0.92 (95% CI (0.86, 0.98)) and I^2 was 52%.



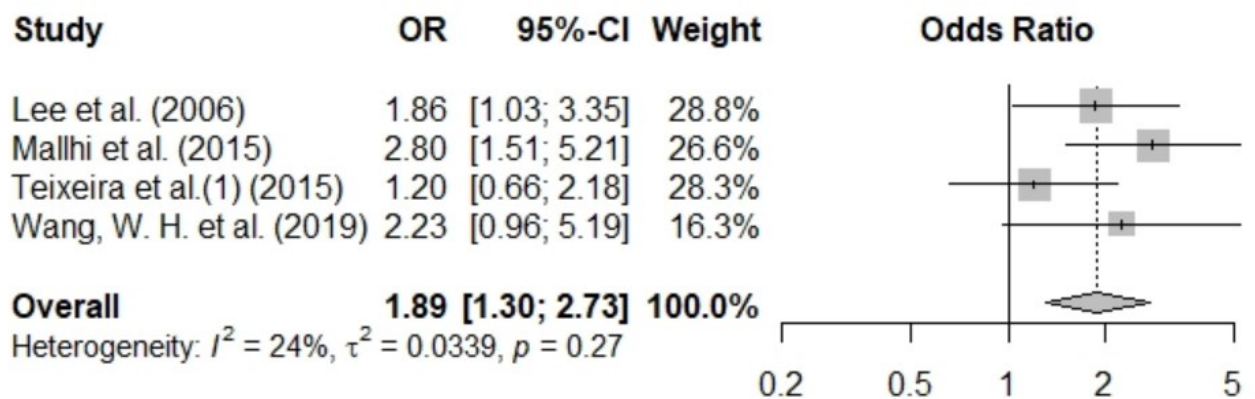
Appendix 6. Association between odd ratios of severe progression (y-axis) and age (x-axis) from dose-response meta-analysis based on linear models (A), linear models with quadratic terms (B), and spline models with knots at age 2, 9, 18, and 50 (C). The blue solid line indicates the central estimate, and the dotted line indicates the 95% CI. Model fitting parameters of linear and non-linear models estimating trends in the risk of developing severe manifestations with age in dose-response meta-analysis.

Appendix 7. Model fitting parameters of linear and non-linear models estimating trends in the risk of developing severe manifestations with age in dose-response meta-analysis.

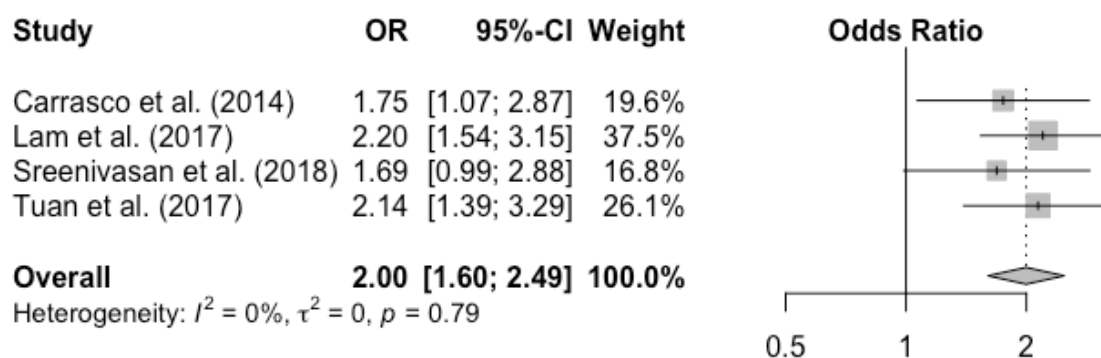
Model	AIC	BIC
Linear model	65.8	67.6
Linear model with quadratic terms	83.1	87.3
Spline model with 4 knots at age 2, 9, 18, and 50	71.5	78.4



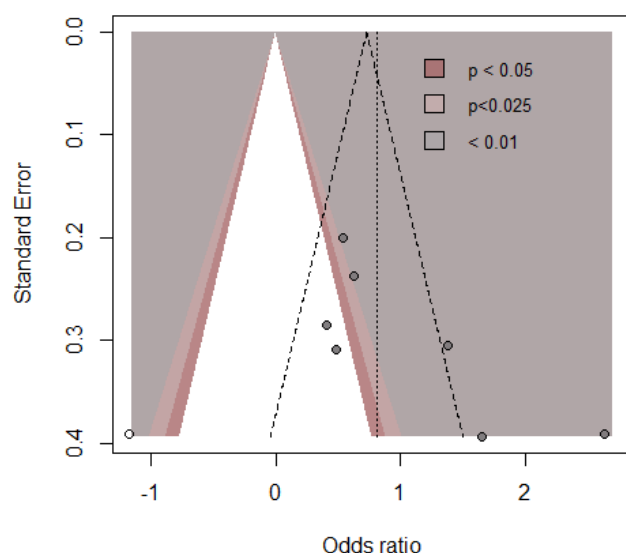
Appendix 8. Funnel plot for association between sex and severe progression. The grey dots indicate association of each study, and the clear dots indicate the imputed associations.



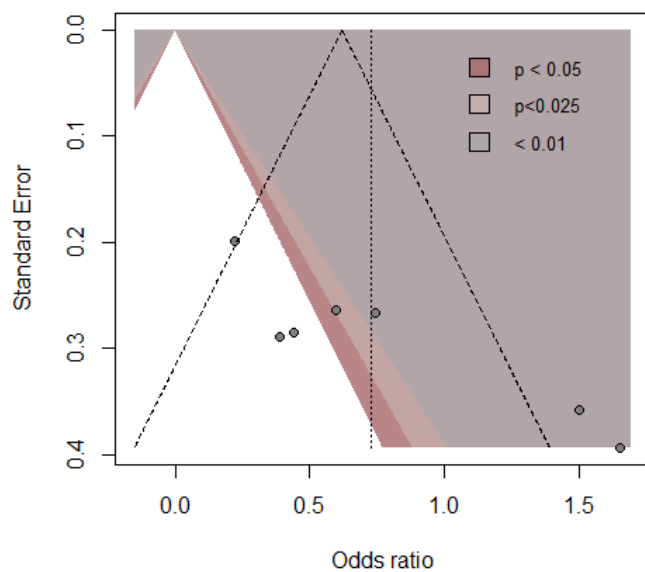
Appendix 9. Forest plot of associations from multivariable models between pre-existing diabetes mellitus and severe progression. Four studies were included in the meta-analysis. The pooled OR was 1.89 (95% CI (1.30, 2.73)) and I^2 was 24%.



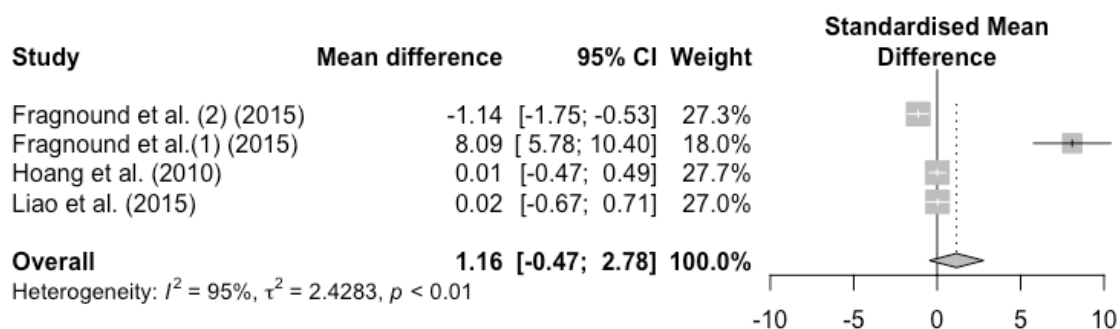
Appendix 10. Forest plot of associations from multivariable models between vomiting and progression to severe disease. Four studies were included in the meta-analysis. The pooled OR was 2.00 (95% CI (1.60, 0.2.49)) and I^2 was 0%.



Appendix 11. Funnel plot for association between AST and progression to severe disease. The grey dots indicate association of each study and the clear dots indicate the imputed associations. Dots on the coloured background indicate statistically significant ORs (red: p-value < 0.05, pink < 0.025 and brown p-value < 0.01); dots on the white background indicate non-significant ORs (p-value ≥ 0.05).



Appendix 12. Forest plot of associations between viraemia and progression to severe disease. Four studies were included in the meta-analysis. The pooled OR was 1.16 (95% CI (-0.47 to 2.78)) and I^2 was 95%.

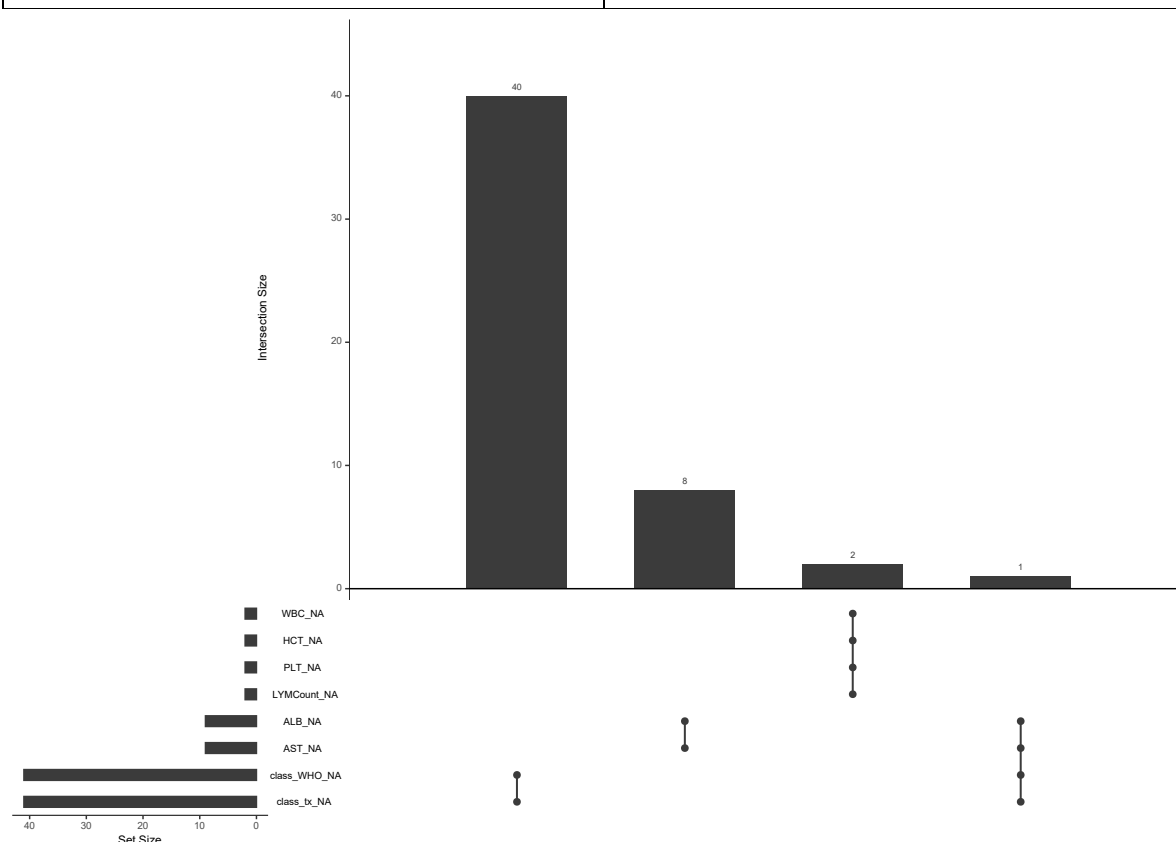


Appendix 13. Funnel plot for association between ALT and progression to severe disease. The grey dots indicate association of each study. While dots on the colour background indicate statistically significant ORs (red: p-value < 0.05, pink < 0.025 and brown p-value < 0.01), on white indicates non-significant.

Appendix 14. Associated factors with progression to severe dengue illness reported in studies with children and adults.

Associated factors reported in studies with children during the febrile phase of disease	Associated factors reported in studies with adults during the febrile phase of disease
1. Age	1. Age
2. Sex	2. Sex

Associated factors reported in studies with children during the febrile phase of disease	Associated factors reported in studies with adults during the febrile phase of disease
3. Nutritional status 4. Weight 5. Presence of vomiting 6. Presence of abdominal pain and tenderness 7. Presence of bleeding 8. Platelet count 9. Haematocrit 10. Aminotransferase levels 11. Immune status 12. Serotypes	3. Mixed comorbidity 4. Diabetes mellitus 5. chronic kidney disease 6. cardiovascular disease 7. Platelet count 8. Haematocrit 9. Immune status 10. Serotype



Appendix 15. Patterns of missingness of candidate predictors. The main bar chart demonstrates the number of missing values of the set of candidate predictors written below the bar. The side bar chart demonstrates the number of missing values of each candidate predictors. WBS: white blood cell count, HCT: haematocrit, PLT: platelet count, LYMCout: lymphocyte count; ALB: serum albumin, AST: aspartate transaminase, class_WHO:

WHO classification, class_tx: a combined endpoint of moderate plasma leakage or severe dengue defined in this study

Appendix 16. Approval of the scientific and ethical committees of collaborating hospitals and the Oxford University Tropical Research Ethical Committee (NCT01421732) and data sharing agreement between Imperial College London.

Oxford Tropical Research Ethics Committee

University of Oxford
Room 8, Manor House
The John Radcliffe, Headington, Oxford OX3 9DZ
tel. +44 (0) 1865 743005, fax +44 (0) 1865 743 002
e-mail: Fiona.Goulthorp@admin.ox.ac.uk



Dr C Simmons
OUCRU
190 Ben Ham Tu, Dist 5
Ho Chi Minh City
Vietnam

5th October 2010

- Dear Dr Simmons

Full Title of Study: Diagnosis and prognosis of severe dengue

OXTREC Reference: 35-10

Thank you for your letter dated 30th September in response to the committee's request for further clarification and amendments to your application and the inclusion of revised protocol, now v5.0, dated 28.07.10, and ICF v4.0, dated 01.10.10.

I am happy as Chairman for OXTREC to give approval for this study.

OXTREC are grateful for annual and end of study reports for this study.

Yours sincerely,

A handwritten signature in black ink, appearing to read 'Richard Mayon-White'.

Dr Richard Mayon-White

OXTREC Chair

Data security

The Chief Investigator will ensure confidentiality of participants taking part in the study in accordance with the Data Protection Act. The research dataset will be fully anonymised, stored as a Microsoft Excel 2016 format and subsequently encrypted with a password, which only specific IT personnel at OUCRU and the Chief Investigator will know. Any original identification data will be stored in a separate document, encrypted and accessible only by authorised staff at OUCRU in accordance with prior ethics approval and researchers will not have access to this document.

Dissemination

No identifiable personal data will be held. We ensure that there is no risk of general attributes disclosure for outputs at aggregate level. The methodology and results of this study will be presented at national and international conferences. The target groups in disseminating the study findings are healthcare personals and health related policy makers who have a role in patients care and healthcare management. Peer reviewed publications and open-access publications are considered to enable wide engagement of researchers and health related policy makers

Timeline

The primary researcher agrees to provide the cohort data within 2 years of the date of signing this MOU.

machine learning techniques incorporating both baseline and dynamic clinical parameters. The models are expected to gain better predictive performances.

In order to facilitate this project, we wish to establish an international collaboration between Health Protection Research Unit Imperial College London and Oxford University Clinical Trial Unit in Vietnam and invite you to join this group.

Agreement:

Statements

The statements below constitute an agreement for access to the datasets for the project identified above.

The principal investigators understand that:

1. The dataset(s) will only be used for inclusion in the above listed analysis.
2. The dataset(s) will not be used to conduct new primary analyses.
3. The dataset(s) will only be available for use by the principal investigators
4. Upon the completion of the project, the primary dataset(s) will be archived with the principal investigator, but no additional analyses completed unless prior consent and involvement of the primary researchers is agreed.
5. The principal investigator will not distribute the dataset(s) or any of its components to any other individual or researcher.

Authorship of manuscripts

Authorship and reporting of this work will follow international guidelines. Investigators from the original study at each site will be offered co-authorship of any new publications arising from this work.

Intellectual property rights

No commercial exploitation of the developed models is anticipated or will be sought.

Primary researchers retain all ownership of primary data.

Data Sharing Agreement

Principal Investigators:

- Dr Sorawat Sangkaew (Imperial College London)
- Dr Sophie Yacoub (Imperial College London and Oxford University Clinical Research Unit, Vietnam)
- Prof Alison Holmes (Imperial College London)
- Dr Ilaria Dorigatti (Imperial College London)

Project:

Developing and determining appropriate prognostic models for dengue deterioration during the febrile phase in a primary care setting

Dataset Requested from Study:

Data set reported in Nguyen MT, Ho TN, Nguyen VV, Nguyen TH, Ha MT, Ta VT, et al. An Evidence-Based Algorithm for Early Prognosis of Severe Dengue in the Outpatient Setting. *Clinical Infectious Diseases*. 2017;64(5):656-63 (Study code:13dx)

Project's purpose:

Since no effective vaccine or specific treatment exists for dengue, the early detection of the infection would allow improved triage and effectively monitoring of patients. In the present study, we aim to develop and determine appropriate clinical prognostic models for dengue deterioration in the early febrile phase at a primary care setting. These models will be constructed using existing data obtained from a prospective cohort data of children with dengue infection presenting within 72 hours of fever. We intend to provide the final model to clinicians in a primary care setting to assist them in triage patients for appropriate medical care and hospitalisation.

A number of clinical prediction models for dengue deterioration have been studied and published. Most of them made a prediction on a basis of a single set of clinical parameters (i.e. clinical information at enrolment) and using conventional statistical models, which based on limited predefined model assumptions. We will develop prognostic models using

The primary researcher provides the principal investigator access to the following dataset(s)
for developing prognostic models for dengue deterioration: (Completed by the primary
researcher)

_____ 13dx

Signatures:

This data sharing agreement is entered into on _____22/5/2019_____ (date)

Primary researcher(s)

Dr Sorawat Sangkaew

Sorawat Sangkaew. (print name)

Principal Investigator

Prof Cameron Simmons

Cameron Simmons (print name)

_____ (signature)

On behalf of all principle investigators in the13dx study.....

Appendix 17. Approval of the Research Ethics Review Committee of Queen Sirikit National Institute of Child Health (QSNICH) (REC.082/2562).

EC-QSNICH

EC-QSNICH

EC-QSNICH

EC-QSNICH

EC-QSNICH

EC-QSNICH

REC-QSNICH.36E



**The Research Ethics Review Committee of
Queen Sirikit National Institute of Child Health**

March 22, 2019

Title of Project : Determining appropriate clinical prediction models for dengue deterioration during the febrile phase at a primary care setting

Principal Investigator : Dr. Sorawat Sangkaew

Place of Proposed Study : Queen Sirikit National Institute of Child Health (Children's Hospital)

Project Time line : October 1, 2018 – September 30, 2021

Reviewed Document :

1. Curriculum Vitae

Approval Document :

1. Thai Protocol (REC-QSNICH.03) (version 3.0 date March 11th, 2019 : Thai Version)
2. English Protocol (version 3.0 date March 11th, 2019 : English Version)
3. Data collection form
4. Certificate

Have been reviewed by The Research Ethic Review Committee of Queen Sirikit National Institute of Child Health based on the Declaration of Helsinki and ICH Good Clinical Practice Guidelines (ICH GCP). Therefore the Thai protocol will be mainly conduct. Please submit the progress report every 12 months.

(Dr.Ratanotai Plubrukarn)

Chairman of The Research Ethics Review Committee of
Queen Sirikit National Institute of Child Health

No. : REC.082/2562 (Expedited)

Protocol Number : 62-021 **Date of Approval** : March 20, 2019 **Date of Expiration** : March 19, 2020

Reference No : REC.082/2562 **Monthly Conference** 4/2019 **Date of Meeting** March 4, 2019

OFFICE OF RESEARCH ETHIC COMMITTEE
Queen Sirikit National Institute of Child Health
Floor 12th 420/8 Ratchawithi Road,
Ratchathewi, Bangkok 10400, Thailand
Tel. 0-2-644-8943

Institutional Review Board Number; IRB00007346
Federal Wide Assurance; FWA00002250

EC-QSNICH

EC-QSNICH

EC-QSNICH

EC-QSNICH

EC-QSNICH

EC-QSNICH

Appendix 18. P-value of the coefficient of an endpoint variable on the specified day of illness obtained from the multivariable models with adjustment of age and sex.

DSS	Day of illness	Parameter	p-value	Combined	Day of illness	Parameter	p-value
	1	HCT	-		1	HCT	-
	2	HCT	0.0000		2	HCT	0.0930
	3	HCT	0.3630		3	HCT	0.1626
	4	HCT	0.9017		4	HCT	0.8184
	5	HCT	0.3700		5	HCT	0.0000
	6	HCT	0.0015		6	HCT	0.0000
	7	HCT	0.0000		7	HCT	0.0000
	1	PLT	-		1	PLT	-
	2	PLT	0.0000		2	PLT	0.1696
	3	PLT	0.9150		3	PLT	0.0006
	4	PLT	0.5578		4	PLT	0.0000
	5	PLT	0.0002		5	PLT	0.0000
	6	PLT	0.0000		6	PLT	0.0000
	7	PLT	0.0000		7	PLT	0.0000
	1	WBC	-		1	WBC	-
	2	WBC	0.0000		2	WBC	0.2961
	3	WBC	0.5052		3	WBC	0.9543
	4	WBC	0.7086		4	WBC	0.7549
	5	WBC	0.9422		5	WBC	0.7218
	6	WBC	0.2791		6	WBC	0.0033
	7	WBC	0.0000		7	WBC	0.0001
	1	NC	-		1	NC	-
	2	NC	0.0000		2	NC	0.6377
	3	NC	0.7235		3	NC	0.4357
	4	NC	0.1619		4	NC	0.0000
	5	NC	0.0013		5	NC	0.0006
	6	NC	0.0034		6	NC	0.6025
	7	NC	0.0100		7	NC	0.2877
	1	LC	-		1	LC	-
	2	LC	0.0000		2	LC	0.7419
	3	LC	0.9318		3	LC	0.7579

DSS	Day of illness	Parameter	p-value	Combined	Day of illness	Parameter	p-value
	4	LC	0.3330		4	LC	0.0000
	5	LC	0.0133		5	LC	0.0000
	6	LC	0.0003		6	LC	0.0056
	7	LC	0.0005		7	LC	0.1015
	1	AST	-		1	AST	-
	2	AST	0.0000		2	AST	0.1051
	3	AST	0.8809		3	AST	0.0001
	4	AST	0.0006		4	AST	0.0000
	5	AST	0.0000		5	AST	0.0000
	6	AST	0.0000		6	AST	0.0000
	7	AST	0.0000		7	AST	0.0046
	1	ALT	NA		1	ALT	NA
	2	ALT	0.0000		2	ALT	0.1544
	3	ALT	0.9779		3	ALT	0.0145
	4	ALT	0.0001		4	ALT	0.0063
	5	ALT	0.0004		5	ALT	0.0001
	6	ALT	0.0000		6	ALT	0.0007
	7	ALT	0.0001		7	ALT	0.0096
	1	ALB	NA		1	ALB	NA
	2	ALB	0.0000		2	ALB	0.9874
	3	ALB	0.4243		3	ALB	0.0709
	4	ALB	0.6186		4	ALB	0.6089
	5	ALB	0.0758		5	ALB	0.0002
	6	ALB	0.0000		6	ALB	0.0000
	7	ALB	0.0000		7	ALB	0.0000

Appendix 19. Coefficients and 95% confidence interval of clinical parameters obtained from the selected marginal models.

Laboratory variable	parameters	Estimate	SE	p-value	95%CI
Outcome: dengue shock syndrome					
HCT	Intercept	33.36	0.68	0.00	(32.02, 34.69)
HCT	Age	0.52	0.05	0.00	(0.42, 0.62)
HCT	Sex: male	0.02	0.27	0.95	(-0.51, 0.55)

Laboratory variable	parameters	Estimate	SE	p-value	95%CI
HCT	DOI	0.14	0.25	0.58	(-0.35, 0.62)
HCT	DSS	-1.67	1.73	0.33	(-5.06, 1.71)
HCT	Quadratic term of DOI	-0.02	0.03	0.43	(-0.08, 0.03)
HCT	Interaction term with DOI	0.40	0.40	0.32	(-0.38, 1.18)
PLT	Intercept	333779.00	14531.41	0.00	(305298, 362260.1)
PLT	Age	-4305.55	903.76	0.00	(-6076.89, -2534.21)
PLT	Sex: male	12492.93	5316.42	0.02	(2072.94, 22912.92)
PLT	DOI	-37973.40	5703.12	0.00	(-49151.3, -26795.4)
PLT	DSS	15343.00	23644.32	0.52	(-30999, 61685.02)
PLT	Quadratic term of DOI	1015.75	642.38	0.11	(-243.3, 2274.8)
PLT	Interaction term with DOI	-14034.10	4701.37	0.00	(-23248.6, -4819.55)
WBC	Intercept	14436.24	632.70	0.00	(13196.16, 15676.31)
WBC	Age	-149.99	34.49	0.00	(-217.59, -82.39)
WBC	Sex: male	431.24	157.02	0.01	(123.48, 739)
WBC	DOI	-3945.97	277.54	0.00	(-4489.94, -3402.01)
WBC	DSS	-814.42	1062.75	0.44	(-2897.38, 1268.53)
WBC	Quadratic term of DOI	374.27	31.69	0.00	(312.17, 436.38)
WBC	Interaction term with DOI	233.32	241.90	0.33	(-240.79, 707.44)
NEU	Intercept	11955.61	577.22	0.00	(10824.28, 13086.94)
NEU	Age	-17.07	19.91	0.39	(-56.1, 21.96)
NEU	Sex: male	273.60	113.62	0.02	(50.9, 496.29)
NEU	DOI	-3828.06	229.93	0.00	(-4278.71, -3377.41)
NEU	DSS	-279.72	792.24	0.72	(-1832.49, 1273.05)
NEU	Quadratic term of DOI	334.01	24.48	0.00	(286.04, 381.98)
NEU	Interaction term with DOI	163.84	172.42	0.34	(-174.1, 501.78)
LYM	Intercept	1275.61	219.22	0.00	(845.96, 1705.27)
LYM	Age	-112.26	15.90	0.00	(-143.42, -81.1)
LYM	Sex: male	125.65	60.23	0.04	(7.6, 243.69)
LYM	DOI	274.81	111.20	0.01	(56.86, 492.76)
LYM	DSS	-415.18	352.77	0.24	(-1106.59, 276.23)
LYM	Quadratic term of DOI	-16.22	13.20	0.22	(-42.08, 9.64)
LYM	Interaction term with DOI	64.54	84.69	0.45	(-101.45, 230.53)
AST	Intercept	20.56	20.98	0.33	(-20.55, 61.67)
AST	Age	-0.83	0.94	0.38	(-2.68, 1.02)

Laboratory variable	parameters	Estimate	SE	p-value	95%CI
AST	Sex: male	23.16	6.96	0.00	(9.51, 36.81)
AST	DOI	3.80	9.63	0.69	(-15.07, 22.67)
AST	DSS	-31.69	56.68	0.58	(-142.78, 79.4)
AST	Quadratic term of DOI	2.01	1.17	0.09	(-0.28, 4.3)
AST	Interaction term with DOI	23.27	14.42	0.11	(-4.99, 51.53)
ALT	Intercept	11.21	10.05	0.27	(-8.5, 30.91)
ALT	Age	0.64	0.60	0.29	(-0.54, 1.81)
ALT	Sex: male	12.61	4.77	0.01	(3.26, 21.96)
ALT	DOI	-1.60	4.36	0.71	(-10.16, 6.95)
ALT	DSS	-3.24	24.62	0.90	(-51.49, 45.02)
ALT	Quadratic term of DOI	1.22	0.52	0.02	(0.2, 2.25)
ALT	Interaction term with DOI	8.53	6.06	0.16	(-3.35, 20.41)
ALB	Intercept	5.07	0.04	0.00	(4.98, 5.15)
ALB	DOI	-0.15	0.01	0.00	(-0.17, -0.13)
ALB	DSS	0.83	0.33	0.01	(0.18, 1.48)
ALB	Interaction term with DOI	-0.26	0.09	0.00	(-0.43, -0.09)
Outcome: combined endpoint of moderate plasma leakage and/or DSS					
HCT	Intercept	34.23	0.68	0.00	(32.9, 35.57)
HCT	Age	0.51	0.05	0.00	(0.42, 0.61)
HCT	Sex: male	0.03	0.27	0.91	(-0.49, 0.55)
HCT	DOI	-0.09	0.25	0.70	(-0.58, 0.39)
HCT	Combined endpoint	-2.46	0.55	0.00	(-3.54, -1.39)
HCT	Quadratic term of DOI	-0.02	0.03	0.53	(-0.07, 0.04)
HCT	Interaction term with DOI	0.64	0.12	0.00	(0.4, 0.87)
PLT	Intercept	330626.20	15292.71	0.00	(300653, 360599.3)
PLT	Age	-3892.12	858.82	0.00	(-5575.37, -2208.86)
PLT	Sex: male	14022.52	5172.43	0.01	(3884.75, 24160.29)
PLT	DOI	-34931.70	5866.49	0.00	(-46429.8, -23433.6)
PLT	Combined endpoint	784.63	11332.86	0.94	(-21427.4, 22996.62)
PLT	Quadratic term of DOI	893.37	634.28	0.16	(-349.8, 2136.54)
PLT	Interaction term with DOI	-9247.28	2423.48	0.00	(-13997.2, -4497.35)
WBC	Intercept	14693.90	663.15	0.00	(13394.15, 15993.65)
WBC	Age	-152.19	35.23	0.00	(-221.23, -83.15)
WBC	Sex: male	426.13	156.85	0.01	(118.72, 733.55)

Laboratory variable	parameters	Estimate	SE	p-value	95%CI
WBC	DOI	-4024.03	285.06	0.00	(-4582.74, -3465.33)
WBC	Combined endpoint	-601.99	555.64	0.28	(-1691.02, 487.04)
WBC	Quadratic term of DOI	378.47	31.52	0.00	(316.68, 440.26)
WBC	Interaction term with DOI	160.46	110.37	0.15	(-55.86, 376.77)
NEU	Intercept	12020.22	610.73	0.00	(10823.21, 13217.23)
NEU	Age	-19.56	20.13	0.33	(-59.02, 19.9)
NEU	Sex: male	261.89	113.71	0.02	(39.02, 484.75)
NEU	DOI	-3856.83	237.51	0.00	(-4322.34, -3391.32)
NEU	Combined endpoint	-71.58	470.77	0.88	(-994.27, 851.12)
NEU	Quadratic term of DOI	336.32	24.56	0.00	(288.18, 384.45)
NEU	Interaction term with DOI	59.38	90.07	0.51	(-117.15, 235.91)
LYM	Intercept	1356.01	222.40	0.00	(920.12, 1791.9)
LYM	Age	-111.77	16.21	0.00	(-143.55, -79.99)
LYM	Sex: male	128.53	60.53	0.03	(9.88, 247.17)
LYM	DOI	255.46	113.55	0.02	(32.9, 478.01)
LYM	Combined endpoint	-228.34	191.79	0.23	(-604.24, 147.55)
LYM	Quadratic term of DOI	-15.14	13.17	0.25	(-40.95, 10.68)
LYM	Interaction term with DOI	32.71	41.73	0.43	(-49.08, 114.5)
AST	Intercept	38.33	22.41	0.09	(-5.6, 82.26)
AST	Age	-1.26	0.93	0.17	(-3.08, 0.56)
AST	Sex: male	19.58	6.71	0.00	(6.43, 32.74)
AST	DOI	-4.44	10.45	0.67	(-24.92, 16.04)
AST	Combined endpoint	-26.14	19.51	0.18	(-64.38, 12.11)
AST	Quadratic term of DOI	2.50	1.26	0.05	(0.02, 4.97)
AST	Interaction term with DOI	17.48	5.97	0.00	(5.77, 29.18)
ALT	Intercept	16.22	10.07	0.11	(-3.52, 35.96)
ALT	Age	0.44	0.61	0.47	(-0.76, 1.63)
ALT	Sex: male	11.00	4.61	0.02	(1.97, 20.03)
ALT	DOI	-4.19	4.47	0.35	(-12.96, 4.58)
ALT	Combined endpoint	-5.22	9.03	0.56	(-22.91, 12.47)
ALT	Quadratic term of DOI	1.36	0.54	0.01	(0.29, 2.42)
ALT	Interaction term with DOI	6.33	2.81	0.02	(0.82, 11.85)
ALB	Intercept	5.03	0.05	0.00	(4.94, 5.13)
ALB	DOI	-0.13	0.01	0.00	(-0.15, -0.11)

Laboratory variable	parameters	Estimate	SE	p-value	95%CI
ALB	Combined endpoint	0.24	0.10	0.02	(0.05, 0.44)
ALB	Interaction term with DOI	-0.10	0.02	0.00	(-0.14, -0.05)

HCT: haematocrit; PLT: platelet count; WBC: white blood cell count; NEU: neutrophile count; LYM: lymphocyte count; AST: aspartate aminotransferase; ALT: alanine aminotransferase; ALB: serum albumin

Appendix 20. Predictive performance (Brier Score, AUROC, CITL, Calibration slope, accuracy, sensitivity, specificity, PPV, and NPV) of the risk prediction models for DSS and the combined endpoints of moderate plasma leakage and/or DSS using logistic regression (LR), artificial neural network (ANN), and recurrent neural network (RNN)-based models (RNN, gate recurrent unit (GRU), and long-short term memory (LSTM))

Metric	Model type	Model	Estimate	Lower 95%CI	Upper 95%CI	Prediction outcome
BS	ANN	DOI-2	0.180	0.168	0.266	The combined endpoint
	ANN	DOI-3	0.191	0.172	0.240	The combined endpoint
	ANN	DOI-4	0.221	0.203	0.244	The combined endpoint
	LR	DOI-2	0.244	0.237	0.259	The combined endpoint
	LR	DOI-3	0.241	0.234	0.257	The combined endpoint
	LR	DOI-4	0.219	0.206	0.239	The combined endpoint
	LR	Enrolment	0.222	0.214	0.239	The combined endpoint
	ANN	Enrolment	0.222	0.202	0.245	The combined endpoint
	RNN	RNN	0.220	0.197	0.235	The combined endpoint
	RNN	GRU	0.219	0.192	0.247	The combined endpoint

Metric	Model type	Model	Estimate	Lower 95%CI	Upper 95%CI	Prediction outcome
	RNN	LSTM	0.220	0.192	0.249	The combined endpoint
	ANN	DOI-2	0.058	0.048	0.077	DSS
	ANN	DOI-3	0.059	0.042	0.079	DSS
	ANN	DOI-4	0.055	0.037	0.075	DSS
	LR	DOI-2	0.058	0.045	0.074	DSS
	LR	DOI-3	0.059	0.042	0.078	DSS
	LR	DOI-4	0.055	0.042	0.073	DSS
	LR	Enrolment	0.057	0.042	0.074	DSS
	ANN	Enrolment	0.059	0.046	0.081	DSS
	RNN	RNN	0.057	0.041	0.067	DSS
	RNN	GRU	0.056	0.040	0.072	DSS
	RNN	LSTM	0.056	0.039	0.071	DSS
AUC	ANN	DOI-2	0.601	0.381	0.804	The combined endpoint
	ANN	DOI-3	0.652	0.494	0.697	The combined endpoint
	ANN	DOI-4	0.691	0.643	0.727	The combined endpoint
	LR	DOI-2	0.548	0.474	0.577	The combined endpoint
	LR	DOI-3	0.579	0.501	0.627	The combined endpoint
	LR	DOI-4	0.688	0.637	0.731	The combined endpoint
	LR	Enrolment	0.676	0.628	0.715	The combined endpoint
	ANN	Enrolment	0.688	0.616	0.722	The combined endpoint

Metric	Model type	Model	Estimate	Lower 95%CI	Upper 95%CI	Prediction outcome
	RNN	RNN	0.687	0.602	0.768	The combined endpoint
	RNN	GRU	0.691	0.569	0.780	The combined endpoint
	RNN	LSTM	0.684	0.571	0.736	The combined endpoint
	ANN	DOI-2	0.627	0.450	0.765	DSS
	ANN	DOI-3	0.587	0.367	0.659	DSS
	ANN	DOI-4	0.678	0.528	0.728	DSS
	LR	DOI-2	0.629	0.536	0.690	DSS
	LR	DOI-3	0.581	0.459	0.633	DSS
	LR	DOI-4	0.707	0.615	0.762	DSS
	LR	Enrolment	0.700	0.592	0.750	DSS
	ANN	Enrolment	0.690	0.510	0.747	DSS
	RNN	RNN	0.633	0.456	0.860	DSS
	RNN	GRU	0.630	0.445	0.877	DSS
	RNN	LSTM	0.638	0.477	0.908	DSS
CITL	ANN	DOI-2	-0.010	-0.142	0.159	The combined endpoint
	ANN	DOI-3	0.063	-0.027	0.134	The combined endpoint
	ANN	DOI-4	0.004	-0.031	0.027	The combined endpoint
	LR	DOI-2	-0.093	-0.093	-0.093	The combined endpoint
	LR	DOI-3	0.006	0.006	0.006	The combined endpoint
	LR	DOI-4	0.003	0.003	0.003	The combined endpoint

Metric	Model type	Model	Estimate	Lower 95%CI	Upper 95%CI	Prediction outcome
	LR	Enrolment	-0.005	-0.005	-0.005	The combined endpoint
	ANN	Enrolment	-0.019	-0.053	0.007	The combined endpoint
	RNN	RNN	-0.021	-0.058	0.219	The combined endpoint
	RNN	GRU	-0.011	-0.032	0.221	The combined endpoint
	RNN	LSTM	0.003	-0.017	0.225	The combined endpoint
	ANN	DOI-2	0.009	-0.077	0.104	DSS
	ANN	DOI-3	-0.012	-0.063	0.046	DSS
	ANN	DOI-4	-0.016	-0.208	0.066	DSS
	LR	DOI-2	0.009	0.009	0.009	DSS
	LR	DOI-3	0.012	0.012	0.012	DSS
	LR	DOI-4	0.011	0.011	0.011	DSS
	LR	Enrolment	-0.001	-0.001	-0.001	DSS
	ANN	Enrolment	0.010	-0.057	0.133	DSS
	RNN	RNN	-0.048	-0.287	0.307	DSS
	RNN	GRU	-0.059	-0.336	0.318	DSS
	RNN	LSTM	-0.071	-0.287	0.405	DSS
CS	ANN	DOI-2	-0.259	-23.757	4.155	The combined endpoint
	ANN	DOI-3	1.308	-3.751	2.049	The combined endpoint
	ANN	DOI-4	1.686	-2.794	2.318	The combined endpoint
	LR	DOI-2	0.524	0.524	0.524	The combined endpoint

Metric	Model type	Model	Estimate	Lower 95%CI	Upper 95%CI	Prediction outcome
	LR	DOI-3	0.767	0.767	0.767	The combined endpoint
	LR	DOI-4	0.963	0.963	0.963	The combined endpoint
	LR	Enrolment	0.920	0.920	0.920	The combined endpoint
	ANN	Enrolment	1.502	-2.280	1.952	The combined endpoint
	RNN	RNN	1.817	-1.590	2.371	The combined endpoint
	RNN	GRU	1.778	-4.325	2.540	The combined endpoint
	RNN	LSTM	1.481	-3.078	1.858	The combined endpoint
	ANN	DOI-2	1.903	-9.478	4.816	DSS
	ANN	DOI-3	0.235	-21.472	1.365	DSS
	ANN	DOI-4	1.162	-4.561	1.506	DSS
	LR	DOI-2	0.677	0.677	0.677	DSS
	LR	DOI-3	0.325	0.325	0.325	DSS
	LR	DOI-4	0.821	0.821	0.821	DSS
	LR	Enrolment	0.792	0.792	0.792	DSS
	ANN	Enrolment	1.185	-4.859	2.801	DSS
	RNN	RNN	0.736	-1.098	4.375	DSS
	RNN	GRU	0.688	-0.603	6.339	DSS
	RNN	LSTM	0.401	-2.236	4.405	DSS
ACC	ANN	DOI-2	0.656	0.476	0.829	The combined endpoint
	ANN	DOI-3	0.674	0.558	0.724	The combined endpoint

Metric	Model type	Model	Estimate	Lower 95%CI	Upper 95%CI	Prediction outcome
	ANN	DOI-4	0.652	0.598	0.686	The combined endpoint
	LR	DOI-2	0.557	0.493	0.572	The combined endpoint
	LR	DOI-3	0.584	0.521	0.634	The combined endpoint
	LR	DOI-4	0.659	0.605	0.695	The combined endpoint
	LR	Enrolment	0.648	0.599	0.680	The combined endpoint
	ANN	Enrolment	0.658	0.590	0.685	The combined endpoint
	RNN	RNN	0.666	0.591	0.720	The combined endpoint
	RNN	GRU	0.667	0.544	0.720	The combined endpoint
	RNN	LSTM	0.666	0.564	0.706	The combined endpoint
	ANN	DOI-2	0.630	0.472	0.754	DSS
	ANN	DOI-3	0.602	0.349	0.788	DSS
	ANN	DOI-4	0.672	0.449	0.782	DSS
	LR	DOI-2	0.637	0.468	0.759	DSS
	LR	DOI-3	0.606	0.378	0.705	DSS
	LR	DOI-4	0.712	0.504	0.876	DSS
	LR	Enrolment	0.700	0.551	0.804	DSS
	ANN	Enrolment	0.715	0.555	0.827	DSS
	RNN	RNN	0.717	0.556	0.961	DSS
	RNN	GRU	0.692	0.568	0.926	DSS
	RNN	LSTM	0.679	0.503	0.882	DSS

Metric	Model type	Model	Estimate	Lower 95%CI	Upper 95%CI	Prediction outcome
SN	ANN	DOI-2	0.662	0.475	0.901	The combined endpoint
	ANN	DOI-3	0.695	0.561	0.791	The combined endpoint
	ANN	DOI-4	0.661	0.561	0.768	The combined endpoint
	LR	DOI-2	0.553	0.412	0.642	The combined endpoint
	LR	DOI-3	0.580	0.445	0.706	The combined endpoint
	LR	DOI-4	0.668	0.543	0.758	The combined endpoint
	LR	Enrolment	0.649	0.569	0.745	The combined endpoint
	ANN	Enrolment	0.656	0.533	0.720	The combined endpoint
	RNN	RNN	0.686	0.610	0.788	The combined endpoint
	RNN	GRU	0.686	0.533	0.759	The combined endpoint
	RNN	LSTM	0.683	0.542	0.746	The combined endpoint
	ANN	DOI-2	0.626	0.467	0.759	DSS
	ANN	DOI-3	0.599	0.321	0.813	DSS
	ANN	DOI-4	0.672	0.440	0.799	DSS
	LR	DOI-2	0.635	0.456	0.766	DSS
	LR	DOI-3	0.605	0.354	0.714	DSS
	LR	DOI-4	0.717	0.500	0.903	DSS
	LR	Enrolment	0.701	0.541	0.820	DSS
	ANN	Enrolment	0.718	0.555	0.839	DSS

Metric	Model type	Model	Estimate	Lower 95%CI	Upper 95%CI	Prediction outcome
	RNN	RNN	0.724	0.553	0.983	DSS
	RNN	GRU	0.696	0.566	0.950	DSS
	RNN	LSTM	0.681	0.486	0.893	DSS
SP	ANN	DOI-2	0.662	0.475	0.901	The combined endpoint
	ANN	DOI-3	0.695	0.561	0.791	The combined endpoint
	ANN	DOI-4	0.661	0.561	0.768	The combined endpoint
	LR	DOI-2	0.553	0.412	0.642	The combined endpoint
	LR	DOI-3	0.580	0.445	0.706	The combined endpoint
	LR	DOI-4	0.668	0.543	0.758	The combined endpoint
	LR	Enrolment	0.649	0.569	0.745	The combined endpoint
	ANN	Enrolment	0.656	0.533	0.720	The combined endpoint
	RNN	RNN	0.686	0.610	0.788	The combined endpoint
	RNN	GRU	0.686	0.533	0.759	The combined endpoint
	RNN	LSTM	0.683	0.542	0.746	The combined endpoint
	ANN	DOI-2	0.626	0.467	0.759	DSS
	ANN	DOI-3	0.599	0.321	0.813	DSS
	ANN	DOI-4	0.672	0.440	0.799	DSS
	LR	DOI-2	0.635	0.456	0.766	DSS
	LR	DOI-3	0.605	0.354	0.714	DSS

Metric	Model type	Model	Estimate	Lower 95%CI	Upper 95%CI	Prediction outcome
	LR	DOI-4	0.717	0.500	0.903	DSS
	LR	Enrolment	0.701	0.541	0.820	DSS
	ANN	Enrolment	0.718	0.555	0.839	DSS
	RNN	RNN	0.724	0.553	0.983	DSS
	RNN	GRU	0.696	0.566	0.950	DSS
	RNN	LSTM	0.681	0.486	0.893	DSS
PPV	ANN	DOI-2	0.518	0.204	0.769	The combined endpoint
	ANN	DOI-3	0.502	0.310	0.626	The combined endpoint
	ANN	DOI-4	0.576	0.494	0.620	The combined endpoint
	LR	DOI-2	0.474	0.399	0.500	The combined endpoint
	LR	DOI-3	0.506	0.423	0.566	The combined endpoint
	LR	DOI-4	0.583	0.495	0.638	The combined endpoint
	LR	Enrolment	0.571	0.494	0.625	The combined endpoint
	ANN	Enrolment	0.583	0.493	0.627	The combined endpoint
	RNN	RNN	0.589	0.498	0.673	The combined endpoint
	RNN	GRU	0.589	0.414	0.649	The combined endpoint
	RNN	LSTM	0.586	0.443	0.648	The combined endpoint
	ANN	DOI-2	0.122	0.000	0.177	DSS
	ANN	DOI-3	0.107	0.000	0.148	DSS

Metric	Model type	Model	Estimate	Lower 95%CI	Upper 95%CI	Prediction outcome
	ANN	DOI-4	0.131	0.000	0.164	DSS
	LR	DOI-2	0.110	0.021	0.155	DSS
	LR	DOI-3	0.099	0.002	0.133	DSS
	LR	DOI-4	0.157	0.000	0.208	DSS
	LR	Enrolment	0.142	0.012	0.177	DSS
	ANN	Enrolment	0.161	0.000	0.234	DSS
	RNN	RNN	0.168	0.000	0.298	DSS
	RNN	GRU	0.184	0.000	0.354	DSS
	RNN	LSTM	0.157	0.000	0.281	DSS
NPV	ANN	DOI-2	0.857	0.717	0.901	The combined endpoint
	ANN	DOI-3	0.822	0.720	0.848	The combined endpoint
	ANN	DOI-4	0.721	0.644	0.760	The combined endpoint
	LR	DOI-2	0.638	0.549	0.685	The combined endpoint
	LR	DOI-3	0.661	0.593	0.709	The combined endpoint
	LR	DOI-4	0.726	0.656	0.776	The combined endpoint
	LR	Enrolment	0.720	0.649	0.755	The combined endpoint
	ANN	Enrolment	0.729	0.654	0.780	The combined endpoint
	RNN	RNN	0.728	0.652	0.772	The combined endpoint
	RNN	GRU	0.731	0.633	0.791	The combined endpoint

Metric	Model type	Model	Estimate	Lower 95%CI	Upper 95%CI	Prediction outcome
	RNN	LSTM	0.732	0.632	0.782	The combined endpoint
	ANN	DOI-2	0.968	0.948	0.978	DSS
	ANN	DOI-3	0.962	0.931	0.970	DSS
	ANN	DOI-4	0.967	0.944	0.977	DSS
	LR	DOI-2	0.966	0.950	0.979	DSS
	LR	DOI-3	0.960	0.938	0.977	DSS
	LR	DOI-4	0.967	0.948	0.975	DSS
	LR	Enrolment	0.970	0.953	0.981	DSS
	ANN	Enrolment	0.970	0.950	0.979	DSS
	RNN	RNN	0.966	0.946	0.990	DSS
	RNN	GRU	0.966	0.947	0.995	DSS
	RNN	LSTM	0.969	0.951	0.997	DSS

BS: the Brier Score; AUC: Area Under the curve; CITL: Calibration in the large; CS: Calibration slope; ACC: accuracy; SN: sensitivity; SP: specificity; PPV: positive predictive value; NPV: negative predictive value; DOI: day of illness.

