# A Three-Component Biomarker Panel for Prediction of Dengue Hemorrhagic Fever

Allan R. Brasier,\* Hyunsu Ju, Josefina Garcia, Heidi M. Spratt, Sundar S. Victor, Brett M. Forshey, Eric S. Halsey, Guillermo Comach, Gloria Sierra, Patrick J. Blair, Claudio Rocha, Amy C. Morrison, Thomas W. Scott,

Isabel Bazan, Tadeusz J. Kochel, and the Venezuelan Dengue Fever Working Group†

Department of Internal Medicine, University of Texas Medical Branch, Galveston, Texas; Sealy Center for Molecular Medicine, UTMB, Institute for Translational Sciences, UTMB, the United States Naval Medical Research Unit-6, Lima, Peru; Department Preventive Medicine and Community Health, UTMB, Laboratorio Regional de Diagnostico e Investigacion del Dengue y otras Enfermedades Virales (LARDIDEV), Instituto de Investigaciones Biomedicas de la Universidad de Carabobo (BIOMED-UC), Maracay, Venezuela; Department of Entomology, University of California, Davis, California

*Abstract.* Dengue virus infections are a major cause of morbidity in tropical countries. Early detection of dengue hemorrhagic fever (DHF) may help identify individuals that would benefit from intensive therapy. Predictive modeling was performed using 11 laboratory values of 51 individuals (38 DF and 13 DHF) obtained on initial presentation using logistic regression. We produced a robust model with an area under the curve of 0.9615 that retained IL-10 levels, platelets, and lymphocytes as the major predictive features. A classification and regression tree was developed on these features that were 86% accurate on cross-validation. The IL-10 levels and platelet counts were also identified as the most informative features associated with DHF using a Random Forest classifier. In the presence of polymerase chain reaction-proven acute dengue infections, we suggest a complete blood count and rapid measurement of IL-10 can assist in the triage of potential DHF cases for close follow-up or clinical intervention improving clinical outcome.

## INTRODUCTION

Dengue viruses (DENV) are members of the flavirirus family that are transmitted by Aedes sp. mosquitoes and may produce a clinically significant disease in humans. Because of a number of factors, including urbanization, globalization of travel, and lack of efficient chemical pesticide-based vector control interventions, DENV infections have re-emerged as a significant international public health problem. Worldwide, an estimated 2.5 billion people in tropical and subtropical regions are at risk of infection. In the Americas alone, an estimated 890,000 cases of dengue fever (DF) were reported in 2007 representing a significant increase from historical levels.<sup>1</sup> The DENV infections produce a graded spectrum of disease severity ranging from asymptomatic infection to a flulike state (DF) to a hemorrhagic form (dengue hemorrhagic fever [DHF]), characterized by plasma leakage and bleeding, representing a life-threatening complication.<sup>2</sup>

Dengue hemorrhagic fever is the result of a complex interplay of host immunologic and genetic factors with DENV serotypes and genotypes. Epidemiological studies indicate a 40-80-fold increased risk of DHF after a second infection with a different serotype.3-5 The "antibody-dependent enhancement" theory proposes that neutralizing antibodies generated during the adaptive immune response to an infecting serotype increases viral burden during infection with a second, different DENV serotype.<sup>6</sup> Virus contained within non-neutralizing heterotypic immune complexes enter immunocytes (particularly monocytes and dendritic cells) through the Fc receptor. Hyper-stimulated cells release enhanced cytokines and other factors leading to the pathophysiological manifestations of vascular leakage and coagulopathy. Other evidence points to DHF being the result of an interplay between viral and host factors, including cell-mediated immunity.<sup>2,7,8</sup>

The mortality of DHF is age-dependent, primarily affecting both children and the elderly.<sup>4</sup> In Southeast Asia, a disproportionate amount of DHF hospitalizations are of children, whereas in the Americas, there is a more even age distribution. Although DHF fatality rates can exceed 20%, early identification and intensive supportive therapy can reduce the rate to 1% or less<sup>9</sup>; consequently, there is clinical need to identify predictive features of DHF early in the course of infection.

In this study, we sought to evaluate whether combinations of clinical and accessible laboratory tests could be used as a surrogate for DHF. These data were analyzed by Bayesian inference methods and a robust predictive logistic regression model was developed incorporating IL-10 concentrations, platelet counts, and lymphocyte counts. A classification and regression tree (CART) was evaluated. Although CART classification was less accurate than that produced by logistic regression, IL-10 and platelet counts were identified as the most informative predictive features.

### MATERIALS AND METHODS

**Ethics statement.** This informed consent study was conducted under a human subjects study protocol no. NMRCD.2005.0007 (Active Dengue Surveillance and Predictors of Disease Severity in Maracay, Venezuela) approved by the Centro de Investigaciones Biomedicas de la Universidad de Carabobo (BIOMED-UC), Maracay, Venezuela, and the Naval Medical Research Center institutional review boards in compliance with all applicable federal regulations governing the protection of human subjects.

**Study population.** Active surveillance for people with dengue infection was conducted in Maracay, Venezuela. Febrile subjects with signs and symptoms consistent with DENV infection who presented at participating clinics and hospitals,<sup>10</sup> or who were identified by community-based active surveillance, were included in the study (Table 1). On the day of presentation, a blood sample was collected for dengue virus reverse transcription-polymerase chain reaction (RT-PCR) confirmation and clinical testing.<sup>11</sup> Individuals with confirmed

<sup>\*</sup>Address correspondence to Allan R. Brasier, MRB 8.128, 301 University Blvd, UTMB, Galveston, TX 77555-1060. E-mail: arbrasie@utmb.edu †Other members of the Venezuelan Dengue Fever Working Group are Iris Villalobos and Carlos Espino.

	Stu	dy pop	ulation f	or the dengue f	ever (DF) and deng	gue l	hem	orrha	igic f	ever	(DI	HF) g	roup	5			
Gender (no.)				Fever (duration, d)				Diarrhea		Dengue virus (DENV) serotype							
Outcome	Age (y)	Males	Females	Platelets $(10^3/\mu L)$	Lymphocytes $(10^3/\mu L)$	2d	3d	4d	5d	6d	8d	Ν	Y	1	2	3	4
DF $(N = 38)$	$15.76\pm7.82$	19	19	$159 \pm 41*$	$46.31 \pm 13.48$	2	5	14	14	3	0	33	5	20	6	8	4
DHF $(N = 13)$	$19\pm13.4$	3	10	$105\pm33$	$36.9 \pm 14.86$	0	0	2	8	2	1	7	6	2	8	2	1
*P < 0.01.																	

TABLE 1
tudy population for the dengue fever (DF) and dengue hemorrhagic fever (DHF) group

Note: Shown are the mean age, gender, platelet and lymphocyte counts, duration of fever before study enrollment (in days, based on subjects recall), presence of diarrhea, and dengue serotype obtained upon entry into the study.

DENV infections were consented and monitored for clinical outcome, and DF and DHF cases were scored following World Health Organization (WHO) case definitions.<sup>12</sup>

**Complete blood cell count.** Whole anticoagulated venous blood upon presentation is obtained from each volunteer. Complete blood cell count was performed with the QBC automated system according to the manufacturer's instructions (Becton-Dickinson, Franklin Lakes, NJ).

**RT-PCR.** Viral RNA was prepared from 140  $\mu$ L sera using QIAamp Viral RNA Mini Kits following the manufacturer's instructions (Qiagen Inc., Valencia, CA). Nested dengue virus RT-PCR was performed on serum samples for virus detection as described previously.<sup>11</sup>

**Multiplex bead-based cytokine measurements.** Plasma samples were analyzed for the concentrations of 9 human cytokines (IL-6, IL-10, IFN- $\gamma$ , IP-10, MIP-1 $\alpha$ , TNF $\alpha$ , IL-2, vascular endothelial growth factor [VEGF], and TNF-related apoptosis-inducing ligand [TRAIL]) according to the manufacturer's recommendations (Bioplex, Bio-Rad, Hercules, CA). For each analyte, a standard curve was generated using recombinant proteins to estimate protein concentration in the unknown sample. For the purposes of modeling, the cytokine values were logarithm base 2(log2)-transformed to approximate a normal distribution.

**Bayesian variable selection for generalized additive models.** To select the models of predictors between smoothing nonlinear terms and linear effects, we performed Bayesian variable selection in generalized additive models (GAM, implemented in the R package spikeSlabGAM<sup>13</sup>).

**Statistical analysis.** The best subsets logistic regression procedure was performed to eliminate covariates and select the best list of predictive variables for model generation. The criteria for selection were based on minimizing the mean squared error of prediction using the Akaike information criteria (AIC)/Bayesian information criteria (BIC). The AIC/ BIC was used in our modeling approach because the results for AIC/BIC converge with those produced by leave-one-out cross validation in larger data sets<sup>14</sup> and because of its ease of implementation. For this modeling SAS, version 9.1.3 (SAS Institute, Inc., Cary, NC) was used.

**Generalized additive models (GAM).** Generalized additive models were estimated by a backfitting algorithm within a Newton-Raphson technique. We used SAS 9.2 PROC GAM (SAS Institute, Inc.) and STATISTICA 8.0 (StatSoft Inc., Tulsa, OK) to fit the GAM fittings with the binary logit link function that provided multiple types of smoothers with automatic selection of smoothing parameters.

**Classification and regression tree modeling.** Decision tree model building was performed with CART (Salford Systems, San Diego, CA). The CART is an iterative classification method for variable selection and predicting categorical response variables that uses a splitting rule to identify a predictive variable and a cutoff that best breaks the population into homogenous classes. The splitting rule used was entropy. Because of the study design, modeling was performed assuming equal likelihood of the DF or DHF classification (equal priors). The model was tested using 10-fold cross-validation to prevent overfitting.

**Random forests classifiers.** Random forests (Salford Systems) built an ensemble of CART trees using a bagging (bootstrap aggregation) technique. A large collection of decorrelated trees are produced and averaged. Variable importance plots are used to evaluate the information content of individual features.

#### RESULTS

Study population. Subjects were enrolled in the study who presented with a new fever (oral) equal to or greater than 38°C, accompanied by two or more of the following manifestations: myalgia, arthralgia, leukopenia, rash, headache, lymphoadenopathy, nausea, vomiting, positive tourniquet test, thrombocytopenia, or hepatomegaly.<sup>10</sup> The clinical characteristics of the study population are shown in Table 1. The individuals that developed DF were aged  $15.8 \pm 7.8$  y (N = 38), whereas the individuals that developed DHF were aged  $19 \pm 13.4$  y (N = 13; not significant). Using the 2009 WHO criteria,15 all 13 DHF had dengue with warning signs; of these three were classified as severe dengue (C) caused by plasma leakage and severe bleeding. None of the severe dengue cases had evidence of shock. There was no significant difference between the groups by duration of fever before study enrollment. On the day patients presented at clinics, platelet counts were lower in the DHF group  $(105 \pm 33 \times 10^3/\mu L)$  than DF group (159  $\pm$  41 × 10<sup>3</sup>/µL; P < 0.001), and the frequency of diarrhea was greater in the DHF group (P = 0.021 Fisher's exact test). All four dengue serotypes were represented in the study (distribution shown in Table 1).

**Multivariate logistic regression modeling for DHF-Bayesian feature reduction.** Within the study population, we selected a feature set of 11 parameters including gender, clinical signs (days of fever, diarrhea), laboratory measurements (lymphocyte/ neutrophil/platelet counts, hemoglobin concentration, red blood cell count) and cytokine concentrations (IL-10, IL-6, TRAIL). Because the underlying data structures of the feature set dictates the selection of an appropriate modeling approach, we analyzed the contributions of parametric (linear) or nonparametric (spline) features using Bayesian variable selection.<sup>13</sup> This method produces a hierarchy of structured model selections for parametric and nonparametric relationships to the outcome for each feature. The posterior probabilities for the linear and spline components are shown in Table 2. The linear component of the log2-transformed IL-10 8

1\*

8

	TABLE Z							
Marginal posterior inclusion probability and term importance								
Coefficients	P (gamma = 1)	Pi	Dimension					
Linear (LIL10)	0.751	0.773	1**					
Spline (LIL10)	0.001	0.000	8					
fct (sex)	0.392	0.070	1*					
Linear (platelets)	0.136	0.059	1					

0.053

0.458

0.011

0.001

0.098

0.000

TADLE 2

\* P (gamma = 1) > 0.25; \*\* P (gamma = 1) > 0.5.

Spline (platelets)

Linear (lymphocytes)

Spline (lymphocytes)

Note: Shown are the posterior model probabilities from the MCMC 8,000 samples from 8 chains, each ran 5,000 iterations after a burn-in of 500.

(L-IL10) had a marginal inclusion probability (P\*[gamma = 1]) of > 0.5, indicating L-IL10 could be considered as a parametric feature. Similarly, the linear component of gender (P\*[gamma = 1 > 0.25) and the lymphocyte count (P\*[gamma = 1] > 0.25) all had high posterior probabilities that were related to the disease outcome.

The Bayesian feature selection approach suggested that linear components of the feature set were related to outcome. The feature set was therefore analyzed by  $\chi^2$  analysis, an approach that assumes the features have a linear relationship with outcome. The rank-ordered list of features included plasma IL-10 ( $\chi^2 = 17$ ), platelet count ( $\chi^2 = 14.2$ ), lymphocyte count ( $\chi^2 = 5$ ), IL-6 ( $\chi^2 = 6.602$ ), presence of diarrhea ( $\chi^2 = 6.234$ ), days of fever ( $\chi^2 = 5.938$ ), hemoglobin concentration  $(\chi^2 = 5.210)$ , and lymphocyte count  $(\chi^2 = 5.056)$  were the features with the largest  $\chi^2$  values. A box-plot presentation of these features by outcome is shown in Figure 1, where log2transformed IL-10, log2-TRAIL, log2-IL-6, and circulating neutrophils concentrations were increased in the DHF population relative to that of the DF population, with the remainder being decreased.

Logistic regression modeling of DHF. Because the feature reduction suggested that the clinical and laboratory data were linearly related with outcome, we used a logistic regression modeling approach for the prediction of DHF. Model building was performed using best subsets selection starting with the entire feature list (Table 2). Of the input variables, initial platelet and lymphocyte counts and log2-IL-10 concentration were features retained in the model. The odds ratio and 95% confidence limits are shown in (Table 3). Increases in IL-10 concentrations were associated with an increased probability of DHF, whereas decreases in platelet and lymphocyte counts were associated with an increased probability of DHF.

Model diagnostics. The receiver operating characteristic (ROC) curve was used to evaluate the model performance.<sup>16</sup> The area under the ROC (AUC) is equivalent to the probability that two cases, one chosen at random from each group, are correctly ordered by the classifier.<sup>17</sup> In the DHF logistic regression model, an AUC of 0.9615 was obtained (Figure 2). Overall, these findings confirmed the excellent performance of the logistic regression model on this data set.

To confirm the logistic regression model, we conducted a Bayesian variable selection, model choice, and regularized estimation in GAM fits separately modeling the parametric and nonparametric components of the features. The  $\chi^2$  statistic in the linear component analysis of deviance was statistically significant (Table 4, top), where the parametric (linear) components for L-IL10, platelets, and lymphocytes were highly significant with P values, whereas the nonlinear "smoothing" components of the IL-10, platelets, and lymphocytes are not significant at the level alpha of 0.05 with the GAM fit (using 3 degrees of freedom). Because the nonlinear components were not independently contributing to the model, the GAM analysis was repeated for the linear components only (Table 4, bottom). Here, P values of 0.037, 0.022, and 0.035 were obtained for L-IL10, platelets and lymphocytes. These values were equivalent to the results produced by logistic regression. Additionally, because the P value equaled 0.895 using the Hosmer and Lemeshow goodness-offit test, we concluded that the logistic regression response function is appropriate. Together, these data further validate the parametric modeling approach using linear regression.

Finally, we examined the distribution of residuals for the logistic regression model. Residual plots of the predictors are useful for examining if individual points are not well fit or influence model performance. We used the deviance residual, partial residual, influence on single fitted value, and influence on the regression coefficients to identify influential observations (not shown). Two outlier/influential data points were identified. The logistic regression model building, including or excluding these observations, produced no significant difference of P values for the IL-10, lymphocytes, and platelet coefficients indicating that the model was robust.

CART decision tree. To aid in the clinical application of the logistic regression model, we sought to represent the three predictive features as a decision tree. The CART is a machine learning tool that seeks to identify the best cutoff for each analyte that produces the most accurate classification of DHF or DF. Performing 10 trials using 10-fold cross-validation resulted in the best model with an average accuracy of 84.6% for DHF and 84.0% for DF (Figure 3). Here, four terminal nodes (indicated in red) are produced by the CART classifier. Twenty-five of the DF cases were predicted on the basis of platelet count alone; another six were identified on the basis of low L-IL10 concentrations and low lymphocyte counts. The AUC for the test data was 0.87.

Random forest classification. To extend the CART analysis, a Random Forest classifier was constructed. Random Forest generates an ensemble of regression trees, and from this, variable importance can be calculated. Variable importance is the relative measure of the influence of the variable on model accuracy. Analysis of the variable importance confirms that L-IL10 and platelets are the two most important variables in the classification, whereas hematocrit and age were least important (Figure 4). We interpret the finding that both the logistic regression and random forest models converge on IL10 and platelets indicate that these features are robust predictors of disease.

Correlations with clinical symptoms. Finally, we examined the correlation between the prominent dengue symptoms (headache, chills, rash, myalgia, cough, and diarrhea) with these two highly informative features (platelet count and IL-10). Weak, but significant correlations were between the presence of rash and platelet count (r = 0.311, P < 0.027), and diarrhea with IL-10 (*r* = 0.280, *P* < 0.047).

#### DISCUSSION

Because previous work has shown that the mortality of DHF is improved with early detection and intensive treatment,<sup>9</sup> the identification of predictive models that aid in early

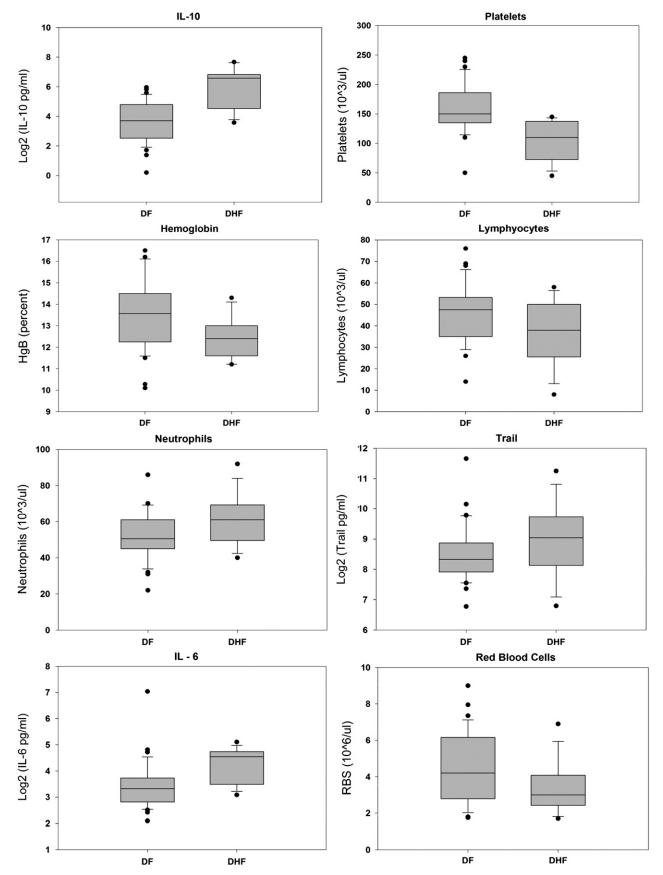


FIGURE 1. Box plots of differentially expressed laboratory values. Shown are the data for cytokines and laboratory measurements for dengue fever (DF) and dengue hemorrhagic fever (DHF) groups. Horizontal bar, median value; shaded box, 25-75% interquartile range (IQR); error bars, median  $\pm 1.5$  (IQR); black circle, outlier.

Linear (lymphocytes)

TABLE 3 Adjusted odds ratios for dengue hemorrhagic fever (DHF) logistic regression model

Effect	Point estimate	95% Wald confidence limits				
Platelet count	0.964	0.934	0.994			
Lymphocytes	0.890	0.802	0.989			
IL-10	5.944	1.172	30.136			

detection of DHF will have an important translational impact in the clinic. Identification of single predictive biomarkers has been elusive; however, the combinations of clinical features and laboratory tests may be informative. Here, we sought to identify predictive models based on assessable measurements that would be currently available in the clinic and laboratory. A major challenge in multivariate modeling is to identify the appropriate modeling approach; the performance of various modeling approaches is highly dependent on the underlying data structures. We have approached this problem using Bayesian modeling, a powerful technique that can model both parametric and nonparametric components of predictive features. This analysis suggested to us that parametric modeling approaches could be applied to these clinical features, a finding that was confirmed by the  $\chi^2$  analysis. A logistic regression model was produced using best subsets selection, which seeks to identify the most meaningful covariates in the group of features related to the probability of DHF. The IL-10, platelet, and lymphocyte concentrations were retained in the model with statistical significance. A random forest classifier also identified IL-10 and platelet count as the major informative features for DHF. These laboratory values have special relevance to the pathogenesis of DHF.

Previous studies have shown that clinical laboratory measurements obtained at the time of initial clinical presentation can partially predict DHF. In a cohort of Thai children presenting within 72 h of symptoms, CART analysis was developed to identify features that predict severe dengue illness.<sup>18</sup> The best CART model produced had a 97% sensitivity for predicting severe dengue illness, but only correctly excluded 48% of non-severe cases.<sup>18</sup> The splitting variables

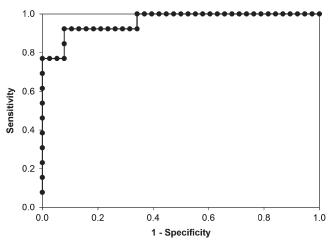


FIGURE 2. Receiver operating characteristic (ROC) curve for the LR model of dengue hemorrhagic fever (DHF). Shown is an ROC curve for the LR predictive model for DHF. Y axis, sensitivity; X axis, 1-specificity.

TABLE 4 Model analysis of deviance tests

GAM analysis incorpo	orating both linea	r and smoothing compo	nents	
Parameters	df	$\chi^2$	Pr > chisq	
Linear (L-IL10)	1	3.460	0.071*	
Linear (platelets)	1	6.000	0.019**	
Linear (lymphocytes)	1	3.725	0.060*	
Spline (L-IL10)	2	5.577	0.162	
Spline (platelets)	2	2.562	0.278	
Spline (lymphocytes)	2	3.341	0.188	
GAM analysis	s incorporating of	nly linear components		
Linear (L-IL10)	1	3.460	0.037**	
Linear (platelets)	1	6.000	0.022**	

Note: Two generalized additive model (GAM) analyses are shown. For each, the dengue fever (DF), degrees of freedom (df), and dominant factors significant at the level = 0.1 (\*) and 0.05 (\*\*) are shown for each parameter. L-IL10, Log2-transformed IL-10 concentration.

3.725

1

identified included white blood cell, monocytes, platelet counts, and hematocrit.<sup>18</sup> Another CART model that differentiated acute dengue from non-dengue febrile illness identified thrombocytopenia as an important discriminating variable.<sup>19</sup> Our study confirms the prognostic importance of reduced platelet counts for DHF. We observe here that the addition of systemic IL-10 levels provide much more predictive accuracy than that produced by circulating cell counts. Of importance, our logistic regression modeling outperformed the CART approach, with an overall AUC of 0.9615. Moreover, logistic regression is a classification approach that provides probabilitistic information for development of DHF, a feature that will aid in future validation studies. Because of the spectrum of DHF severity in our study, we are not able to evaluate the performance of our classifier on the most severe forms of DHF. Larger studies including cases with more severity will be required to resolve this issue.

Our study indicates that IL-10 measured approximately within 1 week of fever is highly related to the risk of developing DHF. We note that IL-10 is an immunosuppressive cytokine<sup>20</sup> secreted by primary monocytes in response to DENV infection mediated by antibody-dependent enhancement.<sup>21,22</sup> Interestingly, IL-10 production reflects active infection, because its secretion by monocytes is replication dependent, and not induced by Fc receptor ligation.<sup>21</sup> The detection of increased IL-10 in our DHF samples from acutely infected patients is consistent with this observation in vitro. Previous work has shown that IL-2, IL-4, IL-6, IL-10, IL-13, and IFN-γ are found in plasma in increased concentrations in patients with severe dengue infections.<sup>23</sup> Moreover, in a prospective study of a single serotype outbreak in Cuba, IL-10 was observed to be higher in severe dengue infections.<sup>24-26</sup> The appearance of IL-10 in our predictive model therefore has biological plausibility. It is important to point out that the predictive information of IL-10 appears to be constrained to times of initial presentation. In data not shown, IL-10 measurements obtained in the same individuals 5 days after study enrollment were not significantly different between DF and DHF. This finding suggests that different protein profiles are produced as the infection evolves.

Reduced platelet concentrations were also identified as being associated with DHF in this study. Thrombocytopenia is a well-established feature of DHF, responsible in part for

0.035\*\*

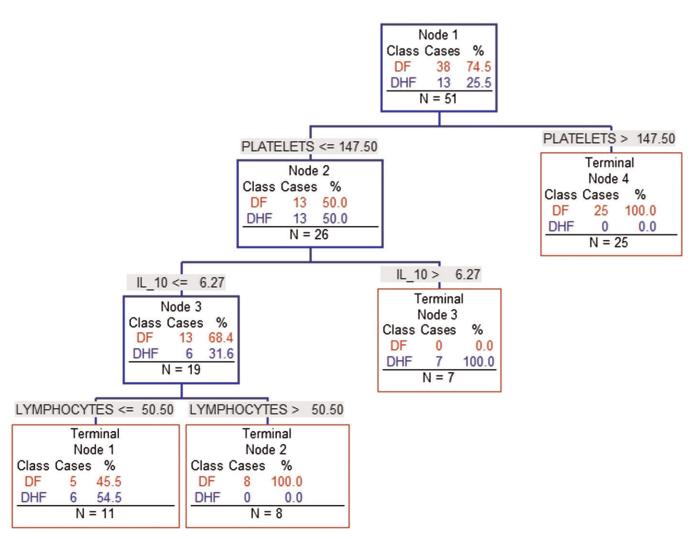


FIGURE 3. Classification and regression tree (CART) for prediction of dengue hemorrhagic fever (DHF). Shown is a CART decision tree for classification of DHF. The number of dengue fever (DF) and DHF cases in the study population and their percentage is indicated in each node. Terminal nodes indicated by red.

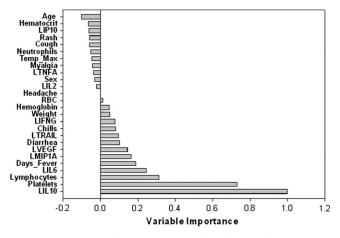


FIGURE 4. Feature importance analysis using Random Forests. Shown is a variable importance plot for the prediction of dengue hemorrhagic fever (DHF) using Random Forest classifier. Variable importance (X axis) is computed as 100% times the change in the margins averaged over all cases. Abbreviation: L = log2-transformed cytokine concentration.

an increased tendency for cutaneous hemorrhages. The origin of thrombocytopenia in DHF is multifactorial, because of the consequence of both bone marrow depression and accelerated antibody-mediated platelet sequestration by the liver.<sup>27</sup>

Cell-mediated immunity is an important protective immune mechanism to dengue infections. Although circulating lymphocyte counts are not reflective of cellular activation, our study indicates that patients who develop DHF have reduced lymphocyte concentrations at presentation. A prospective study of 91 subjects with DENV infection in Taiwan described that a lower percentage of "typical" lymphocytes were observed in subjects with severe dengue infection.<sup>28</sup> Our findings here indicate that lymphocyte counts are also reduced in DHF, but lymphocyte concentrations are not as informative as IL-10 concentrations and platelet counts with disease outcome.

Our study will need to be replicated and validated in a larger study population. Replication will need to include both testing whether IL-10 measurements using a clinical laboratory-based enzyme-linked immunosorbent assay is discriminatory of DHF and re-establishing the optimal cutoff value of IL-10 for a larger population. To aid in these potential validation studies, we have subjected the three features identified by probabilistic logistic regression modeling to CART analysis. The CART trees are readily human interpretable as simple decisions that result in a classification. We emphasize that the CART model here does not quite perform as well as the logistic regression model (in terms of AUC). Nevertheless, the CART analysis suggests that the subjects with dengue infections have specific characteristics; those with high platelet counts (> 147.5 × 10<sup>3</sup>/µL) are very likely to have uncomplicated DF (represented as Node 4, Figure 3), whereas those with low platelet counts (< 147.5 × 10<sup>3</sup>/µL) and high IL-10 are likely to have DHF (represented as Node 3, Figure 3). The group with low platelet counts, low IL-10 concentrations, and low lymphocyte counts are equally represented by DF and DHF outcomes.

In summary, parametric modeling approaches using accessible laboratory data (IL-10, platelets and lymphocyte counts) from patients acutely presenting with RT-PCR-confirmed dengue infections show promise for the early detection of DHF. These predictive models will require further validation on independent study populations.

Received July 19, 2011. Accepted for publication November 18, 2011.

Financial support: This work was supported by the NIH/NIAID Clinical Proteomics Center, HHSN272200800048C (ARB), 1U54RR02614 UTMB CTSA (ARB) and the Military Infectious Diseases Research Program work unit 6000RAD1.S.B0302.

Disclosure: Josefina Garcia, Eric S. Halsey, Patrick J. Blair, Claudio Rocha, Isabel Bazan, and Tadeusz J. Kochel are military service members or employees of the U.S. Government. This work was prepared as part of their official duties. Title 17 U.S.C. §105 provides that 'Copyright protection under this title is not available for any work of the United States Government.' Title 17 U.S.C. §101 defines U.S. Government work as a work prepared by a military service member or employee of the U.S. Government as part of that person's official duties.

Disclaimer: The views expressed in this article are those of the author and do not necessarily reflect the official policy or position of the Department of the Navy, Department of Defense, or the U.S. Government.

Authors' addresses: Allan R. Brasier, Department of Internal Medicine, University of Texas Medical Branch, Galveston, TX, E-mail: arbrasie@utmb.edu. Hyunsu Ju and Heidi M. Spratt, Department of Preventive Medicine and Community Health, University of Texas Medical Branch, Galveston, TX, E-mails: hyju@utmb.edu and hespratt@utmb.edu. Josefina Garcia, Brett M. Forshey, Eric S. Halsey, and Claudio Rocha, U.S. Naval Medical Research Unit Six Unit 3230, DPO AA, FL, E-mails: Josefina.Garcia@med.navy.mil, brett.forshey@ gmail.com, claudio.rocha@med.navy.mil, and Eric.halsey@med.navy .mil. Sundar S. Victor, Department of Biochemistry and Molecular Biology, University of Texas Medical Branch, Galveston, TX, E-mail: ssvictor@utmb.edu. Guillermo Comach, Gloria Sierra, and Isabel Bazan, Instituto de Investigaciones Biomedicas de la Universidad de Carabobo, Estado Aragua, Venezuela, E-mails: gcomach@cantv .net, gmsierrac@yahoo.com, and isabelbazana@gmail.com. Patrick J. Blair, Naval Health Research Center, Respiratory Diseases, San Diego, CA, E-mail: Patrick.Blair@med.navy.mil. Amy C. Morrison and Thomas W. Scott, Department of Entomology, University of California, Davis, CA, E-mails: amy.aegypti@gmail.com and twscott@ ucdavis.edu. Tadeusz J. Kochel, Viral and Rickettsial Diseases Department, Naval Medical Research Center, Silver Spring, MD, E-mail: Tad .kochel@med.navy.mil.

## REFERENCES

 San Martin JL, Brathwaite O, Zambrano B, Solórzano JO, Bouckenooghe A, Dayan GH, Guzmán MG, 2010. The epidemiology of dengue in the Americas over the last three decades: a worrisome reality. *Am J Trop Med Hyg 82*: 128–135.

- Martina BE, Koraka P, Osterhaus AD, 2009. Dengue virus pathogenesis: an integrated view. *Clin Microbiol Rev* 22: 564–581.
- Graham RR, Juffrie M, Tan R, Hayes CG, Laksono I, Ma'roet C, Erlin, Sutaryo, Porter KR, Halstead SB, 1999. A prospective seroepidemiologic study on dengue in children four to nine years of age in Yogyakarta, Indonesia I. studies in 1995–1996. *Am J Trop Med Hyg 61:* 412–419.
- Guzman MG, Kouri G, Bravo J, Valdes L, Vazquez S, Halstead SB, 2002. Effect of age on outcome of secondary dengue 2 infections. *Int J Infect Dis 6*: 118–124.
- Thomas L, Verlaeten O, Cabie A, Kaidomar S, Moravie J, Najioullah F, Plumelle Y, Fonteau C, Dussart P, Césaire R, 2008. Influence of the dengue serotype, previous dengue infection, and plasma viral load on clinical presentation and outcome during a dengue-2 and dengue-4 co-epidemic. *Am J Trop Med Hyg* 78: 990–998.
- Kliks S, 1990. Antibody-enhanced infection of monocytes as the pathogenetic mechanism for severe dengue illness. *AIDS Res Hum Retroviruses 6*: 993–998.
- Green S, Rothman A, 2006. Immunopathological mechanisms in dengue and dengue hemorrhagic fever. *Curr Opin Infect Dis* 19: 429–436.
- Rothman AL, 2011. Immunity to dengue virus: a tale of original antigenic sin and tropical cytokine storms. *Nat Rev Immunol* 11: 532–543.
- Ranjit S, Kissoon N, Jayakumar I, 2005. Aggressive management of dengue shock syndrome may decrease mortality rate: a suggested protocol. *Pediatr Crit Care Med 6:* 6.
- Forshey BM, Guevara C, Laguna-Torres VA, Cespedes M, Vargas J, Gianella A, Vallego E, Madrid C, Aguayo N, Gotuzzo E, Suarez V, Morales AM, Beingolea L, Reyes N, Perez J, Negrete M, Rocha C, Morrison AC, Russell KL, Blair PS, HMRCD Febrile Surveillance Working Group, 2010. Arboviral etiologies of acute febrile illnesses in western South America, 2000–2007. *PLoS Negl Trop Dis 4*: e787.
- Lanciotti RS, Calisher CH, Gubler DJ, Chang GJ, Vorndam AV, 1992. Rapid detection and typing of dengue viruses from clinical samples by using reverse transcriptase-polymerase chain reaction. J Clin Microbiol 30: 545–551.
- 12. World Health Organization (WHO), 1997. *Dengue Haemorrhagic Fever: Diagnosis, Treatment, and Control.* Geneva: World Health Organization.
- Iswaharan H, Rao J, 2005. Spike and slab variable selection: frequentist and Bayesian strategies. *Ann Stat 33*: 730–773.
- Stone M, 1977. An asymptotic equivalence of choice of model by cross-validation and Akaike's criterion. J R Stat Soc, B 39: 44–47.
- Hassa PO, Hottiger MO, 2006. PARP-1 as novel coactivator of NF-kB in inflammatory disorders. Burkle A, ed. *PolyADP Ribosylation*. Austin, TX: Landes Bioscience, 75–87.
- 16. Fawcett T, 2006. An introduction to ROC analysis. *Pattern Recognit Lett 27:* 861–874.
- 17. Hanley JA, McNeil BJ, 1982. The meaning and use of the area under a receiver operating characteristic curve. *Radiology* 143: 29–36.
- Potts JA, Gibbons RV, Rothman AL, Srikiatkhachorn A, Thomas SJ, Supradish PO, Lemon SC, Libraty DH, Green S, Kalayanarooj S, 2010. Prediction of dengue disease severity among pediatric Thai patients using early clinical laboratory indicators. *PLoS Negl Trop Dis 4*: e769.
- Tanner L, Schreiber M, Low JG, Ong A, Tolfvenstam T, Lai YL, Ng LC, Leo YS, Thi Poung L, Vasudevan SG, Simmons CP, Hibberd ML, Ooi EE, 2008. Decision tree algorithms predict the diagnosis and outcome of dengue fever in the early phase of illness. *PLoS Negl Trop Dis 2*: e196.
- Saraiva M, O'Garra A, 2010. The regulation of IL-10 production by immune cells. *Nat Rev Immunol 10*: 170–181.
- Boonnak K, Dambach KM, Donofrio GC, Tassaneetrithep B, Marovich MA, 2011. Cell type specificity and host genetic polymorphisms influence antibody-dependent enhancement of dengue virus infection. J Virol 85: 1671–1683.
- 22. Chareonsirisuthigul T, Kalayanarooj S, Ubol S, 2007. Dengue virus (DENV) antibody-dependent enhancement of infection upregulates the production of anti-inflammatory cytokines, but suppresses anti-DENV free radical and pro-inflammatory cytokine production, in THP-1 cells. J Gen Virol 88: 365–375.

- Bozza F, Cruz O, Zagne S, Azeredo EL, Nogueria RM, Assis EF, Bozza PT, Kubelka CF, 2008. Multiplex cytokine profile from dengue patients: MIP-1beta and IFN-gamma as predictive factors for severity. *BMC Infect Dis 8*: 86.
- 24. Chen LC, Lei HY, Liu CC, Shiesh SC, Chen SH, Liu HS, Lin YS, Wang ST, Shyu HW, Yeh TM, 2006. Correlation of serum levels of macrophage migration inhibitory factor with disease severity and clinical outcome in dengue patients. *Am J Trop Med Hyg* 74: 142–147.
- Chen RF, Yang KD, Wang L, Liu JW, Chiu CC, Cheng JT, 2007. Different clinical and laboratory manifestations between dengue hemorrhagic fever and dengue fever with bleeding tendency. *Trans R Soc Trop Med Hyg 101*: 1106–1113.
- Perez AB, Garcia G, Sierra B, Alvarez M, Vázquez S, Cabrera MV, Rodriguez R, Rosario D, Martinez E, Denny T, Guzmán MG, 2004. IL-10 levels in dengue patients: some findings from the exceptional epidemiological conditions in Cuba. *J Med Virol* 73: 230–234.
   Mitrakul C, Poshyachinda M, Futrakul P, Sangkawibha N,
- Mitrakul C, Poshyachinda M, Futrakul P, Sangkawibha N, Ahandrik S, 1977. Hemostatic and platelet kinetic studies in dengue hemorrhagic fever. *Am J Trop Med Hyg* 26: 975–984.
- Eu-Ahsunthornwattana N, Eu-ahsunthornwattana J, Thisyakorn U, 2008. Peripheral blood count for dengue severity prediction: a prospective study in Thai children. *Pediatrics* 121: S127–S128.