An App Detecting Dengue Fever in Children: Using Sequencing Symptom Patterns for An Online Assessment

Tsair-Wei Chien, MBA¹, Julie Chi Chow, MD² Willy Chou, MD, MBA³,⁴

¹Medical Research Department, Chi-Mei Medical Center, Tainan, Taiwan
²Department of Paediatrics, Chi Mei medical center, Tainan, Taiwan
³Department of Recreation and Health-Care Management & Institute of recreation
⁴Industry Management, Chia Nan University of Pharmacy, Tainan, Taiwan

Taiwan

Running title: An App used for detecting dengue fever
Word count: 2,185
Table: 3 Figure: 3

Corresponding Author
Willy Chou
Chi-Mei Medical Center, 901 Chung Hwa Road, Yung Kung Dist., Tainan 710, Taiwan
E-mail: hsienyiwang@gmail.com

Tsair-Wei Chien (TW):smile@mail.chimei.org.tw
Julie Chi Chow (CC): jcchow2@yahoo.com.tw
Willy Chou(WC): ufan0101@ms22.hinet.net
Abstract

Background: Dengue fever (DF) is one of the most common arthropod-borne viral diseases worldwide, particularly in South East Asia, Africa, the Western Pacific, and the Americas. However, DF symptoms are usually assessed using a dichotomous (i.e., absent versus present) evaluation. There has been no published study that has reported using the specific sequence of symptoms to detect DF. An App is required to help patients/their family members or clinicians to identify DF at an earlier stage.

Objective: We developed an App examining symptoms to effectively predict DF.

Methods: We extracted statistically significant features from 17 DF-related clinical symptoms in 177 pediatric patients (69 diagnosed with DF) using (1) the unweighted summation score and (2) the non-parametric $H^T$ person fit statistic, which can jointly combine (3) the weighted score (yielded by logistic regression) to predict DF risk.

Results: Six symptoms (Family History, Fever $\geq 39^\circ$C, Skin Rash, Petechiae, Abdominal Pain, and Weakness) significantly predicted DF. When a cutoff point of 1.03 ($p = 0.26$) suggested combining the weighted score and the $H^T$ coefficient, the sensitivity was 0.91, and the specificity was 0.76. The area under the ROC curve was 0.88, which was a better predictor: specificity was 5.56% higher than for the traditional logistic regression.

Conclusions: Six simple symptoms analyzed using logistic regression were useful and valid for early detection of DF risk in children. A better predictive specificity
Increased after combining the non-parametric $H^T$ coefficient to the weighted regression score. A self-assessment using patient smartphones is available to discriminate DF and may eliminate the need for a costly and time-consuming dengue laboratory test.

**Key Words:** dengue fever, $H^T$ person mapping statistic, logistic regression, score summation, receiver operating characteristic curve

**Introduction**

Dengue fever (DF) is one of the most common arthropod-borne viral diseases worldwide[1], especially in South East Asia, Africa, the Western Pacific, and the Americas[2,3].

There is, however, no accurate and speedy diagnostic screening test for DF at an early stage because its signs and symptoms—e.g., fever, headache, and myalgia—are similar to those of other illnesses [4-6]. Some studies [4,5] that used a univariate analysis report that the presumptive diagnosis of DF is imprecise. Multivariate logistic regressions also do not significantly distinguish patients with dengue from those with other febrile illnesses [7]. The multivariate discrimination analyses reported sensitivity and a specificity 0.76, and an area under the receiver operating characteristic (ROC) curve (AUC) of 0.93, but costly laboratory tests (Dengue Duo IgM & Rapid Strips; Panbio, Queensland, Australia)[8-11] were needed before DF was serologically confirmed.

DF symptoms are usually assessed using a dichotomous (i.e., absent versus
present) evaluation. The dependent variable (DF$^+$ versus DF$^-$) predicted using independent evaluations with a weighted summation score is more accurate than that using simple evaluations with an unweighted summation score. So far, there has been no published study that has reported using the specific sequence of symptoms reported or observed in specific patients suspected of having DF. All published studies to date still report results using only a standard group of symptoms with an unweighted summation score and merely apply their results to a general group of patients that might have DF.

The non-parametric $H^T$ fit statistic has been used in education and psychometrics to identify aberrant test respondents[12,13]. It is a transposed formulation of a scalability coefficient for items (e.g., symptoms in this study) and the best among 36-person fit statistics for detecting abnormal behaviors [14].

In the present study, we used the $H^T$ coefficient combined with weighted and unweighted variables to examine whether these combinations provide a valid and reliable approach for the early detection of DF in children.

Methods

Sample and clinical symptoms

The sample of 177 pediatric patients ($\leq$ 16 years old; DF$^+$: 69; DF$^-$: 108) was the same as in our previous paper. Guided by the literature, we collected nineteen DF-related clinical symptoms from the patients’ medical records to develop the initial set
85 of items—designated as 0 = “absent” or 1 = “present”—to screen for DF infection: (i) personal history of DF, (ii) family history of DF, (iii) mosquito bites within the previous 2 weeks, (iv) fever ≥ 39°C, (v) biphasic fever, (vi) rash, (vii) petechiae, (viii) retro-orbital pain, (ix) bone pain (arthralgia), (x) headache, (xi) myalgia, (xii) abdominal pain, (xiii) anorexia, (xiv) occult hematuria, (xv) stool occult blood, (xvi) cough, (xvii) sore throat, (xviii) soft (watery) stool, and (xix) flushed skin. Data from these patients’ charts were obtained and approved by the Research Ethics Review Board of the Chi-Mei Medical Center.

93 The $H^T$ fit statistic

$H^T$ is defined for the persons of a dichotomous dataset with $L$ items (in columns) and $N$ persons (in rows), where $X_{ni}$ is the scored (0,1) response of person $n$ to item $i$, and $P_n = S_n/L$. Here, $S_m$ is the raw score for person $m$, and $S_n$ is the raw score for person $n$.

$$H^T(n) = \frac{\sum_{m=1, m \neq n}^{N} (\frac{\sum_{i=1}^{L} X_{ni} X_{mi}}{L} - p_n p_m)}{\sum_{m=1, m \neq n}^{N} (\min[p_n (1 - p_m), p_m (1 - p_n)])}$$

$H^T$ is the sum of the covariances between person $n$ and the other persons divided by the maximum possible sum of those covariances so that the range of $H^T$ is $-1$ to $+1$. When the responses by person $n$ are positively correlated with those of all the other persons, then $H^T(n)$ will be positive. In contrast, when the responses by person $n$
are negatively correlated with those of all the other persons, then $H^T(n)$ will be negative. When person $n$’s responses are random, $H^T(n)$ will be close to zero\[11\]. We hypothesized that DF$^+$ patients have different $H^T$ coefficients than do DF$^-$ patients. All DF$^+$ group members were sequenced to the DF$^-$ group members to obtain an $H^T$ coefficient using formula (1).

Selecting symptoms and determining predictor variables

All symptoms were examined by the probability of Type I error using the following three steps in Figure 1 to determine predictor variables. First, each symptom was separately examined by the univariate approach using a $\chi^2$ test and logistic regression, respectively, for identifying a significant association with DF. Second, two models (i.e., the univariate and the multivariate approaches) were investigated for determining valid predictor variables associated with DF when the probability of Type I error is less than 0.05. Third, the predictor variables were used in a weighted combination for discriminating patients suspected with dengue virus infection.
Detecting dengue fever: a comparison of three models

The efficacy of three models (A, B, and C) for detecting dengue fever was examined: (i) A comparison was made using univariate logistic regression in Model A.

Figure 1 Overall study concept and the flow chart

A. Score summation:
   (1) Unweighted
   (2) Weighted

B. Ht Mapping

Prediction

End
scores, Weighted (i.e., logistic regression) scores, and $H^T$ coefficients, respectively, (ii)

Multivariate logistic regression with the three aforementioned factors combined was

used in Model B, (iii) after selecting the significant variables in Model B, the

combined predictive variables were analyzed using multivariate logistic regression in

Model C to obtain effective weighted coefficients, and (iv) finally, we wanted to use a

single continuous variable yielded by the combined predictive variables in Model C to

compare the AUC with the counterparts in Model A and C.

Statistical tools and data analyses

SPSS 15.0 for Windows (SPSS Inc., Chicago, IL) and MedCalc 9.5.0.0 for

Windows (MedCalc Software, Mariakerke, Belgium) were used to calculate (i) the

probability of false positives (Type I error) using a $\chi^2$ test and logistic regression, (ii)

Youden J index (the higher, the better), AUC (area under the ROC curve), sensitivity,

specificity, and the cutoff point at maximal summations of specificity and sensitivity,

(iii) correlation coefficients among variables of unweighted, weighted, and $H^T$ scores.

Results

Sixty-nine pediatric patients clinically diagnosed with DF and 108 with no

evidence of DF infection were included in this study (Table 1). A $\chi^2$ test and logistic

regression analyses showed that only six symptoms (Family History, Fever $\geq 39^\circ$C,

Skin Rash, Petechiae, Abdominal Pain, and Weakness) were significant for assessing

the likelihood of DF (Table 2).
Table 1. Demographic characteristics of the study sample

<table>
<thead>
<tr>
<th>Demographical Variables</th>
<th>DF(-)</th>
<th>DF(+)</th>
<th>Total</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>%</td>
<td>n</td>
<td>%</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>4</td>
<td>43.5</td>
<td>2</td>
<td>9</td>
</tr>
<tr>
<td>Male</td>
<td>6</td>
<td>56.5</td>
<td>4</td>
<td>10</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0-4</td>
<td>4</td>
<td>8</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>5-9</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>9-16</td>
<td>3</td>
<td>6</td>
<td>3</td>
<td>54.4</td>
</tr>
</tbody>
</table>

146DF(+): patients with a positive dengue fever strip test
147DF(-): patients with a negative dengue fever strip test
148P-values were determined by the $\chi^2$ test
149

Table 2 Logistic analysis of symptoms for the patients suspected with dengue virus infection using the univariate approach

<table>
<thead>
<tr>
<th>Symptom</th>
<th>DF(-)</th>
<th>DF(+)</th>
<th>Total</th>
<th>Chi-square test</th>
<th>$\chi^2$</th>
<th>P-value</th>
<th>Logistic regression.</th>
<th>B</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variable</td>
<td>Prese</td>
<td>n</td>
<td>%</td>
<td>n</td>
<td>%</td>
<td>n</td>
<td>%</td>
<td>$\chi^2$</td>
<td>P-value</td>
</tr>
<tr>
<td>Family history</td>
<td>No</td>
<td>79</td>
<td>73.1</td>
<td>40</td>
<td>58.0</td>
<td>119</td>
<td>67.2</td>
<td>3.74</td>
<td>0.053</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>29</td>
<td>26.9</td>
<td>29</td>
<td>42.0</td>
<td>58</td>
<td>32.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>High fever of 39°C</td>
<td>No</td>
<td>87</td>
<td>80.6</td>
<td>37</td>
<td>53.6</td>
<td>124</td>
<td>70.1</td>
<td>13.30</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>21</td>
<td>19.4</td>
<td>32</td>
<td>46.4</td>
<td>53</td>
<td>29.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Skin rash</td>
<td>No</td>
<td>82</td>
<td>75.9</td>
<td>20</td>
<td>29.0</td>
<td>102</td>
<td>57.6</td>
<td>36.09</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>26</td>
<td>24.1</td>
<td>49</td>
<td>71.0</td>
<td>75</td>
<td>42.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Petechiae</td>
<td>No</td>
<td>106</td>
<td>98.1</td>
<td>60</td>
<td>87.0</td>
<td>166</td>
<td>93.8</td>
<td>7.29</td>
<td>0.007</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>2</td>
<td>1.9</td>
<td>9</td>
<td>13.0</td>
<td>11</td>
<td>6.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Abdominal pain</td>
<td>No</td>
<td>104</td>
<td>96.3</td>
<td>53</td>
<td>76.8</td>
<td>157</td>
<td>88.7</td>
<td>14.03</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>2</td>
<td>3.7</td>
<td>9</td>
<td>13.2</td>
<td>11</td>
<td>6.2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Comparisons of the AUCs for the three study models (A, B, and C) showed that the weighted variable (derived by the Logistic regression) and the $H^T$ coefficient could be jointly used for predicting DF risk using equation (2):

$Logit = -3.32 + 0.93 \times \text{weighted score} + 1.92 \times H^T_{\text{coefficient}}$.

The risk probability can be computed using the transformed formula (3):

$p = \frac{\exp(\text{logit})}{1 + \exp(\text{logit})}$, \hspace{1cm} (3)

where logit denotes a unit of log odds.

A cutoff point of $-1.03$ ($p = 0.26$) was determined using the combined predictive variables in Model C: sensitivity = 0.91, specificity = 0.76, and AUC = 0.88 (Figure 2 and Table 3). Predictive power was better: specificity was 5.56% (i.e., 75.93-70.37 shown in Table 3) higher than when using traditional logistic regression; however, the AUC was slightly lower (0.72) than when using the unweighted (0.84) and the weighted (0.87) variables (Table 2). The $H^T$ coefficients related to the weighted and unweighted scores were 0.26 and 0.22, respectively. The weighted score has a higher correlation coefficient than does the unweighted score to the $H^T$ coefficients.
Figure 2 Four models plotted by ROC curves

### Table 3: Comparisons of AUC for the study models

<table>
<thead>
<tr>
<th>Approach</th>
<th>Logistic regression</th>
<th>ROC curve analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Steps</td>
<td>B, p-value, AUC, Youden J, Cut point, Sensitivity, Specificity</td>
<td></td>
</tr>
</tbody>
</table>

(1) Model A: Univariate approach with a single variable compared to the DF using Logistic regression and ROC analysis

| Unweight^c | 1.60* | <0.001 | 0.84 | 0.58 | >1.00 | 79.70 | 78.70 |
| Weight^d   | 0.97* | <0.001 | 0.87 | 0.61 | >-0.93| 91.30 | 69.40 |
| H^e coeff. | 3.75* | <0.001 | 0.72 | 0.53 | >0.15 | 65.20 | 88.00 |

Unweight

(2) Model B: Multivariate approach with combined these three variables in regressing the DF using Logistic regression

| Weight 0.31 | 0.77* | 0.014 |
| 0.595       |
| H^e coeff.  | 3.08* | 0.001 |
| Constant    | -1.03 | 0.350 |
(3) Model C: 
Combined these two significant predictor variables using Logistic regression

<table>
<thead>
<tr>
<th></th>
<th>Coef.</th>
<th>SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weight H^t</td>
<td>2.962*</td>
<td>0.01</td>
</tr>
<tr>
<td>coeff.0.919*</td>
<td></td>
<td>0.001</td>
</tr>
<tr>
<td>&lt;0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>-0.463</td>
<td>0.75</td>
</tr>
</tbody>
</table>

(4) A single continuous variable yielded by the combined predictor variables

Note. a: coefficient of Logistic regression, b: Youden J index, c: item-score summation method, d: multiplying item-score with the weighted regression coefficient, e: the H^t coefficient, f: using the two combined variables to predict patient’s DF, *:p<.05
Figure 3 Snapshots on a smart phone responding questions (top) and the result (bottom) for assessing the patient DF.
A snapshot on a smart phone responding to questions (Figure 3, top) was generated and the results for assessing whether the patient has DF (Figure 3, bottom) were determined, which indicated that patients suspected of having DF could directly scan the QR-code to obtain their DF logit scores (or the risk probability) and examine whether these 6 symptoms are useful for predicting a high DF risk ($> -1.03 \text{ logits}$ or $p \geq 0.26 = \exp(-1.03 \text{ logits})/(1+\exp(-1.03 \text{ logits}))$).

**Discussion**

We found that using the $H^T$ coefficient yielded predictions that were 5.56% more specific (i.e., 75.93-70.37 shown in Table 3) than those of traditional logistic regression. The $H^T$ index is promising when the patient sequence symptom pattern is compared with the DF$^+$ group to detect dengue fever in children. It can be combined with the weighted summation score to jointly predict the DF risk and then to report that risk on smartphones.

The $H^T$ coefficient has been used in education and psychometrics to identify aberrant test respondents [12,13]. Although some have used item response theory (IRT) fit statistics (e.g., outfit mean square error > 2.0) to select abnormal responses that indicate cheating, careless responding, lucky guessing, creative responding, or random responding[15], our literature review revealed no published papers that reported using the $H^T$ coefficient in medical settings, especially for detecting
individual aberrant response patterns different from the study reference sample, or,
like the current study, identifying the DF risk by comparing their sequence symptom
pattern to that of the DF+ group.

A diagnosis of DF is usually confirmed by three steps: (i) observing DF-related
symptoms, (ii) testing laboratory data such as white blood cells (WBCs) and platelets
(PLTs), and (iii) serologically verifying DF using dengue IgM and IgG antibodies,
polymerase chain reaction (PCR) analysis, and virus isolation tests. The latter two are
relatively expensive. It is needed to develop a self-assessment approach (e.g.,
scanning QR-code, responding questions, and obtaining the DF risk on his/her smart
phone) (1) helping patients for consultation at an earlier stage, (2) prompting doctors
sampling patient laboratory data when he/her DF risk reaches a cutpoint of $P = 0.26$
($=\exp(-1.03 \text{ logits})/(1+\exp(-1.03 \text{ logits}))$).

We found that the weighted score was a better predictor than was the unweighted
score (see Model A and Model B in Table 3). However, we still see so many scales in
a medical setting using unweighted summation scores to determine the presence or
absence of disease. Along with the smartphones popularly used in the technical age,
the way of obtaining the DF risk on smartphones using the combined $H^T$ coefficient
and weighted scores is available and worth recommending to healthcare providers to
use for detecting the risk for DF.
This study has some limitations. First, the DF cutpoint based on the symptoms of our study sample might be biased toward that population. Moreover, we did not remove abnormal data when the $H^r$ coefficient was less than the critical value of 0.22, which best identifies aberrantly responding examinees[14]. Second, although the sample size was small, using the Rasch $H^r$ coefficient combined with the AUC yielded highly accurate discriminatory screening. This finding, however, requires confirmation in prospective studies of other regions with a substantial incidence of DF.

Conclusions

Analyzing six simple symptoms using logistic regression is useful and valid for the early detection of DF risk in children. Combining the Rasch $H^r$ coefficient with the weighted score yields a prediction that is 5.56% more specific than does traditional logistic regression. A self-assessment app using patient smartphones is available to help people suspected of having DF, and it might eliminate the need for costly and time-consuming laboratory tests.

List of abbreviations

AUC: Area under ROC curve
DF: Dengue fever
ROC: Results: a receiver-operating characteristic

Competing interests

The authors declare that they have no competing interests.
Authors’ contributions

TWC conceived and designed the study, performed the statistical analyses and were in charge of recruiting study participants. CC and TWChelped design the study, collected information and interpreted data. WC monitored the research. All authors read and approved the final article. This research was supported by grant Chi-Mei Foundation Hospital research CMFCR10593 from the Chi-Mei Medical Center. The authors have no other funding or conflicts of interest to disclose.

REFERENCES

6. Hammond SN, Balmaseda A, Pérez L, et al. Differences in dengue severity in


Figure Legends

Figure 1 Overall study concept and the flow chart.

Figure 2 Four models plotted by ROC curves.

Figure 3 Snapshots on a smart phone responding questions (top) and the result (bottom) for assessing the patient DF.