Evaluation of Possible Dengue Outbreak Detection Methodologies for Thailand, which one should be implemented?

by

Supharerk Thawillarp

A thesis submitted to Johns Hopkins University in conformity with the requirements for the degree of Master of Health Sciences Informatics-Research

Baltimore, Maryland

May, 2017

© 2017 Supharerk Thawillarp

All Rights Reserved
Evaluation of Possible Dengue Outbreak Detection Methodologies for Thailand, which one should be implemented?

Supharerk Thawillarp¹, Carlos Castillo-Salgado², Harold P. Lehmann¹

¹ Division of Health Sciences Informatics, Johns Hopkins University School of Medicine
² Department of Epidemiology, Johns Hopkins University Bloomberg School of Public Health

Abstract

Dengue has become a more impact vector-borne disease than malaria globally in both morbidity distribution and economic resources. The burdens of Dengue in Thailand seem to be one of the highest in the South-East Asia Region of the World Health Organization (WHO). Current WHO’s Dengue Strategic Plan is to improve the countries’ capacity for early detection to allow timely outbreak prevention and control. However, there are only few published research papers that address the use, implementation and evaluation of novel early detection methodologies. This is a first national operational research project with objectives to study the feasibility of implementing modern dengue outbreak detection methodologies and to evaluate these methods at better alternatives to the current outbreak detection method (median-5-years) in Thailand. We conducted a descriptive Ecological study from a complete Dengue dataset retrieved from the Department of Epidemiology, Ministry of Public Health, Thailand. During 2003-2015, there were 1,014,201 visits and 13 outbreaks of Dengue virus with the largest attack in 2013. While each of the studied detection methods displayed unique characteristics, we observed similar values of the averages and medians, upper Confidence Intervals (CI) and percentiles across three methods. Same period media-5-years might be good for alert threshold. EARS methods were able to detect every outbreak but they did not provide information on outbreak long-term trend or magnitude. Moving percentiles or upper CI could provide information for long-term trends and epidemic thresholds. Off-seasonal median or average might be suitable for seasonal thresholds. Each detection method has its own strengths and weaknesses, thus implementing these methodologies could be of great epidemiological assistance local public health surveillance systems and for early Dengue outbreak detection.

Keywords: Dengue, Public Health Informatics, early disease detection, Surveillance systems, Disease Notification, Thailand
Preface

This thesis is the result of my hard work at Division of Health Sciences Informatics (DHSI), Johns Hopkins University as part of my master program. Before I was here, I was thinking about using sophisticate algorithm, innovation and complex way to mitigate the problem. After reviewing many literature just to find out that most modern methods required more comprehensive data which is not available in my country. I keep digging for a couple of months but none of them works.

After a period of confusion, I think there should be another solution that is simple but feasible. With this alternative approach, I found an answer I was looking for. I realized that doing that will require an advance mathematics and statistics, which is wrong. The simple answer was not taken lightly from the researcher, instead because they have work intensive and thoroughly to make it simple and applicable. It is like we change our position and the way we look at the data also change. I have an important lesson that will change my life forever.

I am grateful to all people from both Department of Epidemiology and DHSI that were involved in this study especially Dr. Carlos Salgado Castillo and Dr. Harold P. Lehmann for the support during this research. I also thanks Hojoon Daniel Lee for helping me with the analyzed. I also would like to thank my friends fellow students for their support.

I hope you will enjoy reading the thesis.

Supharerk Thawillarp

Baltimore, May 2017
Table of Contents

Abstract ................................................................................................................................. ii

Preface ................................................................................................................................. iii

Introduction ............................................................................................................................ 1

Objectives............................................................................................................................... 2

Literature Review .................................................................................................................... 2

  Dengue ................................................................................................................................. 2

  Clinical ................................................................................................................................. 4

Case definitions ....................................................................................................................... 5

History of Dengue ................................................................................................................... 6

Mosquito Control ..................................................................................................................... 8

  Patient ................................................................................................................................. 9

  Contact ............................................................................................................................... 9

  Environment ..................................................................................................................... 9

Surveillance system ............................................................................................................... 10

  Thailand reporting system .............................................................................................. 10

  Thailand Electronic Health Record (EHR) and HIS ............................................................ 12

Surveillance Algorithms ....................................................................................................... 14

Study Design .......................................................................................................................... 15

Study Population .................................................................................................................... 15

Data sources .......................................................................................................................... 16

Detection methodologies ...................................................................................................... 16
List of Table and Figures.

Figure 1 Dengue Serotypes Circulation in Thailand (MOPH Thailand) ................................................................. 3
Figure 2 Aedes aegypti life cycle.................................................................................................................................. 6
Figure 3 Dengue, countries or areas at risk .................................................................................................................. 7
Figure 4 R506 Reporting System .................................................................................................................................. 11
Figure 5 R506 Application (MOPH) .......................................................................................................................... 13
Figure 6 Comparison of the EARS Algorithms and Moving-Median-5-years .............................................................. 23
Figure 7 Moving threshold methods (a) Median 5 years ............................................................................................... 26
Figure 8 Same period moving threshold methods ........................................................................................................ 27
Figure 9 Off-seasonal threshold methods ................................................................................................................... 28
Figure 10 Sensitivity and False Positive Rate (1-Specificity) among every alert threshold compared to same period
moving median 5 years (Baseline) ......................................................................................................................... 29

Table 1 Thailand Clinical Case Definitions ................................................................................................................ 7
Table 2 Methods for determining seasonal thresholds Epidemiological Surveillance Standards for Influenza ......... 14
Table 3 Characteristics of DF/DHF/DSS visit and incidence rate per 100,000 person-year, Ministry of Public Health
Thailand, 2003-2015 ............................................................................................................................................. 21
Table 4 Positive Predictive Value, Sensitivity, and Specificity of Detection Algorithms ............................................... 24
**Introduction**

Dengue fever is a vector-borne disease caused by Dengue virus which is transmitted by mosquitoes (*Aedes aegypti* and *Aedes albopictus*) and is commonly found in tropical countries. Clinical presentation can vary from mild fever to shock. There is no definite treatment or vaccination for Dengue, only supportive treatment is available. As of 2012, Dengue becomes more impact vector-borne disease than malaria globally in both morbidity and economically (1). Among them, South-East Asia region seems to face the largest impact (2). Thailand, a tropical country in South East Asia with an approximate 65 million people and faces a nation-wide Dengue Outbreak every 1-2 years resulting in many live and economic burdens. In 2015, Ministry of Public Health, Thailand (MOPH) reported 144,952 cases (222.58 per 100,000) including 147 deaths (0.23 per 100,000) from Dengue (3). The burden of Dengue in Thailand seems to be the highest globally (4) and It was estimated that Dengue has placed a total burden of US $485 million including illness, treatment, and vector control. (5)

The current World Health Organization (WHO) Dengue Strategic Plan in South-East Asia Region is centered in improving their ability to improves its early detection and timely outbreak control (2). These interventions could significantly reduce fatality rate (6). One intervention to control Dengue transmission is to eliminate its vector, mosquitoes by applying insecticide in water containers, and insecticide spraying. Combining those with medical resources, they require a huge amount of resources. With limited budget, if the authorities of the Ministry of Health know accurately when Dengue outbreak is approaching, they can manage resources more effectively and allows better Dengue control (7).

Up until now, there is a large number of published papers on outbreak detection, however most of them are about Influenza. Even WHO did publish Interim Global Epidemiological Surveillance Standards specifically on Influenza (8). In the contrary, there are only few papers on Dengue outbreak detection. Runge-Ranzinger et al. mentioned in his Dengue disease surveillance systematic literature review that none of reviewed studies evaluated the threshold for initiate outbreak response and distinguish the outbreak from expecting seasonal Dengue (9). They also stated that researching among this topic for Dengue is the highest priority.

The Thailand Ministry of Public Health (MOPH) is using median-5-years as a threshold for outbreak detection of infectious diseases including Dengue for a long time. Median-5years threshold was calculated from finding median visit
number among same period 5 years historically. For example, median-5-years for May, 2015 will be calculated from May, 2010-2014. Despite being attacked by many large outbreaks for a long time, there is no study assessing its performance against other modern detection methodologies and no quantitative threshold to distinguish outbreak from seasonal Dengue as well. It is very important since this threshold was implemented in every public health office Health Information System (HIS) application “R506” throughout the country for data cleaning, managing and analyzing secondary data from hospitals to provide recommendation for health policy, disease control and administrative use. Implementing more outbreak detection methods could improve Dengue surveillance system with a very low cost.

This is the first national research to study the feasibility of implementing modern outbreak detection methodologies and will compare modern detection methodologies against current outbreak detection method (median-5-years) in Thailand. Possible recommendation would be modifying or implementing new detection methodologies in every public health office HIS if proved to be feasible and better than current method. This can allow the government to prepare itself earlier by stocking essential resources and for early intervention. Having more essential resources for disease control intervention, more lives can be saved.

**Objectives**

1. To assess the feasibility of modern surveillance algorithm and threshold implementation for Dengue in Thailand.


**Literature Review**

**Dengue**

Dengue is cause by a Dengue virus. It’s in genus Favivirus, Family Flavivirdae. Dengue virus has four serotypes; DEN-1, DEN-2, DEN-3, and DEN-4.

DEN-1 patient will have higher risk of Dengue Hemorrhagic Fever, and most patients have more severe clinical conditions in primary infection (10) (11).
DEN-2 patients have overall more severe clinical manifestations. Patients usually present with abdominal pain and hepatomegaly. DEN-2 severity will increase if case of secondary infection. This serotype is the most common in secondary infection patients with Dengue Hemorrhagic Fever (12) (13) (14).

DEN-3 patients have a higher incidence of musculoskeletal and Gastro Intestinal track manifestations such as nausea and vomiting. DEN-3 is responsible for most of severe Dengue outbreak in Thailand (1987, 1995-1999) and it is the most common serotype in primary Dengue infection. Some literature reported that DEN-3 patients also have severe clinical manifestations in secondary infection (13) (14) (15).

DEN-4 patients have higher incidence of respiratory and cutaneous manifestations. DEN-4 is the most common serotype for secondary infection (13) (16).

Dengue serotype circulation in Thailand was not consistent as they continuously shift from one to another serotype each year, most of the time with a mixture serotype. DEN-1 seems to continuously increase its contribution during last 40 years in contrary to DEN-4 which is decreasing.

**Figure 1 Dengue Serotypes Circulation in Thailand (MOPH Thailand)**
It transmitted to humans through the bites of the female *Aedes aegypti* mosquito (The male mosquito is not blood feeding and doesn’t bite). *Aedes aegypti* is a tiny, black mosquito with banded legs. It is a very common mosquito in tropical areas as they cannot survive a colder environment. They usually bite human preferably indoor for approximately two hours after sunrise and several hours before sunset. Aside from Dengue, *Aedes aegypti* is also a primary vector for many viral diseases such as chikungunya, yellow fever, Zika fever and others diseases. Because of that, it’s considered as one of the deadliest insect in the world (17).

This mosquito usually lays its eggs in man-made water containers and natural containers such as plant leaves, tree holes under dark or shady environment. The container with organic material such as leaves, algae is more preferred by the mosquito as it contains nutrient for the larva. (16)

Three days after they bite humans, the mosquito will start to lay its eggs. Because of their short flight range (200 meters), egg sites are usually close to location where mosquito was found. Their eggs are very durable and can survive in open environments for more than six months. Then eggs expose to water, it become hatched. The larva feeds on organic substances in the container. A week later, they become adult mosquito. Their life span is around three weeks.

### Clinical

After being bitten by the mosquito, 5-8 days later the patients start to develop symptoms. Patients symptoms are unspecific ranging from fever, headache, myalgia, arthralgia, orbital pain, anorexia, nausea, vomiting and rash. In overall, Dengue case-fatality rate is 0.025. There is no specific treatment available. There are three type of Dengue; Dengue Fever, Dengue Hemorrhagic Fever and Dengue Shock Syndrome.

Dengue Fever (DF) patient usually develops fever, headache, vomiting, muscle pain, and skin rash after 3-4 days after infection. Most patient recover within 2-7 days later, however, small number of patients progress in to Dengue Hemorrhagic Fever. As Thailand is also a Chikungunya endemic area, it’s important to distinguish DF from Chikungunya as both present similar clinical in early stages (18).

Dengue Hemorrhagic Fever (DHF) is characterized by its clinical symptoms; high fever, hemorrhagic phenomena (positive tourniquet test, petechial, easily bruising and sometimes, epitaxies), hepatomegaly and cardiovascular failure. Its unique laboratory finding is thrombocytopenia with hemoconcentration. In this stage, patients usually have plasma leakage which is the case of hemoconcentration (increase hematocrit). Plasma leakage is the main process that is
responsible to DHF severity and distinguishes it from DF. At the end of the febrile phase (2\textsuperscript{nd} - 7\textsuperscript{th} days), patient temperature starts to drop and often present with circulatory disturbance. In this stage patients, may become sweat, restless or have cool extremities, increase in pulse rate and decrease in blood pressure might be seen. Most patients spontaneously recover after with or without a supportive treatment of Intravenous Fluid and electrolyte. In more severe cases, patients might develop a shock or even death if not properly treated. Early diagnosis and treatment largely improve the outcome of DHF.

Dengue Shock Syndrome (DSS) is the condition among patient who develops shock after the febrile phase. Patient may present cool skin, cyanosis, rapid pulse, lethargy and acute abdominal pain. DSS is characterized by rapid pulse and narrow pulse pressure less than 20 mmHg. DSS has a very high mortality rate and patients die within 12-24 hours without appropriate treatment. If patients are able to survive this period, they will start to recover within 2-3 days. In total, DHF clinical course is 7-10 days (19).

Case definitions

There are two types of case definitions, clinical and surveillance. The clinical case definition is the definition that physicians use for patient care, while the surveillance definition is used by the Bureau of Epidemiology to collect and verify the available data on the disease. For Dengue, there are two types of surveillance case definitions; suspected and confirmed cases. The MOPH decision maker will decide about the type of management and of interventions based on the different surveillance case definitions.
**Figure 2 Aedes aegypti life cycle.**

(https://www.cdc.gov/dengue/entomologyecology/m_lifecycle.html)

**History of Dengue**

The first Dengue outbreak was discovered in 1779 in Asia, Africa and North America. After World War II, a Dengue Epidemic was reported in the Pacific, America and South-East Asia regions. *Aedes aegypti* was discovered as a transmission vector in 1903 and a virus was reported to be a Dengue causative agent in 1907. The Dengue virus was first isolated in 1943 by Ren Kimura and Susumu Hotta (20).

The first epidemic of the 20th century was reported from the Philippines in 1954. Since then Dengue has been reported in tropical countries in Africa, America, East-Mediterranean, South-East Asia and the Pacific region.
The first Dengue outbreak in Thailand was discovered in 1949 and the first epidemic was in 1958 in the Bangkok region. It has continuously increasing since then. Its pattern usually included an outbreak every 1-2 years, however, during the last 15 years, its pattern has becoming more unpredictable.

**Table 1 Thailand Clinical Case Definitions**

<table>
<thead>
<tr>
<th>Category</th>
<th>Clinical</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dengue Fever (DF)</td>
<td>✓ Fever with at least 2 of the following symptoms or signs</td>
</tr>
<tr>
<td></td>
<td>Headache</td>
</tr>
<tr>
<td></td>
<td>Muscle pain</td>
</tr>
<tr>
<td></td>
<td>Joint pain</td>
</tr>
<tr>
<td></td>
<td>Orbital pain</td>
</tr>
<tr>
<td></td>
<td>Rash</td>
</tr>
<tr>
<td></td>
<td>Tourniquet Test positive</td>
</tr>
<tr>
<td></td>
<td>Clinical bleeding (hypermenorrhea, epistaxis, petechiae)</td>
</tr>
<tr>
<td></td>
<td>✓ Leukopenia (WBC &lt; 5,000)</td>
</tr>
<tr>
<td></td>
<td>✓ Lymphocyte predominated</td>
</tr>
</tbody>
</table>

**Figure 3 Dengue, countries or areas at risk**

*(World Health Organization, 2004)*
Dengue Hemorrhagic Fever (DHF) ✓ Fever with compatible clinical or tourniquet test positive with any of the following signs
Hemoconcentration
Thrombocytopenia
Chest-x-ray: pleural effusion

Dengue ✓ Dengue Hemorrhagic Fever with hypotension or pulse pressure < 20 mmHg

Shock Syndrome (DSS) mmHg

Table 2. Thailand Surveillance Case Definition (21)

<table>
<thead>
<tr>
<th>Category</th>
<th>Case Definition Type</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dengue Fever (DF) and Dengue</td>
<td>Suspected case</td>
<td>Compatible with Clinical Case</td>
</tr>
<tr>
<td>Hemorrhagic Fever (DHF)</td>
<td>Confirmed case</td>
<td>Compatible with Clinical Case</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Definition with any of the following</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Four-fold rising of Dengue Antibody from Paired serum</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Dengue Antibody &gt; 1:1,280 or Dengue IgM &gt; 1:40 from single serum</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Dengue PCR positive</td>
</tr>
<tr>
<td>Dengue Shock Syndrome (DSS)</td>
<td>Confirmed case</td>
<td>Compatible with Clinical Case</td>
</tr>
</tbody>
</table>

Mosquito Control

Mosquito control plays an important role in controlling Dengue transmission. *Aedes aegypti* can lay its eggs in many places such as man-made water storage, plants, discarded food, beverage containers, tires and decorative plants. The
aim is to eliminate larva and adult mosquitoes as much as possible. The Bureau of Epidemiology of Thailand has adapted the WHO Dengue control manual (14) to be compatible with the Thailand situation. There are three main aspects of disease control; patient, contact and environment (21).

Patient

No quarantine is needed. Patients who still have fever would require bed nets and mosquito repellant to prevent Dengue transmission through mosquito bites.

Contact

Community Active case finding is needed in order to find other Dengue patients who did not go to the hospital, especially those whom have clinical symptoms of Dengue during 14 days before the patient clinical onset. Also, it is important to follow their clinical progress for 28 days after the last patient presented any Dengue clinical symptom. Public education is needed regarding self-protection from mosquito. Also, it is important to inform about signs and symptoms that require hospital care. In addition, it is important to provide information about the activities to reduce the mosquito larva (such as emptying outdoor water containers).

Environment

One of the most important preventive action is to eliminate the eggs and water containers serving as breeding sites as frequently as possible and to cover all water containers. Also, to dispose of outdoor trash containers that could become a breeding site. Another intervention is the use of insecticide fogging and using Temephos, an organophosphate larvicide to kill the mosquito larva. With appropriate use, Temephos can be used even in drinking water (22). Further environmental follow-up will be conducted by local health officers or village health volunteers. They will record the percentage of containers with larvae and the percentage of households with larva containers to evaluate the impact of the Dengue control measures.

To determine which are the most appropriate disease control interventions or the need of combination of interventions, the health officers would make a decision regarding important factors such as local environment, resources, and the cultural context that will improve the intervention feasibility and effectiveness as a result.
Surveillance system

Surveillance is “an ongoing, systemic data collection including data analysis and timely dissemination relevant to disease control and prevention”. In short, surveillance is information for action. A surveillance system is a key component for detection of sudden changes or disease occurrence and it provides information to decision makers to take action and develop specific health policies. Surveillance data can also be a source for health research. There are three main categories of surveillance.

**Passive surveillance** is similar to routine data reporting system. It gathers disease data from the reports from health care providers. Passive surveillance is the most common type surveillance. In this type, the main responsibility will fall upon the health care provider to report data using standardized forms to local or central health departments. The common problem of passive surveillance is underreporting and the late reports.

**Active surveillance** opposed to passive surveillance, health departments are responsible to collect directly the data. Even though active surveillance provides better data quality and faster than passive surveillance, it’s much more expensive in terms of labor and cost. Therefore, active surveillance usually is implemented in deadly situations that require fast and accurate data.

**Sentinel surveillance** is a mixture of passive and active surveillance by selecting specific areas or specific health providers to gather a high-quality data that could not be obtained by passive system while maintaining low cost. However, as sentinel surveillance take place in specific areas, some rare diseases or diseases that are not under surveillance will not be detected by them would be missed.

Thailand reporting system

Thailand has a National Reporting System which was established mainly for reimbursement in its Universal Health-care Coverage. However, to make Disease Control timelier, Thailand MOPH decided to establish a separated passive surveillance called “R506” system. This system requires every public hospital to report every disease included in the priority list to the Bureau of Epidemiology, MOPH. Its reporting process is similar to National Electronic Disease Surveillance System (NEDSS) of US-CDC. This system was implemented in all government hospitals and it is a voluntary report from private hospitals. In special situations, R506 system can switch to active surveillance therefore
offering the MOPH a very flexible data source. Currently there are 52 conditions that need to be reported to the Bureau of Epidemiology.

When physicians make a diagnosis that is in the priority diseases list, the data will go to the Hospital’s Family or Community Medicine. They will verify the information and they may interview the patients to obtain more information using a predefined semi-structured questionnaire specific for each disease surveillance before being submitted to the District Health Office, Regional Disease Control Office and Ministry of Public Health Department of Disease Control. Each office will clean and analyze the data for their own use such as public health planning and management.

If more information is needed, depending on the situation, they will individually or jointly investigate and provide the disease control together. The report will be submitted to MOPH by email and it will be entered to the database manually.

*Figure 4 R506 Reporting System*

(Source: Introduction to Public Health Surveillance, Darin Areechokchai, Ministry of Public Health Thailand)
Thailand Electronic Health Record (EHR) and HIS

Up until now, there is no official survey or research project on the extent of the implementation of the EHR in Thailand. The only information available so far is the Thai Medical Informatics Association annual meeting report (23) and the results of one Thai student’s PhD Dissertation (24). In overall, most hospitals in Thailand use best of breed software, the market share was dominated by HOSxP higher than 50%, only 15.87% still internally is developing their EHR.

The Thailand Health Office does have its own HIS application named “R506 application”, while it doesn’t have any EHR features, the purpose of this application is the same as those in the District Health Information Software (DHIS2). Normally, each health officer will import data from EHR to manage cleaning and analyzing before submitting the data to R506 system in the CSV format. This application was developed by the Department of Disease Control based on Microsoft Access 2003. R506 was designed to reduce the burden on the health officers by offering in one click clean and converted data, generating the disease report and common data analysis such as a GIS spot map, and an outbreak detection and descriptive analysis. The R506 application is available for free at Department of Disease Control website (available only in Thai language).
As of 2016, Thailand still doesn’t have any national syntactic and security standard, or the Health Information Exchange (HIE) infrastructure. The only available standard is the ICD-10 TM, a semantic standard. Because of this, most electronic data transmission including R506 system is still mainly relying on sending Coma Separated Values (CSV) file.

The new HIE system is under development by the Healthcare Information System Standards and Processing Administration (HISPA), MOPH. Syntactic standards are being developed by the Thai Health Information Standard Development Center (THIS) using SNOMED CT and LOINC. Thailand is also planning to implement country Health Level 7 Clinical Document Architecture (CDA) in the future. However, there is still no consensus plan on developing security standards at the moment. Even though there is still a long way to go for the Thailand HIE system, it is moving forward.
Surveillance Algorithms

After an Influenza Epidemic in 2009, WHO realized that many countries were still lacking of well-established surveillance systems and there were problems in recognizing, monitoring, evaluating and predicting disease outbreak situations. Therefore, WHO proposed a surveillance manual in order to assist in standardizing the collection and interpretation of the data. This standardization will allow better interpretation of seasonal and future outbreaks (8). There are several surveillance algorithms developed for outbreak detection by calculating the needed thresholds with various levels of complexity. These models range from inspection, regression, to time series analysis. WHO described the advantages and disadvantages of these three methods.

Several Bayesian statistics based surveillance algorithms such as Naïve Bayes Classifier, Efficient Bayesian Multivariate Classification (EBMC) and Bayesian Network with K2 were proposed recently, however they required Natural Language Process for data extraction which is currently not available for local language (Thai) (25) (26). Moreover, MOPH reporting system did not require to submit chief complaint and presentation therefore Bayesian based surveillance algorithms were not feasible for Thailand and were not included in this study.

<table>
<thead>
<tr>
<th>Method description</th>
<th>Examples</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Visual</strong></td>
<td>Graphically based Model based (Time series, regression)</td>
<td>Very simple to implement and understand</td>
<td>Overly simplified, will not capture any trend changes over time.</td>
</tr>
<tr>
<td>Based on a visual analysis of past data, define baseline, off seasonal baseline, threshold and seasonal threshold values</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Averaging</strong></td>
<td>Moving Epidemic Methods (MEM)</td>
<td>Simple to implement</td>
<td>Can allow a past season’s or week’s aberrant values to influenza.</td>
</tr>
</tbody>
</table>
Usually involves calculating a median or mean of data

**Process control**

Based on similar processes to those used in detecting anomalies in industrial production. Most methods rely on some method of setting an upper control limit. Some methods also involve looking at the rate of change in the data series.

<table>
<thead>
<tr>
<th>Method</th>
<th>Shewhart charts</th>
<th>Best for detection of start of season and unusual patterns.</th>
<th>Not as accurate as time series methods.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cumulative sum (CUSUM) charts</td>
<td></td>
<td>Works well in situations where rates are low.</td>
<td>May be sensitive to small changes in reporting efficiency</td>
</tr>
<tr>
<td>Exponentially weighted moving average charts</td>
<td></td>
<td>Good at detecting the start of the season when the start is slow</td>
<td></td>
</tr>
</tbody>
</table>

**Study Design**

We conducted a descriptive ecological study based on secondary data from National Surveillance System (R506), Bureau of Epidemiology, Ministry of Public Health of Thailand. The ecological unit is the Dengue outcome by (Country-month and Country-week).

**Study Population**

The R506 captures official health data from every government hospitals in Thailand that matches specific ICD-10 compatible for the Disease Prevention and Epidemiologic study, similar to National Electronic Disease Surveillance System (NEDSS) of US-CDC. These data were submitted to provincial health offices on weekly basis for validation and cleaning by local public health staff before submitted to the Bureau of Epidemiology, Ministry of Public Health.
population is the Thai patients using public hospitals with a diagnosis of any dengue condition. The study did not include private hospital patient data since the private sector is not required to provide that information.

**Data sources**

An entire Dengue patient dataset (2003-2015) with ICD-10 codes A90, A91, and A99 were obtained from the Bureau of Epidemiology, Ministry of Public Health, Thailand website. This dataset is from an open-data source available for public use without restrictions or limitations. We obtained approximately 1 million de-identified individual Dengue Fever records. Derived relevant variables included the individual visit record with local diagnosis code, gender, age, nationality, occupational, location, hospital class, patient type, outcome, time of diagnosis, time of visit and time of report.

**Detection methodologies**

At the time of study, there is no official Dengue Surveillance Guideline available. We used the WHO Interim Global Epidemiological Surveillance Standards for Influenza for our approach in this study recognizing the potential differences with Dengue surveillance. The WHO’s manual specifies three categories for determining outbreak thresholds: Visual inspection, Averaging, and Process control. Since Thailand does not have officially Dengue outbreak criteria for thresholds, we included only algorithms that do not need pre-determined thresholds. Therefore, other more advanced statistical algorithms such as regression models, cumulative sum (CUSUM), and exponentially weighted moving average were excluded from this study (27).

**Visual Inspection**

This is the simplest method of outbreak identification by inspecting of several historical data and determining the threshold. This method is widely used in Influenza Surveillance to determine thresholds and may be used to compare to more complex statistical techniques (28)(29)(30). Visual inspection was currently used in our study to determine whether the case number goes over Moving-5-years-median threshold in Thailand as well. Therefore, we implemented the visual inspection as one of the detection methods in this study.

**Averaging**

Since WHO Surveillance Manual did not specify exactly on how to calculate averaging seasonal thresholds, used the suggestion from Ee Laine et al. (29) for our threshold calculation. These authors used 90 and 95 percentiles and averages with upper 95 and 90 Confidence Intervals. These thresholds were calculated based on 5 years’ weekly historical
data as is the minimal data available for local health office is 5 years. The Thailand population increased only 4% during 2003-2015, we used number of visits instead of incidence per 100,000 population in this study (31).

We also included 5-years-medians as they are widely used in Thailand Public Health for several years. This approach is widely recognized by Department of Disease Control and it was implemented as a default detection threshold for those in the local public health Information System. However, until now there is no study assessing its sensitivity, specificity, timeliness and its comparison to other early detection algorithms.

In this study, we categorized averaging thresholds calculations into three types; moving averages, same period moving and off-seasonal thresholds.

Moving thresholds

Moving thresholds were calculated from previous 5 years weekly data directly. For example, we calculated the average threshold of May 2015 by finding the median value from May 2010- April 2014 consecutively.

Same period moving thresholds

A same period moving thresholds were calculated from previous 5 years of the same week. For example, we calculated the 5-year-median threshold of May 2015 by finding the median value from May 2010, May 2011, May 2012, May 2013 and May 2014. MOPH use this method to calculate the Dengue outbreak threshold.

Off-seasonal threshold

The thresholds were calculated from previous 5 years off-seasonal Dengue (September to April) historical data. For example, we calculated off-seasonal average 2015 threshold by averaging September to April 2010-2014 data.

Process control

Early Aberration Reporting System Algorithms

The Early Aberration Reporting System (EARS) algorithms, developed by Centers for Disease Control and Prevention (US-CDC) were designed to signal early warning of upcoming outbreaks and allowing more time for Ministry of Public Health for preparation to achieve better Dengue control and prevention. EARS are currently used nationally and international in several large public events such as the US Democratic National Convention (Boston, MA), Republican National Convention (Ney York, NY), G8 Submit (Sea Island, GA), and the 2004 Summer Olympics (Athens Greece) (32). EARS were successfully producing early alerts with sensitivity of more than 80% and specificity more than 90% in many important infectious diseases especially Influenza (33)(34).
Until now there are very few studies evaluating the implementation of these important algorithms on Dengue early outbreak detection (35)(36).

The EARS objective is to provide an early detection and warning signals for syndromic and bio-surveillance information. When EARS signals occur, it means that there is an early signal for the next outbreak. This will support the need for further investigation from the public health officers.

The EARS consist of three methods, “C1”, “C2”, and “C3”. C1 and C2 implement moving averages and standard deviations based on the previous 7 days. The only difference of the C2 is that this method has 2 days lag. In other words, while signal of day 1 will contribute to day 2 baseline on C1, signal on day 1 and 2 won’t contribute to day 3 baseline on C2. (37)(38) C1 and C2 will alert when it goes beyond 3 standard deviations. C3 is calculated by using previous two days in C2, and will alert when C3 is more than 2 standard deviations. While we categorized Averaging thresholds calculations into 3 categorize, EARS may use one single of these calculation methods.

\[
C_1(t) = \frac{Y(t) - \bar{Y}_1(t)}{S_1(t)}
\]

\[
C_2(t) = \frac{Y(t) - \bar{Y}_1(t)}{S_3(t)}
\]

\[
C_3(t) = \sum_{i=t}^{t-2} \max[0, C_2(i) - 1]
\]

Where S and \(\bar{Y}_1(t)\) are moving average and standard deviation, respectively.

\[
\bar{Y}_1(t) = \frac{1}{7} \sum_{i=t-1}^{t-7} Y(i) \quad S_1^2 = \frac{1}{6} \sum_{i=t-1}^{t-7} [Y(i) - \bar{Y}_1(i)]^2
\]

\[
\bar{Y}_3(t) = \frac{1}{7} \sum_{i=t-3}^{t-9} Y(i) \quad S_3^2 = \frac{1}{6} \sum_{i=t-3}^{t-9} [Y(i) - \bar{Y}_1(i)]^2
\]

Data Analysis

Descriptive statistics was conducted using means, medians, maximums, minimums, percentages and ranges. Visualization was used in this study to compare among detection surveillance algorithms and thresholds. All data analyses were conducted in R 3.3.2 software.
Results

During 2003-2015, there were 1,014,201 visits that were a combination of Dengue Fever (DF), Dengue Hemorrhagic Fever (DHF) and Dengue Shock Syndrome (DSS) and including 1,122 deaths. There were 120 visits per 100,000 person-year. DF contributed the highest of the visits (52%) followed by DHF (46%) and DSS (2%). Most of them were in the Thai population (87.2%) followed by the Burmese (12.8%). The median age was 15 (IQR 10-24) and a male to female ratio was 1:1. The age group with the highest incidence were 5-12-year-old (28.6 per 100,000 person-year), 13-18-year-old (25.3 per 100,000 person-year) and 26-45 years old (18.8 per 100,00 person-year) respectively. Approximately half of the visits were students (48.3%), elementary service workers (16.4%) and farmers (6.1%).

Thailand has been affected by Dengue outbreaks every year, 13 outbreaks in total in the study period. Most of the outbreaks show a seasonal pattern which consistently occurred and subsided at the same period (May to August) during 2003-2015 (Figure 6). The strongest outbreak was in 2013 and the weakest outbreak was in 2014.

Process Control threshold

EARS were the only algorithms in this study that were able to detect every outbreak during 2008-2015 including the 2014 outbreak which was undetected by other detection methods in this study.

This method does not require pre-determined outbreak thresholds as shown in Table 5. Moreover, EARS alert signals seem to be the earliest compared to other detection methods. Overall, EARS often provide the first early signal in C3, and then followed by C2 and C1. When looking from the first signal to outbreak peak the duration of all three signals are not much different. When approaching the outbreak peak, we were more likely to see C1, C2 and C3 altogether, while C3 usually is observed earlier at the outbreak as indicated before. However, C1, C2, and C3 disappear after the outbreak peak. In Figure 6 we observed that C3 sent more signals than C2 and C2 sent more signals than C1. Most of the time C3 overlaps C2 and C2 overlaps C1.

EARS algorithms were able to detect the upcoming second peak as well, by signal alerts during outbreaks as shown in 2012, 2013 and 2015 outbreaks. Normally, EARS stop sending signals after the outbreak peaks and start declining. However, among the 2012-2013 and 2015 outbreaks EARS unusually sent signals after the peak. Those alerts were followed by a second outbreak peak. This is a prominent feature as none of other detection methods were able to
provide this information. However, EARS didn’t provide any information about the outbreak severity. We observed no unique outbreak change pattern during the three largest outbreaks in our data (2010, 2013 and 2015).

EARS successfully detected every single outbreak during 2003-2007. These outbreaks were not identified by other methods, since they required 5 years of historical data for the calculation of the thresholds. EARS proved to retain its warning performance even without historical data. When considering Sensitivity and False-Positive as shown in Figure 10. EARS outperformed every other algorithm in this study and EARS C1 is very close to our baseline (same period moving median).

**Moving threshold**

In this category, we clearly saw that average and median measurements showed similar results as did the percentile and upper confidence interval (CI). (Figure 7.)

Averages and median measurements successfully detected every outbreak except for 2014. Unfortunately their signals appeared much later than those of the EARS algorithms.

Percentile and upper confidence interval thresholds were very high and thus failed to detect most outbreaks and they gave the slowest signals in this study. However, those outbreaks signals were the four largest in our data (2008, 2010, 2013, and 2015).

While having many drawbacks, an important finding of this category is that percentile and CI clearly illustrated the Dengue trends better than other methods.

**Same period Moving threshold**

Unlike the moving threshold method, this threshold method represents the seasonal pattern from historical data (Figure 8.). Again, in this category, we observed resembled appearance between the average and median, and the percentile and upper Confidence Interval.

Percentile and Upper confidence intervals were able to detect the four largest outbreaks as well. However, they detected earlier than those of Moving threshold percentile and Confidence intervals. Both thresholds became lower and close to the mean and median during off-seasonal period.

However, there are some difficulties in interpreting the signals as the visit numbers are very close to the threshold line. Considering the fact that the thresholds need to signal the outbreak as early as possible in order to be useful, deciding to consider the alert as detectable might be controversial. Considering that the visit numbers nearly exceed the threshold
they will alert the MOPH to start, preparing the control measures. We considered 2009, 2011, 2012 and 2015 outbreaks as successfully detected. None of the methods in this category were able to send the signals earlier than EARS.

**Off-seasonal threshold**

Overall, the appearances of all methods in this category were the same. Their appearances did not show any seasonal pattern with their calculation threshold methods. The only difference with these methods is their magnitude. Percentiles and Confidence Intervals did not have much difference from the median and the mean, as Thailand non-seasonal number of Dengue visits is very low and this low number has not been changed during last 10 years, except for the 2012-2013 season. Because of that, every method in this category successfully detected more outbreaks in comparison to last two methods. Unlike the other two methods, the percentiles and upper CI were able to capture every outbreak except the 2014 that was one of the largest four outbreaks.

The means and medians were useful to detect every outbreak especially the 2014 outbreak which was not detected by previous two categories. This method allows showing the Dengue long-term trends. (Figure 9.) However, the Off-seasonal threshold signals were still later than those of EARS.

---

**Table 3** Characteristics of DF/DHF/DSS visit and incidence rate per 100,000 person-year, Ministry of Public Health Thailand, 2003-2015

<table>
<thead>
<tr>
<th>Gender</th>
<th>Visit (%)</th>
<th>Incidence per 100,000 person-year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>516,921 (51.6)</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>483,993 (48.4)</td>
<td></td>
</tr>
<tr>
<td>Occupation</td>
<td>Number</td>
<td>Percentage</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>------------</td>
<td>------------</td>
</tr>
<tr>
<td>Students</td>
<td>489,746</td>
<td>48.9%</td>
</tr>
<tr>
<td>Elementary Service Workers</td>
<td>164,430</td>
<td>16.4%</td>
</tr>
<tr>
<td>Farmer</td>
<td>61,994</td>
<td>6.2%</td>
</tr>
<tr>
<td>Unemployed</td>
<td>25,081</td>
<td>2.5%</td>
</tr>
<tr>
<td>Merchant</td>
<td>16,386</td>
<td>1.6%</td>
</tr>
<tr>
<td>Others</td>
<td>243,227</td>
<td>24.4%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Age Group</th>
<th>Number</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-4</td>
<td>50,196</td>
<td>5.0%</td>
</tr>
<tr>
<td>5-12</td>
<td>283,197</td>
<td>28.3%</td>
</tr>
<tr>
<td>13-18</td>
<td>266,846</td>
<td>26.7%</td>
</tr>
<tr>
<td>19-25</td>
<td>153,129</td>
<td>15.3%</td>
</tr>
<tr>
<td>26-45</td>
<td>182,211</td>
<td>18.2%</td>
</tr>
<tr>
<td>46-60</td>
<td>47,558</td>
<td>4.7%</td>
</tr>
<tr>
<td>&gt;60</td>
<td>17,569</td>
<td>1.8%</td>
</tr>
</tbody>
</table>
Figure 6  Comparison of the EARS Algorithms and Moving-Median-5-years

Solid black bars represent alert signals from C1, C2, and C3. As Median-5-years required 5 years historical data, the threshold was available only after 2008, while EARS does not need any historical data and immediate available since 2003. EARS signals appear before every outbreak and disappear after the visit number start to rise.
<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Total Signal (week)</th>
<th>Total Outbreaks (week)</th>
<th>Positive Predictive Value (PPV)</th>
<th>Negative Predictive Value (NPV)</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>EARS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C1</td>
<td>23</td>
<td>291</td>
<td>86.96%</td>
<td>31.04%</td>
<td>100.00%*</td>
<td>97.60%</td>
</tr>
<tr>
<td>C2</td>
<td>80</td>
<td>291</td>
<td>78.75%</td>
<td>32.14%</td>
<td>100.00%*</td>
<td>86.40%</td>
</tr>
<tr>
<td>C3</td>
<td>122</td>
<td>291</td>
<td>78.69%</td>
<td>33.67%</td>
<td>100.00%*</td>
<td>79.20%</td>
</tr>
<tr>
<td>Same period moving threshold</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median (Baseline)</td>
<td>291</td>
<td>291</td>
<td>100.00%</td>
<td>100.00%</td>
<td>100.00%</td>
<td>100.00%</td>
</tr>
<tr>
<td>Average</td>
<td>237</td>
<td>291</td>
<td>98.73%</td>
<td>68.16%</td>
<td>80.41%</td>
<td>97.60%</td>
</tr>
<tr>
<td>Upper 95% CI</td>
<td>151</td>
<td>291</td>
<td>100.00%</td>
<td>47.17%</td>
<td>51.89%</td>
<td>100.00%</td>
</tr>
<tr>
<td>Upper 90% CI</td>
<td>162</td>
<td>291</td>
<td>100.00%</td>
<td>49.21%</td>
<td>55.67%</td>
<td>100.00%</td>
</tr>
<tr>
<td>95 Percentile</td>
<td>128</td>
<td>291</td>
<td>100.00%</td>
<td>43.40%</td>
<td>43.99%</td>
<td>100.00%</td>
</tr>
<tr>
<td>90 Percentile</td>
<td>114</td>
<td>291</td>
<td>100.00%</td>
<td>41.39%</td>
<td>39.18%</td>
<td>100.00%</td>
</tr>
<tr>
<td>Moving threshold</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>244</td>
<td>291</td>
<td>85.25%</td>
<td>51.74%</td>
<td>71.48%</td>
<td>71.20%</td>
</tr>
<tr>
<td>Average</td>
<td>195</td>
<td>291</td>
<td>89.74%</td>
<td>47.51%</td>
<td>60.14%</td>
<td>84.00%</td>
</tr>
<tr>
<td>Upper 95% CI</td>
<td>47</td>
<td>291</td>
<td>100.00%</td>
<td>33.88%</td>
<td>16.15%</td>
<td>100.00%</td>
</tr>
<tr>
<td>Upper 90% CI</td>
<td>80</td>
<td>291</td>
<td>100.00%</td>
<td>37.20%</td>
<td>27.49%</td>
<td>100.00%</td>
</tr>
<tr>
<td>95 Percentile</td>
<td>63</td>
<td>291</td>
<td>100.00%</td>
<td>35.41%</td>
<td>21.65%</td>
<td>100.00%</td>
</tr>
<tr>
<td>Percentile</td>
<td>Value</td>
<td>N</td>
<td>Median</td>
<td>Average</td>
<td>Upper 95% CI</td>
<td>Upper 90% CI</td>
</tr>
<tr>
<td>-----------------</td>
<td>-------</td>
<td>----</td>
<td>--------</td>
<td>---------</td>
<td>--------------</td>
<td>--------------</td>
</tr>
<tr>
<td>90</td>
<td>53</td>
<td>291</td>
<td>79.93%</td>
<td>86.10%</td>
<td>99.09%</td>
<td>89.57%</td>
</tr>
<tr>
<td>Off-seasonal</td>
<td></td>
<td></td>
<td>79.93%</td>
<td>86.10%</td>
<td>99.09%</td>
<td>89.57%</td>
</tr>
<tr>
<td>threshold</td>
<td></td>
<td></td>
<td>79.93%</td>
<td>86.10%</td>
<td>99.09%</td>
<td>89.57%</td>
</tr>
<tr>
<td>Median</td>
<td>299</td>
<td>291</td>
<td>79.93%</td>
<td>86.10%</td>
<td>99.09%</td>
<td>89.57%</td>
</tr>
<tr>
<td>Average</td>
<td>259</td>
<td>291</td>
<td>86.10%</td>
<td>56.69%</td>
<td>40.52%</td>
<td>42.69%</td>
</tr>
<tr>
<td>Upper 95% CI</td>
<td>110</td>
<td>291</td>
<td>99.09%</td>
<td>40.52%</td>
<td>37.46%</td>
<td>50.17%</td>
</tr>
<tr>
<td>Upper 90% CI</td>
<td>163</td>
<td>291</td>
<td>89.57%</td>
<td>42.69%</td>
<td>50.17%</td>
<td>86.40%</td>
</tr>
<tr>
<td>95 Percentile</td>
<td>117</td>
<td>291</td>
<td>97.44%</td>
<td>40.80%</td>
<td>39.18%</td>
<td>97.60%</td>
</tr>
<tr>
<td>90 Percentile</td>
<td>105</td>
<td>291</td>
<td>98.10%</td>
<td>39.55%</td>
<td>35.40%</td>
<td></td>
</tr>
</tbody>
</table>

*As the purpose of the EARS algorithms is to provide an early warning not an outbreak threshold, their sensitivity was calculated based on whether their signals appear before the outbreak.
Figure 7  

Moving threshold methods (a) Median 5 years  

(a) Median 5 years, (b) 95 and 90 Percentile, (c) Average, Upper 90 and 95% Confidence Interval and EARS algorithm (C1, C2, and C3), 2008-2015. This threshold category clearly shows an increasing Dengue visit in long-term. While, the four largest outbreak were detected, many outbreaks were not.
Figure 8 Same period moving threshold methods

(a) Median 5 years, (b) 95 and 90 Percentile, (c) Average, Upper 90 and 95% Confidence Interval and EARS algorithm (C1, C2, and C3), 2008-2015. This category clearly describes data seasonal pattern. Percentile and Upper CI, Median and average shows similar threshold. A small outbreak in 2014 were not detected by any threshold in this category.
Figure 9 Off-seasonal threshold methods

(a) Median 5 years, (b) 95 and 90 Percentile, (c) Average, Upper 90 and 95% Confidence Interval and EARS algorithm (C1, C2, and C3), 2008-2015. This category appearance is similar to Figure 7, with more sensitivity as it almost detects 2014 peak. However, their sensitivity is still lower than EARS.
EARS C1, C2 and C2 outperform other algorithms. C1 is very close to baseline (Same period moving median)

Discussion

Thailand has been affected by Dengue outbreaks every year, thus detection methodologies that are able to provide information on severity and the upcoming outbreak is necessary. While there are only few published references on Dengue surveillance detection methodologies, numerous literature and an official guideline from WHO are available for influenza surveillance. This is the first national study to evaluate various Dengue outbreak detection methods in Thailand.

As shown in the results section, Dengue outbreaks in Thailand clearly show a seasonal pattern. Generally, the Dengue outbreaks starting and ending periods are very consistence. Even though MOPH can anticipate the rising of Dengue during the rainy season, having an early warning is a very important public health advantage. At the beginning of the outbreak, the number of Dengue visits will increase rapidly that passive Dengue surveillance system could not produce signals early enough to allow preparation for the control. Therefore, having an early warning will provides more time for MOPH to prepare for the upcoming outbreaks.
Overall, we obtained similar information with average and median, percentile and upper CI across three methods. This finding is consistence other South East Asia region study (39). Taking together, each method provides different interpretation.

Same period moving method clearly illustrated the Dengue seasonal pattern and it has been used by the MOPH for the Dengue threshold calculation. When the number of visits goes beyond the threshold, MOPH will re-prioritize its action and start to allocate more resources for Dengue outbreak control. However, during non-Dengue season, the threshold become very low and not appropriate, especially percentile and upper CI threshold. The reason is the non-seasonal period threshold become very low. A small peak during this period could goes beyond the threshold and inappropriately trigger epidemic alert, as happened during 2012-2013 outbreaks.

Off-seasonal methods seem to be the most sensitive. It also illustrates Dengue trend information. While median and average were able to provide earlier outbreak warning, adding percentile and confidence interval threshold doesn’t add much information. The possible explanation is they are not much different from average and median unlike in the other methods. In this method, average or median provide the best sensitivity among every methodology in this study, second to only EARS. Its potential application is to implement off-seasonal average or median as a threshold to determine the beginning and the end of seasonal Dengue outbreak.

Moving threshold methods illustrate the Dengue trend the best. Its percentile and upper CI were not affected by off-seasonal periods and were able to detect all four largest outbreaks in our data. This is better than the same period moving threshold as it won’t provide a false epidemic alert during off-seasonal period. Thus, this threshold proves to be better for the epidemic threshold. However, mean and median are not as sensitive as off-seasonal methods’ therefore, not suitable for seasonal or alert threshold.

EARS algorithms outperform other methods in this study by being able to detect even the smallest outbreak year (2014) which could not be detected by other thresholds. In terms of timeliness, EARS provide the earliest outbreak signals in this study without any historical data needed. However, EARS do not provide outbreak severity information nor illustrate long term trends as does for other methods.

The primary drawback of EARS algorithms is it provides too many alerts before an outbreak actually occurs. Their signal disappears when outbreak reach its peak and declining. Even of that, it was able to detect a second outbreak
peak during season as well (2012-2013 and 2015). Therefore, EARS implementation might prove to be of great benefit for Thailand as a signal for upcoming Dengue outbreak peak and encourage MOPH to start its control preparation. Our finding confirms recent studies in China that EARS was able to provide an early signal of upcoming outbreak and its upcoming peak (40).

Even though EARS provide invaluable outbreak information, it can cause confusion among decision makers whether they should to take action since there are too many alert signals. Many of the EARS algorithms on Influenza and Influenza Like Illness (ILI) also faced this problem as well (41). This can result in putting more burden to local health worker as the MOPH will depends on them to validate the false outbreak signals. To filter EARS signals and reduce the confusion, some study recommends combining EARS signals together (42). However, in this study we suggest to the MOPH to rely mainly on C1. As we did not observe any information gain in adding C2, C3 or any combination of these methods. Another possible approach is to combining EARS C1 with other thresholds reviewed in this study to reduce false signal as much as possible.

During non-Dengue outbreak season, MOPH does not have many resources left to maintain the Dengue surveillance system as much as they have during the outbreaks. The reason is that the MOPH has limited resources and they need to relocate resources to address other public health problems when Dengue outbreak subsides. However, MOPH could not completely ignore the Dengue outbreak detection. The reason is the Dengue outbreaks during non-Dengue season is uncommon and usually follow by a severe outbreak as occurred in 2012-2013 outbreaks. A possible approach is for the health authorities to decide on what level of sensitivity they can tolerate based on current resources available and the potential outbreak situation. For example, they might decide to tolerate more false signals during non-Dengue season.

Overall, the methods were able to produce outbreak signals from our data, thus implementation is feasible. We found that each threshold and its algorithms have its own strengths and weaknesses, in order to get the appropriate thresholds there are several options to be explored. For example, the MOPH can implement multiple thresholds using EARS for early warning, median-5-years for alert thresholds, percentile and upper CI for epidemic thresholds for the off-seasonal periods. For the off-seasonal periods we may use the mean or median to determine the thresholds. Further study will be needed in order to evaluate the usefulness and validity of these possibilities.
Since every health office is using a HIS developed by MOPH, implementation of these thresholds and algorithms would be feasible by adding them in the next release version. Some training might be needed in order to clarify threshold interpretation and make a decision. A pilot implementation could be conducted in central MOPH and regional health office to gather feedback before national implementation.

We did not evaluate several modern surveillance algorithms based on Bayesian statistics such as Naïve Bayes Classifier, Efficient Bayesian Multivariate Classification (EBMC) and Bayesian Network with K2 in this study. The reason is that they required Natural Language Process for data extraction which is currently not available for local language (Thai) (25) (26). Moreover, MOPH reporting system did not collect chief complaint and clinical presentation, therefore Bayesian based surveillance algorithms were not feasible for Thailand.

This study was conducted using weekly dengue information from Thailand, a tropical country in Southeast Asia. Applying the study results to other countries might need to take consideration of their difference in the data reporting system, the Dengue outbreak characteristics and the available public health infrastructure. In spite of the fact that Dengue is still an important public health threat for many tropical countries, there are only few studies available on Dengue early warning and detection methods. We would like to encourage other public health authorities and researchers from tropical countries to review and evaluate these innovative early detection methods to continuously improve the public health surveillance and control.

**Limitations**

There were some limitations in this study. Our data sources were collected from government hospitals only. Private hospitals and clinics are not required in Thailand to submit health data to MOPH. However, this situation doesn’t have much impact on the study results; as their capacity is much less than the government hospitals, thus contributing to very small proportion of the cases.

Instead of using the number of illness episodes, we used the number of visits as numerator to calculating the incidence rate because of lacking computation resources for linking population information and disease data. Merging disease surveillance and population identity database required a very high-performance computer, as both databases are very large and MOPH does not have budget to support that. However, those data were validating and de-duplicated from
local and regional health offices. Thus, using number of visits should be acceptable for the estimation of the incidence rate in this study.

The study implements visual inspection as one of the standards which may introduce some bias in determining outbreaks. However, this method has been widely used in many surveillance detection investigations and currently it is one of the current methods of determining outbreaks in Thailand, thus this method may represent the practical judgment of the local health officers.

**Conclusion**

In summary, implementing the originally proposed WHO surveillance methodologies (for Influenza) for Dengue outbreak detection is feasible. However, there is no one-fit-all solution for the early outbreak detection for Dengue. Same period median-5-years might be good for an alert threshold. EARS methods were able to detect every outbreak but they don’t provide information on trend and severity. Moving percentile or upper CI could provide information for long-term trend and the epidemic thresholds. Off-seasonal median or average might suitable for seasonal thresholds. As this study was specific on Thailand climate and reporting system, implementing our recommendation for other country might consider the specific context of their local public health surveillance systems and the epidemiological risk factors of their Dengue outbreak situation. However, there are several early detection methods available and we encourage other tropical countries to explore their Dengue data and epidemiological situation in order to improve the public health surveillance system.

**Acknowledgement**

We thank Bureau of Epidemiology, Department of Disease Control, Ministry of Public Health of Thailand for their contribution and funding support to make this research successful. We also would like to express our appreciation towards Hojoon Daniel Lee from Depart of Epidemiology, Johns Hopkins University Bloomberg School of Public Health for his suggestion and comments.

**References**


27. Steiner SH, Grant K, Coory M, Kelly HA. Detecting the start of an influenza outbreak using exponentially weighted moving average charts. BMC Medical Informatics and Decision Making. 2010 June.


Curriculum Vitae

Supharerk Thawillarp, M.D.

Personal information

Workplace

Bureau of Epidemiology, Department of Disease Control, Ministry of Public Health, Thailand

Current positions

Medical Doctor, Professional Level
Graduate student, Division of Health Sciences Informatics, Johns Hopkins University School of Medicine

Telephone: 14434805518, +66819100146

Email: sthawil1@jhmi.edu, raynus.blueray@gmail.com

Education

August 2015-Present   Master of Science Expected May, 2017( 
Division of Health Sciences Informatics  
School of Medicine, Johns Hopkins University

Thesis: Evaluation of Possible Dengue Outbreak Detection Methodologies for Thailand, which one should be implemented?

June 2013 -May 2015  Field Epidemiology Training Program)FETP(, 
Bureau of Epidemiology, Ministry of Public Health, Thailand

Capstones

- Situation of Heat Related Illness in Thailand
- Situation of Lead Exposure among pre-school Children in Bangkok, Thailand

June 2004-May 2010   Doctor of Medicine  
Faculty of Medicine Prince of Songkla University, Songkhla, Thailand
**Professional Experience**

October 2016-May 2017  
Teaching Assistant  
PH 340.770 Public Health Surveillance  
PH 340.770 Special Studies in Advanced Public Health Surveillance  
PH 315.709 Health Sciences Informatics, Knowledge Engineering and Decision Support  
Johns Hopkins Bloomberg School of Public Health

April 2015-May 2015  
Guest Researcher  
Centers for Disease Control and Prevention (USCDC), Atlanta, Georgia, USA

June 2013-June 2015  
Resident (Board of Preventive Medicine)  
Field Epidemiology Training Program (FETP)  
Bureau of Epidemiology, Department of Disease Control  
Ministry of Public Health Thailand

September 2011-June 2013  
General Practitioner (Intern)  
Pua Crown Hospital  
Nan, Thailand

September 2010-September 2011  
General Practitioner (Intern)  
Nan Provincial Hospital  
Nan, Thailand

**Certificate of Achievement**

May, 2016  
Certified HL7® V3 RIM Specialist  
Health Level Seven International
March, 2016  Certified HL7® V2.7 Control Specialist
            Health Level Seven International

February, 2016 Certified HL7 CDA® Specialist
            Health Level Seven International

February, 2016 Machine Learning Foundations :A Case Study Approach
            University of Washington on Coursera

January, 2016 Machine Learning course
            Stanford University on Coursera

December, 2015 Health Informatics on FHIR
            Georgia Institute of Technology on Coursera

October, 2015 SNOMED CT Foundation course,
            International Health Terminology Standards Development Organization
            IHTSDO

November, 2016 SNOMED CT Implementation course,
            International Health Terminology Standards Development Organization
            IHTSDO

March, 2017 SNOMED CT Content Development Theory course,
            International Health Terminology Standards Development Organization
            IHTSDO

May, 2017 Database Management Essentials
            University of Colorado on Coursera

May, 2017 Relational Database Support for Data Warehouses
            University of Colorado on Coursera
June, 2017  Global Health Informatics
Massachusetts Institute of Technology on Edx

May 2017  Data Warehouse Concept, Design and Data Integration
University of Colorado on Coursera

June, 2016  Introduction to Big Data
University of California, San Diego on Coursera

June, 2016  Data Structures
University of California, San Diego on Coursera

June, 2016  Algorithms on Graphs
University of California, San Diego on Coursera

May, 2016  Algorithmic Toolbox
University of California, San Diego on Coursera

May, 2016  Hadoop Platform and Application Framework
University of California, San Diego on Coursera

December, 2016  Big Data Modeling and Management Systems
University of California, San Diego on Coursera

January, 2017  Big Data Integration and Processing
University of California, San Diego on Coursera

March, 2016  Exploratory Data Analysis
John Hopkins University on Coursera

March, 2016  Getting and Cleaning Data
Skills

Programs: Stata, Epi-info, Epidata, QGIS,

Computer Languages: R, Java, Python

Publications

International Publication


**Thai Publication**


**Other Publications**
1. **Supharerk Thawillarp**, Chumkasean P. Summary situation of Brucellosis in Thailand, Annual Epidemiological Surveillance Report AESR(, 2013;83-84

2. **Supharerk Thawillarp**, Siripanij S. Summary situation of Heavy Metal poisoning in Thailand, Annual Epidemiological Surveillance Report AESR(, 2013;154


**Manual Production**


- Literature Review
- Co-author of Suspect Case Report Process Chapter
- Co-author of Thailand Situation Summary Chapter
- Co-author of Investigation Process Chapter

**Presentation**

**International conference**

**Oral presentation**

Presented “Fatal Injuries from Boat Travelling in Pattaya, Thailand, November 2013 ” at The 7th Asian Conference on Safe Communities in Busan, South Korea, 13 May 2014.


**Thailand conference**

**Poster presentation**

Presented “Dengue cluster investigation in two districts, Ubon Ratchathani, January-July 2013 : Epidemiological characteristics and key vector containers” at The 22nd National Epidemiology Seminar, February 2015.
Professional Membership

Epidemiology Association of Thailand (EpiAT) February 2015-present

Thai Medical Informatics (TMI) November 2014-present

Field Epidemiology Association of Thailand (FEAT) December 2014-present

The Medical Council of Thailand November 2010-present

Professional Service

Project

1. Public Health Informatics training program
   Ministry of Public Health, Thailand and Centers for Disease Control and Prevention (USCDC)
   - Curriculum Design
   - Instructor
   - Teaching Assistant

   The Johns Hopkins Center for Bioengineering Innovation & Design (CBID)
   - Research Design and Methodology Consultant
   - Data analysis
   - Android Application Developer

3. Lead exposure surveillance, prevention and eradication in preschool child committee, 2015, Department of Disease Control, Ministry of Public Health, Thailand
   - Medical Consultant
   - Committee Member

4. Driver license examination revision committee 2015, Department of Land Transport, Ministry of Transport, Thailand
   - Medical Consultant

Surveillance system evaluation
Data analysis and management team leader


Outbreak Investigator

Principal Investigator

1. Explosion in chemical substance warehouse at Samut Prakan
2. Diphtheria death in Satun September 2013
3. Food Poisoning, Mae Rim, Chiang Mai

Co-Principal Investigator

1. The outbreak of Staphylococcus aureus in Neonatal unit, Bangkok
2. Vancomycin Resistance Enterococci (VRE) outbreak in Ramahibodi hospital
3. An Investigation of Influenza A H1N1 2009 (deaths, Chiang Dao District, Chiang Mai
4. Hepatitis A Outbreak in Chon Buri – Rayong
5. Food poisoning, Chiang Rai
6. Meningococcemia in Prison, Nonthaburi

Teaching Assistance) Short Course and Seminar

Data analysis workshop, The 22\textsuperscript{nd} National Epidemiology Seminar, 2-3 Feb 2015

- Prepare Assignment and Application
- Help Participant for Technical Problem

Invited instructor

Presented the example of injury investigation to Marine Officers, Kanchanaburi. 7\textsuperscript{th} January, 2014

Presented the current project of Ministry of Public Health in controlling Blood Lead Level Project, Kanchana Buri, Thailand, 2014
Invited speaker

Heat stroke in summer, Good Morning Thailand, People Share, Mono29 channel 17th April 2015