Hand, Foot, and Mouth Disease in Thailand: a comprehensive modelling of epidemic dynamics

Suraj Verma¹, M. A Razzaque¹*, U Sangtongdee¹, C Arpnikanondt², B Tassaneetrithep³, D Arthan⁴, C Paratthakonkun⁵ and N Soonthornworasiri⁶

Abstract

Background
Hand Foot and Mouth Disease (HFMD) is a highly contagious disease and has become an epidemic in many Asian-Pacific countries, including Thailand. With such epidemic characteristics and potential economic impact, HFMD is a significant public health issue. Comprehensive modelling of HFMD’s epidemic dynamics can be useful in understanding and predicting any potential outbreak of it, and manage its impact efficiently and effectively. Generally, the transmission dynamics of infectious diseases vary across geolocations due to different socio-economic situations, demography, and people’s lifestyles. However, there is no nation-wide and comprehensive (i.e., the inclusion of reinfections in the model) modelling of HFDM dynamics in Thailand. We aim to develop a nation-wide comprehensive modelling of HFMD’s epidemic dynamics and understand the reinfection cases in Thailand.

Methods
We have formulated Susceptible - Exposed - Infectious - Recovered - Susceptible (SEIRS) epidemiological model with dynamic vitals, including reinfections, to investigate the transmission of this disease in Thailand. We also introduced periodic seasonality to model the seasonal effect. According to the model, the spread of this disease is uneven throughout the provinces in Thailand. So, we have grouped the provinces into three clusters (i.e., highly, moderately and least affected provinces) using K-means unsupervised machine learning algorithm for better estimation of the parameters and fitting the model. We collected data from three local hospitals in Thailand to analyze the reinfection cases.

Results
The result from the analysis of HFMD recorded cases from three hospital (years 2012 to 2016) shows that 11% (approximately) are reinfections. By fitting the model with HFMD confirmed cases (years 2011 to 2019) and considering the reinfections, the basic reproductive number \( R_0 \) was estimated to be 2.643, 1.91 and 3.246 for three clustered provinces.

Conclusion
In a conclusion, it is found that HFMD is re-infectious disease in Thailand. It is also found that the spread of HFMD is not uniform across the provinces in Thailand. The basic reproductive number \( R_0 \) was estimated to be greater than 1 for all the three clusters. This indicates that under the same social and environmental condition, this disease will persist in coming years.
Keywords: Hand, Foot, Mouth Disease; HFMD; SEIRS; Thailand; Epidemiological model; KMeans

Background

HFMD is a highly contagious pediatric disease that became an epidemic in the Asia-pacific region by the 1990s. This disease has spread in several Asian-Pacific countries, including Australia, Cambodia, China, Japan, Malaysia, Taiwan, Vietnam, and Thailand [1]. In Thailand, the first outbreak was recorded in 2012 with a total number of 45464 infected people which was the largest in a decade. Thailand experienced the second outbreak of this disease in 2016 [2]. With such frequent outbreaks, HFMD has been a critical health issue and has tremendous socio-economic effects. This disease affects the health of children, their social life and their education as they need to keep them isolated from healthy children. Though it is a pediatric disease, even parents are affected as they have to take special care of their children during this period. HFMD outbreaks have caused a considerable economic strain in the communities and healthcare systems of the affected nation.

For healthcare and financial planning, knowing the current spread of the disease and forecasting the epidemic is very important. Mathematical models of epidemiology explaining the population dynamics of contagious diseases have played a significant role in better understanding of epidemiological patterns and control of diseases [3]. As the patterns of transmission of infectious diseases vary through geolocations due to various socioeconomic conditions, demographics and people’s lifestyles [4], a nation-wide and/ or regional and comprehensive model is required to understand the transmission of this disease. Several mathematicians and researchers have proposed various epidemiological models for better understanding of the transmission of contagious diseases. Chuo, F. et. al [5] used a SIR model to predict the number of infected people during the outbreak in Sarawak of Malaysia. However, this model did not consider any seasonality effect and reinfection cases in the spread of this disease. Later on, a periodic model considered with quarantine effect was proposed by Liu, J. [6]. Another approach was adopted by Du, Zhicheng et al. [7] to model this disease using Time-series Susceptible Infected Recovered model. In this model, a time scale for this disease was considered to be two weeks based on the characteristics of HFMD where the duration of the transition from infection to recovered and temporary immunity is about 7 to 10 days. Very few studies have been conducted for modelling the outbreak of this disease in Thailand. Chadsuthii S, Wichapeng S [8] conducted research on the outbreak of HFMD in Bangkok 2016 and proposed customised SEIR model. The customised model considered Susceptible (S), Exposed but not yet infectious (E), Infectious with symptoms (I), hospitalised (H) and Recovered (R). The basic reproductive number ($R_0$) was predicted to be 1.441. However, the study was conducted for the outbreak of 2016 for Bangkok only, and no seasonal effect was considered in this model. Furthermore, in a study conducted by Huang, Jiao, et al. [9] found that multiple viruses, Enterovirus-71 (EV-71) and Coxsackievirus A16 (CA 16) are the causative agents of HFMD. Children infected with one virus are likely to be infected with other viruses that cause the same disease. As per their study conducted in China, 1.9% were reinfected at 12 months, 3.3% at 36 months and 4.0% at 38.8
months. However, there is no study that is taking care of the cases of reinfection in Thailand. Importantly, to date, no nation-wide and comprehensive model has been developed for modelling the transmission of HFDM that considers seasonal effect and reinfection cases.

This research aims to develop comprehensive modelling of HFMD’s epidemic dynamics in Thailand. First, we analysed the existing HFMD datasets (2011-2020) and grouped the affected provinces/zones into three clusters (i.e., highly, moderately and least affected provinces). Then, we developed a nation-wide epidemiological model of the disease to generate an understanding of the transmission of this disease and to predict the number of infected people considering the reinfection cases in Thailand. This paper plans to answer the following questions through the analysis of the HFMD model and a case study conducted at three local hospitals in Thailand.

- What is the status of HFMD disease in Thailand?
- What are the chances that this disease will outbreak as an epidemic in the year 2020?
- Are all the provinces of Thailand equally affected with HFMD?
- What is the probability of reinfection cases in Thailand?

Section 2 presents the materials and methods, including the data collection and preprocessing, and the construction of theoretical SEIRS model. The results from the study of the collected data are summarised in Section 3. This section also demonstrates the results obtained from clustering and fitting the clusters in the model with the use of $R^2$ to check goodness of fit. In Section 4, we discussed the findings from the analysis and case study performed and the impact of COVID-19 in the spread of HFDM. Section 5 concludes the work.

**Methods**

**Data collection**

For the purpose of this study, we collected province-wise monthly reported data of HFMD cases from National Disease Surveillance (report 506), Bureau of Epidemiology (BoE) (January 2011 to July 2020). We also gathered data for HFMD cases for the different age group of people in Thailand for the year 2019 from BoE [2]. The population of children under 14 years for years 2019 was collected from The World Bank Group [10]. Moreover, to understand the reinfection cases patients’ records were collected from three local hospitals, namely Samut Sakon, Suan Pueng, and Ranong.

The three healthcares are district hospitals which aim to offer standard medical services for Thai and non-Thai resident. Those local datasets were derived from a retrospective study associated with a situation and trend of hospital patients in western borders between 2012 and 2016 [11]. All personal identifiers were not obtained. It ensures that confidentiality became extreme strictly enforced throughout the study. Ethical approval for the retrospective study which obtained local datasets was provided from the Faculty of Tropical Medicine, Mahidol University.

**Data Preprocessing**

The raw data collected from local hospitals were preprocessed to get the total number of reinfection cases reported in the selected three hospitals (Figure 2). In
first step, patients’ personal details were de-identified to maintain their privacy. Then, patients with record count ≥2 were identified. After that, followed up records were removed to get the visit of cases with this disease but not a follow up.

![Figure 1: National data preparation](image1)

![Figure 2: Local data preprocessing](image2)

**Model**

We modelled the trajectory of the infection of HFMD disease in Thailand from 2011 to 2019 using Susceptible - Exposed - Infectious - Recovered (SEIR) model with vital dynamics which is a compartmental model. This model considers the incubation period of the disease [12]. This model was customised as SEIRS to address the reinfection cases where a fraction of the recovered patients get reinfected with the same disease. The population was divided into following four compartments:

- **Susceptible (S):** denotes people who are not infected but can get infected.
- **Exposed (E):** denotes people who are exposed but not yet infected.
- **Infected (I):** denotes people who are infected and can pass it on
- **Removed (R):** denotes people who have recovered from this disease or died.

Let us assume the total population of children who can be infected with HFMD be N. Then, the total population can be represented as, \( N = S + E + I + R \). The equations for SEIRS model are as follows.

![Figure 3: SEIRS Model with vital Dynamics.](image3)

\[
\frac{dS}{dt} = \mu(N - S) - \beta \frac{I}{N} S + \omega R \tag{1}
\]

\[
\frac{dE}{dt} = \beta \frac{I}{N} S - (\sigma + \mu) E \tag{2}
\]
\[ \frac{dI}{dt} = \sigma E - (\gamma + \mu)I \] (3)

\[ \frac{dR}{dt} = \gamma I - (\mu + \omega)R \] (4)

Here, \(\mu\) represents the birth and mortality rates per capita. Let us assume that the birth rate and the death rate is the same giving the constant population throughout the year. \(\beta\) represents the contact/transmission rate at which any infected individual contacts with susceptible people to spread the disease. If \(I\) infected people, contact \(S\) susceptible people and the rate at which disease can spread is \(\beta\) then \(\beta \frac{I}{S}\) represents the average number of infections. \(\frac{1}{\sigma}\) and \(\frac{1}{\gamma}\) represents the mean latent period and infectious period respectively. The rate at which recovered patients become susceptible is represented by \(\omega\). The basic reproductive number for this model is given by

\[ R_0 = \frac{\beta}{(\sigma + \mu)(\gamma + \mu)} \] (5)

**Seasonality of transmission**

HFMD in Thailand is a recurrent epidemic disease. Historical data illustrates that the disease outbreaking seasonally during the rainy season each year. The transmission rate varies seasonally giving a spike during the rainy season. The seasonal transmission was considered in our model focusing on the rate of transmission \(\beta\) as \(\beta = \beta(t)\). A sinusoidal trigonometric function was introduced to estimate \(\beta\) as follows.

\[ \beta = \beta_0 [1 + \sin(\frac{2\pi(t + \alpha)}{T})] \] (6)

Where, \(\beta_0\), \(\alpha\), \(t\) and \(T\) represents the baseline of transmission rate, time constant, time and time span of seasonal cycle respectively [13]. The main advantage of introducing this trigonometric function is to explore the qualitative characteristics of seasonally varying disease transmission and varying transmission rate.

**Data Fitting and Mathematical Simulation**

We applied the K-means algorithm with cluster size 3 to cluster the provinces based on the number of confirmed HFMD infection cases. The provinces were clustered into highly, moderately and least affected provinces throughout the year 2011 to 2019. These clustered provinces were fitted in the SEIRS model. We used the data collected from the BoE to simulate the model. The time of each cycle (T) was considered as 12 months. We estimated the rate of reinfection (\(\omega = 0.119\)) from the dataset gathered from the three local hospitals. We used Scipy optimiser’s curve fitting tool to fit the model with the data and estimate the parameters. The curve-fitting tool uses least-squares minimisation technique to fit the data in the model. We adjusted the model parameters to fit the model to match data closely without any overfitting. The goodness of fit of the simulated model was estimated using the determination coefficient (\(R^2\)).
Results

Analysis of spread of Disease

We analysed the collected data to understand the spread of HFMD in different provinces throughout the year. We found that the disease generally outbreaks in June, July and August of each year throughout Thailand. Two outbreaks were seen in 2012 and 2016, where the number of infected cases during the rainy season was very high compared to other seasons (Figure 4). We plotted a heatmap using the HFMD cases to visualise the spread of this disease across provinces during the year 2011 to 2019. We found from the heatmap that provinces could be grouped into three clusters as per the severity of infection (Figure 5). The K-means algorithm clustered the provinces based upon the number of infection cases from the year 2011 to 2019. Based on the severity of infection of the disease as observed in the heatmap, we identified three different clusters of provinces in Thailand. The provinces are clustered into highly, moderately and least affected provinces (Table 2, Figure: 8).

To analyse the age group of infected people in Thailand, we plotted a bar graph for the year 2019. As expected, we observed that children under 14 years are more infected than older people (Figure 6). We used Paintmaps online tool to plot the heatmap in Thailand map and visualise the geographical spread of this disease for the year 2019 (Figure ??). The geographical map illustrates that HFMD spreads in most of the provinces of Thailand, and Bangkok is the most affected province. From the case study performed in three different local hospitals in Thailand, we found that there is a significant number of patients who visited the hospital were reinfected with HFMD. For the selected three hospitals (2012-2016), we found different cumulative reinfection cases or rate are not insignificant, and it is increasing with the inclusion of an additional year (Table 1). For example, the reinfection rate or percentage for the year 2012-2015 is 9.39%, and this has changed to 11.92 % just adding the year 2016.

<table>
<thead>
<tr>
<th>Year</th>
<th>Reinfection Cases</th>
<th>Patients</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>2012</td>
<td>14</td>
<td>383</td>
<td>3.66</td>
</tr>
<tr>
<td>2012 to 2013</td>
<td>42</td>
<td>696</td>
<td>6.03</td>
</tr>
<tr>
<td>2012 to 2014</td>
<td>91</td>
<td>1151</td>
<td>7.90</td>
</tr>
<tr>
<td>2012 to 2015</td>
<td>147</td>
<td>1565</td>
<td>9.39</td>
</tr>
<tr>
<td>2012 to 2016</td>
<td>249</td>
<td>2089</td>
<td>11.92</td>
</tr>
</tbody>
</table>

Figure 4: HFMD confirmed cases 2011 - 2019.
Fitting the Model
SEIRS model was fitted using curve fitting and the results obtained from the fitting for each cluster are illustrated in figures 9, 10 and 11. For better fitting, the model was fitted for each year and values for the parameters $\alpha, \beta_0, \sigma, \mu, \gamma$ were estimated.

Estimation of parameters
For the model simulation, the value of $N$ was considered as the population of children under 14 years. The initial infected population, $I0$ was taken from the
Table 2: Clustering of provinces based on the number of infections.

<table>
<thead>
<tr>
<th>Clusters</th>
<th>Severity of Infection</th>
<th>Provinces</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cluster-1</td>
<td>Low</td>
<td>Lampang, Lamphun, Mae Hong Son, Nan, Phayao, Phrae, Phetchabun, Sukhothai, Tak, Uttaradit, Chai Nat, Kamphaeng Phet, Phichit, Uthai Thani, Ang Thong, Nakhon Nayok, Nonthaburi, N. Nakhon S. Ayutthaya, Pathum Thani, Sing Buri, Kanchanaburi, Nakhon Pathom, Phetchaburi, Prachuap Khiri Khan, Ratchaburi, Samut Sakhon, Samut Songkhram, Suphan Buri, Chachoengsao, Chanthaburi, Prachin Buri, Sa Kaeo, Trat, Kalasin, Maha Sarakham, Roi Et, Bungkan, Loei, Nakhon Phanom, Nong Bua Lam Phu, Nong Khai, Sakon Nakhon, Udon Thani, Chaiyaphum, Amnat Charoen, Mukdahan, Yasothon, Chumphon, Krabi, Nakhon Si Thammarat, Phangnga, Phuket, Ranong, Narathiwat, Pattani, Phatthalung, Satun, Songkhla, Trang, Yala</td>
</tr>
<tr>
<td>Cluster-2</td>
<td>High</td>
<td>Bangkok</td>
</tr>
<tr>
<td>Cluster-3</td>
<td>Average</td>
<td>Chiang Mai, Chiang Rai, Phitsanulok, Nakhon Sawan, Lop Buri, Saraburi, Chon Buri, Rayong, Samut Prakan, Khon Kaen, Buri Ram, Nakhon Ratchasima, Surin, Si Sa Ket, Ubon Ratchathani, Surat Thani</td>
</tr>
</tbody>
</table>

BoE. The parameters $\alpha$, $\beta_0$, $\sigma$, $\mu$, $\gamma$ were estimated by fitting the infected data in the model using curve-fit.

Table 3: Estimation of parameters for provinces in cluster 1.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$N$</td>
<td>6563521</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$I_0$</td>
<td></td>
<td></td>
<td>134</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$E_0$</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$T$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\omega$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.119</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha$</td>
<td>-0.41</td>
<td>-0.577</td>
<td>0.054</td>
<td>0.337</td>
<td>1.578</td>
<td>-0.256</td>
<td>0.483</td>
<td>-0.083</td>
<td>-0.575</td>
</tr>
<tr>
<td>$\beta_0$</td>
<td>3.061</td>
<td>1.835</td>
<td>1.964</td>
<td>2.379</td>
<td>1.539</td>
<td>2.007</td>
<td>2.076</td>
<td>2.204</td>
<td>2.186</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>0.739</td>
<td>40.018</td>
<td>2.175</td>
<td>3.015</td>
<td>0.71</td>
<td>11.818</td>
<td>2.836</td>
<td>3.85</td>
<td>4.02</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>3.083</td>
<td>0.919</td>
<td>0.469</td>
<td>0.631</td>
<td>-0.024</td>
<td>0.928</td>
<td>0.384</td>
<td>0.716</td>
<td>0.762</td>
</tr>
<tr>
<td>$\mu$</td>
<td>-0.331</td>
<td>1.261</td>
<td>0.81</td>
<td>1.067</td>
<td>0.398</td>
<td>1.151</td>
<td>0.996</td>
<td>1.033</td>
<td>1.065</td>
</tr>
<tr>
<td>$R_0$</td>
<td>15.261</td>
<td>3.877</td>
<td>1.829</td>
<td>2.984</td>
<td>0.369</td>
<td>3.802</td>
<td>2.121</td>
<td>3.038</td>
<td>3.246</td>
</tr>
</tbody>
</table>

Validity of Model

The fitting of the model was tested by regression analysis. The determination coefficient ($R^2$) was calculated using the following formula. $R^2$ for each cluster is
Table 4: Estimation of parameters for provinces in cluster 2.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>779326</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$I_0$</td>
<td>39</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$R_0$</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\gamma$</td>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\omega$</td>
<td>0.119</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\rho$</td>
<td>-0.512</td>
<td>-0.781</td>
<td>0.879</td>
<td>0.813</td>
<td>1.438</td>
<td>-0.001</td>
<td>1.115</td>
<td>0.998</td>
<td>-0.317</td>
</tr>
<tr>
<td>$\beta$</td>
<td>4.749</td>
<td>4.821</td>
<td>1.488</td>
<td>1.216</td>
<td>0.68</td>
<td>1.388</td>
<td>1.522</td>
<td>1.673</td>
<td>1.812</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>1.103</td>
<td>74.438</td>
<td>0.761</td>
<td>4.683</td>
<td>4.253</td>
<td>29.975</td>
<td>2.134</td>
<td>1.436</td>
<td>7.488</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>5.861</td>
<td>3.539</td>
<td>-0.623</td>
<td>-0.052</td>
<td>-0.825</td>
<td>0.552</td>
<td>-0.111</td>
<td>-0.099</td>
<td>0.624</td>
</tr>
<tr>
<td>$\mu$</td>
<td>-0.477</td>
<td>2.98</td>
<td>0.947</td>
<td>0.846</td>
<td>1.066</td>
<td>0.877</td>
<td>0.797</td>
<td>0.717</td>
<td>1.036</td>
</tr>
<tr>
<td>$R_0$</td>
<td>45.07</td>
<td>30.221</td>
<td>0.215</td>
<td>0.818</td>
<td>0.151</td>
<td>1.927</td>
<td>0.761</td>
<td>0.689</td>
<td>2.643</td>
</tr>
</tbody>
</table>

Table 5: Estimation of parameters for provinces in cluster 3.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>360432</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$I_0$</td>
<td>107</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$R_0$</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\gamma$</td>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\omega$</td>
<td>0.119</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\rho$</td>
<td>-0.633</td>
<td>0.264</td>
<td>0.936</td>
<td>0.271</td>
<td>-0.435</td>
<td>-0.194</td>
<td>1.104</td>
<td>-0.177</td>
<td>-0.092</td>
</tr>
<tr>
<td>$\beta$</td>
<td>2.802</td>
<td>0.634</td>
<td>0.825</td>
<td>1.896</td>
<td>2.022</td>
<td>1.632</td>
<td>1.572</td>
<td>2.16</td>
<td>1.907</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>0.622</td>
<td>83.55</td>
<td>2.907</td>
<td>5.728</td>
<td>1.752</td>
<td>23.559</td>
<td>2.208</td>
<td>2.916</td>
<td>3.074</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>2.279</td>
<td>0.34</td>
<td>-0.578</td>
<td>0.616</td>
<td>0.368</td>
<td>0.674</td>
<td>-0.382</td>
<td>0.73</td>
<td>0.384</td>
</tr>
<tr>
<td>$\mu$</td>
<td>-0.174</td>
<td>0.131</td>
<td>0.891</td>
<td>1.033</td>
<td>0.866</td>
<td>1.054</td>
<td>1.055</td>
<td>0.84</td>
<td>0.917</td>
</tr>
<tr>
<td>$R_0$</td>
<td>8.189</td>
<td>0.298</td>
<td>0.198</td>
<td>2.649</td>
<td>1.67</td>
<td>2.699</td>
<td>0.715</td>
<td>2.632</td>
<td>1.91</td>
</tr>
</tbody>
</table>

illustrated in table 6. The value of $R^2$ is greater than 0.6 which illustrates that at least 60% of the recorded data fitted well in the model.
Discussion
The main finding of this study is the reinfection cases or rate of HFMD in Thailand. HFMD recovered patients become immune with the infected virus and the chances of getting reininfected with the same virus are very rare. However, as the disease is caused by multiple viruses (Entrovirus-71 (EV-71) and Coxsackievirus A16 (CA 16)) [14], there is a chance of getting infected with other virus causing the same disease. From the result of the case study performed at three local hospitals in Thailand, it is found that 11.92% (approximately) of the recorded cases are reinfection cases. The cases of reinfection are significantly high, especially during the rainy seasons.

HFMD is most prevalent spreading in younger ages compared to the older groups of patients across the country. During the mid of each year, the total number of patients increases consistently in a similar pattern of the previous. One of the reasons is that staring from early May to late July is the rainy season in Thailand and this disease transmits quite rapidly during the rainy season. Moreover, primary and secondary schools begin the first semester in the middle of May [15] which leads to the increase in physical contact between infected and susceptible children. Also, at the beginning of the semester, children are not provided with full support, especially wellness and health checks as well as physical education. These may cause the infection to reach a peak during the mid of each year. The reinfection is also the supporting cause of outbreak during rainy seasons each year.

\[
SS_{\text{tot}} = \sum_i (y_i - \bar{y})^2
\]
\[
SS_{\text{res}} = \sum_i (y_i - f_i)^2
\]
\[
R^2 = 1 - \frac{SS_{\text{res}}}{SS_{\text{tot}}}
\]

<table>
<thead>
<tr>
<th>Clusters</th>
<th>(R^2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cluster 1</td>
<td>0.63</td>
</tr>
<tr>
<td>Cluster 2</td>
<td>0.777</td>
</tr>
<tr>
<td>Cluster 3</td>
<td>0.756</td>
</tr>
</tbody>
</table>

Table 6: Calculation of \(R^2\).
Figure 4 illustrates that the number of infected cases is very high during rainy seasons and the outbreak of this disease occurred in 2012 and 2016. The heatmap (Figure: 5) and geographical map (Figure: 7) help us to visualise the spread of this disease across all the provinces. We found that all provinces are not equally affected. Bangkok is observed to be the most affected province. Since the transmission of this disease is uneven across the provinces, it is important to identify the provinces which have similar severity of transmission of the disease. We applied the K-means algorithm to cluster the provinces based on the severity of the infection cases. The resulting cluster grouped the provinces into three categories, highly, moderately and least affected provinces. The main advantage of clustering the provinces based on the number of infections is that highly affected provinces can get more attention and the healthcare facilities and human resources to take appropriate actions to control the disease.

Mathematical models have always been an important tool for understanding the spread of epidemiological diseases. Till to date, no comprehensive model has been developed to address the nation-wide transmission of this disease with reinfection cases in Thailand. To understand and predict the spread of this disease in Thailand, we have built a comprehensive compartmental SEIRS model with vital dynamics by incorporating the data from the BoE surveillance data and reinfections rate from three hospitals to estimate the $R_0$ of HFMD. We also have introduced periodic seasonality in this model to address the seasonal effect in the transmission of this disease in Thailand. Generally, the time from initial infection to the occurrence of signs and symptoms (incubation period ) of HFDM is 3 to 7 days [14]. During this period, infected people are not yet infectious. In the proposed model, $E$ (Exposed) addresses the incubation (latent) period. This model also addresses the rate of reinfection where a fraction of the recovered patient become susceptible because of different virus type. $\omega$ is the rate at which recovered patients become susceptible. The proposed SEIRS model was fitted with recorded HFMD cases based upon the provinces in each cluster to get a better estimation of the parameters. The basic reproductive number($R_0$) obtained from the simulation of three clustered provinces is greater than 1 which indicates that under the same social and environmental condition, this disease will persist in coming years.

Effect of COVID-19

HFMD is transmitted primarily through direct contact with contaminated discharge, saliva or stool from infected patients or contaminated objects. Moreover, attending kindergarten or child-care centres or schools increases the risk factor for the transmission of this disease [16]. Avoiding physical interaction with people who have HFMD, such as kissing, embracing, and sharing cups and eating utensils plays a significant role in minimising the spread of this disease. Because of COVID-19 pandemic, a nationwide lock-down and quarantine of infected patients are implemented. The schools are closed and no physical contact is made with infected patients who have played a significant role in decreasing the number of HFMD patients in 2020. For each month of this year, the number of infected cases is decreasing. The predicted number of HFMD infected patients and the recorded number of cases for year 2020 for each cluster is illustrated in figures 11a, 11b and 11c. These figures illustrate that there is a significant decrease in the number of infected cases and even
during the rainy season, the cases are decreasing which is because of the impact of COVID-19 lock-down imposed.

![HFMD Prediction for year 2020](image)

(a) Provinces in cluster 1.  
(b) Provinces in cluster 2.  
(c) Provinces in cluster 3.

Figure 12: HFMD Prediction for year 2020

Conclusion

In this paper, we have developed a nation-wide and comprehensive model of HFMD’s of epidemic dynamics in Thailand using SEIRS epidemiological model with dynamic vitals, including reinfections, to investigate the transmission of this disease in Thailand. From the experiment performed using the model and study conducted at three local hospitals in Thailand, we can conclude that HFMD is a critical health issue in Thailand and the spread of this disease is uneven across the provinces in Thailand. Hence, we clustered the provinces in three clusters based on the number of infections from the year 2011 to 2019 and the basic reproductive number ($R_0$) obtained from the SEIRS model for all these clusters is greater than 1 which indicates that under the same social and environmental condition this disease will persist in coming years.

Moreover, we found that the rate of reinfection, especially during the rainy season is remarkably high in Thailand. This is one of the reasons for the higher infection rate or cases during the rainy season. However, the number of infected cases is significantly low for the year 2020 due to COVID-19 pandemic. The lockdown imposed due to the pandemic help to minimise the direct or indirect physical contact with infected people and spread of COVID-19. This also helped to minimise the spread of HFMD. So, taking proper and informed actions could improve the management of HFMD like diseases in Thailand and globally.

Abbreviations

**BoE** Bureau of Epidemiology. 3, 5, 13  
**HFMD** Hand Foot and Mouth Disease. 1–6, 12  
**SEIR** Susceptible - Exposed - Infectious - Recovered. 4  
**SEIRS** Susceptible - Exposed - Infectious - Recovered - Susceptible. 1, 3–5, 11–13

Source of Map

Two maps which illustrated HFMD spreading over Thailand was created with Paintmaps an online tool. To Paintmaps, a Free and Open Source Geographic Information to use under an open license, it was an original source as a basemap before shading colors to define clusters.
Figure Legends
Figure 1 BoE data preprocessing steps for fitting the SEIRS model.
Figure 2 is the preprocessing steps for raw data collected from local hospital to estimate the reinfection rate of this disease. Those data were obtained from Ranong Hospital at Southern borders, Samutsakorn Hospital at central area, Rachaburi at western border.
Figure 3 represents the SEIRS epidemiological model which illustrates the compartmental division of population and the rate of infection and recovery.
Figure 4 illustrates HFMD confirmed cases from year 2011 to 2019 throughout Thailand. This figure shows that there is sharp peak in the number of HFMD cases during rainy season of each year.
Figure 5 is the heatmap to illustrate HFMD cases for year 2011 to 2019 across all the provinces in Thailand. From the figure, we can visualize that the provinces can be grouped into three groups as per the spread of this disease.
Bangkok seen to be highly affected province.
Figure 6 shows that children age group 1, 2, 3 and 4 are highly infected with this disease.
Figure 7 and 8. Shade maps of Thailand entirely with 77 provinces. Each map is illustrated how severity of HFMD infection spreading and reporting over the country in 2019. Two basemaps of "PaintMaps" demonstrate in backgrounds, these are a Free and Open Source Geographic Information which applied under an open license.
Figure 12 illustrates the predicted and actual HFMD cases in 2020. It is seen that the actual HFMD cases in each clustered provinces is significantly low compared to the prediction.

Declaration
Ethics approval and consent to participate
Three datasets which are from local hospitals are secondary data. They were collected as a part of a different research in 2017. The prior authors had conducted the study under their own ethical clearance. It is a certificate of ethical approval (MUTM2017-040-03). The approval was obtained from Ethics Committee of the Faculty of Tropical Medicine at Mahidol University. The committee considered its proposal and study procedures thoroughly. The research is a retrospective study which conducted without human samples and collections. Only periodic clinical data were transfer to this current study. A written consent is not required to conduct a prediction analysis study.

Consent to publish
Not applicable.

Availability of data and materials
The data supportive of the whole study's results have always been available from the corresponding author, MR, upon reasonable request.

Competing interests
The authors declare that they have no competing interests.

Funding
This study was funded by the Newton Fund - Institutional Links from British Council and National Research Council of Thailand with the grant ID is 527643161.

Authors’ contributions
SV and MR participated in the design of the study, interpreted the data, wrote and edited the manuscript. US conducted statistical analysis, interpreted the local data and review partially the manuscript. CA provide feedback in analysing model and computational terms relevant. BT reviewed significantly clinical points associated with disease prevalence and various medical terms. DA, CP, and NS shared and supported practical experience upon a retrospective study of the hospital at western borders and hospital at central area. All authors have read and approved the manuscript.

Acknowledgements
This research was supported by British Council Thailand. Some special thanks go to Charnyote Pluempiiwiriyawej and Naeti Suksomboon for their helps and willingness as domain experts alongside the research project. Thanks also extends for the support of Thai Health Promotion Foundation.

Author details
1 School of Computing, Eng. & Digital Tech., Teesside University, Southfield Rd, Middlesbrough, UK. 2 King Mongkut’s University of Technology Thonburi, Bangkok, Thailand. 3 Center of Research Excellence in Immunoregulation, Research Department, Faculty of Medicine Siriraj Hospital, Bangkok, Thailand. 4 Department of Tropical Nutrition and Food Science, Faculty of Tropical Medicine, Mahidol University, Bangkok, Thailand. 5 College of Sports Science and Technology, Mahidol University, Nakhon Pathom, Thailand. 6 Department of Tropical Hygiene, Faculty of Tropical Medicine, Mahidol University, Bangkok, Thailand.

References


