Parameter estimation of HFMD infection in Malaysia with SIR model

Hanis Nasir [a,1,*], Fuaada Mohd Siam [a]

[a] Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Johor, Malaysia
[1] hanisnasir20@gmail.com*
* corresponding author

ARTICLE INFO

Keywords
HFMD
SIR Model
Data
Public Health

ABSTRACT

Hand, foot and mouth disease (HFMD) is a mildly infectious disease that happens to children due to low body immunity against viruses. In this paper, a simple SIR model is employed to represent the dynamics of HFMD in Malaysia. A weekly reported data of HFMD infection from week-1 of 2017 to week-36 of 2018 is used in the mathematical analysis. The procedure of parameter estimation is implemented in order to identify the parameter of the SIR model that fits the model with the reported data. The fminsearchbnd routine is used to minimize the objective function, which is the sum of absolute error minimization. By the Chi-squared test, the estimated parameters are reasonable. Based on the average value of the basic reproduction number R0, HFMD infection will remain to exist in the population under the current condition. Overall, this finding may be useful for public health personnel to reduce disease infection risk by planning more effective prevention strategies.

This is an open access article under the CC–BY-SA license.

1. Introduction (*Heading 1*) (bold, 11 pt)

Hand, foot and mouth disease (HFMD) is a mild spreading disease that usually affected the children who are below 10 years old [1]. The disease comes from some viruses such as Coxsackievirus A16 (CV-A16) and Enterovirus 71 (EV71) [2], [3]. CV-A16 or coxsackievirus comes from enteroviruses from the family Picornoviridae, while EV71 comes from one of a large family of viruses which build up in the human gastrointestinal tract. EV71 infection can cause serious illness and people infected with this virus may fatal. HFMD can be spread by the fecal-oral route, such as food and water that was prepared in the presence of fecal matter and poor cleaning after handling feces or anything that has been in contact with it. Others, it also spread through contact with mucus, respiratory secretion, saliva, and skin lesions of an infected person. The oral secretions and contact with surfaces that are contaminated with HFMD viruses also contribute to the disease spreading. There are many symptoms of HFMD such as feeling vaguely unwell, ulcers in the mouth, rashes or blisters on the palms and feet, and sometimes on the buttocks. HFMD usually did not cause serious illness because its symptom will usually disappear within 7-10 days without medical treatment. However, in some rare cases, HFMD can cause fatal complications such as polio-like paralysis or encephalitis[4], [5].

10.33292/amm.v3i1.30

Although HFMD is considered as a mild and moderately contagious disease, it gives impacts in social and economic due to no specific treatment to cure HFMD. Owing to that, much attention has been focused on devising methods for preventing the spread of HFMD. This illness commonly occurs in countries with an environmental condition like in summer or early autumn [6]. Based on the World Health Organization [7], Asian countries are very common to HFMD incidence, for example, Malaysia, China, Japan, and Vietnam. In Malaysia, the first outbreak of HFMD was in 1997 that affected a young child. Since that time, many prevention protocols have been taken into action, which includes a public reminder about the precautions to be taken to ensure the effectiveness of control measures of HFMD. Figure 1 shows a reported data on weekly incidence cases of HFMD in Malaysia from January 2017 to September 2018.

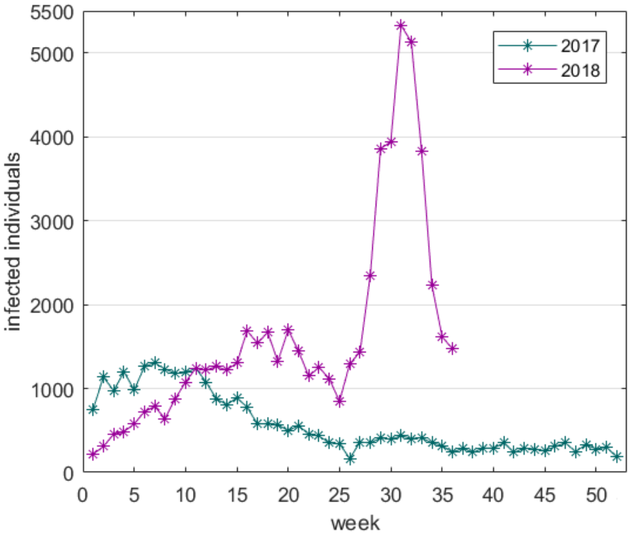


Fig. 1. Weekly data of HFMD incidence in Malaysia (News Desk, 2018).

The famous SIR epidemic model was firstly introduced by Kermack and McKendrick in 1927. SIR model describes the population into susceptible, infected, and recovered individuals in the population. In this model, susceptible individuals in S-state have possibilities to get infected and make a progression to infection I-state until recovery and goes into R-state.

In 2008, used SIR model to study the HFMD spreading among children in Sarawak, which is one of the states in East Malaysia [8]. The authors reported that there are 31 deaths happened in 1997 and 13 deaths in 2006. In their model, susceptible individuals are the category of children who is below ten years old. The model simulation is compared with the HFMD outbreak data of Sarawak in 2006. Again in Sarawak, studied the HFMD infectious disease by using time series modelling [9]. They used an auto-regression moving average (ARMA) model to fit with the reported case of HFMD from 2006 to 2012. This is the first time series analysis that applied to HFMD in Sarawak and the ARMA model has successfully fit with the trend of the disease. The ARMA model has over 90% forecast accuracy. Hence, the weekly prediction from the ARMA model in the first 16 weeks of 2013 gives some advantages for the future prevention strategies in Sarawak.

In 2016, studied HFMD in Sarawak by using age-dependent mixing patterns [10]. They have reported that the HFMD viruses may not only affect the children but adults as well. There is a possibility of HFMD cases among adults to happen in the future due to the changes in the climate and environment. With the emergence of these factors, the HFMD reported cases from 2011 to 2013 with the patients' age is used to determine the suitable value of each age group in their model. From the result, they suggested that by reducing the contact between young children will reduce the infection. In 2017, studied the HFMD outbreaks by considering the incubation and post-infection shedding period [11]. They improved the SIR model in by adding the variable of exposed individuals and clinically recovered individuals where the symptoms have subsided [8]. Performed the model simulation against the HFMD outbreak data of Sarawak in 2006 [11]. In comparison with simulation obtained, the model simulation [8]. Provided a good fit in the first ten weeks, while model simulation only able to fit with the HFMD data that have been push one week in advance[11].

It is important to continue the study of HFMD infection due to the disease is still occurs in the population. The scope of this paper is in Malaysia and this paper uses an epidemic-type structure model, which is the same [8]. Here, the procedure of parameter estimation is implemented, which has not been [11]. The procedure of parameter estimation is a systematic way of determining the unknown parameters of the model based on the actual data (see Figure 1). There is a need to estimate some parameters in the model due to these parameters are difficult to be measured directly.

In this work, a traditional SIR model is used to investigate the number of new reported cases of HFMD infection in Malaysia. The weekly data shown in Figure 1 is used to determine the parameter in the SIR model and the basic reproduction number, R0. Understanding the dynamics of HFMD is needed so it could help to suggest possible ways to control the spreading of HFMD. Besides that, the results obtained can be useful in predicting the future trend of HFMD in Malaysia based on the value of R0.

The content of this paper begins with a brief introduction and review of some previous analyses of HFMD. Next, the methodology section continued with the modelling formulation of an epidemic-type model for HFMD and the discussion on parameter fitting procedure. Then, the results are presented in the results and discussion section, which included the model simulation by using the estimated parameters. Lastly, the findings of this paper are briefly discussed in the concluding section.

2. Methodologu

2.1. Mathematical model

The SIR model categorized the total population, N into three compartments. Even though the SIR model is oversimplified, the paper can give insight into future work of suitable epidemic-type structure for HFMD in Malaysia once recent clinical data are available. In this paper, we use the SIR model of the form

$$\left.\begin{aligned}\frac{dS}{dt} &= \mu N - f(I)S - \mu S + \lambda R,\\ \frac{dI}{dt} &= f(I)S - \rho I - (\mu + \delta)I,\\ \frac{dR}{dt} &= \rho I - \lambda R - \mu R,\end{aligned}\right\} \tag{1}$$

with the initial condition of $S(0) > 0$, $I(0) \geq 0$ and $R(0) \geq 0$. The parameters on the model are: μ is the population growth rate, f(I) is the force of infection (per capita rate), λ is the rate of recovered individuals loss the immunity and become susceptible individuals, ϱ is the recovery rate, and δ is the mortality rate due to severe illness condition. The force of infection function, f(I) is written as $f(I) = \alpha\gamma\frac{I}{N}$, where α is the total number of contacts between susceptible and infected individuals (including effective or not) and γ is the infection risk that is the probability of effective contact and gets affected [12].Besides, the basic reproduction number, R0 is given as follows:

$$R_o = \frac{\alpha\gamma}{\rho+\mu+\delta}. \tag{2}$$

R_0 is also called as the basic reproductive ratio. R_0 is the essential theory of any infectious epidemic-type mathematical model. R_0 is defined as an expected number of secondary cases produced by a single infection into a completely susceptible population. Note that, if $R_0 < 1$, the infected individuals will die out and the disease eliminated. Otherwise, if $R_0 > 1$, infected individuals grow and spread in a population. Besides, if $R_0 = 1$, the epidemic will become endemic in the population as every infected person will infect one person before recovering [8].

2.2. Initial condition and parameter assumption

The interaction of each individual in the population is not the same all over time. For that reason, the parameter is investigated in a 26-weeks interval (half-year). For the initial number of susceptible $S(0)$, the subject of the investigation is fixed to 10% of the total population in Malaysia. Based on the reported data by the Department of Statistics Malaysia [13], the estimated total of the Malaysian population in 2018 is 32.385 million. Thus, $S(0) = 3.2385 \times 10^6$. Next, based on the data given in Figure 1, the initial number of the infected individuals $I(0)$ is set according to the respective week that is investigated. As for the number of recovered individuals, $R(0) = 0$ in each new computation of simulation in which none of the infected individuals has recovered in the beginning.

For parameter μ, it is assumed that the birth and natural death rate is the same. As estimated by the Department of Statistics Malaysia [13], the annual population growth for the year 2017 and 2018 is 1.2% and 1.1% per year, respectively. On average of these two years of population growth rate, the parameter μ is fixed and set as 2.2055×10^{-4} per week. Following, the transmission rate is set within 0.0001 to 50 per day, which means there is almost no contact per day and can go for up to 50 contacts per day [14]. Thus, for parameter α, it is assumed to have at least one contact person and a maximum of 50 contacts person per day. For a probability-type parameter, γ is always bounded in [0, 1]. Once the infected individuals get the treatment, they get immunity from that specific HFMD virus. However, another type of different HFMD virus can still affect them, hence become susceptible individuals again. Both parameter ρ and λ depend on the effectiveness of the treatment [14]. As put in the boundaries, ρ is within 0.1 to 1 per day and λ is within 0.0001 to 1 per day. For the mortality rate, it is significantly small ($\delta \approx 0$ per week) due to no new HFMD death case has been reported. The death caused by EV71 HFMD virus remains two cases, which are one in Sarawak and one in Penang (News Desk, 2018).

All these model assumptions and parameter boundaries will be used in the procedure of parameter fitting for the SIR model (1) with the reported data shown in Figure 1.

2.3. The procedure of parameter estimation

To estimate the parameter using the HFMD cases in Malaysia, a minimizing objective function (or minimizing the sum of absolute error (SAE)) is formulated. The SAE function that will fit the simulation data with the reported data is defined as follows:

$$\text{SAE} = \sum_{i=1}^{n} |log_2(1+y(i)) - log_2(1+\tilde{y}(i))|, \tag{3}$$

where y is the infected individuals at i-th week and $\tilde{y}$ is the model simulation of infected individuals at i-th week. This objective function is minimized by using *fminsearchbnd* built-in MATLAB function, which is based on Nelder-Mead Simplex optimization. Some recent works used *fminsearchbnd* syntax in searching the value of parameters of their models that fits the simulation data and the actual (or experimental) data[15],[16].

There are four parameters, $\mathbf{X} = [\alpha, \gamma, \lambda, \rho]$ that will be fitted using SAE minimization and the other parameters are fixed. It is essential to identify the value of these four parameters because they represent human interaction and body immunity. These factors are varied depends on the type of environment and the value needs to be identified by parameter estimation. Table 1 summarized the value and boundary of parameters and variables. The parameters are in per week unit since the reported data is available weekly and the solution of model (1) is solved numerically using *ode45* MATLAB syntax.

Table 1. List of parameter boundary and the fixed parameter.

Parameter	Boundary/Value	Source
α	[7, 350] week $^{-1}$	SAE
γ	[0, 1]	SAE
μ	2.2055×10^{-4} week $^{-1}$	Fixed
λ	[0.0007, 7] week $^{-1}$	SAE
ρ	[0.7, 7] week $^{-1}$	SAE
δ	0 week $^{-1}$	Fixed
S(0)	3.2385×10^6	Fixed
I(0)	Week 1 or 27	From real data
R(0)	0	Fixed

3. Result and Discussion

3.1. The estimated parameter and data fitting

Using the weekly data given in Figure 1, the analysis of data estimation is divided into four intervals: Week 1-26 (2017), Week 27-52 (2017), Week 1-26 (2018), and Week 27-36 (2018). The initial guess for *fminsearchbnd* syntax in each duration is set as $\mathbf{X}_0 = [7, 0.5, 0.0007, 7]$ and the optimal solution is obtained when the value of SAE is converged.

Table 2. The estimated value of the parameters, R_0, and SAE.

Duration:		Week 1-26 (2017)	Week 27-52 (2017)	Week 1-26 (2018)	Week 27-36 (2018)
Para:	α	7.0203	7.9526	7.0880	7.5853
	γ	0.9882	0.6249	0.9833	0.9599
	λ	0.0206	0.3003	0.0228	0.0007
	ρ	6.8288	4.9762	6.7723	7.0000
R_0		1.0159	0.9986	1.0291	1.0401
SAE		4.2735	4.9029	4.9081	5.5130

The result of estimated parameters for each duration is presented in Table 2 together with the value of R_0 and SAE. As shown in Figure 2, the model simulation provides a good fit towards the reported data of HFMD with the average value of SAE is 4.8988.

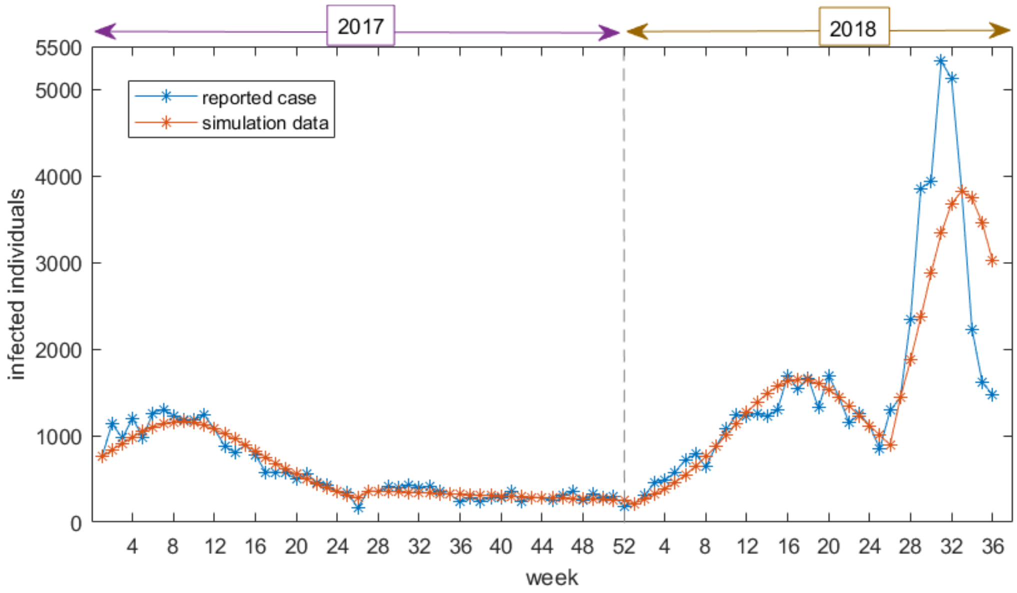


Fig. 2. The data fitting of weekly incidence data and model simulation of infected individuals.

3.2. Chi-square test of goodness of fit

A chi-squared test is employed to test how well the model reflects the data. By considering the same hypothesis used by [14] , the null- and alternative-hypothesis is as follows:

- Null-hypothesis, H_0: The estimated parameters are equal to actual values.
- Alternative-hypothesis, H_1: The estimated parameters are not equal to actual values.

Based on [17], the Chi-squared test statistics is defined as $\chi_v^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$, where E_i is the observation of HFMD data, O_i is the model observation of HFMD data, and v is the degree of freedom. The degree of freedom, $v = n - p - 1$ where n is the number of observations and p is the number of parameters of the hypothesized distribution estimated. The decision rule is to reject H_0 if $\chi_v^2 > \chi_{v,\alpha}^2$; otherwise do not reject H_0. $\chi_{v,\alpha}^2$ is the critical value of the Chi-squared test with v degree of freedom and α-level of significance, which is obtained from the Chi-squared distribution table.

By using a 0.05 level of significance, the result is presented in Table 3. The decision is the null-hypothesis, H_0 cannot be rejected at 5% significant level by Pearson's chi-squared test.

Table 3. Chi-squared value and degree of freedom of each interval.

Duration	**Week 1-26 (2017)**	**Week 27-52 (2017)**	**Week 1-26 (2018)**	**Week 27-36 (2018)**
v	21	21	21	5
χ_v^2	0.1608	0.1715	0.1600	0.3886
$\chi_{v,0.05}^2$	32.671	32.671	32.671	11.070

3.3. Sensitivity analysis of parameter in R0

The analysis is continued with the sensitivity of R_0 towards some critical parameters, which is α, γ, and ρ. Sensitivity analysis provides an index value on changes of the output variable with respect to the input parameter [16]. The index value will indicate which parameter is the most sensitive towards the changes of the output variable. By considering R_0 as the output variable and parameter α, γ, and ρ as the input parameters, the partial derivative of R_0 is derived as follows:

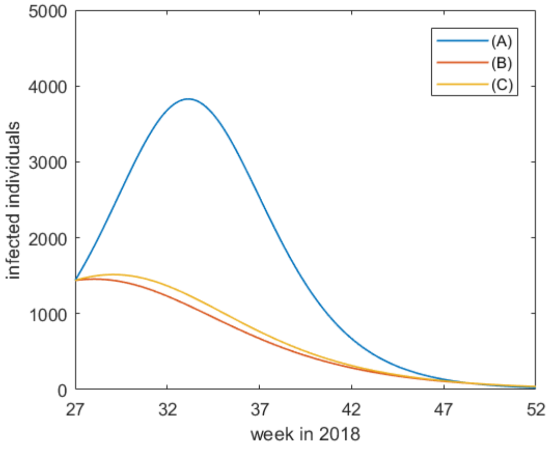
$$\frac{\partial R_0}{\partial \alpha} = \frac{\gamma}{\rho+\mu\delta}, \frac{\partial R_0}{\partial \gamma} = \frac{\alpha}{\rho+\mu\delta}, \frac{\partial R_0}{\partial \rho} = -\frac{\alpha\gamma}{(\rho+\mu\delta)^2}. \tag{4}$$

According to the sign shown by the derivative, parameter ρ is the only parameter that will decrease the value of R_0. In theory, if some parameters are controlled such that $R_0 <1$, then the disease will die out [14]. Hence, some suggestions and prevention protocols based on these parameters will be discussed further in the following section.

3.4. Discussion

From Table 2, no significant change of α can be observed in each duration, which means the people's behaviour and interaction are almost the same over the time. For the infection risk probability, most of the estimated value of parameter γ is close to 1, except Week 27-52 (2017). The value of γ that is close to 1 indicates that every contact of a susceptible person with an infected person will cause infection. For the recovery rate ρ, there are approximately 5-7 infected individuals discharged and enter the category of recovered person in one week. The average value of R_0 is 1.0209 which indicates that the HFMD infection in Malaysia will occur and consistent. By taking the definition of $R_0 > 1$, the HFMD infection will grow, but considering a less strict value of average R_0 that is very close and approximately 1, the infected individuals will at least infect one person before recovering.

By using the estimated parameters in Table 2(last column), model simulations are presented in Figure 3(A) for the Week 27-52 and hence, some prevention strategies are proposed.



4.

Fig. 3. (A) A model simulation of infected individuals using parameter: α = 7.5853 week $^{-1}$, γ = 0.9599, μ = 2.2055 $\times 10^{-4}$ week $^{-1}$, λ = 0.0007 week $^{-1}$, ρ = 7.0000 week $^{-1}$ and δ = 0 week $^{-1}$. (B) If α decreases, such that α = 7.3198 (= 7.5853 × 0.9650). (C) If ρ increases, such that ρ = 7.2275 (= 7.0000 × 1.0325).

According to the model simulation shown in Figure 3, reducing the contact rate between susceptible persons able to reduce the infection cases. These efforts may include a short-duration closure of kinder-garden and school during sudden increases in infection cases. Not only that but the increase of recovery rate also able to effectively reduce the infection cases. As the infected person gets the treatment immediately after got the symptoms of HFMD, it can significantly reduce the possibility of contact with other susceptible people. Besides, preventive measures such as quarantine and personal protection can be provided to the infected persons and against the exposure to other people.

Again, by using the estimated value of parameters in Table 2(last column) into Equation (4), the sensitivity index of R_0 for parameter α, γ, and ρ is 0.1371, 1.0836, and -0.1486, respectively. The sensitivity index for parameter α and γ shows that it is essential not to only decrease α, but parameter γ as well so that R_0 will be decreasing. Also, the control will be more effective by increasing parameter ρ because a negative index value indicates that the increment of ρ will decrease the R_0 value.

The model presented in this paper is a step in investigating the HFMD infection in Malaysia's environment. To develop this work further, in-depth details of the formulation of an epidemic-type structure model of HFMD infection should be done in the future. This paper provides a preliminary optimization analysis of data fitting to the reported cases of HFMD in Malaysia. This modelling enables us to compute the value R_0 from the procedure of parameter fitting and then predict the future trend of HFMD infection outbreak.

4. Conclusion

In this paper, the weekly HFMD data in Malaysia is used in data fitting of the HFMD infection using SIR model. By using the SAE minimization, the estimated parameters are obtained and the R0 value is calculated. The value of R0 is approximately 1 which indicates the HFMD infection persists in the population. Thus, the prevention strategies discussed in this paper should be taken into account to prevent a large-scale infection disease. From the analysis, it can be seen that the SIR model able to fit with the number of weekly reported cases of HFMD in Malaysia. However, for the sudden increase that happens in Week 24 (2018) onwards, the model is not being able to accurately fits the data. Due to that, an improvement for the model is needed to provide a better epidemic-type structure in representing the HFMD infection. For further development, the SIR model should be improved by considering more variables to make it more realistic with real-life human interactions.

Acknowledgment

This research was supported by Ministry of Education Malaysia and Universiti Teknologi Malaysia. The second author would like to thank Universiti Malaysia Terengganu for the research facilities.

References

[1] X.-Y. Wang, W. Ma, S. Ma, and H. Huo, “Modelling HFMD epidemic with periodic infection rate based on data analysis,” in *The 27th Chinese Control and Decision Conference (2015 CCDC)*, 2015, pp. 843–846.

[2] C. Song, X. Shi, Y. Bo, J. Wang, Y. Wang, and D. Huang, “Exploring spatiotemporal nonstationary effects of climate factors on hand, foot, and mouth disease using Bayesian Spatiotemporally Varying Coefficients (STVC) model in Sichuan, China,” *Sci. Total Environ.*, vol. 648, pp. 550–560, 2019.

[3] W. M. Koh, H. Badaruddin, H. La, I. Mark, C. Chen, and A. R. Cook, “Severity and burden of hand, foot and mouth disease in Asia: a modelling study,” *BMJ Glob. Heal.*, vol. 3, no. 1, p. e000442, 2018.

[4] S. Chadsuthi and S. Wichapeng, “The modelling of hand, foot, and mouth disease in contaminated environments in Bangkok, Thailand,” *Comput. Math. Methods Med.*, vol. 2018, 2018.

[5] T. P. Van Boeckel *et al.*, "Hand, foot, and mouth disease in China: critical community size and spatial vaccination strategies," *Sci. Rep.*, vol. 6, p. 25248, 2016.

[6] N. Roy, "Mathematical modeling of hand-foot-mouth disease: quarantine as a control measure," *Int. J. Adv. Sci. Eng. Technol. Res.*, vol. 1, no. 2, 2012.

[7] WHO, "Hand, Foot and Mouth Disease Information Sheet," 2012. .

[8] F. Chuo, S. Tiing, and J. Labadin, "A simple deterministic model for the spread of hand, foot and mouth disease (HFMD) in Sarawak," in *2008 Second Asia International Conference on Modelling & Simulation (AMS)*, 2008, pp. 947–952.

[9] N. M. Sham, I. Krishnarajah, M. Shitan, and M.-S. Lye, "Time series modelon hand, foot and mouth disease in Sarawak, Malaysia," *Asian Pacific J. Trop. Dis.*, vol. 4, no. 6, pp. 469–472, 2014.

[10] M. S. Noraishah and I. Krishnarajah, "AGE-DEPENDENT MIXING PATTERNS OF HAND, FOOT AND MOUTH DISEASE (HFMD) IN SARAWAK, MALAYSIA," *Far East J. Appl. Math.*, vol. 94, no. 6, p. 437, 2016.

[11] S. J. Chan, J. Labadin, and Y. Podin, "A dynamic SEIPR model for the spread of hand, foot and mouth disease in Sarawak," *J. Telecommun. Electron. Comput. Eng.*, vol. 9, no. 3–10, pp. 125–129, 2017.

[12] J. C. Blackwood and L. M. Childs, "An introduction to compartmental modeling for the budding infectious disease modeler," *Lett. Biomath.*, vol. 5, no. 1, pp. 195–221, 2018.

[13] D. of S. Malaysia, "Current population estimates, Malaysia, 2018-2019." Department of Statistics Malaysia Putrajaya, 2019.

[14] Y. Li, J. Zhang, and X. Zhang, "Modeling and preventive measures of hand, foot and mouth disease (HFMD) in China," *Int. J. Environ. Res. Public Health*, vol. 11, no. 3, pp. 3108–3117, 2014.

[15] F. M. Siam, M. Grinfeld, A. Bahar, H. A. Rahman, H. Ahmad, and F. Johar, "A mechanistic model of high dose irradiation damage," *Math. Comput. Simul.*, vol. 151, pp. 156–168, 2018.

[16] M. H. Nasir and F. M. Siam, "Simulation and sensitivity analysis on the parameter of non-targeted irradiation effects model," *J. Teknol.*, vol. 81, no. 1, 2019.

[17] D. C. Montgomery and G. C. Runger, *Applied statistics and probability for engineers*. John wiley & sons, 2010.