Med & Health Dec 2023; 18(2): 578-588 https://doi.org/10.17576/MH.2023.1802.18

ORIGINAL ARTICLE

Prediction of Influenza-Like Illness Incidence in Cheras, Malaysia based on Environmental Data using the Generalised Additive Model

KIRAN GB[1], AFIQAH SYAMIMI M[2], KAMARUL IM[2], MOHD RIZAL AM[1], MOHD ROHAIZAT H[1], AZIZUL RAHMAN MJ[1], UMI KALSOM SA[3], PETRICK P[4], RAHMAN J[5], MOHD SHAHRUL MN[6]

[1]Department of Public Health Medicine, [3]Department of Medical Microbiology and Immunology, [4]Department of General Medicine, Faculty of Medicine, Universiti Kebangsaan Malaysia, Jalan Ya'acob Latiff, Bandar Tun Razak, 56000 Cheras, Kuala Lumpur, Malaysia
[2]Department of Community Medicine, School of Medical Sciences, Universiti Sains Malaysia, Kubang Kerian, Kelantan, Malaysia
[5]UKM Medical Molecular Biology Institute (UMBI), Universiti Kebangsaan Malaysia, Cheras, Kuala Lumpur, Malaysia
[6]Department of Earth Sciences and Environment, Faculty of Science and Technology, Universiti Kebangsaan Malaysia, Bangi, Selangor, Malaysia

Received: 11 October 2023 / Accepted: 22 November 2023

ABSTRAK

"Influenza-like illness" (ILI) menyumbang kepada 500,000 kematian di seluruh dunia setiap tahun. Kajian ini menerangkan aliran epidemiologi ILI di Cheras, Kuala Lumpur, Malaysia dan mencari bukti saintifik mengenai aktiviti ILI dan faktor alam sekitar. Daerah Kesihatan Cheras mengumpul data pengawasan ILI dan stesen cuaca terdekat yang menyediakan data persekitaran mengenai suhu purata, kelembapan, kelajuan angin, curah hujan kumulatif, "particulate matter" (PM)10 dan PM2.5 tahap. Sebanyak 51,245 kes ILI dilaporkan pada 1 Disember 2021 hingga 30 April 2023. Perhubungan non-linear antara insiden ILI dan pembolehubah persekitaran telah dimodelkan menggunakan Model Tambahan Umum (GAM). Tempoh penundaan untuk suhu purata, kelembapan rata-rata, curah hujan kumulatif, kelajuan angin purat, tahap PM 10 dan tahap PM 2.5 ialah 2, 3, 3, 3, 0, dan 3 hari, masing-masing. Model dengan lag optimum lebih baik menggambarkan varians kes ILI (R^2 = 0.5, penyimpangan yang dijelaskan = 58%) daripada model tanpa pemilihan penundaan (R^2 = 0.5, penyimpang yang

Address for correspondence and reprint requests: Mohd Rohaizat Hassan / Mohd Rizal Abd Manaf. Department of Public Health Medicine, Faculty of Medicine, Universiti Kebangsaan Malaysia, Jalan Yaacob Latiff, Bandar Tun Razak, 56000 Cheras, Kuala Lumpur, Malaysia. Tel: +6012-6343303 / +6019-3208925 Email: rohaizat@ppukm.ukm.edu.my / mrizal@ppukm.ukm.edu.my

dijelaskan = 57.2%). Model Lag menunjukkan nilai p yang signifikan untuk PM10 tetapi tidak mempunyai keserasian yang penting antara variabel prediktor. Oleh itu, model akhir (R^2 = 0.5, penyimpangan yang dijelaskan = 59.2%) mempunyai k = 15. Hujan yang lebih tinggi, kelembapan relatif, suhu yang lebih sejuk, dan kelajuan angin yang berkurangan meningkatkan kes ILI. PM2.5 dan PM10 juga menyumbang kepada ILI.

Kata kunci: *Alam sekitar; influenza; model ramalan*

ABSTRACT

Influenza-like illness (ILI) accounts for 500,000 fatalities worldwide yearly. Environmental factors contribute to spread of respiratory infections. This study describes the epidemiological trends of ILI in Cheras, Kuala Lumpur, Malaysia and seeks scientific evidence on ILI activity and environmental factors. Cheras Health District collected ILI surveillance data and the nearest weather station which provided environmental data on mean temperature, humidity, wind speed, cumulative rainfall, particulate matter (PM)10 and PM2.5 levels. A total of 51,245 ILI cases were reported from 1st December 2021 to 30th April 2023. The non-linear relationship between ILI incidence and environmental variables was modelled using the Generalised Additive Model (GAM). The lag time for mean temperature, mean humidity, cumulative rainfall, mean wind speed, PM 10 level, and PM 2.5 level was 2, 3, 3, 3, 0, and 3 days, respectively. The model with the optimal lag better describes ILI case variance (R^2=0.5, explained deviance=58%) than the model without lag selection (R^2=0.5, explained deviance=57.2%). The Lag Model indicates a significant p-value for PM10 but no significant concurvity between predictor variables. Thus, the final model (R^2=0.5, explained deviance=59.2%) has k=15. Higher rainfall, relative humidity, colder temperature, and decreased wind speed increased ILI incidence. PM2.5 and PM10 also contributes to ILI.

Keywords: Environmental; influenza; prediction model

INTRODUCTION

Influenza-like illness (ILI) refers to a collection of symptoms that resemble those brought on by influenza virus, but may also be brought on by other respiratory viruses. In public health and epidemiology, "ILI" refers to a cluster of symptoms that resemble influenza but have not been validated by laboratory testing. Each year, the influenza virus is responsible for millions of respiratory illnesses and approximately 500,000 fatalities worldwide. In developed countries, the majority of influenza-related fatalities occur among those aged 65 and older (WHO 2020). In Malaysia, genotypes of the Influenza A virus have been identified, and influenza surveillance is conducted to track the virus's spread (Rahman et al. 2014). In 2009-2010, Malaysia also experienced an epidemic of influenza caused by the A (H1N1) virus with 2253 confirmed cases and 78 fatalities (Hashim et al. 2021). ILI has been an infectious disease that is being monitored since the early 2000 in Malaysia and even more vigilantly after the COVID-19 pandemic. This surveillance is done by the Communicable Disease Unit of the Kuala Lumpur Health Department according to the ILI Surveillance program of Ministry of Health Malaysia which follows the Influenza Surveillance program of WHO (WHO 2020).

Respiratory virus infections are mostly believed to spread through large respiratory droplets, small particle nuclei (aerosol) and also direct contact (CDC 2022; Cowling et al. 2013; Rahman 2013). There are factors that have been found to contribute to the spread of infection among humans which include environmental and non-environmental factors. Environmental factors that contribute to the increase in spread of the respiratory viral infections are rainfall, humidity, ground temperature, particulate matter (PM) and airflow (Pica & Bouvier 2012; Saha et al. 2016).Unlike European countries, Malaysia is a tropical country with almost no seasonal change. Influenza and other viral infections are seen all year round in Malaysia. There are times of the year with heavier rainfall which is from October till January (Meteorology Department of Malaysia 2022), which have shown a higher number of cases (Oong et al. 2015; Wong et al. 2020). This finding is also similar to a study done in New Delhi, India, whereby cases were higher during the monsoon season and higher humidity (Saha et al. 2016).

Modelling ILI cases using environmental data is crucial for a number of reasons. Identifying the environmental factors that have the most significant association with the incidence of ILI is the first step. A study conducted in Huludao, China, for instance, determined that low air temperature and low relative humidity played a significant role in the epidemic pattern of ILI cases (Bai et al. 2019). Secondly, modelling can aid in the development of effective prevention and control strategies by predicting the spread of ILI. Using surveillance, weather, and Twitter data, one study utilised a model based on long-term short-term memory to predict the spread of ILI (Athanasiou et al. 2023). However, there has been lack of study on ILI and environmental factors in Malaysia. Air quality models also use environmental data such as PM to determine relationship with air pollution and incidence of diseases (United States Environmental Protection Agency 2022). Lastly, modelling can help to establish the relationships between environmental variables and ILI cases, thereby shedding light on the disease's underlying mechanisms.

Regression analysis is a common statistical technique for determining the relationship between two variables, such as environmental factors and ILI cases (PennState College of Earth and Mineral Sciences 2022). Some studies using environmental data did not account non-linearity, which introduce bias. Generalised Additive Model (GAM) is used to model ILI cases based on environmental data which can provide valuable insights into the disease's epidemiology and enlighten public health strategies for preventing and controlling its spread. As regional climate patterns and seasonal influenza epidemics vary by latitude, the influenza control policy should be determined locally.

The objective of this study was to describe the epidemiological patterns of ILI in Cheras, Kuala Lumpur, Malaysia and model their relationship with the ILI incidences. Specifically, to study the relationship between environmental factors with the incidence of ILI and to predict the incidence of ILI cases using environmental factors at a national ILI sentinel surveillance site of Malaysia in the capital city with a warm temperate climate.

MATERIALS AND METHODS

This study involved data ILI census from the Communicable Disease Unit of the Kuala Lumpur Health Department which served a total of 2.1 million population (Department of Population Statistics 2021). Cheras (N 3°6′ and E 101°43′) is a district in Kuala Lumpur which is the capital of Malaysia. This ILI census included all ILI cases seen by clinicians from all primary health clinics under the Cheras Health District and these data were collected on a daily basis from 1st December 2021 till 30th April 2023. A total of 516 days of ILI surveillance data were collected, as a minimum of 365 days of data were needed to run a prediction analysis in an infectious disease due to possible seasonal variations (Chae et al. 2018). All patients who presented to government clinics in Cheras Health District were included in the ILI census. Patients who lived in Cheras but presented to other clinics outside of Cheras Health District were not included in the ILI census, thus were excluded from this study. Individual patient data in ILI census were anonymised with no personal identifying information. The case definition of ILI used for this study was patients having fever 38°C, cough, and onset within the last ten days (WHO 2020).

The environmental data were gathered from the Department of Environment Malaysia and they were collected during the same time period. The collected data were from Petaling Jaya Weather Station: the closest station to the Cheras population, which is 8 km of the population-weighted centre of the city. The environmental factors collected were daily mean temperature (°C), mean humidity (%), cumulative rainfall (mm), mean wind speed (m/s), PM10 level (µg/m3) and PM2.5 level (µg/m3). The ILI cases surveillance data and environmental variables were retrieved in Microsoft Excel format and was subsequently exported into SPSS and R software IDE (Posit Software, PBC formerly RStudio 2023). For continuous variables, descriptive statistics such as mean, standard deviation and median were used. Spearman correlation analysis was conducted to explore the relationships between environmental factors. To examine the association between the number of ILI cases and environmental factors, a GAM was employed.

Using the GAM function from the "mgcv" R package, the impact of the environmental variables on ILI cases were determined. The cubic smoothing function and the Poisson family were implemented. On the basis of the dispersion of newly estimated data around a y=x line (R^2), forward and backward sequential variable selection was used to build the model. The importance of the spline parameters was evaluated and incorporated into the model. The lag effects of weather on ILI cases were investigated for one, two and three days based on previous literature (Davis et al. 2012; Zhao et al. 2018). The most parsimonious model was chosen based on its R^2 score. We then explored to determine the optimal model with cubic regression splines for mean temperature, mean humidity, cumulative rainfall, mean wind speed, PM10 level and PM2.5 level.

RESULTS

A total of 51,245 ILI cases were reported from 1st December 2021 to 30th April 2023 in Cheras city. The average number of ILI cases was 99 (SD=40.43) cases per day, showing that there was a seasonality in ILI cases in Cheras. Most weather variables from Petaling Jaya weather station showed readings that were of tropical climate with little variability per day (Table 1). However, the daily cumulative rainfall variable had a high variability (SD=20.86), and this pattern corresponded to the monsoon season in west coast of Malaysia during the month October and November.

The Spearman correlation analysis (we did not assume bivariate normal distribution) was done to study the relationship between each environmental variable. The daily mean temperature had a strong negative monotonic relationship with daily mean relative humidity (-0.83) and moderately negative with daily cumulative rainfall (-0.5). The relationship between daily mean

TABLE 1: Distribution of ILI cases and environmental variables in Cheras during the study period

Variables	Mean (SD)	Minimum	Maximum	Percentiles		
				25th	50th	75th
Total ILI cases	99.31 (40.43)	16.00	356.00	73.00	98.00	1210
Daily mean temperature (°C)	28.05 (1.23)	24.30	31.60	27.30	28.00	28.90
Daily mean relative humidity (%)	76.09 (7.11)	56.00	98.00	71.80	75.80	80.40
Daily cumulative rainfall (mm)	12.06 (20.86)	0	187.00	0	1.20	16.45
Daily mean wind speed (m/s)	1.08 (0.24)	0	2.00	0.90	1.10	1.20
Daily PM10 (ug/m^3)	26.24 (10.06)	0	57.30	18.40	27.10	32.90
Daily PM2.5 (ug/m^3)	17.14 (9.58)	0.75	48.90	9.35	15.60	23.66

temperature and daily mean wind speed was moderate positive (0.43) and weak positive relationship with daily PM10 level (0.33) and daily PM2.5 level (0.33). There was a moderate positive relationship (0.56) between daily mean relative humidity and daily cumulative rainfall level (Table 2).

The lag effect was incorporated into GAMs for environmental variables, which permitted a more comprehensive comprehension of the relationship between variables over time, capturing temporal dynamics and enhancing predictive accuracy in situations where time dependencies or delayed effects are relevant. The correlation analysis between environmental variables and the ILI cases found that the optimal time was 3 days for daily mean relative humidity, daily cumulative rainfall, daily mean wind speed and daily PM2.5 level. It was found that the lag time for daily mean temperature was 2 days and for daily PM10 level was 0 day (Table 3).

Following this, univariate analysis was done and it was found that all environmental variables and their lag values were significant. GAM was done to model the relationship between the incidence of ILI (count) and the environmental factors (independent variables). A model was first done using Lag 0 values and then a second model was done using the identified optimal lag values. We observed that the model with the optimal lag better explained the variation of ILI cases (R^2=0.5, explained deviance=58%) as compared to the model with no lag selection (R^2=0.5, explained deviance=57.2%) (Table 4). Model checking of the Lag Model showed significant k-value of the PM10 variable but with no significant concurvity between all predictor variables. Therefore, the k value was adjusted to 15. The final model (R^2=0.5, explained deviance=59.2%) was depicted below:

Total ILI cases ~
s(mean temperature Lag 2) + s(mean humidity Lag 3) + s(cumulative rainfall Lag 3) + s(mean wind Lag 3) + s(PM10, k=15) + s(PM2.5 Lag3)

As the mean temperature fell below

TABLE 2: Spearman correlation coefficients between environmental variables in Cheras

	Daily mean temperature	**Daily mean relative humidity**	**Daily cumulative rainfall**	**Daily mean wind speed**	**Daily PM10**	**Daily PM2.5**
Daily mean temperature (°C)	1.00					
Daily mean relative humidity (%)	-0.83	1.00				
Daily cumulative rainfall (mm)	-0.50	0.56	1.00			
Daily mean wind speed (m/s)	0.43	-0.43	-0.20	1.00		
Daily PM10 (ug/m^3)	0.33	-0.30	-0.27	0.09	1.00	
Daily PM2.5 (ug/m^3)	0.33	-0.31	-0.07	0.05	0.21	1.00

TABLE 3: Lag effect selection for environmental variables

	Lag day 0	Lag day 1	Lag day 2	Lag day 3
Daily mean temperature (°C)	-0.131	-0.153	-0.164	-0.144
Daily mean relative humidity (%)	0.098	0.119	0.131	0.140
Daily cumulative rainfall (mm)	0.029	-0.021	-0.015	0.031
Daily mean wind speed (m/s)	-0.076	-0.084	-0.064	-0.086
Daily PM10 (ug/m^3)	-0.129	-0.101	-0.088	-0.125
Daily PM2.5 (ug/m^3)	-0.088	-0.142	-0.090	-0.157

TABLE 4: Summary of different GAM models

Model	Variable	Effective degree of freedom	Reference number of degrees of freedom	Chi-square	P-value	Adjusted R^2	Explained deviance (%)
Crude Model	Daily mean temperature (°C)	7.408	8.353	74.23	<0.001	0.485	57.2
	Daily mean relative humidity (%)	8.196	8.813	92.39	<0.001		
	Daily cumulative rainfall (mm)	7.407	8.315	56.16	<0.001		
	Daily mean wind speed (m/s)	7.948	8.653	108.16	<0.001		
	Daily PM10 (ug/m^3)	8.842	8.991	385.90	<0.001		
	Daily PM2.5 (ug/m^3)	8.398	8.895	112.10	<0.001		
Model with Lag	Daily mean temperature (Lag 2)	7.738	8.538	170.79	<0.001	0.512	58.0
	Daily mean relative humidity (Lag 3)	8.195	8.813	97.18	<0.001		
	Daily cumulative rainfall (Lag 3)	5.320	6.345	70.00	<0.001		
	Daily mean wind speed (Lag 3)	7.973	8.669	76.24	<0.001		
	Daily PM10 (Lag 0)	8.849	8.991	417.50	<0.001		
	Daily PM2.5 (Lag 3)	7.777	8.609	63.80	<0.001		
Model with Lag and PM10k (k=15)	Daily mean temperature (Lag 2)	7.802	8.577	152.98	<0.001	0.527	59.2
	Daily mean relative humidity (Lag 3)	7.811	8.617	72.44	<0.001		
	Daily cumulative rainfall (Lag 3)	5.484	6.519	68.61	<0.001		
	Daily mean wind speed (Lag 3)	7.786	8.551	69.89	<0.001		
	Daily PM10 (Lag 0)	13.315	13.903	512.39	<0.001		
	Daily PM2.5 (Lag 3)	7.615	8.509	50.29	<0.001		

25°C, the relative risk of ILI cases increased. No significant variation in the relative risk for ILI was seen within the 25°C to 30°C range. However, as the temperature rose further, the relative risk for ILI subsequently decreased. ILI had a non-linear relationship (edf = 7.81) with relative humidity. Most ILI cases were concentrated within the 65 to 85% humidity range with a slight peak around 75%. At the extreme ends of relative humidity which was below 60% and above 90%, the relative risk for ILI decreased. ILI also had a non-linear relationship with rainfall. The most incidence of ILI cases were concentrated below 50 mm with a slight peak around 25 mm. On the other hand, the estimated incidence of ILI increased with wind speed less than 0.6 m/s and also when the wind speed was more than 1.7 m/s. The estimated effect of PM10 level and PM2.5 level on the incidence of ILI cases was nonlinear with the peak of incidence was at 28 μg/m^3 (Figure 1).

DISCUSSION

Malaysia is a tropical country with almost no seasonal change, unlike European countries. In this study, influenza and other viral infections are seen all year round. There were periods of the year with heavier rainfall, which was in October till January and it coincided with higher number of ILI cases. This finding was similar to other studies which were done in Malaysia (Meteorology Department of Malaysia 2022; Oong et al. 2015; Wong et al. 2020). As described in the result section, the incidence of ILI cases was seen to increase with lower temperature especially below 25°C. Higher humidity and rainfall

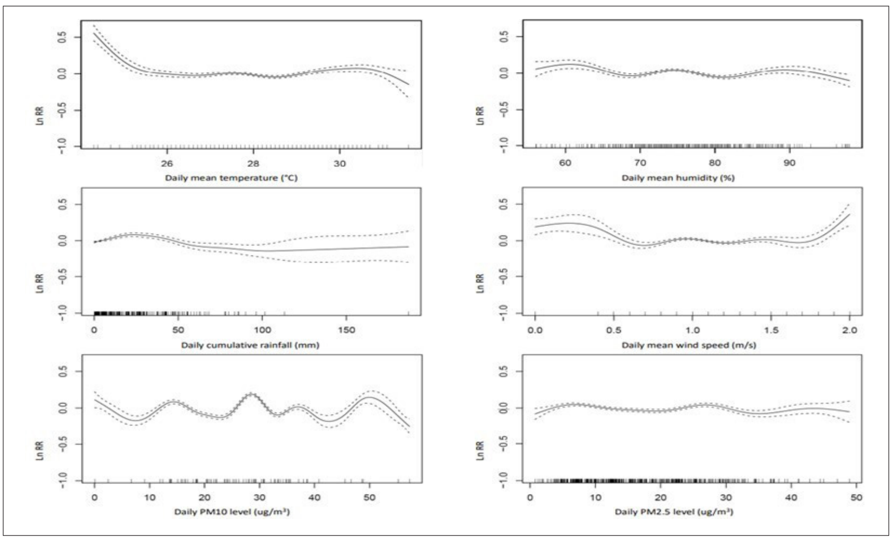

FIGURE 1: Relationship between ILI activity and environmental variables obtained from smooth functions in GAM *The solid line indicated the relative risk's natural logarithm of the environmental factor, and the two dotted lines represented the standard error

were showing an increase in number of ILI cases. These findings were also supported by studies that showed the risk of seasonal respiratory virus transmission was higher with higher amount of rainfall, higher relative humidity and colder temperature (Oong et al. 2015; Saha et al. 2016; Sam et al. 2015). This finding was also similar with other studies on Influenza A, adenovirus, RSV and rhinoviruses; whereby lower ground temperature and higher humidity, rainfall and rain days associated with higher respiratory infection rates (Paynter 2014; Shaman & Kohn 2009). The non-environmental factors were not explored in this study. There might be confounders that are indirectly affected by weather, whereby when it is too hot or too cold, people tend to stay indoors and it might be crowded areas which contributes to higher risk of exposure to an infection. When it is raining heavily, less patients will seek treatment at healthcare facilities leading to a wider standard error, possibly contributing to the almost constant number of cases when the rainfall is more than 70mm in this study. During a major influenza outbreak in China in 2017-2018, it was also seen that the number of ILI cases were negatively correlated with heavy rainfall (Zhu et al. 2023).

Airflow and ventilation also play a role in spread of respiratory virus infection. Airflowis the speed of air flowing through indoor and ventilation is the degree of mixing of air between indoor and outdoor. Rooms with lower air flow and poorer ventilation have a higher chance of retaining the virus in aerosol and thus, higher infectivity rate (Pica & Bouvier 2012; Shaman & Kohn 2009). On the other hand, the presence of PM in the air contributes to an increase in ILI cases, though not in a great degree. Another study was done in Poland showed decrease in PM 2.5 level reduced the transmission of Covid-19 and other respiratory viruses (Toczylowski et al. 2021). The possible explanation to PM contributing to incidence of ILI cases is by increasing virus survivability by lengthening the protective effect for the viruses that are trapped in aerosols or fomites (Paynter 2014; Shaman & Kohn 2009).

The model developed here had 59.2% explained deviance and this can be further improved by incorporating more predictors like non-environmental variables: patient's sociodemographic factors and vaccinations rates. The exploration of interactions between environmental and non-environmental variables will be able to improve the model's performance. This study used GAM to model ILI cases based on environmental data and it had several advantages as it provided greater flexibility in analysing environmental data. Because of the data was not normally distributed, it provided an effective fit by allowing specification of error pattern and captured nonlinear relationships. The GAM in this study also improved prediction of incidence of ILI cases based on environmental variables which was useful for effective prevention and controlling strategies. This study had a longitudinal study design which allowed the evaluation of ILI trends across different time of the year. The findings of this study were also consistent to other studies which provided a greater confidence on observed association between incidence of ILI and environmental factors.

The limitations were the findings of this study might differ slightly at countries with different climate, as Malaysia has tropical climate. Data used in this study was of 18 months duration and not more as the documentation of ILI census were affected during the COVID-19 pandemic. The environmental data was obtained from Petaling Jaya weather station, which is eight kilometres distance and the nearest station to Cheras's population weighted center. There were other variables that were not accounted for in this study, such as the patient's sociodemographic factor, vaccination rates and Cheras residents who sought treatment from primary health clinics outside of Cheras vicinity. This was purely a cross sectional study, thus, and no control group was studied to strengthen the argument for a potential association between ILI cases and environmental factors.

CONCLUSION

The incidence of ILI was found to be higher with higher amount of rainfall (slight peak around 25 mm), higher relative humidity (65 to 85% humidity range with a slight peak around 75%), colder temperature (below 25°C) and lower wind speed (less than 0.6 m/s). The presence of PM 2.5 and PM 10 also contributed to the incidence of ILI. The GAM models supported the non-linear relationship between incidence of ILI and the environmental variables such as mean temperature, mean humidity, cumulative rainfall, mean wind speed, PM 10 level and PM2.5 level. To control the number of ILI cases, targeted interventions can be done, such as vaccination campaigns for the high risk population namely children, elderly, patients with underlying chronic lung disease and immunocompromised individuals; hand hygiene promotions; respiratory hygiene etiquette; indoor air quality management by ensuring proper ventilation and air filtration systems; regular cleaning and disinfection of frequently touched area or objects in public space and public awareness campaigns which help in behavioural changes among population.

ACKNOWLEDGEMENT

The author would like to thank all healthcare providers of government health clinics of Cheras Health District for their involvement in data collection.

REFERENCES

Athanasiou, M., Fragkozidis, G., Zarkogianni, K. , Nikita, K.S. 2023. Long short-term memory–based prediction of the spread of influenza-like illness leveraging surveillance, weather, and twitter data: Model development and validation. *J Med Internet Res* **25**: e42519.

Bai, Y.L., Huang, D.S., Liu, J., Li, D.Q., Guan, P. 2019. Effect of meteorological factors on influenza-like illness from 2012 to 2015 in Huludao, a northeastern city in China. *PeerJ* **7**: 1-14.

CDC. 2022. Influenza CDC. https://www.cdc.gov/flu/about/disease/spread.htm [15 August 2022].

Chae, S., Kwon, S., Lee, D. 2018. Predicting infectious disease using deep learning and big data. *Int J Environ Res Public Health* **15**(8): 1596.

Cowling, B.J., Ip, D.K.M., Fang, V.J., Suntarattiwong, P., Olsen, S.J., Levy, J., Uyeki, T.M., Leung, G.M., Malik Peiris, J.S., Chotpitayasunondh, T., Nishiura, H., Mark Simmerman, J. 2013. Aerosol transmission is an important mode of influenza A virus spread. *Nat Commun* **4**: 1935.

Davis, R.E., Rossier, C.E., Enfield, K.B. 2012. The impact of weather on influenza and pneumonia mortality in New York City, 1975–2002: A retrospective study. *PLoS One* **7**(3): e34091.

Department of Population Statistics. 2021. Population census of Malaysia. https://www.dosm.gov.my/portal-main/landingv2 [25 June 2023].

Hashim, J.H., Adman, M.A., Hashim, Z., Mohd Radi, M.F., Kwan, S.C. 2021. COVID-19 Epidemic in Malaysia: Epidemic progression, challenges, and response. *Front Public Heal* **9**(May): 1-19.

Meteorology Department of Malaysia. 2022. Climate and season in Malaysia. https://www.met.gov.my/ [1 August 2022].

Oong, X.Y., Ng, K.T., Lam, T.T.Y., Pang, Y.K., Chan, K.G., Hanafi, N.S., Kamarulzaman, A., Tee, K.K. 2015. Epidemiological and evolutionary dynamics of influenza B viruses in Malaysia, 2012-2014. *PLoS One* **10**(8). e0136254.

Paynter, S. 2014. Humidity and respiratory virus transmission in tropical and temperate settings. *Epidemiol Infect* **143**(6): 1110-8.

PennState College of Earth and Mineral Sciences. 2022. Review of basic statistical analysis-methods for analysis. https://www.e-education.psu.edu/meteo469/node/209 [16 June 2023].

Pica, N., Bouvier, N.M. 2012. Environmental factors affecting the transmission of respiratory viruses. *Curr Opin Virol* **2**(1): 90-5.

Posit Software, PBC formerly RStudio, P. 2023. R Studio Software. https://posit.co/products/open-source/rstudio/ [12 June 2023].

Rahman, M.M., Wong, K.K., Isahak, I., Rashid, Z.Z., Alfizah, H. 2014. Identification of genotypes of Influenza A virus in Malaysia. *Pak J Med Sci* **30**(5): 1068-71.

Rahman, M.M., Wong, K.K., Hanafiah, A., Isahak, I. 2013. Influenza and respiratory syncytial viral infections in Malaysia: Demographic and Clinical perspective. *Pak J Med Sci* **30**(1): 161-5.

Saha, S., Chadha, M., Shu, Y. 2016. Divergent seasonal patterns of influenza types A and B across latitude gradient in Tropical Asia. *Influenza Other Respir Viruses* **10**(3): 176-84.

Sam, I.C., Su, Y.C.F., Chan, Y.F., Nor'E, S.S., Hassan, A., Jafar, F.L., Joseph, U., Halpin, R.A., Ghedin, E., Hooi, P.S., Fourment, M., Hassan, H., AbuBakar, S., Wentworth, D.E., Smith, G.J. 2015. Evolution of influenza B virus in Kuala Lumpur, Malaysia, between 1995 and 2008. *J Virol* **89**(18): 9689-92.

Shaman, J., Kohn, M. 2009. Absolute humidity modulates influenza survival, transmission, and seasonality. *Proc Natl Acad Sci* **106**(9): 3243-8.

Toczylowski, K., Wietlicka-Piszcz, M., Grabowska, M., Sulik, A. 2021. Cumulative effects of particulate matter pollution and meteorological variables on the risk of influenza-like illness. *Viruses* **13**(4): 1-14.

United States Environmental Protection Agency (EPA). 2022. Meteorology data and Processors. https://www.epa.gov/scram/meteorological-data-and-processors [16 June 2023].

WHO. 2020. World Life Expectancy. https://www.worldlifeexpectancy.com/country-health-profile/malaysia [1 August 2022].

Wong, P.L., Sii, H.L., P'ng, C.K., Ee, S.S., Yong Oong, X., Ng, K.T., Hanafi, N.S.,Tee, K.K., Tan, M.P. 2020. The effects of age on clinical characteristics, hospitalization and mortality of patients with influenza-related illness at a tertiary care centre in Malaysia. *Influenza Other Respir Viruses* **14**(3): 286-93.

Zhao, N., Cao, G., Vanos, J.K., Vecellio, D.J. 2018. The effects of synoptic weather on influenza infection incidences: A retrospective study utilizing digital disease surveillance. *Int J Biometeorol* **62**(1): 69-84.

Zhu, A.Q., Li, Z.J., Zhang, H.J. 2023. Spatial timing of circulating seasonal influenza A and B viruses in China from 2014 to 2018. *Sci Rep* **13**(1): 7149.