



OPEN

Association Between Seasonal Influenza and Absolute Humidity: Time-Series Analysis with Daily Surveillance Data in Japan

Keita Shimmei^{1,2✉}, Takahiro Nakamura², Chris Fook Sheng Ng³, Masahiro Hashizume⁴, Yoshitaka Murakami⁵, Aya Maruyama⁶, Takako Misaki⁶, Nobuhiko Okabe⁶ & Yuji Nishiwaki²

Seasonal influenza epidemics are associated with various meteorological factors. Recently absolute humidity (AH) has garnered attention, and some epidemiological studies show an association between AH and human influenza infection. However, they mainly analyzed weekly surveillance data, and daily data remains largely unexplored despite its potential benefits. In this study, we analyze daily influenza surveillance data using a distributed lag non-linear model to examine the association of AH with the number of influenza cases and the magnitude of the association. Additionally, we investigate how adjustment for seasonality and autocorrelation in the model affect results. All models used in the study showed a significant increase in the number of influenza cases as AH decreased, although the magnitude of the association differed substantially by model. Furthermore, we found that relative risk reached a peak at lag 10–14 with extremely low AH. To verify these findings, further analysis should be conducted using data from other locations.

Seasonal influenza epidemics are a serious public health concern, as they are associated with increased hospital admissions and an estimated 291,000–646,000 seasonal influenza-associated respiratory deaths annually worldwide¹. Influenza epidemics are also an economic problem—financial losses attributable to influenza in the United States amounted to around US \$90 billion in². It is therefore essential to identify factors that contribute to the development of influenza epidemics.

During the past few decades, studies have examined associations of influenza epidemics with meteorological factors^{3–11}. Recently, the effect of absolute humidity (AH; i.e. the density of water mass in air [g/m³]) on the survival of influenza virus has been gained attention. Evidence from laboratory experiments suggest that influenza virus survival is more closely associated with AH than with temperature or relative humidity^{12,13}. Additionally, a few epidemiological studies, which analyzed surveillance data, shows negative association of AH with influenza infection^{14–17}. However, these studies mainly investigated weekly surveillance data. Analysis with more granular data, such as daily data, has obvious possibility to bring accurate estimates in the association, but remain largely unexplored.

In such studies investigating associations of infectious diseases with meteorological factors, time-series analysis is widely used^{18–29}. This type of analysis must address two issues: control for autocorrelation (i.e. how to account for the effect of previously infected people) and control for seasonality. Although the approaches selected to address these problems greatly affect study findings, few studies have examined discrepancies in results obtained with differing methods.

In this study, we conducted time-series analysis of daily influenza surveillance data to examine the association of AH with the number of influenza cases and the strength of associations. In addition, this study examined how adjustment for seasonality and autocorrelation affected the results of the analysis.

¹The Poverty and Equity Global Practice, The World Bank, Washington, DC, US. ²Department of Environmental and Occupational Health, Toho University, Tokyo, Japan. ³School of Tropical Medicine and Global Health, Nagasaki University, Nagasaki, Japan. ⁴Department of Global Health Policy, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan. ⁵Department of Medical Statistics, Toho University, Tokyo, Japan. ⁶Kawasaki City Institute for Public Health, Kawasaki, Japan. ✉e-mail: shimmeikeita@gmail.com

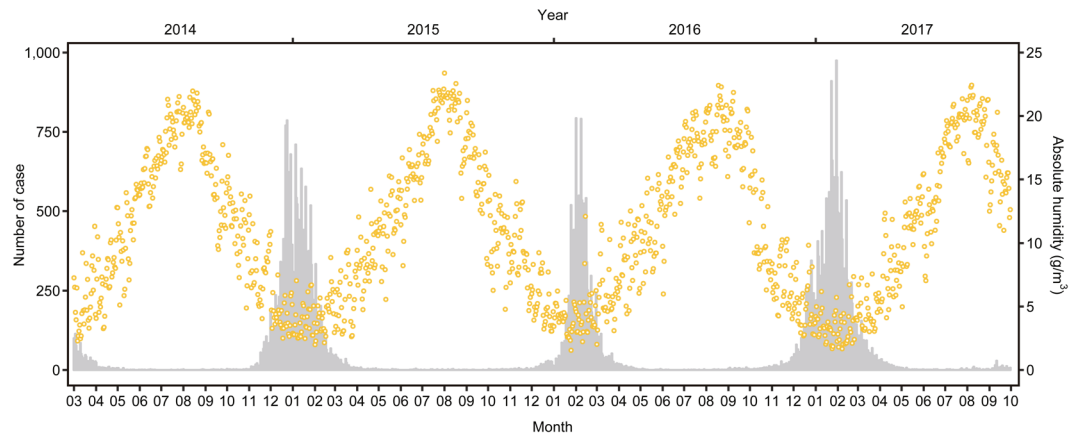


Figure 1. Epidemic curve (grey bar) and daily AH (yellow plot) from March 2014 to October 2017. $N = 1310$ days.

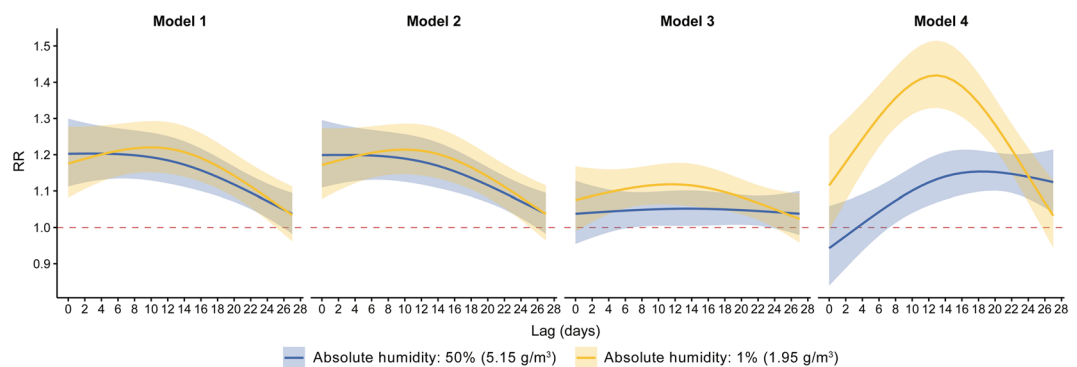


Figure 2. Lag-response relationship at lag 50th (blue) and 1st (yellow) percentile of AH by four models. Solid line and shaded area represent relative risk (RR) and its 95% confidence interval respectively. Horizontal dotted line in red show $RR = 1$ representing no difference in risk.

Results

To identify any trend in AH and influenza cases, we graphed the epidemic curve and plotted daily AH during the entire study period, from March 2014 through October 2017 (Fig. 1). Three obvious influenza epidemics were revealed during this period. The number of influenza cases usually started to increase in November or December, reached a peak in January or February, and decreased to almost zero by the end of May. Few cases were reported during May through September. The AH scatterplot shows periodic change: the greatest increase was in August and greatest decrease was in February. In the time-series analysis, we excluded data from May 1 through September 30 in each year and observed a total of 544 days in our sample. The mean (SD) number of influenza cases in the sample was 5.79 (2.64), and mean (SD) AH was 176.06 (126.26).

Figure 2 shows the RR-lag relationship for the 4 models at 50th and 1st percentiles of AH. RR (relative risk) refers to the multiple of influenza cases at a given AH percentile, as compared with the reference value—the 95th percentile of AH. In all models, a curve with a clear peak and inverted U-shape appears at the 1st percentile of AH. The RR approaches its maximum at lag 10–14 days. A comparison of the models showed that Models 1 and 2 had almost identical graphs and associations. At 1st percentile of AH, model 4 had the highest peak RR, 1.42 (95% CI, 1.33–1.51), while Model 3 had a moderate RR, 1.12 (95% CI, 1.06–1.18).

Figure 3 shows the exposure-response relationship for the 4 models at lag 12. In all models, RR increased as AH decreased. A comparison of the models showed, again, that Models 1 and 2 yielded almost identical associations. As was the case for the RR-lag relationship (Fig. 2), at lag 12 and a 1st percentile of AH, Model 4 had the highest RR and Model 3 had the lowest RR among four models.

Figure 4 shows the overall effects of AH on influenza cases, including all delayed effects up to lag 27. In all models, RR increased as AH decreased. A comparison of the models revealed 3 key findings. First, Models 1 and 2 yielded almost identical associations. Second, RRs increased logarithmically in Models 1 and 2 and exponentially in Models 3 and 4. Third, the magnitude of RR estimates at the 1st percentile of AH differed largely between Models 1/2, 3, and 4. The respective values were 74.2 (95% CI, 18.0–306.4), 10.9 (95% CI, 3.6–32.9), and 789.2 (95% CI, 198.1–3142.8) (See Supplementary Information).

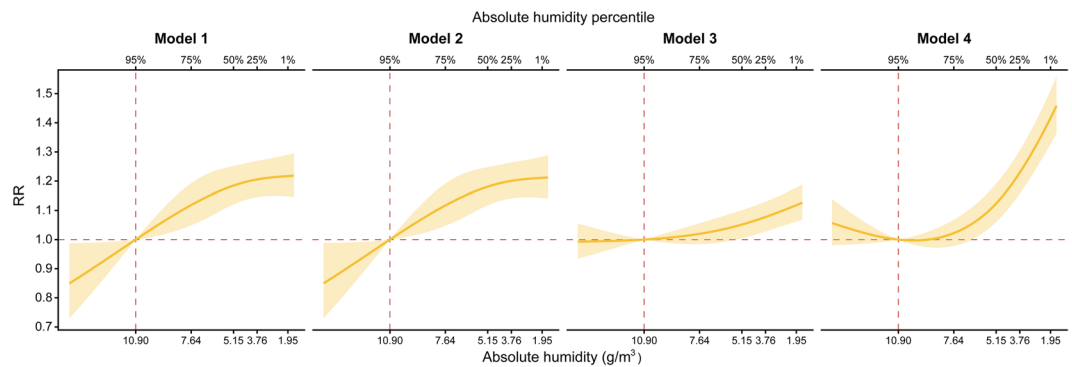


Figure 3. Exposure-response relationship at lag 12 by four models. Yellow solid line and shaded area represent relative risk (RR) and its 95% confidence interval respectively. Horizontal and vertical dotted lines in red show reference value of AH and RR = 1 representing no difference in risk.

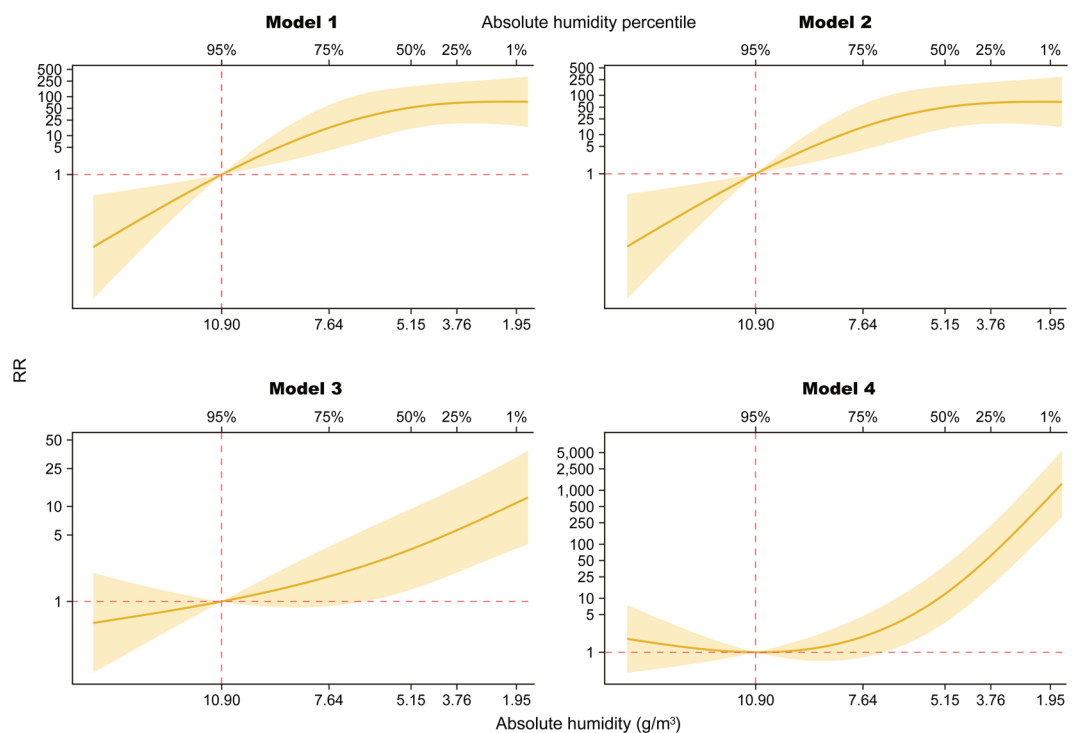


Figure 4. Overall cumulative exposure-response relationship by four models. Yellow solid line and shaded area represent relative risk (RR) and its 95% confidence interval respectively. Horizontal and vertical dotted lines in red show reference value of AH and RR = 1 representing no difference in risk. Scale size in vertical axis is transformed into common logarithm.

Discussion

This study analyzed daily influenza surveillance data with distributed lag non-linear model to investigate associations of AH with influenza and the strength of associations. To the author's knowledge, this is the first study to compare different adjustment for seasonality and autocorrelation in regression model. The results for all 4 models clearly show a significant increase in the number of influenza cases as AH decreased, although the magnitude of the association differed substantially by model. The lag between extremely low AH (1st percentile) and the greatest number of persons diagnosed at clinics was 10–14 days.

The present findings are consistent with those of previous laboratory studies^{12,13}, which showed a high correlation between influenza virus survival and AH. Past and present evidence suggests that increased survival of influenza virus at low AH increases virus transmission and the number of infected persons. In addition, our results are in line with previous epidemiological study¹⁴, which analyzed weekly surveillance data in Hong Kong, confirming the negative association of AH with influenza infection both in temperate climate and subtropical climate. However, comparison of the strength of present and past associations is difficult because the reference values used in prior studies using weekly data differed from those used in this study. In addition, no previous study

reported lag for peak RR at low AH, so direct comparison is not possible. Although the present results showing a lag of 10–14 days from the day of low AH to the day of diagnosis were consistent among models, a previous study estimated that the incubation period for influenza was 1.5–3 days^{30–32}. Future studies should attempt to resolve this discrepancy by analyzing daily data from other areas.

The objective of the article is to compare and discuss the results from the 4 different models rather than to choose one best model. First, the results of Models 1 and 2 were almost identical, which indicates that variation in the number of influenza cases can be adequately explained with all variables except for the term of controlling autocorrelation. The results of Models 3 and 4 differed; thus, the smoothing splines used to control for seasonality in Model 1 and 2, $ns(date)$, have largely accounted for the autocorrelation term. Second, the RR scale was substantially lower when the autocorrelation term (the logarithm of lagged outcome counts, $\log(Y_{t-1})$) was incorporated into the model (Model 3 v.s. 4). This suggests that a large proportion of cases occurring at time t was dependent on the size of the infected population, i.e. the number of cases in the recent past, and that inclusion of past cases caused a downward bias in the total effect of AH exposure, if we assume that part of the effect was attributable to a causal path involving the more immediate impact of that exposure, which caused higher counts and more infected persons at time $t - 1$ ³³. Third, the seasonality control affected the RR differently depending on the autocorrelation control. The RR decreased in the model with autocorrelation by changing seasonality control function from $ns(date)$ to $ns(epidemic\ year) + ns(day\ of\ epidemic\ year)$, i.e. from Model 1 to Model 3. On the other hand, the RR increased in the model without autocorrelation by the switch of seasonality control, i.e. from Model 2 to Model 4. Thus, future studies analyzing daily influenza data should recognize how different specifications for seasonality control yield different results.

Daily rather than weekly influenza data were analyzed, and this granularity of data is a strength of this study. Influenza surveillance data are frequently available only on a weekly basis; daily data are not easy to obtain. Although daily meteorological data are available, most previous studies analyzed weekly influenza data and thus converted daily meteorological data to weekly averages. Obviously, such conversion results in loss of information on small daily changes in meteorological factors. Thus, the availability of daily data increases granularity. Another strength of this study is its comparison of approaches to adjusting for seasonality and autocorrelation. We showed that varied approaches indicated presense of associations, although the strength of the associations varied greatly between models. These findings might assist in the development of models in future research.

Our study had several limitations. First, the misclassification of the outcome, influenza cases, might occur. Although the real-time surveillance system requests all city medical facilities to report influenza cases, reporting is not mandatory; therefore, data are missing for some influenza cases. If nonreporting of influenza cases is independent of AH, misclassification would be nondifferential and might lead to underestimation of the association. However, the nature and severity of any bias cannot be accurately determined in the current system, because we have no means to ascertain whether the absence of an observation of an influenza case was due to the lack of a diagnosis or to a reporting failure. Second, the potential measurement errors in the exposure, meteorological factors, might be another limitation. Data on relative humidity and temperature, which are used for calculating AH, were collected outdoors in the city. However, during winter - the influenza season - most people spend more time indoors than outdoors. Thus, the present AH values might not represent actual exposures. Nonetheless, a previous study reported that indoor AH was strongly correlated with outdoor AH ($\rho = 0.96$)³⁴. Thus, any bias resulting from nondifferential misclassification is likely to be limited. Third, the study results were based on the data from only one area, Kawasaki city. The models should be verified with a dataset from different area. Fourth, the lack of other meteorological factors in our regression model, which previous studies reported as factors associated with influenza infection, is another limitation. Temperature was excluded due to strong collinearity with AH ($\rho = 0.93$, $p = 0.01$) to avoid multicollinearity. Solar radiation was not included due to unavailability of data in Kawasaki city. Future studies should deal with this issue by incorporating these factors into the model.

Methods

This study was approved by the ethics committees of School of Medicine, Toho University (Approval No.A17022), located in Tokyo, Japan.

Study area and characteristics. Kawasaki City was selected as the study area. The city is located in north-east Kanagawa Prefecture, which is adjacent to Tokyo (Fig. 5). It is divided into 7 administrative districts and has an area of 144.35 km². The population was 1,503,690 in 2017, making it the seventh largest city in Japan. Kawasaki has a temperate climate with 4 distinct seasons; it is hot and humid in summer and cold and dry in winter.

The Kawasaki City Infectious Disease Surveillance System collects real-time data on influenza infections³⁵. Since March 2014, the system has required all city medical facilities to report the number of daily influenza diagnoses, by sex, age category, and area. Rapid diagnosis kits are used to differentiate influenza types A and B, and surveillance data are available from a city-run website (See Data Availability Statement). For this study, data on the daily total numbers of influenza A cases were obtained.

Daily meteorological data were obtained from a website providing information on air quality in Kawasaki, as were meteorological observational data, such as hourly temperature and hourly relative humidity, in each of the 7 city districts. These data were used to calculate hourly AH in each district, according to the following formula³⁶:

$$AH(g/m^3) = \frac{VP \times 216.5}{273.15 + T} \quad (1)$$

where T is temperature and VP is vapor pressure, which is used to define relative humidity (RH) as

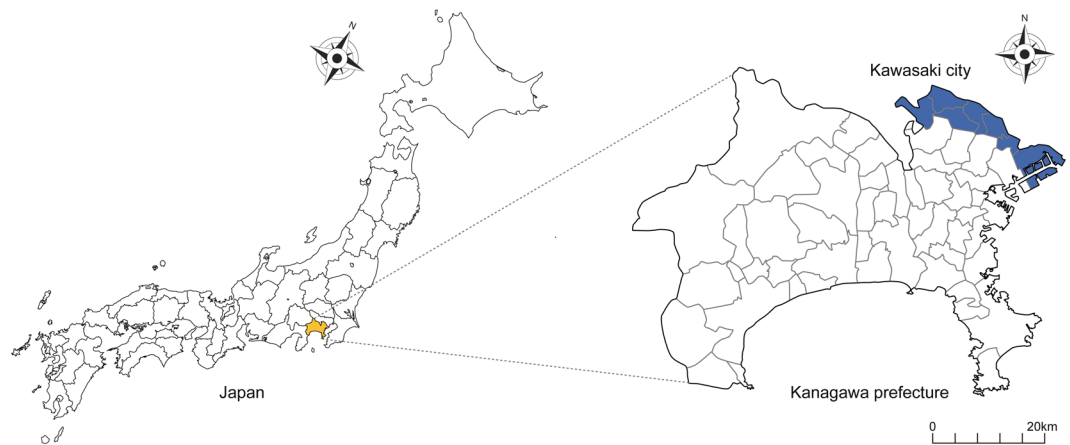


Figure 5. Map of Kanagawa prefecture and Kawasaki city.

$$VP(hPa) = 6.11 \times 10^{\frac{7.59 \times T}{T+240.72}} \times \frac{RH}{100}. \quad (2)$$

Average hourly AH was then calculated in all 7 districts, after which average daily values were obtained.

Statistical analysis. Time-series analysis was used to examine the association between daily AH and number of influenza cases. Quasi-Poisson regression was used to address overdispersion in count data^{22,33,37–39}. To investigate the nonlinear and delayed effects of AH on the daily number of influenza cases, we used a distributed lag nonlinear model⁴⁰, expressed as

$$\begin{aligned} Y_t &\sim \text{Poisson}(\mu_t) \\ \log(\mu_t) &= \alpha + cb(AH) + \beta \cdot DOW_t + \gamma \cdot \text{holiday}_t + \text{offset}(\log(\text{clinic})) \\ &\quad + f(\text{seasonality}) + g(\text{autocorrelation}) \end{aligned} \quad (3)$$

where subscript t is the day of observation, Y_t is the observed number of influenza cases on day t , and α is the intercept. The main exposure, AH, is transformed by function $cb(\cdot)$ —a cross-basis function to represent the 2-dimensional relationship between exposure and its lagged effects; natural cubic splines were used for exposure response and lag response. Two control variables were used: DOW_t is a vector for dummy variables representing the day of the week on day t , and holiday_t is a dummy variable that takes the value 1 if day t is a holiday. An offset term for the total number of medical facilities by day, $\text{offset}(\log(\text{clinic}))$, is added to adjust for daily variation in the number of medical facilities reporting influenza cases.

We investigated how adjustments for seasonality and autocorrelation might affect the estimation of associations. Although many previous studies used a single natural cubic spline of the date for seasonal adjustment, some recent studies have attempted to use decomposed seasonality terms, i.e. natural cubic splines of the epidemic year and day of epidemic year⁴¹. Some studies using autocorrelation adjustment reported that adding an autoregressive term increased goodness of fit, but others found that including such past cases in the model could cause a downward bias because past cases are an intermediate variable between past exposure and present cases^{42,43}. Thus, on the basis of these previous findings, we constructed 4 models with 2 patterns for each adjustment, represented by $f(\text{seasonality})$ and $g(\text{autocorrelation})$, as follows:

- Model 1: $f(\text{seasonality}) = ns(\text{date}, df = 4/\text{year})$, $g(\text{autocorrelation}) = \log(Y_{t-1})$
- Model 2: $f(\text{seasonality}) = ns(\text{date}, df = 4/\text{year})$, $g(\text{autocorrelation}) = 0$
- Model 3: $f(\text{seasonality}) = ns(\text{epidemic year}, df = 2) + ns(\text{day of epidemic year}, df = 2)$, $g(\text{autocorrelation}) = \log(Y_{t-1})$
- Model 4: $f(\text{seasonality}) = ns(\text{epidemic year}, df = 2) + ns(\text{day of epidemic year}, df = 2)$, $g(\text{autocorrelation}) = 0$

where epidemic year indicates the period from October 1 to April 30, and ns indicates natural cubic spline.

This analysis has other potential confounders. Temperature and solar radiation were reported as confounders in a previous study^{44–46}. However, because temperature was highly correlated with AH and solar radiation was not available from the website, we excluded them from our model.

Maximum lag was set to 28 days in the model. The degrees of freedom in the natural cubic spline for $cb(\cdot)$, $f(\cdot)$ and $g(\cdot)$ were determined by using Akaike's Information Criterion for quasi-Poisson (Q-AIC) values in a grid search. An excess number of outcome zeros could potentially bias estimated values. Therefore, we removed data from May 1 to September 30, during which there was no epidemic of type A Influenza. The period of the regression analysis was thus October 1, 2014 to April 30, 2017. To assess associations, relative risks (RRs) and their confidence intervals (CIs) were estimated, and the 95th percentile of AH in the above period was used as the reference value. The analysis was performed by using the *dlm* package in R, version 3.5.2.

Data availability

The surveillance data and meteorological data analysed during the current study are available in the website repository, <https://kidss.city.kawasaki.jp/en/modules/realsurveillance/> and <http://sc.city.kawasaki.jp/taiki/HOURLY/HT01I20190509.html>.

Received: 13 June 2019; Accepted: 20 March 2020;

Published online: 08 May 2020

References

- Iuliano, A. D. *et al.* Estimates of global seasonal influenza-associated respiratory mortality: a modelling study. *Lancet* **391**, 1285–1300, [https://doi.org/10.1016/S0140-6736\(17\)33293-2](https://doi.org/10.1016/S0140-6736(17)33293-2) (2018).
- Molinari, N.-A. M. *et al.* The annual impact of seasonal influenza in the US: Measuring disease burden and costs. *Vaccine* **25**, 5086–5096, <https://doi.org/10.1016/j.vaccine.2007.03.046> (2007).
- Loosli, C. G., Robertson, O. H. & Puck, T. T. The production of experimental influenza in mice by inhalation of atmospheres containing influenza virus dispersed as fine droplets. *The Journal of Infectious Disease* **72**, 142–153 (1943).
- Edward, D. G., Elford, W. J. & Laidlaw, P. P. Studies on air-borne virus infections: I. experimental technique and preliminary observations on influenza and infectious ectromelia. *The Journal of Hygiene* **43**, 1–10 (1943).
- Lester, W. The influence of relative humidity on the infectivity of air-borne influenza A virus, PR8 strain. *The Journal of experimental medicine* **88**, 361–8, <https://doi.org/10.1084/jem.88.3.361> (1948).
- Hemmes, J. H., Winkler, K. C. & Kool, S. M. Virus survival as a seasonal factor in influenza and poliomyelitis. *Nature* **188**, 430–431 (1960).
- Harper, G. Airborne micro-organisms: survival tests with four viruses. *The Journal of Hygiene* **59**, 479–486 (1961).
- Lowen, A. C., Mubareka, S., Steel, J. & Palese, P. Influenza virus transmission is dependent on relative humidity and temperature. *PLoS Pathogens* **3**, 1470–1476, <https://doi.org/10.1371/journal.ppat.0030151> (2007).
- Lowen, A. C., Steel, J., Mubareka, S. & Palese, P. High temperature (30C) blocks aerosol but not contact transmission of influenza virus. *Journal of Virology* **82**, 5650–5652, <https://doi.org/10.1128/JVI.00325-08> (2008).
- Park, J.-E. *et al.* Effects of temperature, humidity, and diurnal temperature range on influenza incidence in a temperate region. *Influenza Other Respir Viruses* **14**, 11–18, <https://doi.org/10.1111/irv.12682> (2019).
- Zhang, Y. *et al.* The complex associations of climate variability with seasonal influenza a and b virus transmission in subtropical shanghai, china. *Science of The Total Environment* **701**, <https://doi.org/10.1016/j.scitotenv.2019.134607> (2020).
- Shaman, J. & Kohn, M. Absolute humidity modulates influenza survival, transmission, and seasonality. *PNAS* **106**, 3243–3248 (2009).
- Grenfell, B. T., Lipsitch, M., Shaman, J. & Pitzer, V. E. Absolute humidity and the seasonal onset of influenza in the continental United States. *PLOS BIOLOGY* **8**, 1–13, <https://doi.org/10.1371/journal.pbio.1000316> (2010).
- Yapeng, L. X. Z. & Wang, X.-L. Impact of weather factors on influenza hospitalization across different age groups in subtropical hong kong. *International Journal of Biometeorology* **62**, 1615–1624, <https://doi.org/10.1007/s00484-018-1561-z> (2018).
- Chong, K. C. *et al.* Association between meteorological variations and activities of influenza a and b across different climate zones: a multi-region modelling analysis across the globe. *Journal of Infection* **80**, 84–98, <https://doi.org/10.1016/j.jinf.2019.09.013> (2019).
- Peci, A. *et al.* Effects of absolute humidity, relative humidity, temperature, and wind speed on influenza activity in toronto, ontario, canada. *Appl Environ Microbiol* **85**, e02426–18, <https://doi.org/10.1128/AEM.02426-18> (2019).
- Ka Chun, C. *et al.* Latitudes mediate the association between influenza activity and meteorological factors: A nationwide modelling analysis in 45 japanese prefectures from 2000 to 2018. *Science of The Total Environment* **703**, 134727, <https://doi.org/10.1016/j.scitotenv.2019.134727> (2020).
- Hashizume, M. *et al.* The effect of rainfall on the incidence of cholera in bangladesh. *Epidemiology* **19**, 103–110, <https://doi.org/10.1097/EDE.0b013e31815c09ea> (2008).
- Onozuka, D. & Hashizume, M. The influence of temperature and humidity on the incidence of hand, foot, and mouth disease in Japan. *Science of the Total Environment* **410–411**, 119–125, <https://doi.org/10.1016/j.scitotenv.2011.09.055> (2011).
- Eisenberg, M. C., Kujbida, G., Tuite, A. R., Fisman, D. N. & Tien, J. H. Examining rainfall and cholera dynamics in Haiti using statistical and dynamic modeling approaches. *Epidemics* **5**, 197–207, <https://doi.org/10.1016/j.epidem.2013.09.004> (2013).
- Zhao, X., Chen, F., Feng, Z., Li, X. & Zhou, X.-H. The temporal lagged association between meteorological factors and malaria in 30 counties in south-west China: a multilevel distributed lag non-linear analysis. *Malaria Journal* **13**, 1–12 (2014).
- Imai, C. & Hashizume, M. A systematic review of methodology: time series regression analysis for environmental factors and infectious diseases. *Tropical Medicine and Health* **43**, 1–9, <https://doi.org/10.2149/tmh.2014-21> (2015).
- Xu, M. *et al.* Non-linear association between exposure to ambient temperature and children's hand-foot-and-mouth disease in Beijing, China. *PLoS ONE* **10**, 1–15, <https://doi.org/10.1371/journal.pone.0126171> (2015).
- Xu, H. Y. *et al.* Statistical Modeling Reveals the Effect of Absolute Humidity on Dengue in Singapore. *PLoS Neglected Tropical Diseases* **8**, <https://doi.org/10.1371/journal.pntd.0002805> (2014).
- Lee, W.-h. *et al.* An investigation on attributes of ambient temperature and diurnal temperature range on mortality in Five East-Asian countries. *Scientific Reports* **7**, 1–9, <https://doi.org/10.1038/s41598-017-10433-8> (2017).
- Xiao, Y. *et al.* The influence of meteorological factors on tuberculosis incidence in Southwest China from 2006 to 2015. *Scientific Reports* **1–8**, <https://doi.org/10.1038/s41598-018-28426-6> (2018).
- Yu, X., Feric, Z., Cordero, J. F., Meeker, J. D. & Alshawabkeh, A. Potential influence of temperature and precipitation on preterm birth rate in Puerto Rico. *Scientific Reports* **8**, 1–9, <https://doi.org/10.1038/s41598-018-34179-z> (2018).
- He, F. *et al.* Construction and evaluation of two computational models for predicting the incidence of influenza in Nagasaki Prefecture. *Scientific Reports* **1–9**, <https://doi.org/10.1038/s41598-017-07475-3> (2017).
- Yin, F. *et al.* Analysis of the effect of PM10 on hand, foot and mouth disease in a basin terrain city. *Scientific Reports* **1–6**, <https://doi.org/10.1038/s41598-018-35814-5> (2019).
- Lessler, J. *et al.* Incubation periods of acute respiratory viral infections: a systematic review. *The Lancet Infectious Diseases* **9**, 291–300, [https://doi.org/10.1016/S1473-3099\(09\)70069-6](https://doi.org/10.1016/S1473-3099(09)70069-6) (2009).
- Reich, N. G., Lessler, J., Cummings, D. A. T. & Brookmeyer, R. Estimating incubation period distributions with coarse data. *Statistics in Medicine* **28**, 2769–2784, [10.1002/sim](https://doi.org/10.1002/sim) (2009).
- Nishiura, H. & Inaba, H. Estimation of the incubation period of influenza A (H1N1-2009) among imported cases: addressing censoring using outbreak data at the origin of importation. *Journal of Theoretical Biology* **272**, 123–130, <https://doi.org/10.1016/j.jtbi.2010.12.017> (2011).
- Imai, C., Armstrong, B., Chalabi, Z., Mangtani, P. & Hashizume, M. Time series regression model for infectious disease and weather. *Environmental Research* **142**, 319–327, <https://doi.org/10.1016/j.envres.2015.06.040> (2015).
- Nguyen, J. L., Schwartz, J. & Dockery, D. W. The relationship between indoor and outdoor temperature, apparent temperature, relative humidity, and absolute humidity. *Indoor Air* **24**, 103–112, <https://doi.org/10.1111/ina.12052> (2014).
- Kawasaki-City. Real-time Surveillance Website <https://kidss.city.kawasaki.jp/en/modules/realsurveillance/> (2020).
- Oyj, V. Calculation Formulas for Humidity - Humidity Conversion Formulas. Tech. Rep., Vaisala, Helsinki (2013).

37. Bhaskaran, K., Gasparrini, A., Hajat, S., Smeeth, L. & Armstrong, B. Time series regression studies in environmental epidemiology. *International Journal of Epidemiology* **42**, 1187–1195, <https://doi.org/10.1093/ije/dyt092> (2013).
38. Hedlund, C., Blomstedt, Y. & Schumann, B. Association of climatic factors with infectious diseases in the Arctic and subarctic region a systematic review. *Global Health Action* **7**, 1–16 (2014).
39. Wedderburn, R. W. M. Quasi-likelihood functions, generalized linear models, and the Gauss-Newton method. *Biometrika* **61**, 439–447 (1974).
40. Gasparrini, A., Armstrong, B. & Kenward, M. G. Distributed lag non-linear models. *Statistics in Medicine* **29**, 2224–2234, <https://doi.org/10.1002/sim.3940> (2010).
41. Gasparrini, A. *et al.* Mortality risk attributable to high and low ambient temperature: a multicountry observational study. *The Lancet* **386**, 369–375, [https://doi.org/10.1016/S0140-6736\(14\)62114-0](https://doi.org/10.1016/S0140-6736(14)62114-0) (2015).
42. Barnett, A. G., Stephen, D., Huang, C. & Wolke, M. Time series models of environmental exposures: Good predictions or good understanding. *Environmental Research* **154**, 222–225, <https://doi.org/10.1016/j.envres.2017.01.007> (2017).
43. Schisterman, E. F., Cole, S. R. & Platt, R. W. Overadjustment bias and unnecessary adjustment in epidemiologic studies. *Epidemiology* **20**, <https://doi.org/10.1097/EDE.0b013e3181a819a1> (2009).
44. Dowell, S. F. Seasonal variation in host susceptibility and cycles of certain infectious diseases. *Emerging Infectious Diseases* **7**, 369–374 (2001).
45. Yamshchikov, A. V. *et al.* Vitamin D for treatment and prevention of infectious disease: a systematic review of randomized controlled trials. *Endocr Pract* **15**, 438–449, <https://doi.org/10.4158/EP09101.ORR.VITAMIN> (2009).
46. Juzeniene, A. *et al.* The seasonality of pandemic and non-pandemic influenzas: the roles of solar radiation and vitamin D. *International Journal of Infectious Diseases* **14**, e1099–e1105, <https://doi.org/10.1016/j.ijid.2010.09.002> (2010).

Acknowledgements

This study was supported by the Research Complex Promotion Program; Create new values through technological and social innovation Wellbeing Research Campus “Tonomachi” and Grant-in-Aid JP17H07136 from the Ministry of Education, Culture, Sports, Science, and Technology, Japan.

Author contributions

K.S. performed the study, including the statistical analysis and writing the manuscript. M.H. and C.F.S.N. supervised the statistical analysis and helped writing and editing the manuscript. A.M., T.M. and N.O. collected the data. T.N. and Y.M. reviewed the manuscript. Y.N. supervised the projects. All authors read and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-020-63712-2>.

Correspondence and requests for materials should be addressed to K.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020