# Measuring the seasonality of human contact patterns and its implications for the spread of respiratory infectious diseases

Allisandra G. Kummer [a,1] , Juanjuan Zhang [b,1], Maria Litvinova [c], Alessandro Vespignani [d], Hongjie Yu [b,2*] , Marco Ajelli [a,2*]

a. Laboratory for Computational Epidemiology and Public Health, Department of Epidemiology and Biostatistics, Indiana University School of Public Health, Bloomington, IN, USA
b. School of Public Health, Fudan University, Key Laboratory of Public Health Safety, Ministry of Education, Shanghai, China
c. Department of Epidemiology and Biostatistics, Indiana University School of Public Health, Bloomington, IN, USA
d. Laboratory for the Modeling of Biological and Socio-technical Systems, Northeastern University, Boston, MA, USA

[1]Co-first author, contributed equally
[2]Co-senior author, contributed equally

*Corresponding authors: Hongjie Yu & Marco Ajelli, **Email:** yhj@fudan.edu.cn & marco.ajelli@gmail.com

**Keywords:** Seasonality; human contact patterns; mathematical modeling; respiratory pathogens

**Abstract**

Considerable uncertainties surround the seasonality of respiratory infectious diseases. To which extent the observed seasonality is associated with biological reasons (e.g., virus survival rates, host immune dynamics) or human behavior remains unclear. Here, we investigate the association between temperature and human contact patterns using data collected through a contact diary-based survey between December 24, 2017 and May 30, 2018 in Shanghai, China. We found a significant inverse relationship between number of contacts and temperature seasonal trend (p=0.003) and temperature daily variation (p=0.009), with contacts increasing from 19.6 (95%CI: 14.9-22.2) in December to 24.4 (95%CI: 19.0-28.0) in January and declining to 10.9 (95%CI: 10.1-11.9) in May. This seasonal trend in number of contacts translates into a seasonal trend in the basic reproduction number – mean number of secondary cases generated by a typical infector in a fully susceptible population. By setting the basic reproduction number at 1.4 on December 24, weekly mean estimates showed a clear increasing trend during the fall, beginning at 1.14 (95%CI: 0.78-1.39) in October and reaching 2.02 (95%CI: 1.60-1.35) in February and then remaining below 1 in the summer. Epidemic dynamics comparable with those of seasonal influenza are obtained through model simulations when the infection is seeded during the fall; however, their dynamics become more complex when seeded after February (e.g., double peaks or no epidemic until after the summer). Our findings indicate a distinct seasonal trend among human contact patterns and highlight a behavioral mechanism contributing to the seasonality of respiratory infectious diseases.

## Introduction

Several respiratory infectious diseases, including influenza, respiratory syncytial virus (RSV), measles, and coronavirus disease 2019 (COVID-19), show clear seasonal trends and cyclic epidemics [1-5]. Specifically, temperate regions experience highest incidence of these diseases during the winter seasons with fewer cases occurring during the summer months whereas other places with tropical climates, such as Singapore, may observe higher incidence of disease in warmer months or year-round [1, 5-8]. Possible explanations for the seasonality of respiratory infectious diseases include the variations in meteorological conditions (e.g., absolute humidity, temperature) that influence virus transmission, survival, and host susceptibility [7, 9, 10]. Previous studies found that transmission rates for influenza A increased when relative humidity was lower, suggesting that dynamics of relative humidity allow the airborne virus to remain viable for longer periods of time [9, 10]. Additionally, research on absolute humidity has indicated that influenza virus survival is associated with water vapor in the air, suggesting that absolute humidity might be better measure for seasonality because it is a measurement of water vapor regardless of temperature [9, 10]. Seasonal changes have also been linked to changes in host immune function, particularly decreases in mucosal integrity during dry seasons may increase susceptibility to infection [6, 8]. However, the underlying mechanisms that drive seasonality remain unclear.

A large amount of work has been done identifying human contact patterns as a key determinant for infectious disease transmission [7, 11-14]. For example, contacts in locations such as schools and workplaces tend to have the highest rates of transmission due to close contacts [7, 15-17]. While contacts made in schools and workplaces are relatively consistent, human behavior adapts to contextual changes due to working days, weekend days, holidays, and weather conditions [7]. Seasonal trends in human behavior could then be one of the drivers underlying the seasonality of disease transmission. School openings and closures during holiday breaks reduce the number of contacts among school-age children which may explain why influenza is reduced during holiday breaks [11, 12, 15, 18]. Furthermore, people tend to spend more time indoors when temperature drops [7, 19], increasing an individual's proximity to others thus impacting their likelihood of contracting an infectious disease [20, 21].

In this work, we investigate whether human contact patterns follow seasonal trends and to what extent these trends can contribute to shaping the seasonal trends observed for respiratory infectious diseases. To this aim, a regression analysis using contact survey data and meteorological data from Shanghai, China before the COVID-19-pandemic was conducted to assess the relationship between temperature seasonal trend and its daily variations with contact patterns. Results showed significant associations between contact patterns and both the temperature seasonal trend and daily variations. We then leveraged the obtained results to calibrate a mathematical model of the transmission of a respiratory pathogen. The performed simulations provide a mechanistic explanation of the observed influenza seasonality based on human contact patterns data.

## Results

**Contact Patterns.** Analyses were conducted for 965 participants using data collected through a diary-based contact survey [22] and daily maximum temperatures in Shanghai between December 24, 2017 and May 30, 2018 to investigate seasonality of human contact patterns. A total of 18,116 human contacts (mean = 18.7 per participant, Interquartile Range (IQR): 4.0 to 30.0) were analyzed (Table 1). Most weeks have an average total number of contacts similar to the overall mean with the exception of four weeks where the number of participants was less than 10 (Fig. 1*A*). Seasonality was represented through temperature seasonal trend and temperature daily variation. Seasonal trend was defined as a spline interpolating daily (maximum) temperatures, and temperature daily variation was defined as the difference between daily

(maximum) temperature and seasonal trend (Fig. 1*B*). Details are provided in the Methods section and the Supporting Information (SI) Appendix.

Age, gender, household size, occupation type, number of years lived in Shanghai, day of the week (i.e., weekday or weekend) and type of day (e.g., regular day, vacation) when the contact diary was completed were included in the analyses to adjust for confounding factors. Descriptive statistics for the covariates are shown in Table 1. The study sample included slightly more female (50.9%) than male (49.1%) participants. Most participants were between the ages of 19 and 59 (49.4%), were employed (41.5%), and had lived in Shanghai for more than 10 years or their entire life (89.3%). Adults 60 years and older reported 12.6 contacts on average (IQR: 4.0 to 16.0) whereas those 19-59 years had 21.4 contacts (IQR: 5.0 to 33.0) and participants 0-18 years old had an average of 20.5 contacts (IQR: 4.0 to 34.0). Employed persons reported a number of contacts similar to that of students (22.5 vs. 21.2 on average). Approximately, 73.1% of participants completed the diaries during a weekday (Monday-Friday) and had more contacts on average (mean: 20.3, IQR: 5.0 – 32.0) compared to the weekend days. Most participants (65.8%) completed their contact diaries on a regular day (i.e., a working day not during school vacations).

**Seasonality of contact patterns.** A negative binomial regression model with a log link function was used to estimate the effects of seasonality on human contact patterns while adjusting for the covariates. Estimated effects were considered statistically significant at $p \leq 0.05$. Details are reported in the Methods section and in the SI Appendix.

Seasonal and daily variations in temperature were associated with a significant decrease in the number of contacts. Specifically, for each degree increase in seasonal temperature trend, there is a decrease of 0.987 ($p = 0.003$, 95% CI: 0.978 to 0.996) expected contacts. Likewise, for each degree of daily temperature variation, expected contacts decrease by 0.981 ($p = 0.009$, 95% CI: 0.966 to 0.996) contact. As a result, we estimate an average of 19.6 (95%CI: 14.9-22.2) contacts per day in the month of December that increases to 24.4 (95%CI: 19.0-28.0) in January and declines to 10.9 (95%CI: 10.1-11.9) in May (Fig 2*A*). There is a decreasing trend in number of contacts as daily maximum temperature increases possibly indicating a movement from activities in proximity with others (i.e., indoor aggregations) to more distanced activities with lower number of contacts (i.e., outdoor activities).

**Estimated variation of the reproduction number.** We defined the potential reproduction number at time *t* as the number of secondary cases generated by a typical infector in a completely susceptible population at time *t*. For illustrative purpose, we set the potential reproduction number to be 1.4 (a typical value for seasonal influenza [23]) on December 24, 2017 (the first day of the survey) and estimate its value at any time point of the year. The estimated monthly mean of the potential reproduction number decreases from 1.89 (95%CI: 1.32 to 2.40) in January 2018 to 1.07 (95%CI: 0.81 to 1.33) in May 2018 (Fig. 2*B* inset). Then, the analyses were extended to months beyond the study period to include all seasons of the year (from October 1, 2017 to September 30, 2018). We estimate an increasing trend through the fall and winter seasons, which starts from a value of 1.14 (95%CI: 0.78-1.39) in October, followed by a gradual decrease in average estimated contacts through the spring season, reaching 0.96 (95%CI: 0.78 to 1.03) in June and remaining on average below the epidemic threshold throughout the summer (Fig. 2*B*).

The main analysis used maximum daily temperature as a measurement for temperature seasonal trend and daily variation because individuals are likely to engage in activities during the warmest parts of the day. To assess the robustness of our findings, we carried out two sensitivity analyses for additional meteorological measurements: average daily temperature and absolute humidity. Absolute humidity has been used in previous literature as a measurement for seasonality [9, 10]; therefore, it was included in this sensitivity analysis to assess consistency. Trends using average daily temperature reflected those in the main analysis with the estimated monthly mean potential reproduction number decreasing from 2.12 (95%CI: 1.66 to 2.44) in January to 0.93 (95%CI: 0.86 to 1.02) in May (Fig. S6*B*). Results using absolute humidity showed a similar pattern, although

with lower values, with a decreasing trend from January ($R_0$ = 1.91; 95%CI: 1.50 to 2.09) to May ($R_0$ = 0.61; 95%CI: 0.50 to 0.72) (Fig. S6*C*).

**Modeling influenza spread.** Among the different respiratory infectious diseases showing clear seasonal trends, we select influenza as an illustrative example. We calibrated a simple homogenous-mixing compartment model to obtain $R_0$=1.4 on December 24 [23] and a generation time of 2.8 days [24]. By simulating a set of epidemics, each one seeded on the first day of each month, we estimated seasonal variations in key epidemiological outcomes. Simulations account for the estimated variability in the number of contacts for each day. The model's details are reported in the Methods section.

Median infection attack rates were approximately 30% with the exception of epidemics seeded in February (12.8%; 95%CI: 2.8 to 24.8) and March (17.5%; 95%CI: 0.1 to 44.9), although seeding the epidemic in the spring lead to highly variable results (Fig. 3*A*). Overall, the median infection attack rates range from 10% to 30% which is consistent with previous estimates for influenza [25-27] . As shown in Figure 3*B*, median peak week incidence is highest for an epidemic beginning on October 1 (698 new infections per 10,000 individuals; 95%CI: 3738 to 1,101) and lowest for an epidemic beginning on February 1 (200 new infections per 10,000 individuals; 95%CI: 39 to 442). When we seed the epidemic on October 1 - a typical time for the start of influenza in the Northern hemisphere [25] – the simulations show a bell-shape curve with a likely peak in January-February (Fig. 3*C*). The daily net reproduction number increases from October until mid-December before gradually declining throughout the rest of the year (Fig. 3*D*). Results obtained by seeding the epidemic at different times of the year are reported in the SI Appendix (Fig. S9 & Fig. S10).

To assess variations in these estimates, sensitivity analyses were conducted using two additional $R_0$ values of 1.2 and 1.6. When $R_0$ = 1.2 and the epidemic is seeded in October, infection attack rates ranged from 0.0% to 8.6% with median peak week incidence being highest in November (121 new infections per 10,000 individuals; 95%CI: 32 to 811) and lowest in June (0 new infections per 10,000 individuals; 95%CI: 0 to 327) (Fig. S11*A*-*D*). When $R_0$ = 1.6, infection attack rates ranged from 20.1% to 46.8% with median peak week incidence being highest in November (1,227 new infections per 10,000 individuals; 95%CI: 1,026 to 1,668) and lowest in March (386 new infections per 10,000 individuals; 95%CI: 139 to 585) (Fig. S12*A*-*D*).

To evaluate differences between climates, the same approach and data were used to simulate epidemics in other locations assuming contact patterns being similar to those we estimated for Shanghai, but with changes in temperature. Specifically, Beijing and Guangzhou were chosen because of their climatic conditions being very different from Shanghai. Beijing has a humid, continental climate where summers are hot and humid and winters are cold [28] while Guangzhou has a humid, subtropical climate [29]. Compared to simulation results of Shanghai, both locations showed lower median infection attack rates where rates for Beijing were approximately 25% with the exception of January (16.1%; 95%CI: 3.1 to 39.5) and February (9.1%; 95%CI: 0.1 to 28.8) (Fig. S8*A*). Rates for Guangzhou were slightly lower at around 23% with the exception of January (11.9%; 95%CI: 2.8 to 51.2), February (8.4%; 95%CI: 0.08 to 40.4), and March (15.5%; 95%CI: 0.01 to 36.5) (Fig. S8*A*). Likewise, peak week incidence was lowest for Guangzhou remaining below 500 new infections per 10,000 individuals with slightly higher rates in Beijing and the highest rates in Shanghai (Fig. S8*B*). Interestingly, in Guangzhou (which lies on the Tropics of Cancer), we estimate the potential reproduction number to be above the epidemic threshold year-round (Fig. S7*C*).

## Discussion

Results of the current study demonstrated a strong association between weather and human contact patterns; specifically, these findings suggest that the number of contacts is dependent on the season and variations in daily temperature. Similarly, the influenza transmission model informed with the measured contact patterns showed a marked seasonal trend, further illustrating this relationship as a possible driver for seasonality of respiratory viruses. While there is no

distinction between indoor and outdoor contacts, these patterns indicate marked human behavioral changes driven by seasonal changes. From being in proximity with others during the cold weather season to fewer close contacts during the warm season, suggesting a shift from indoor to outdoor activities. This would likely be the case for contacts made in community locations outside of usual social environments such as homes, schools, and workplaces due to changes in gathering/meeting contexts. To assess these associations, the current study included an additional analysis of the community contacts which reflected the strong associations between seasonal trend and contact patterns, but not for daily temperature variations (Table S3).

In the main analysis, employed persons had significantly more contacts than those who were unemployed, and working days had significantly more contacts than weekend days, suggesting most contacts happen during work-related activities. Inverse associations between number of contacts and weekend days found in the main analysis are supported by claims in previous literature that people have more contacts during working days [7]. In contrast, there are significantly more community contacts on Sundays and Saturdays than during the weekend days, and unemployed persons have more community contacts than employed individuals (Table S3). Variations in weekend schedules and routines for unemployed people may allow for more opportunities to interact in the community. This may explain the significantly lower number of total contacts found for unemployed individuals.

This research has a number of limitations surrounding the nature of the contact survey and the modeling work. First, data for human contact patterns was taken from a single location; therefore, these findings cannot be generalized to areas outside of Shanghai. However, this is one of few studies examining the relationship between seasonality and human contact patterns and can be used to guide future research on this topic. This study also made general conclusions over an entire year despite having only 6 months of data which was addressed by using the regression model to predict estimates but will need to be checked using new data that covers an entire year. The survey was conducted during the pre-COVID era, in the absence of national mitigation interventions. Consequentially, it is unknown whether these findings are upheld during or after the COVID-19 pandemic; therefore, future studies should examine human contact data collected after 2019 to check whether these findings hold. Regarding the transmission model, the current study uses a simple, homogenous compartmental model to assess patterns of an epidemic consistent with influenza. We did not intend to provide a realistic model calibrated on a specific influenza season, considering school calendars, vaccination, age-specific transmission risks and susceptibility to infection.

Despite the above limitations, these results may be instrumental in explaining the seasonality of respiratory diseases such as influenza. Multiple simulations for influenza were compared using different attack rates illustrating consistent variations in incidence of disease over time based on differences in contact patterns and starting dates of the epidemic. When seeding an epidemic on October 1, there is one distinct peak of mean weekly incidence and gradual decreasing trend after the peak date. However, when considering an epidemic that begins on March 1, there is a small peak of infections in June with much larger peak in January-February (Fig. S9*F*). There is a slight increasing trend in the daily net reproduction number through March before declining below 1.0 between May and October followed by a gradual increase through the winter season (Fig. S10*F*). These patterns reflect those seen in the United Kingdom during the 2009 H1N1 pandemic [18].

While this study modeled simulations for influenza, these findings can be used to model any infectious disease that demonstrates a seasonal pattern in transmission such as RSV, measles, and COVID-19 [1-5]. Nonpharmacologic interventions aimed at limiting the number of contacts were able to reduce SARS-CoV-2 transmission [30-33], showing that seasonal trends in contact patterns could contribute to induce seasonality in COVID-19. In fact, despite the implementation of interventions contributed to mask seasonal trends in COVID-19, higher rates of SARS-CoV-2 infection and mortality are, nevertheless, recorded during colder months [20, 34, 35]. The mechanism proposed in this study could potentially improve our understanding of SARS-CoV-2

transmission patterns year-round and help elucidating possible seasonal trends in the post-pandemic phase.

**Methods**

The current study is based on data collected for 965 individuals in Shanghai, China who participated in a diary-based contact survey conducted from December 24, 2017 to May 30, 2018. Individual demographic and socioeconomic information of the study participants were collected along with the number of persons with whom they had contact during a 24-hour period before the interview, the date when contacts occurred, and details of each contact (i.e., relationship, location, duration, and type). Meteorological data for Shanghai was obtained for the Hongqiao International Airport Station using an online historical archive of weather reports [36]. Daily temperatures for the study period were matched to the dates when participants completed their contact diaries.

**Seasonal Trend and Daily Temperature Variation.** Seasonal trend was estimated by fitting a cubic smoothing spline to the maximum daily temperatures between December 24, 2017 and May 30, 2018. Daily temperature variation was defined as the difference between the maximum daily temperatures and the seasonal trend.

**Covariates.** In addition to the temperature seasonal trend and daily variation, several covariates were included in the analysis to adjust for characteristics and sources of potential influence on human contact patterns. These covariates are age, gender, household size, occupation type, live year, weekday, and type of day the diary was completed. Interview responses indicating a participant did not know or was unwilling to answer were recoded as missing.

Age groups were created for the descriptive analysis, stratifying participants into three age groups (0-18, 19-59, and 60+ years old). Age groups were chosen based on dominant social environments (i.e. school, work, and retirement). Gender categories included Male or Female. Occupation type was separated into three categories: Student, Employed, Not Employed. Participants attending preschool were included into the Student category and the retired participants into the Not Employed category. For the regression analysis, age and household size were normalized by calculating their overall means for each category of the occupation type. These means were then subtracted from the age or household size of individuals in that category to determine their normalized age and their normalized household size.

The live year variable was defined as the number of years a participant had resided in Shanghai. Categories for live year included <6years, 6-10 years, or >10 years/entire life. Participants who were younger than the number of years they had resided in Shanghai were recategorized with a live year of >10 years/entire life. The weekday variable included three levels: Monday-Friday, Saturday, and Sunday. Type of day was defined as whether the diary was completed on a regular day, irregular day, or New Year holiday. The irregular day referred to days when a person had significant variations to their normal day schedules (e.g., day off or school holiday). New year holiday types of day included participants who reported completing the diaries on irregular days during the New Year holiday season, from January 26 to February 22, 2018.

A full description of the contact survey can be found in *Zhang et al.* [22] and contact diaries are openly available on Zenodo [37].

**Ethics statement**. Ethics approval was received from the institutional review board of the School of Public Health, Fudan University (IRB no. 2018-01-0659S). Informed consent was obtained from all subjects (from a parent or guardian if the participant was under 18 years of age).

**Regression analysis.** The effect of seasonal trend and daily temperature variation on total contacts was analyzed using negative binomial regression while adjusting for covariates normalized age, normalized household size, occupation type, gender, weekdays, live year, and type of day. Observations with a Cook's Distance greater than 20 times the mean value were

considered outliers and excluded from the analysis. Incidence rate ratios were calculated by exponentiating the coefficients and confidence interval from the regression results (Fig. S3). Diagnostics were performed to assess regression assumptions (Fig. S2).

Two sensitivity analyses were performed using alternative meteorological data in Shanghai. All methods remained the same except seasonal trend and daily variations were calculated using average daily temperatures or the daily maximum absolute humidity.

**Estimated Contacts.** A new dataset was created to include observations for each day between December 24, 2017 and May 30, 2018 (n = 158) using the variables included in the regression. Occupation types, live years, weekdays, types of day, and gender were included as dummy variables for each category. Normalized age and normalized household size were fixed to their mean values, dummy variables for Saturday and Sunday were assigned a value of 1 when corresponding to Saturday and Sunday on the calendar respectively, all other variables were set to their reference group (gender = male, occupation type = Employed, live years = >10 years/entire life), and seasonal trend and daily variation was assigned their true values for each day. Each observation was transformed into a matrix with 1,000 observations, each with the exact same values.

Confidence intervals from the regression output for seasonal trend and daily temperature variation were used to create 1,000 multinomial samples for variable coefficients. These sample values were used to replace the coefficients for seasonal trend and daily variation in the regression model, resulting in 1,000 new model predictions for each day. Daily total contacts were estimated by calculating the means and 95% CI for each day from the predicted values.

**Variation in the reproduction number.** The potential reproduction number for each day was calculated by multiplying the estimated number of contacts in that day by a scaling factor. The scaling factor was chosen so that the basic reproduction number corresponds to 1.4 on December 24, 2017. This process for estimating contacts and the potential reproduction number was repeated for every day from October 1, 2017 to September 30, 2018. Seasonal trends and daily temperature variations were calculated for the entire year.

**Infection transmission model.**

We use a homogenous-mixing transmission model that classified individuals into three compartments: susceptible (*S*), infectious (*I*), and removed (*R*). The following equations were used to simulate the transmission process:

$$\dot{S}(t) = -\beta c(t)\frac{I(t)}{N}S(t)$$
$$\dot{I}(t) = \beta c(t)\frac{I(t)}{N}S(t) - \gamma I(t)$$
$$\dot{R}(t) = \gamma I(t)$$

Where:

- N represents the total number of individuals in the population and it is set to 1,000,000;
- $\gamma$ represetns the recovery rate;
- *c(t)* represents the number of contacts at time *t*;
- $\beta$ represents the per-contact transmission rate.

For the simple SIR model, the recovery rate corresponds to the inverse of the generation time [38], which we set at 2.8 days, in agreement with influenza literature [24]. The contact rate for each day of the study period was derived from the analysis of the contact survey data. The transmission rate was set to obtain $R_0$=1.4 (1.2. and 1.6 were used as sensitivity analyses) on December 24, 2017. The uncertainty on the estimated daily number of contacts was included by

randomly selecting 50 time series of the number of contacts *c(t)* predicted by the regression model.

Infection attack rates were calculated by dividing the number of infected individuals at the end of each simulation by the total simulated population. Weekly incidence of new infections for each simulation was found by calculating the difference between the incidence at the beginning and end of each week and then dividing it by the total simulated population. The mean and 95% IQR across all simulations were calculated to obtain mean weekly incidence. Peak week incidence was defined as the maximum weekly incidence and calculated by obtaining the maximum weekly incidence for each simulation and then choosing the median values and 95% IQR. The daily net reproduction number was calculated as the reproduction number multiplied by the fraction of suscepitle population on that day. Infection attack rates, weekly incidence of new infections, peak week incidence, and daily net reproduction number were calculated for an epidemic beginning on the first day of each month.

In addition to Shanghai, daily maximum temperatures for two additional locations (Guangzhou and Beijing, China) were included to evaluate the differences between climatic conditions. The climate in Shanghai is considered humid, subtropical; however, the city experiences all four seasons with temperate to cold, damp winters [39]. While Guangzhou also has a humid, subtropical climate, the winters are dry and more temperate [29]. In contrast, Beijing's climate is classified as humid continental where the summers are hot and humid, and the winters are dry and cold, albeit brief [28]. Meteorological data was obtained from the Beijing Capital International Airport Station [40] and the Guangzhou Baiyun International Airport Station [41] from an online historical archive of weather reports. For these sensitivity analyses, the regression output from the main analysis was used to estimate the number of contacts and potential reproduction numbers based on the temperature time series in each location. Methods for calculating seasonal trend, estimating contacts, estimating the potential reproduction number, infection attack rates, and peak week incidence remained the same.

Analyses were performed using R software (version 4.1.0).

**Author Contributions:** M.L., H.Y., and M.A. designed research. A.G.K. and J.Z. analyzed data. A.G.K., J.Z., M.L., A.V., H.Y., and M.A. interpreted the results. A.G.K. and M.A. wrote the paper. J.Z., M.L., A.V., and H.Y. edited the paper.

**Competing Interest Statement:** H.Y. has received research funding from Sanofi Pasteur, GlaxoSmithKline, Yichang HEC Changjiang Pharmaceutical Company, and Shanghai Roche Pharmaceutical Company. M.A. has received research funding from Seqirus. None of this funding is related to this research. Other authors have declared that no competing interest exists.

**Acknowledgments:** The authors would like to thank Nicole Samay for her assistance in preparing the figures. H.Y. acknowledges funding from the Key Program of the National Natural Science Foundation of China (82130093). A.G.K., M.L., A.V., and M.A. acknowledge support from the Cooperative Agreement no. NU38OT000297 of the Council of State and Territorial Epidemiologists (CSTE). The findings and conclusions in this study are those of the authors and do not necessarily represent the official position of the funding agencies.

**References**


1. CDC. *How Flu Spreads*. 2018 August 27, 2018; Available from: https://www.cdc.gov/flu/about/disease/spread.htm.
2. CDC, *Respiratory Syncytial Virus Infection (RSV)*. 2020, Centers for Disease Control and Prevention.
3. CDC. *Basics of COVID-19*. 2021 November 4, 2020; Available from: https://www.cdc.gov/coronavirus/2019-ncov/your-health/about-covid-19/basics-covid-19.html.
4. CDC. *Transmission of Measles*. 2020 November 5, 2020; Available from: https://www.cdc.gov/measles/transmission.html.
5. Noah, N.D., *Cyclical patterns and predictability in infection.* Epidemiology & Infection, 1989. **102**(2): p. 175-190.
6. Grassly, N.C. and C. Fraser, *Seasonal infectious disease epidemiology.* Proceedings of the Royal Society B: Biological Sciences, 2006. **273**(1600): p. 2541-2550.
7. Willem, L., et al., *A nice day for an infection? Weather conditions and social contact patterns relevant to influenza transmission.* PloS one, 2012. **7**(11): p. e48695.
8. Fares, A., *Factors influencing the seasonal patterns of infectious diseases.* International journal of preventive medicine, 2013. **4**(2): p. 128.
9. Shaman, J., et al., *Absolute humidity and the seasonal onset of influenza in the continental United States.* PLoS biology, 2010. **8**(2): p. e1000316.
10. Shaman, J. and M. Kohn, *Absolute humidity modulates influenza survival, transmission, and seasonality.* Proceedings of the National Academy of Sciences, 2009. **106**(9): p. 3243-3248.
11. Wallinga, J., P. Teunis, and M. Kretzschmar, *Using data on social contacts to estimate age-specific transmission parameters for respiratory-spread infectious agents.* American journal of epidemiology, 2006. **164**(10): p. 936-944.
12. Mossong, J., et al., *Social contacts and mixing patterns relevant to the spread of infectious diseases.* PLoS medicine, 2008. **5**(3): p. e74.
13. Leung, K., et al., *Social contact patterns relevant to the spread of respiratory infectious diseases in Hong Kong.* Scientific reports, 2017. **7**(1): p. 1-12.
14. Buckee, C., A. Noor, and L. Sattenspiel, *Thinking clearly about social aspects of infectious disease transmission.* Nature, 2021. **595**(7866): p. 205-213.
15. Chao, D.L., M.E. Halloran, and I.M. Longini, *School opening dates predict pandemic influenza A (H1N1) outbreaks in the United States.* The Journal of infectious diseases, 2010. **202**(6): p. 877-880.
16. Ibuka, Y., et al., *Social contacts, vaccination decisions and influenza in Japan.* J Epidemiol Community Health, 2016. **70**(2): p. 162-167.
17. Volz, E.M., et al., *Effects of heterogeneous and clustered contact patterns on infectious disease dynamics.* PLoS computational biology, 2011. **7**(6): p. e1002042.
18. Merler, S., et al., *Determinants of the spatiotemporal dynamics of the 2009 H1N1 pandemic in Europe: implications for real-time modelling.* PLoS computational biology, 2011. **7**(9): p. e1002205.
19. McCurdy, T. and S.E. Graham, *Using human activity data in exposure models: analysis of discriminating factors.* Journal of Exposure Science & Environmental Epidemiology, 2003. **13**(4): p. 294-317.
20. Liu, X., et al., *The role of seasonality in the spread of COVID-19 pandemic.* Environmental research, 2021. **195**: p. 110874.
21. Orenes-Piñero, E., et al., *Evidences of SARS-CoV-2 virus air transmission indoors using several untouched surfaces: A pilot study.* Science of The Total Environment, 2021. **751**: p. 142317.
22. Zhang, J., et al., *Patterns of human social contact and contact with animals in Shanghai, China.* Scientific reports, 2019. **9**(1): p. 1-11.
23. Biggerstaff, M., et al., *Estimates of the reproduction number for seasonal, pandemic, and zoonotic influenza: a systematic review of the literature.* BMC infectious diseases, 2014. **14**(1): p. 1-20.

24. Vink, M.A., M.C.J. Bootsma, and J. Wallinga, *Serial intervals of respiratory infectious diseases: a systematic review and analysis.* American journal of epidemiology, 2014. **180**(9): p. 865-875.
25. Trentini, F., et al., *Characterizing the transmission patterns of seasonal influenza in Italy: lessons from the last decade.* BMC public health, 2022. **22**(1): p. 1-9.
26. Vinh, D.N., et al., *Age-seroprevalence curves for the multi-strain structure of influenza A virus.* Nature Communications, 2021. **12**(1): p. 6680.
27. Kwok, K.O., et al., *Relative incidence and individual-level severity of seasonal influenza A H3N2 compared with 2009 pandemic H1N1.* BMC infectious diseases, 2017. **17**(1): p. 1-12.
28. World Travel Guide, *Weather in Beijing*. 2019.
29. Gong-fu, Z. and P. Kuo, *Guangzhou*. Encyclopedia Britannica.
30. Zhang, J., et al., *The impact of relaxing interventions on human contact patterns and SARS-CoV-2 transmission in China.* Science Advances, 2021. **7**(19): p. eabe2584.
31. Zhang, J., et al., *Changes in contact patterns shape the dynamics of the COVID-19 outbreak in China.* Science, 2020. **368**(6498): p. 1481-1486.
32. Jarvis, C.I., et al., *Quantifying the impact of physical distance measures on the transmission of COVID-19 in the UK.* BMC medicine, 2020. **18**(1): p. 1-10.
33. Quaife, M., et al., *The impact of COVID-19 control measures on social contacts and transmission in Kenyan informal settlements.* BMC medicine, 2020. **18**(1): p. 1-11.
34. IHME Covid- forecasting team, *Modeling COVID-19 scenarios for the United States.* Nature medicine, 2020. **27**: p. 94-105.
35. Hoogeveen, M.J. and E.K. Hoogeveen, *Comparable seasonal pattern for COVID-19 and flu-like illnesses.* One Health, 2021. **13**: p. 100277.
36. Weather Underground, *Shanghai, People's Republic of China Weather History*.
37. Zhang, J., et al., *Data and code for changes in contact patterns shape the dynamics of the novel coronavirus disease 2019 outbreak in China*. 2020. Doi: 10.5281/zenodo.3775672
38. Liu, Q.-H., et al., *Measurability of the epidemic reproduction number in data-driven contact networks.* Proceedings of the National Academy of Sciences, 2018. **115**(50): p. 12680-12685.
39. World Travel Guide, *Weather in Shanghai*. 2019.
40. Weather Underground, *Beijing, People's Republic of China Weather History*. 2022.
41. Weather Underground, *Guangzhou, Guangdong, People's Republic of China Weather History*. 2022.

**Figures and Tables**

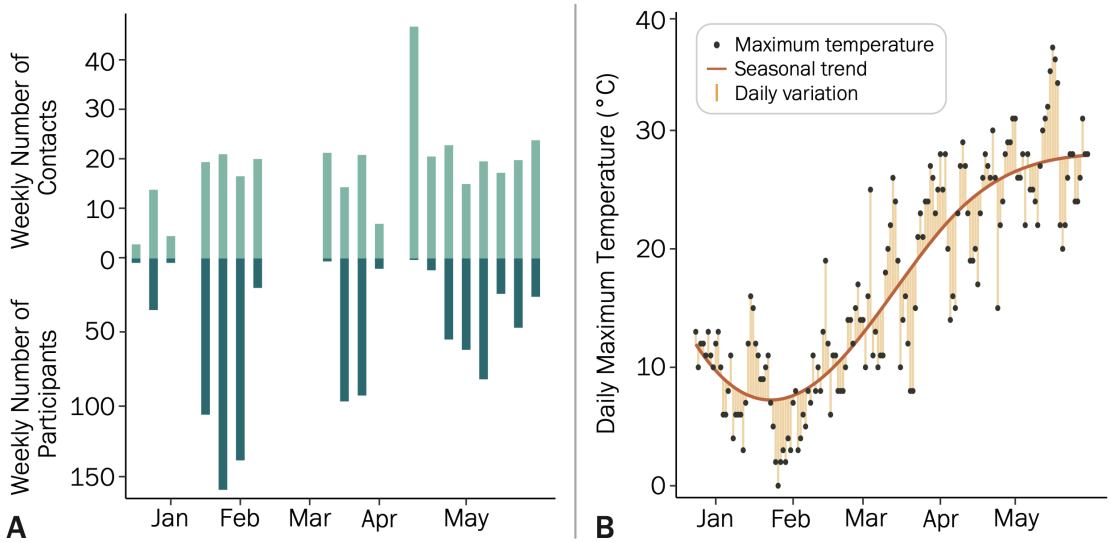


**Figure 1. A.** Average number of contacts and number of participants interviewed in Shanghai, China during each week from December 24, 2017 to May 30, 2018. **B.** Daily maximum temperature (°C) for each day during the study period, the seasonal trend of the temperatures, and the daily variation between the maximum temperature and the seasonal trend.

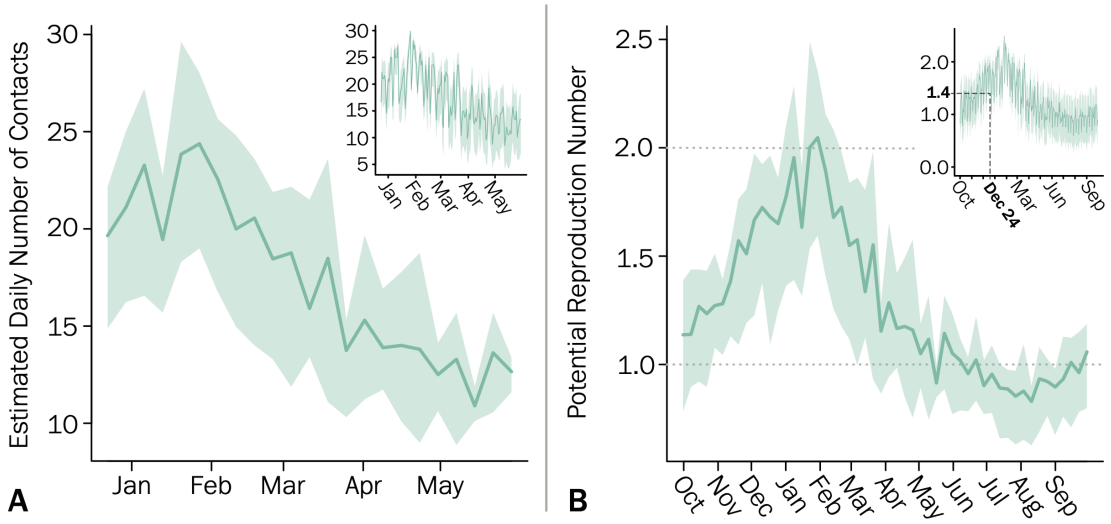


**Figure 2. A.** Estimated daily number of contacts for each week when seasonal trend, daily variation, and weekday vary while all other variables were fixed between December 24, 2017 to May 30, 2018. Inset shows estimated number of contacts for each day. Line and shaded area represent the mean and 95% CI of the daily values, respectively. **B.** Estimated potential reproduction number for each week extended over the whole year without consideration for differences in contact patterns during summer breaks. Inset shows the potential reproduction number for each day, which was set at 1.4 on December 24. Line and shaded area represent the mean and 95% CI of the mean daily values for each season, respectively.

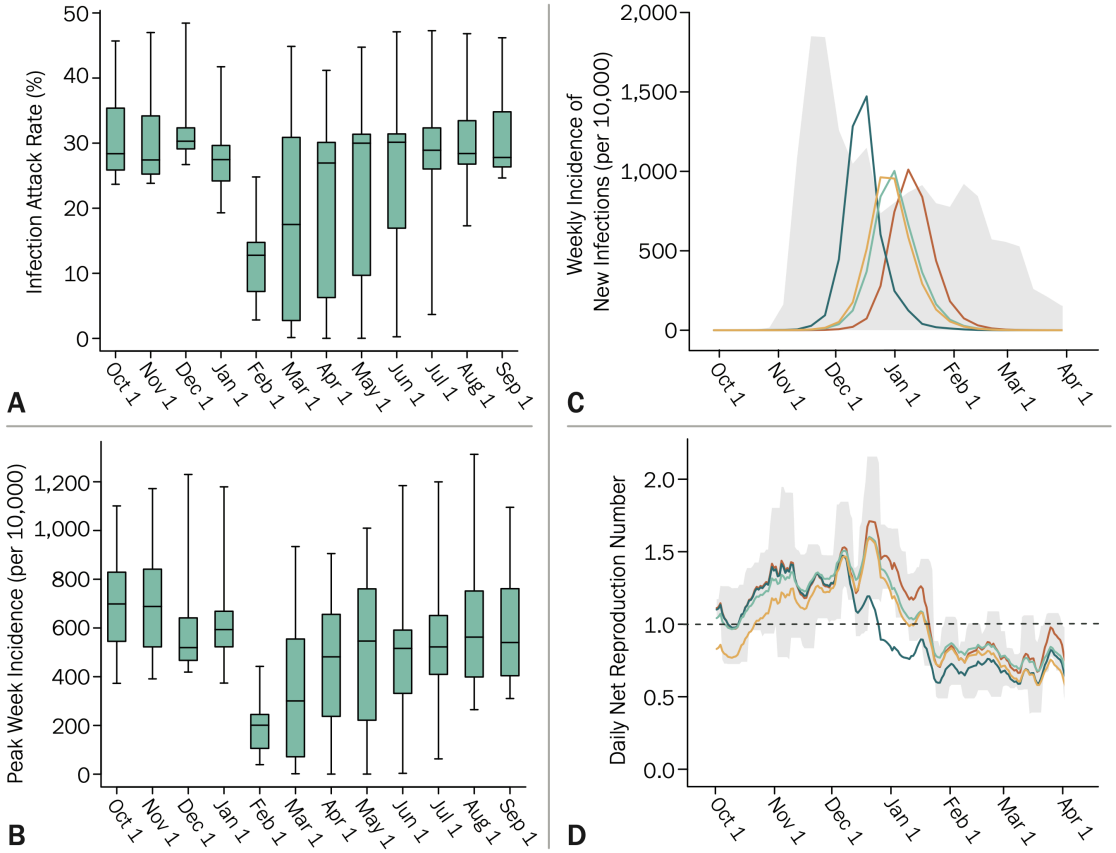


**Figure 3. A.** Infection attack rate for an epidemic beginning on October 1, 2017. The boxplot reports quantiles 0.025, 0.25, 0.5, 0.75, and 0.975 of the distribution. **B.** Peak week incidence for an epidemic beginning on the first day of each month from October 2017 to September 2018. The boxplot reports quantiles 0.025, 0.25, 0.5, 0.75, and 0.975 of the distribution. **C.** Weekly incidence of new infections per 10,000 people for an epidemic seeded on October 1, 2017. Lines represent four randomly selected simulations. The shaded area represents the 95% IQR of the distribution. **D.** Daily net reproduction number for an epidemic beginning on October 1, 2017. Lines represent four randomly selected simulations. The shaded area represents the 95% IQR of the distribution.

**Table 1.** Descriptive statistics for the total contacts.

| | | Number of Contacts | |
|---|---|---|---|
| | N (%) | Mean | IQR |
| **Total** | 965 (100.0) | 18.73 | (4.00, 30.00) |
| **Gender** | | | |
| Female | 491 (50.9) | 18.54 | (4.00, 29.50) |
| Male | 474 (49.1) | 19.01 | (4.25, 31.00) |
| **Age group** | | | |
| 0-18 | 221 (22.9) | 20.49 | (4.00, 34.00) |
| 19-59 | 477 (49.4) | 21.44 | (5.00, 33.00) |
| 60+ | 267 (27.7) | 12.58 | (4.00, 16.00) |
| **Occupation Type** | | | |
| Student | 252 (26.2) | 21.22 | (4.00, 34.00) |
| Employed | 400 (41.5) | 22.49 | (6.00, 34.00) |
| Not Employed | 307 (31.8) | 12.03 | (3.00, 15.00) |
| Missing | 6 (0.6) | 13.00 | (3.00, 9.00) |
| **Live year** | | | |
| < 6 years | 47 (4.9) | 17.68 | (4.00, 22.50) |
| 6-10 years | 52 (5.4) | 20.87 | (5.00, 34.50) |
| >10 years/entire life | 862 (89.3) | 18.73 | (4.00, 29.75) |
| Missing | 4 (0.4) | 13.75 | (3.75, 19.50) |
| **Type of Day** | | | |
| Regular day | 635 (65.8) | 18.72 | (4.00, 31.00) |
| New year holiday | 211 (21.9) | 17.72 | (5.00, 21.00) |
| Irregular day | 96 (9.9) | 22.38 | (6.75, 36.00) |
| Missing | 23 (2.4) | 14.91 | (4.00, 26.00) |
| **Weekday** | | | |
| Monday – Friday | 705 (73.1) | 20.30 | (5.00, 32.00) |
| Sunday | 143 (14.8) | 15.36 | (4.00, 20.5) |
| Saturday | 117 (12.1) | 13.75 | (4.00, 17.00) |

Note. IQR = Interquartile range.

**Table 2.** Negative binomial regression models of the effects of seasonal trend and daily temperature variation on total contacts adjusting for the covariates.

| Effect | Number of Contacts | | |
|---|---|---|---|
| | Estimate | Pr(>\|z\|) | 95% CI |
| **Seasonal trend** | -0.013 | 0.003** | (-0.022, -0.004) |
| **Daily temperature variation** | -0.019 | 0.009** | (-0.034, -0.004) |
| **Normalized age** | 0.005 | 0.114 | (-0.001, 0.011) |
| **Normalized household size** | 0.059 | 0.054 | (-0.002, 0.121) |
| **Occupation Type** | | | |
| Employed | ref | | |
| Not Employed | -0.680 | <0.0001*** | (-0.834, -0.525) |
| Students | -0.010 | 0.198 | (-0.253, 0.055) |
| **Live year** | | | |
| >10 years/entire life | ref | | |
| < 6 years | -0.177 | 0.224 | (-0.459, 0.123) |
| 6-10 years | 0.054 | 0.694 | (-0.209, 0.334) |
| **Weekday** | | | |
| Monday – Friday | ref | | |
| Saturday | -0.395 | <0.0001*** | (-0.584, -0.199) |
| Sunday | -0.202 | 0.024* | (-0.378, -0.021) |
| **Type of Day** | | | |
| Regular day | ref | | |
| Irregular day | 0.162 | 0.121 | (-0.039, 0.372) |
| New year holiday | -0.183 | 0.056 | (-0.372, 0.007) |
| **Gender** | | | |
| Male | ref | | |
| Female | 0.060 | 0.344 | (-0.067, 0.187) |

Note. Total $N$ = 965. d$f$ = 910. CI = Confidence Intervals.
*p < 0.05. ** p < 0.01. ***p < 0.001.