THE ROYAL SOCIETY PUBLISHING

A global assessment of the impact of school closure in reducing COVID-19 spread

Joseph T. Wu[1,2,†], Shujiang Mei[3,†], Sihui Luo[4,5,†], Kathy Leung[1,2,†], Di Liu[1,2], Qiuying Lv[3], Jian Liu[6], Yuan Li[3], Kiesha Prem[7], Mark Jit[1,2,7], Jianping Weng[4,5], Tiejian Feng[3,‡], Xueying Zheng[4,5,‡] and Gabriel M Leung[1,2,‡]

[1]WHO Collaborating Centre for Infectious Disease Epidemiology and Control, School of Public Health, LKS Faculty of Medicine, The University of Hong Kong, Hong Kong
[2]Laboratory of Data Discovery for Health (D[2]4H), Hong Kong Science Park, New Territories, Hong Kong
[3]Department of Communicable Diseases Control and Prevention, Shenzhen Center for Disease Control and Prevention, Shenzhen 518055, People's Republic of China
[4]The First Affiliated Hospital of USTC, Division of Life Science and Medicine, University of Science and Technology of China, Hefei, People's Republic of China
[5]Clinical Research Hospital (Hefei) of Chinese Academy of Science, Hefei, People's Republic of China
[6]Anqing Hospital Affiliated to Anhui Medical University (Anqing Municipal Hospital), Anqing, People's Republic of China
[7]Centre for Mathematical Modelling of Infectious Diseases, Department of Infectious Disease Epidemiology, London School of Hygiene and Tropical Medicine, London, UK

KL, 0000-0003-4777-388X

Prolonged school closure has been adopted worldwide to control COVID-19. Indeed, UN Educational, Scientific and Cultural Organization figures show that two-thirds of an academic year was lost on average worldwide due to COVID-19 school closures. Such pre-emptive implementation was

predicated on the premise that school children are a core group for COVID-19 transmission. Using surveillance data from the Chinese cities of Shenzhen and Anqing together, we inferred that compared with the elderly aged 60 and over, children aged 18 and under and adults aged 19–59 were 75% and 32% less susceptible to infection, respectively. Using transmission models parametrized with synthetic contact matrices for 177 jurisdictions around the world, we showed that the lower susceptibility of school children substantially limited the effectiveness of school closure in reducing COVID-19 transmissibility. Our results, together with recent findings that clinical severity of COVID-19 in children is lower, suggest that school closure may not be ideal as a sustained, primary intervention for controlling COVID-19.

This article is part of the theme issue 'Data science approach to infectious disease surveillance'.

1. Introduction

As the COVID-19 pandemic began to hit different locations during the first wave, many jurisdictions around the world had pre-emptively implemented school closure to suppress and mitigate its local spread [1,2]. The UN Educational, Scientific and Cultural Organization (UNESCO) estimated that during April, May and June 2020, around 1.6, 1.2 and 1.1 billion (90%, 68% and 62%) of enrolled learners around the world were affected by country-wide or localized school closure (https://en.unesco.org/covid19/educationresponse). Such swift and widespread adoption of prolonged school closure (which lasted weeks to months) was largely predicated on the premise that school children would be a core group for COVID-19 transmission as they are for influenza [2]. As countries are reviewing strategies for social and economic resumption, there is a need to evaluate the rationale and effectiveness of school closure against COVID-19 [3].

Reducing contacts made at school has been effective in controlling influenza because (i) children are biologically more or equally as susceptible to infection compared with adults; (ii) children have more contacts with other children at school than adults have with their peers; and (iii) children are less vigilant about personal infection control (e.g. hand hygiene) compared with adults [4,5]. Indeed, previous epidemiological studies in different populations have shown that the reproductive number of seasonal influenza and A/H1N1 pandemic influenza in 2009 (H1N1pdm09) was significantly reduced during school closure or holidays (table 1). While (ii) and (iii) are generally true, it is not clear whether (i) is valid for COVID-19. Although an early analysis of contact tracing and surveillance data from Shenzhen suggested that age had no effect on susceptibility to COVID-19 infection [13], more recent studies indicated that children might only be 35–50% as susceptible as adults [14–16].

On the other hand, there are considerable downsides to prolonged school closure, especially when recent seroprevalence studies suggest that there is not yet sufficient population immunity to avert subsequent waves of infection [17,18] until a safe and effective vaccine becomes widely available to the majority of the global population, and that could be a year or longer away. Governments have a duty of care to school-aged children, both for the infection risk of congregating in schools and also for their developmental and learning needs if schools are indefinitely suspended. School-aged children with special needs are particularly adversely affected. Of course, parental duties for childcare have been non-trivial and are a cause of widening inequity by socioeconomic position.

We aim to assess the potential impact of school closure in reducing the spread of COVID-19 in different parts of the world. To this end, we first inferred the effect of age on COVID-19 susceptibility and infectiousness using detailed surveillance and contact tracing data from Shenzhen and Anqing, China (the raw data from Shenzhen was the same as that in [13]) as well as published estimates from Guangzhou and Hunan, China [15,16]. We then parametrized our previous age-structured transmission model [19] with the inferred epidemiological parameters to simulate the epidemiological impact of school closure in an

Table 1. Comparison between our results and published estimates of reduction in reproductive number for H1N1pdm09 and seasonal influenza during school closure and holidays. In these comparisons, school contacts were assumed to be reduced by 60% during school closure and holidays.

|	published estimates	our estimates
H1N1pdm09 in the UK during summer holidays [6]	35% (30–40%)	37%
H1N1pdm09 in Hong Kong during closure of kindergartens and primary schools [7]	13% (10–15%)	16%
H1N1pdm09 in Hong Kong during summer holidays [7]	35%	37%
H1N1pdm09 in India during summer holidays [8]	14–27%	26%
H1N1pdm09 in China during school holidays [9]	37% (28–45%)	33%
seasonal influenza in France during summer holidays in 1985–2006 [10]	13–17%	17% if no age effect
		35% if H1N1pdm09-like
seasonal influenza B in Hong Kong during closure of kindergartens and primary schools in 2018 [11]	16% (10–26%)	7% if no age effect
		16% if H1N1pdm09-like
seasonal influenza in South Korea during school holidays in 2014–2016 [12]	6–23%	17% if no age effect
		38% if H1N1pdm09-like

otherwise unmitigated epidemic in 177 jurisdictions around the world for which contact matrices have been synthesized [20]. To gauge the relative effectiveness of school closure in comparison with other non-pharmaceutical interventions (NPIs), we estimated the reduction in infectious contacts made in non-school settings that would be needed in order to achieve the reduction in transmissibility or schoolchildren infection attack rate (IAR) conferred by school closure.

2. Methods

(a) Data on within-household transmission

We analysed within-household transmission of COVID-19 in Shenzhen and Anqing based on the COVID-19 line-lists and outbreak investigation records collated and compiled by the Shenzhen and Anhui Centers for Disease Control and Prevention. Details about these databases have been previously reported (figure 1a,b) [13,22]. For this study, we first identified cases who were almost certainly infected outside their household based on their travel and contact history. We then inferred the membership of their households based on outbreak investigation records. After excluding households with more than eight members (because they were rare and we were relatively uncertain whether these clusters were bona fide households), our dataset comprised a total of 182 household clusters (147 from Shenzhen reported between 1 January and 8 March 2020, and 35 from Anhui reported between 9 January and 20 February 2020; see electronic supplementary material, figures S1–S13) with 238 primary/co-primary cases and 599 household contacts.

A close contact was defined to be infected if any of his/her SARS-CoV-2 RT-PCR tests were positive during isolation (at home or designated facilities). Since 20 January 2020, in both Shenzhen and Anqing, individuals who fulfilled one of the following criteria were traced: (1) resided in or had travel history to Wuhan or Hubei in the past 14 days; (2) had exposure to wild animals in the past 14 days; (3) had influenza-like or COVID-19 like symptoms such as fever, dry cough, dyspnoea and diarrhoea and (4) had close contacts with confirmed or suspected COVID-19 patients within 14 days before and after their symptom onset. Individuals who fulfilled any of
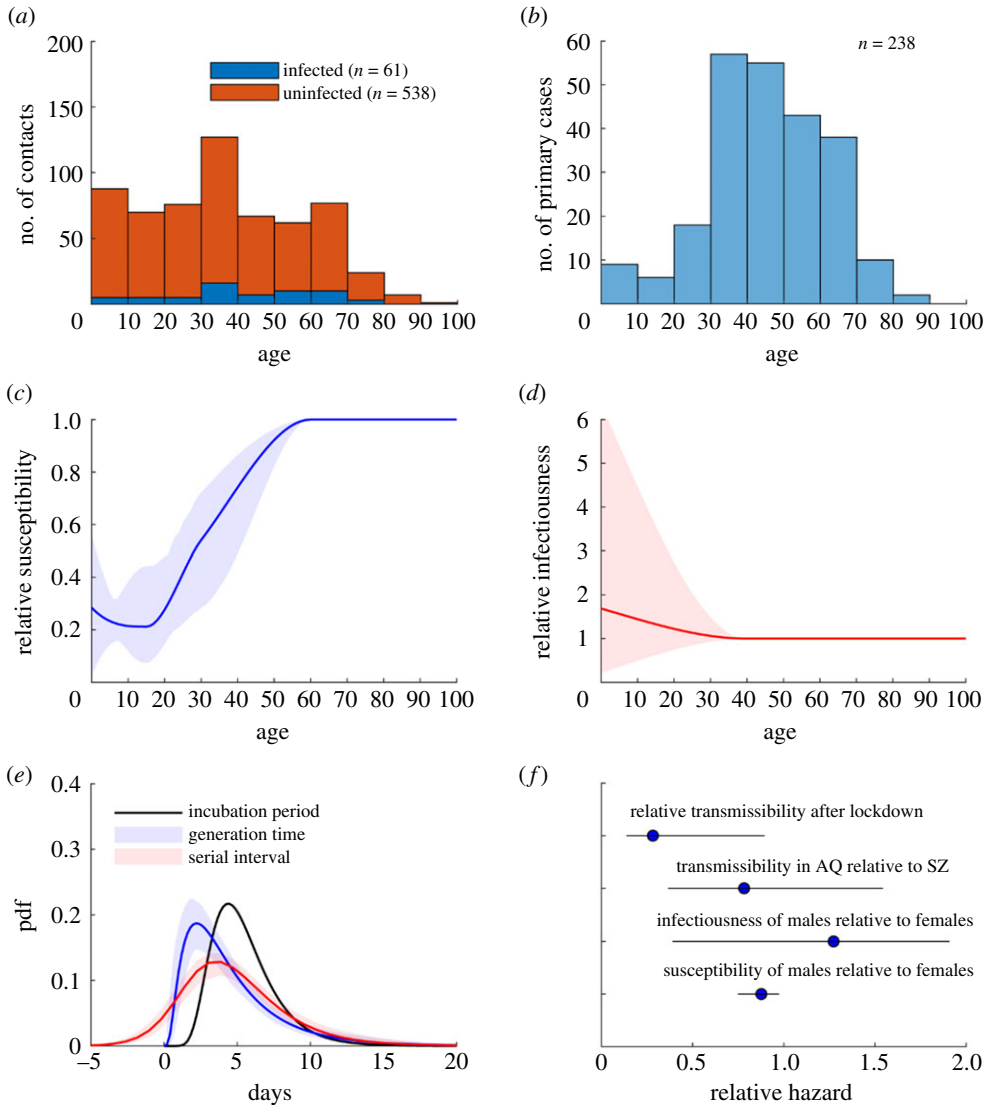
Figure 1. Within-household transmission dynamics of COVID-19 among 182 clusters in Shenzhen and Anqing. Key assumptions included the incubation period distribution was lognormal with median 5.1 days and mean 5.5 days [21]; susceptibility was the same for those aged 60 or above; infectiousness was the same for those aged 40 or above. (*a,b*) Age distribution of household contacts (stratified by infection outcome) and primary cases. (*c–e*) Age-specific susceptibility, age-specific infectiousness, generation time distribution and serial interval distribution. Lines and shades correspond to MAP estimates and 95% CrIs, respectively. (*f*) Susceptibility and infectiousness of males relative to females; transmissibility in Anqing relative to Shenzhen; and relative transmissibility after versus before Wuhan lockdown. Circles and bars correspond to MAP estimates and 95% CrIs, respectively. (Online version in colour.)

the above criteria were identified by the community or local CDC staff in person or by telephone and were placed under close epidemiological investigation. They were isolated at home (before 24 January 2020) or designated facilities (on or after 24 January 2020) for medical surveillance for 14 days and had a SARS-CoV-2 RT-PCR test upon onset of symptoms or every 3 days if there were no symptoms. Individuals were admitted to designated hospitals once their SARS-CoV-2 RT-PCR test was positive.

(b) Household transmission model

We developed a transmission model to simulate the spread of COVID-19 within households as follows. Given that during the study period there were only 417 and 83 cases of COVID-19 in Shenzhen and Anqing, respectively, and that contact tracing was very extensive, the probability of multiple importations within a household (i.e. some of the non-primary cases were infected via community transmission that were not identified by contact tracing) was likely to be negligible in our dataset. As such, we assumed that among household contacts, the risk of community transmission was negligible compared with the force of infection exerted by the primary/co-primary cases. Following previous household transmission modelling studies for H1N1pdm09, we assumed that interactions, case ascertainment and infection prevention measures were homogeneous among members within households (i.e. did not depend on age) [23,24]. We assumed that the probability of case ascertainment was 100% in the base case and 70% in the sensitivity analyses (electronic supplementary material, figure S14).

We inferred the following parameters by fitting the model to the data: (i) age-specific susceptibility to infection; (ii) age-specific infectiousness; (iii) relative susceptibility and infectiousness of males compared with females; (iv) baseline transmission rate for each household size; (iv) the generation time distribution; (v) the transmission rate in Anqing relative to Shenzhen; and (vi) changes in within-household transmission rate after nationwide lockdown measures were implemented. We used θ to denote the set of parameters subject to statistical inference.

(i) Infectiousness profile and generation time distribution

Suppose individual j was infected by individual i. We used $f_{inc}(\cdot)$ to denote the probability density function (pdf) of the incubation period and assumed that their incubation periods were independent and denoted them by X_i and X_j, respectively. Let Y be the time it took individual i to infect individual j after the former had become infected (i.e. the generation time for this pair). We assumed that the pdf of Y was dependent on X_i as follows:

$$P(Y=y \mid X_i=x) = f_{gen}(y \mid \theta, x) = \begin{cases} 0 & \text{if } y < \mu_{inf}x; \\ g(y - \mu_{inf}x \mid a_{inf}, b_{inf}) & \text{if } y \geq \mu_{inf}x, \end{cases}$$

where $g(\cdot \mid a_{inf}, b_{inf})$ was the gamma pdf with shape a_{inf} and scale b_{inf}, $\mu_{inf}x$ was the time at which individual i became infectious which was assumed to increase linearly with the incubation time x. We assumed that the time it took infectiousness to peak after the infected individual had become infectious was $\rho_{inf}x$ (such that $b_{inf} = \rho_{inf}x/(a_{inf}-1)$). The parameters $(\mu_{inf}, a_{inf}, \rho_{inf})$ were included in θ for statistical inference. The unconditional pdf of the generation time for individuals i and j was thus

$$\left. \begin{aligned} f_{gen}(y \mid \theta) = P(Y=y) &= \int_0^\infty P(Y=y \mid X_i=x) f_{inc}(x) \, dx \\ &= \int_0^\infty f_{gen}(y \mid \theta, x) f_{inc}(x) dx. \end{aligned} \right\}$$

The pdf of the serial interval for individuals i and j, namely $Z = Y + X_j - X_i$, was

$$\left. \begin{aligned} f_{SI}(z \mid \theta) = P(Z=z) &= P(Y + X_j - X_i = z) \\ &= \int_0^\infty P(Y + X_j = z + x \mid X_i = x) f_{inc}(x) \, dx, \end{aligned} \right\}$$

where

$$\begin{aligned} P(Y + X_j = z + x \mid X_i = x) &= \int_0^{z+x} P(X_j = z + x - u \mid X_i = x) f_{gen}(u \mid \theta, x) \, du \\ &= \int_0^{z+x} f_{inc}(z + x - u) f_{gen}(u \mid \theta, x) \, du. \end{aligned}$$

We used $F_{SI}(\cdot \,|\, \theta)$ to denote the pdf of the serial interval. In the base case, we assumed that the incubation period distribution was lognormal with median 5.1 days and mean 5.5 days (which corresponded to a coefficient of variation of 0.4) [21]. In the sensitivity analyses, we perturbed the mean incubation period by ±1 day while maintaining the same coefficient of variation (electronic supplementary material, figure S14).

(ii) Statistical inference of transmission parameters

The overall likelihood function was

$$L(\theta) = \left(\prod_k L_k(\theta) \right) L_{SI}(\theta),$$

where $L_k(\theta)$ was the likelihood function for cluster k and

$$L_{SI}(\theta) = \prod_{j \in A} f_{SI}(s_j \,|\, \theta) \prod_{j \in B} F_{SI}(s_j^U \,|\, \theta) - F_{SI}(s_j^L \,|\, \theta),$$

was the likelihood function for the serial interval data collected from outside Anqing and Shenzhen. Here, s_j was the serial interval data point for infector–infectee pair j, and s_j^L and s_j^U were the corresponding lower-bound and upper-bound if there was uncertainty in that data. A and B referred to the sets of serial interval data without and with uncertainty, respectively (see electronic supplementary material, table S2). The likelihood function $L_k(\theta)$ was formulated as follows.

We partitioned each cluster k into three sets of individuals:

— Set A_k comprised all the primary cases (i.e. those who were infected outside and introduced the infection into the cluster).
— Set B_k comprised all members who were ascertained to be infected during the cluster outbreak initiated by the primary cases in A_k.
— Set C_k comprised all other members in the cluster.

Let n_k be the size of cluster k, a_i and g_i be the age and sex of individual i, and T_i and O_i be the time of infection and time of symptoms onset if he/she was infected (i.e. $i \in A_k$ or $i \in B_k$), respectively. Our aim was to infer the following epidemiological characteristics of COVID-19 (the corresponding parameters that were included in θ for statistical inference are indicated; see below for their detailed formulation):

1. The effect of age on susceptibility to infection ($\alpha_{\text{sus}}(0)$, $\alpha_{\text{sus}}(15)$ and $\alpha_{\text{sus}}(30)$).
2. The effect of age on infectiousness ($\alpha_{\text{inf}}(0)$).
3. The effect of sex on susceptibility to infection ($\delta_{\text{sus}}(\text{male})$).
4. The effect of sex on infectiousness ($\delta_{\text{inf}}(\text{male})$).
5. Transmission rate for cluster with sizes $n = 2, \ldots, 8$ ($\beta(2), \ldots, \beta(8)$).
6. The relative within-household transmissibility of COVID-19 in clusters from Anqing compared to that from Shenzhen (ω_{AQ}).
7. The relative within-household transmissibility of COVID-19 after lockdown measures were implemented in Wuhan on Jan 23 followed by all major Chinese cities shortly after (σ).
8. The pdf of the generation time (μ_{inf}, a_{inf}, ρ_{inf}; see the previous subsection for the formulation of the generation time pdf).

Based on the travel and contact history of the infected cases, outbreak investigators were able to specify a range for their time of infection. We denoted this range by $[T_{\text{min},i}, T_{\text{max},i}]$ for individual

$i \in A_k$ or $i \in B_k$, i.e. $T_{\min,i} \le T_i \le T_{\max,i}$. The likelihood for cluster k conditioned on the time of infection of its primary cases (i.e. T_i, $i \in A_k$) was formulated as follows:

$$L_k(\theta|T_i, i \in A_k) = \prod_{m=1}^{3} L_{k,m}(\theta|T_i, i \in A_k),$$

where

1. $L_{k,1}(\theta|T_i, i \in A_k) = \prod\limits_{i \in A_k} f_{inc}(O_i - T_i)$ was the probability that individual $i \in A_k$ (i.e. the primary cases) developed symptoms at time O_i given that he/she was infected at time T_i.
2. Let $\Lambda_j(t)$ be the force of infection that the primary cases exerted on individual j at time t, and φ be the probability of case ascertainment among those infected. If there was only one non-primary case in cluster k, then $L_{k,2}(\theta|T_i, i \in A_k) = \varphi \int_{T_{\min,j}}^{T_{\max,j}} \Lambda_j(T_j) \exp\left(-\int_0^{T_j} \Lambda_j(t)\mathrm{d}t \int_0\right) f_{inc}(O_j - T_j)\mathrm{d}T_j$ was the probability that the non-primary case $j \in B_k$ (i) was infected at time $T_j \in [T_{\min,j}, T_{\max,j}]$, (ii) developed symptoms at time O_j and (iii) was ascertained. If there were more than one non-primary case in cluster k (i.e. some of them might be tertiary cases), then

$$L_{k,2}(\theta|T_i, i \in A_k) = 1 - \prod_{j \in B_k}\left[1 - \varphi\left(1 - \exp\left(-\int_{T_{min,j}}^{T_{max,j}} \Lambda_j(t)\,\mathrm{d}t\right)\right)\right]$$

was the probability that at least one (but not necessarily all) of the non-primary cases $j \in B_k$ was infected by the primary cases. The force of infection $\Lambda_j(t) = \sum_{i \in A_k} \lambda_{ij}(t)$ where $\lambda_{ij}(t)$ was the force of infection exerted by primary case $i \in A_k$ on individual j at time t. We modelled $\lambda_{ij}(t)$ as follows:

$$\lambda_{ij}(t) = \omega(k)L(t)\beta(n_k)\delta_{inf}(g_i)\delta_{sus}(g_j)\alpha(a_j)f_{gen}(t - T_i|\theta, O_i - T_i)Q_{ij}(t),$$

where:

— $\omega(k)$ was the relative transmission rate in Anqing compared to Shenzhen, i.e. $\omega(k) = \omega_{AQ}$ if cluster k was from Anqing and 1 if it was from Shenzhen.
— $L(t)$ was the relative transmission rate after Wuhan was locked down, i.e. $L(t) = \sigma$ if time t was after 23 January 2020 and 1 otherwise.
— $\beta(n)$ was the baseline transmission rate for clusters of size n.
— $\delta_{inf}(\cdot)$ was sex-specific infectiousness (assumed to be 1 for females).
— $\delta_{sus}(\cdot)$ was sex-specific susceptibility (assumed to be 1 for females).
— $\alpha_{inf}(\cdot)$ was age-specific infectiousness which was assumed to be (i) equal to 1 for those aged 40 years or above (this age group accounted for around half of all primary cases in our dataset; figure 1b); and (ii) a piece-wise cubic Hermite interpolating polynomial function (https://www.mathworks.com/help/matlab/ref/pchip.html) of age for those aged under 40 years where the value of $\alpha_{inf}(0)$ was included in θ for statistical inference. In the sensitivity analysis, we considered age 30 and 50 as alternative age cutoffs (electronic supplementary material, figure S14).
— $\alpha_{sus}(\cdot)$ was age-specific susceptibility which was assumed to be (i) equal to 1 for those aged 60 years or above and (ii) a piece-wise cubic Hermite interpolating polynomial function (https://www.mathworks.com/help/matlab/ref/pchip.html) of age for those aged under 60 years where the value of $\alpha_{sus}(\cdot)$ at age 0, 15 and 30 years were included in θ for statistical inference.
— $Q_{ij}(t) = 1$ if individual j was exposed to individual i at time t and 0 otherwise (see electronic supplementary material, figures S1–S13).

3. $L_{k,3}(\theta|T_i, i \in A_k) = \prod_{j \in C_k}\left[1 - \varphi\left(1 - \exp\left(-\int_0^{\infty} \Lambda_j(t)\,\mathrm{d}t\right)\right)\right]$ was the probability that all individuals in C_k had not been ascertained to be infected by the end of the cluster outbreak.

The overall likelihood for cluster k was

$$L_k(\theta) = \int_{\Omega_k} L_k(\theta | T_i, i \in A_k) \prod_{i \in A_k} h_{\text{inf},i}(T_i) \, dT_i,$$

where Ω_k was the Cartesian product of $[T_{\min,i}, T_{\max,i}]$ for all $i \in A_k$ and $h_{\text{inf},i}$ was the pdf of the time of infection for individual $i \in A_k$. Let T_{lockdown} denote the time when Wuhan was locked down (i.e. 23 January 2020). Given that (1) primary cases had history of travelling to or were residents of Hubei, or had contact history with suspected cases or returnees from Hubei; and (2) community transmission of COVID-19 in Shenzhen and Anqing was relatively low during the study time horizon, we assumed that the probability of $T_i = t$ was proportional to the epidemic size in Wuhan at time t if time $t < T_{\text{lockdown}}$ and zero otherwise. That is, $h_{\text{inf},i}(T_i) = e^{rT_i} / \int_{T_{\min,i}}^{T^*_{\max,i}} e^{r\tau} \, d\tau = re^{rT_i}/(e^{rT^*_{\max,i}} - e^{rT_{\min,i}})$ for $T_{\min,i} \leq t \leq T^*_{\max,i}$ where r was the epidemic growth rate in Wuhan (assumed to be 0.133 day^{-1} which corresponded to a doubling time of 5.2 days [19]) and $T^*_{\max,i} = \min(T_{\max,i}, T_{\text{lockdown}})$.

To integrate published estimates of age-specific and sex-specific susceptibility into our inference, we used the estimates from Jing et al. [15] and Zhang et al. [16] as priors for age- and sex-specific susceptibility (see electronic supplementary material, table S3 for details). We used flat non-informative priors (uniform distributions on $[0, \infty)$) for the remaining parameters because there were little published or publicly available evidence on their values. We estimated the model parameters θ using Markov Chain Monte Carlo methods with Gibbs sampling. Point estimates and statistical uncertainty were presented using the maximum posteriori probability (MAP) estimate and 95% credible intervals (CrIs), respectively.

(c) Age-structured transmission model

We partitioned the population of each jurisdiction into the following $m = 15$ age groups: 0–4, 5–12, 13–18, 19–23, 24–29, 30–34, 35–39, 40–44, 45–49, 50–54, 55–59, 60–64, 65–69, 70–74 and 75+. We assumed that age had no effect on infectiousness (which was consistent with the absence of strong evidence for age effect on COVID-19 infectiousness in figure 1d). Let α_i be the relative susceptibility of age group i. For COVID-19, we parametrized α_i's according to figure 1c. For H1N1pdm09, we parametrized α_i's according to a previous household study which indicated those aged 0–18 and ≥ 50 were 2 and 0.2 times as susceptible compared with those aged 19–50 [24].

We used the following SIR model to simulate COVID-19 and influenza epidemics where $S_i(t)$ and $R_i(t)$ were the number of susceptible and recovered individuals in age group i at time t, and $I(t, \tau)$ was the number of infected individuals in age group i at time t who were infected at time $t - \tau$:

$$\frac{dS_i(t)}{dt} = -\alpha_i S_i(t) \pi_i(t),$$

$$\frac{\partial I_i(t, \tau)}{\partial t} + \frac{\partial I_i(t, \tau)}{\partial \tau} = -f_{\text{gen}}(\tau) I_i(t, \tau),$$

$$I_i(t, 0) = \alpha_i S_i(t) \pi_i(t),$$

$$\frac{dR_i(t)}{dt} = \int_0^t f_{\text{gen}}(\tau) I_i(t, \tau) \, d\tau$$

and

$$\pi_i(t) = \sum_{j=1}^m \frac{\beta_{ij}}{N_j} \int_0^t I_j(t, \tau) \, d\tau,$$

where β_{ij} was the contact rate between age group i and j from Prem et al. [20] and $N_1, \ldots, N_m$ was the age distribution obtained from World Population Prospects 2019. (https://population.un.org/wpp/DataQuery/).

The next-generation matrix (NGM) for this SIR model was

$$T_G \begin{bmatrix} \alpha_1\beta_{11} & \cdots & \alpha_1\beta_{1m} \\ \vdots & \ddots & \vdots \\ \alpha_m\beta_{m1} & \cdots & \alpha_m\beta_{mm} \end{bmatrix},$$

where T_G was the mean generation time (figure 1e). The basic reproductive number R_0 was the largest eigenvalue of this matrix.

The synthetic contact matrices from Prem *et al.* [20] could be decomposed by the settings in which contacts were made: household, primary school or below, secondary school or above, workplace and community. Let y_{ijs} be the average number of contacts that an individual from age group i made with age group j in setting s in the synthetic contact matrices. The corresponding average number of contacts between age groups i and j in setting s was $N_i y_{ijs}$, and the proportion of contacts made in setting u compared to all contacts made was $\sum_{i \leq j} N_i y_{iju} / \sum_s \sum_{i \leq j} N_i y_{ijs}$.

In our model, if school closure was implemented, all contacts made at school were removed from the contact matrix (unless specified otherwise) and the resulting NGM was used to calculate the reproductive number. Similarly, when the frequency or transmissibility of contacts made in a given setting (e.g. community) was reduced by x%, the contact matrix of that setting was discounted by x% accordingly.

3. Results

(a) Household transmission dynamics

We estimated that children were substantially less susceptible to COVID-19 infection than adults and the elderly (figure 1c). Specifically, individuals aged 0–12 (primary school age and below), 13–23 (secondary school and higher education age) and 24–59 (adults) were 0.23 (0.15–0.33), 0.28 (0.13–0.45) and 0.68 (0.58–0.85) times as susceptible as those aged 60 or above, respectively. We found that children were less susceptible to COVID-19 infection primarily because of the lower attack rate observed among children in our dataset (figure 1a; electronic supplementary material, table S1). Given that only 15 primary cases were younger than 20 and they were co-primary cases with their parents or caregivers in our data (figure 1b), the posterior distribution of the relative infectiousness of children was not sufficiently precise to refute the null hypothesis that children and adults were equally infectious (figure 1d). The mean generation time and serial interval was 4.8 (4.0–5.9) days and 7% of serial intervals were negative (figure 1e). Infectiousness peaked when infected individuals were 44% (10–66%) through their incubation period. Consequently, 67% (56–78%) of infections occurred before the infectors developed symptoms. Compared with females, males were 0.86 (0.75–0.97) times as susceptible to infection (figure 1f). We estimated that males were 1.31 (0.41–1.92) times more infectious than females, i.e. weak evidence that the difference in infectiousness between males and females (if any) was significant (figure 1f). Within-household transmissibility was similar between Shenzhen and Anqing and dropped by 62% (11–86%) after lockdown measures were implemented in Wuhan on 23 January followed by all major Chinese cities shortly after (figure 1f). These inference results were insensitive to our assumptions regarding the incubation period (mean 5.5 days with a coefficient of variation 0.4 in the base case) except that the proportion of presymptomatic transmission would increase (decrease) by around 10% if the mean incubation period was decreased (increased) by 1 day (electronic supplementary material, figure S14). The results were also insensitive to assumptions regarding the probability of case ascertainment among household contacts (0.7–1) and age cutoff for age-specific infectiousness (electronic supplementary material, figure S14).

(b) School closure model

For COVID-19, we parametrized the transmission model with the posterior estimates of age-specific susceptibility from our within-household transmission inference and assumed that age
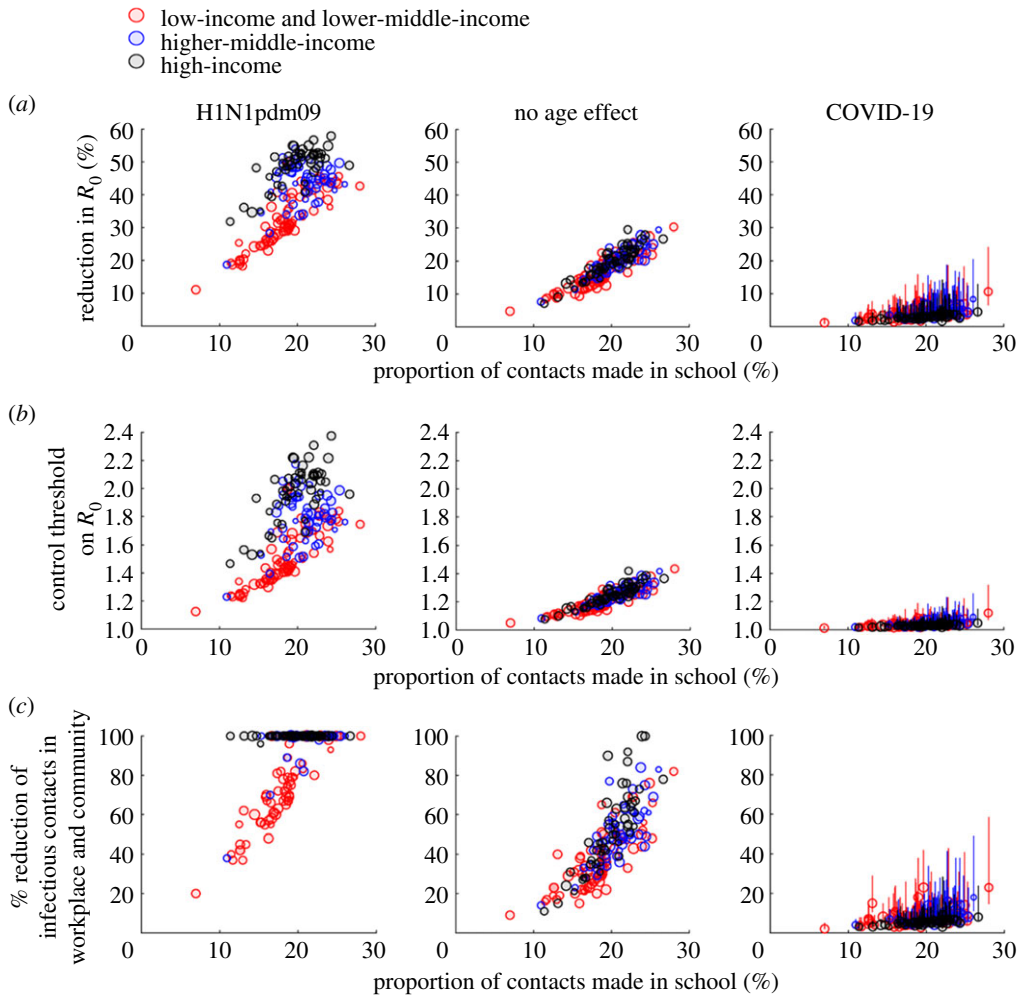
Figure 2. The effectiveness of school closure in reducing transmissibility as a function of proportion of contacts made in school across 177 jurisdictions. See electronic supplementary material, table S4 for the numerical values of all data points shown here. The sizes of data points are proportional to the log of the corresponding population size. Circles and bars (for COVID-19) indicate point predictions and 95% prediction intervals, respectively. (a) Reduction in R_0 conferred by school closure against H1N1pdm09 (left), a pathogen to which individual of all ages were equally susceptible (middle) and COVID-19 (right). (b) Outbreak threshold (in terms of R_0) under school closure. (c) Per cent reduction in workplace and community contacts that would be required in order to achieve the reduction in R_0 conferred by school closure. A required percentage of 100% means that this was impossible. (Online version in colour.)

had no effect on infectiousness (figure 2c,d). For pandemic influenza, we used the H1N1pdm09 parameters from a previous household transmission study which indicated those aged 0–18 and greater than 50 were 2 and 0.2 times as susceptible compared with those aged 19–50 [24]. We also considered seasonal influenza for which age had no clear effect on susceptibility or infectiousness.

To avoid underestimating the impact of school closure, we assumed (unless specified otherwise) that all contacts made at school would be eliminated during school closure and holidays while contacts in all other settings were unchanged. We estimated the reduction in basic reproductive number (R_0) conferred by school closure in each jurisdiction; the reduction in effective reproductive number (R_t) would be similar if the age distribution of the susceptibles was similar to the overall age distribution (e.g. when most of the population were still susceptible). We

estimated the peak prevalence and IAR in each jurisdiction with and without school closure for different values of R_0 (which reflected the baseline transmission dynamics when school closure started, i.e. the combined effect of the intrinsic transmissibility of the virus and the control measures in the population).

(c) Effectiveness of school closure

The relative reduction in R_0 conferred by school closure was primarily driven by the proportion of contacts made in school (figure 2a). Our results were consistent with previous studies that had empirically inferred the reduction of the reproductive number for influenza during school closure and holidays in mainland China, France, Hong Kong SAR, India, South Korea and the UK if 60% of school contacts were eliminated during these times while contacts in all other settings were unchanged (table 1).

Because schoolchildren were twice as susceptible to H1N1pdm09 compared with adults, school closure could generally reduce the R_0 of H1N1pdm09 by 20–58% (except for South Sudan where the out-of-school rates exceed 60%; figure 2a). By contrast, school closure was much less effective at reducing the R_0 of COVID-19, to which children were less susceptible compared with adults (figure 1c), even if all school contacts would be eliminated during school closure. Specifically, school closure could only reduce the R_0 of COVID-19 by $x_R = 2$–5%, 2–8% and 1–11% in high-income, upper-middle-income and other populations, respectively (figure 2a for the sensitivity of these estimates to the parametric uncertainty in age-specific susceptibility). This implies that school closure alone would be unable to control COVID-19 in these populations unless the reproductive number (that corresponded to the baseline contact matrix) was already below $1/(1-x_R) \approx 1 + x_R$, i.e. close to the critical value of 1 (figure 2b). If only kindergartens and primary schools were closed, R_0 of COVID-19 in high-income and other populations would be reduced by 1–2% and 1–5%, respectively. Closing secondary schools and higher education institutions would have similar effects except in low-income populations with high out-of-school rates among adolescents and youth of upper secondary school age or above. Comparable reductions in R_0 conferred by school closure could be achieved by reducing infectious contacts in workplace and community by 3–9% and 2–23% for high-income and other populations, respectively (figure 2c for uncertainty levels). This might be achieved by reducing the frequency or transmissibility (e.g. face mask wearing and heightened personal hygiene) of these contacts. By contrast, this substitution strategy would be too disruptive to be practical or simply impossible for H1N1pdm09 (i.e. the effect of school closure could not be substituted even if all workplace and community contacts were to be eliminated).

For COVID-19, the reduction in morbidity and mortality conferred by school closure would depend on the baseline R_0 as well as the timing and duration of school closure (figure 3). If the baseline R_0 was around 2.5 (e.g. the epidemic was largely unmitigated), implementing school closure throughout the epidemic would reduce peak prevalence in high-income and other populations by 5–17% and 4–28%, respectively (figure 3a for uncertainty levels). The reductions would be higher for school children (because they are the target group of school closure) and slightly lower for older adults (because physical interactions among children and the elderly are relatively weak in general). By contrast, if the baseline R_0 was around 1.5 (e.g. school closure was implemented as a complementary measure in addition to other NPIs such as the use of facemasks), implementing school closure throughout the epidemic would reduce peak prevalence by 9–27% and 7–52% in high-income and other populations, respectively (figure 4a for uncertainty levels). The reduction in final IAR would be lower than that for peak prevalence. For example, the overall IAR would be reduced by 6–19% and 5–33% in high-income and other populations if $R_0 = 1.5$ (figure 4b); the corresponding reductions were 5–14% and 4–16% if $R_0 = 2.5$ (figure 3b). Reducing infectious contacts in workplace and community settings by 3–28% would achieve the reduction in IAR under 18 conferred by school closure if $R_0 = 1.5$; the corresponding reduction requirement would increase to 10–44% if $R_0 = 2.5$.
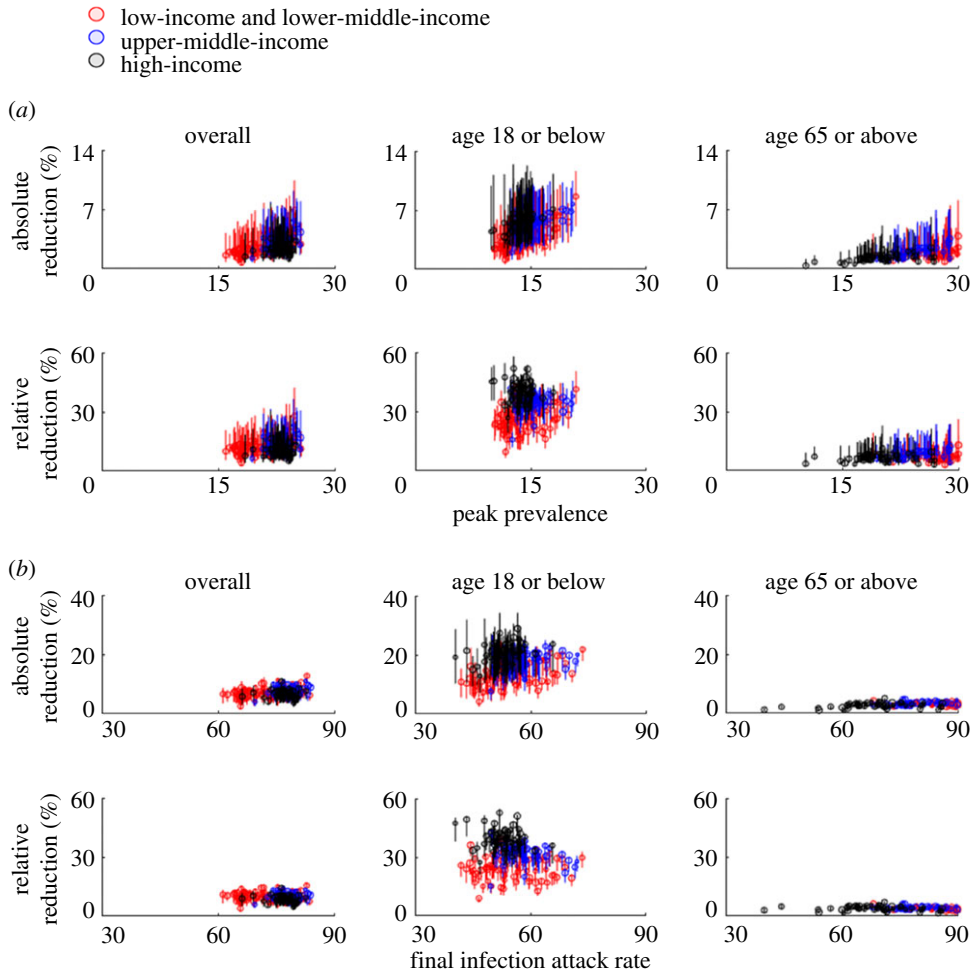
Figure 3. The public health impact of school closure across 177 jurisdictions when $R_0 = 2.5$. See electronic supplementary material, table S4 for the numerical values of all data points shown here. Circles and bars indicate point predictions and 95% prediction intervals, respectively. (*a*) Reduction in peak prevalence for all ages (left), those aged 18 or below (middle) and those aged 65 or above (right). (*b*) Reduction in final IAR for all ages (left), those aged 18 or below (middle) and those aged 65 or above (right). (Online version in colour.)

Note that these comparative outcomes across jurisdictions were conditional on the baseline R_0 which would not necessarily be the same in all populations due to heterogeneity in age structures, contact patterns, climatic factors and intervention policies. Importantly, our results do not suggest that prevalence and IAR would generally be higher or lower in any particular population. By the same token, our results do not suggest that school closure would generally confer greater reductions in peak prevalence and IAR in any particular population.

4. Discussion

The fundamental motivation for school closure presumes schoolchildren are a core transmission group. In concordance with recent studies [14–16], we have shown that this is not the case for COVID-19 because compared to older individuals, school children are substantially less susceptible to COVID-19. Consequently, even prolonged school closure that could hypothetically eliminate all school contacts for weeks to months would have limited effectiveness in reducing transmissibility.
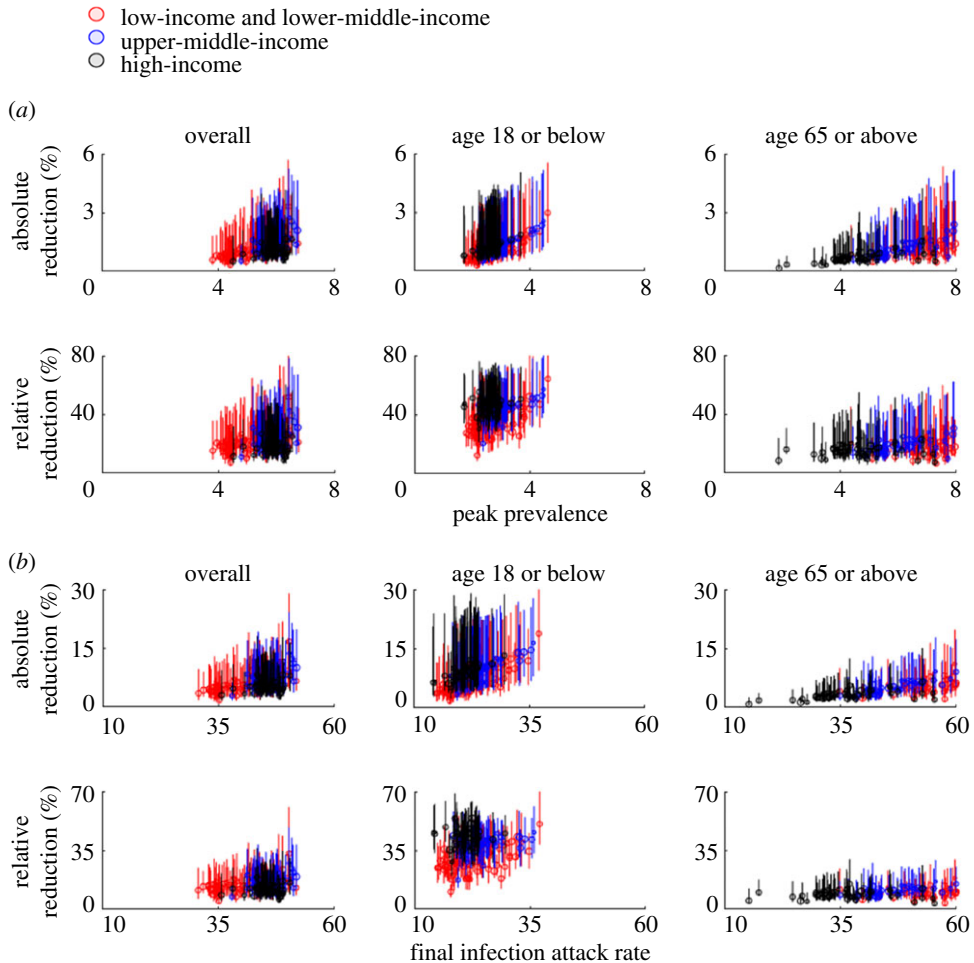
Figure 4. The same layout as in figure 3 when $R_0 = 1.5$. (Online version in colour.)

Our results, together with recent findings that clinical severity of COVID-19 in children is lower than that for adults [19,25] and that prolonged school closure could significantly impair the emotional, mental and physical health of schoolchildren (see below) suggest that school closure may not be ideal as a sustained primary NPI for controlling COVID-19. In fact, the specific objective underlying school closure should be more clearly articulated so that alternative options could be comparatively assessed. For example, the overall disease transmission in the population could be mitigated by reducing infectious contacts (e.g. by enhancing personal infection prevention) which is likely to be more feasible and less disruptive to children's overall wellbeing than school closure (https://www.cdc.gov/coronavirus/2019-ncov/community/schools-childcare/operation-strategy.html). On the other hand, short-term school suspension could and should be considered when other NPIs (e.g. avoidance of mass gatherings, physical distancing in workplace and community, enhanced personal hygiene, etc.) are already in place, where the marginal effect of school closure is critical to achieving the target control levels (e.g. when hospitals have exceeded their surge capacity and clinical outcomes including the case-fatality risk are on the verge of unnecessarily deteriorating simply because the health system is about to implode).

Recent studies have highlighted numerous downsides of prolonged school closure on the mental and physical wellbeing of children, including increased risks of loneliness, academic

achievement gaps, reduction in physical activity, malnutrition (for children who rely on school meals), increased risk of domestic abuse and violence, etc. [26–30]. Even within high-income populations, these adverse effects would likely be disproportionately amplified among socio-economically disadvantaged groups (e.g. lack of or limited access to affordable childcare, technical and financial support for digital learning, online social interactions and physical exercises). In addition, closing schools could inadvertently cause childcare shortages and hence worker absenteeism among parents. The resulting worker absenteeism in the healthcare system (which has been estimated to be around 20% for the US healthcare workforce) could indirectly exacerbate the case-fatality risk of COVID-19 patients and therefore the overall mortality rate of the pandemic due to weakened healthcare [31]. As such, compensatory measures should be proactively devised to ameliorate these adverse events if school closure is employed to control COVID-19. For example, the childcare arrangement should be explicitly considered in order to minimize unintended worker absenteeism and the downstream adverse effect on healthcare quality.

We estimated that within-household transmissibility in Shenzhen and Anqing was reduced by 62% (11–86%) in January–March after nationwide lockdown measures were implemented in China. Given that human mobility in all Chinese major cities was reduced by greater than 90% during this period (all individuals were mandated to stay home and only one household member was allowed to leave home every 2 to 3 days to procure basic (e.g. food) and essential (e.g. medications) supplies, such reduction in within-household transmissibility was likely due to heightened precautions of personal infection prevention. Sustained adoption of such precautions in the 177 jurisdictions could in principle reduce the reproductive number of COVID-19 by 6–41%, i.e. more effective than school closure.

We estimated that compared with females, males were less susceptible to COVID-19 infection but might be more infectious (albeit the uncertainty in our estimate of the latter is high; figure 1*f*). Our estimates are consistent with that from the largest household transmission study in China [32]. However, seroprevalence studies from other countries suggested that males might be more susceptible to infection than females [33,34]. As such, further investigations are needed to elucidate the effect of sex on susceptibility to and infectiousness of COVID-19 infection.

Our study has several important limitations. First, we assumed that interactions, case ascertainment and infection prevention measures were homogeneous among members within households, hence attributing the effect of age on the risk of COVID-19 infection to intrinsic biological susceptibility. In reality, adult–adult and adult–child interactions might be associated with different transmission rates, and case ascertainment and infection prevention measures might be age-dependent (though less likely given that the whole country implemented strict lockdowns at maximal pandemic response level). Furthermore, during the 2009 influenza pandemic, household transmission models similar to ours generated age-specific susceptibility estimates for H1N1pdm09 [23,24] that were consistent with that from community transmission and sero-surveillance models [7]. Taken together, these studies provide external support for our surmise that children were biologically less susceptible to COVID-19 infection and our estimates of age-specific susceptibility are applicable in general transmission settings.

Second, we assumed that our estimates based on data from Chinese cities were applicable to the rest of the world. While our estimates are consistent with similar results from other localities [14,35–37], more region-specific data would be needed to validate this assumption.

Third, we assumed that school closure would eliminate all school contacts without affecting contacts in other settings. In reality, school closure could increase some non-school contacts (e.g. more contacts between children and their caregivers, more contacts among children outside schools), in which case school closure against COVID-19 would be less effective than estimated here. This is particularly true in resource-constrained settings where childcare assistance and personal protective measures (e.g. facemasks and hand hygiene supplies) are limited.

Fourth, we assumed that age had no effect on infectiousness [38]. If children are less infectious than adults, as postulated by recent studies [39], then school closure would be less effective than estimated here.

Fifth, our framework is built on the contact matrices synthesized by Prem *et al.* [20] which have been increasingly used to study infectious diseases including COVID-19 [14,40,41]. These have been compared against empirical surveys, in terms of both raw numbers of contacts and impact of interventions in COVID-19 models. Comparisons indicate good concordance between empirical and synthetic matrices [20]. However, differences are observed in older adults in some countries, which could be partly due to small sample sizes in older adults in empirical studies, particularly in countries with younger populations. This highlights the critical importance of conducting periodic sufficiently powered social contact surveys in different parts of the world for informing epidemic surveillance and control of a wide range of infectious diseases including COVID-19, influenza, TB, RSV, measles, etc.

Sixth, almost all contact studies (whether empirical or synthetic) define contacts as involving either physical touch or two-way conversations (see e.g. Mossong *et al.* [42]). Although students may be in a class of 20 other students, they only have 'contact' with a small subset of them. Indeed, studies have found that when contacts are defined as such, there is no statistically significant association between class size and total number of contacts (see e.g. Béraud *et al.* [43]). This matches guidelines and current (limited) understanding of SARS-CoV-2 transmission; for instance the United States Centers for Disease Control and Prevention define a 'close contact' (mandating quarantine) as someone who was 'within 6 feet of someone who has COVID-19 for a total of 15 min or more' (https://www.cdc.gov/coronavirus/2019-ncov/if-you-are-sick/quarantine.html). However, it cannot be ruled out that transmission can occur between classmates who are not close contacts. This highlights the importance of evidence-based school opening guidelines, such as minimizing movement within the classroom, keeping desks well space, encouraging face covering and frequent handwashing and ensuring good ventilation in classrooms whenever practicable.

References

1. Wang G, Zhang Y, Zhao J, Zhang J, Jiang F. 2020 Mitigate the effects of home confinement on children during the COVID-19 outbreak. *Lancet* **395**, 945–947.
2. Viner RM, Russell SJ, Croker H, Packer J, Ward J, Stansfield C, Mytton O, Bonell C, Booy R. 2020 School closure and management practices during coronavirus outbreaks including COVID-19: a rapid systematic review. *Lancet Child Adolesc. Health* **4**, 397–404. (doi:10.1016/S2352-4642(20)30095-X)
3. Esposito S, Principi N. 2020 School closure during the coronavirus disease 2019 (COVID-19) pandemic: an effective intervention at the global level? *JAMA Pediatr.* **174**, 921–922. (doi:10.1001/jamapediatrics.2020.1892)

4. Fong MW, Gao H, Wong JY, Xiao J, Shiu EY, Ryu S, Cowling BJ. 2020 Nonpharmaceutical measures for pandemic influenza in nonhealthcare settings—social distancing measures. *Emerg. Infect. Dis.* **26**, 976.

5. Xiao J, Shiu EY, Gao H, Wong JY, Fong MW, Ryu S, Cowling BJ. 2020 Nonpharmaceutical measures for pandemic influenza in nonhealthcare settings—personal protective and environmental measures. *Emerg. Infect. Dis.* **26**, 967.

6. Birrell PJ *et al.* 2011 Bayesian modeling to unmask and predict influenza A/H1N1pdm dynamics in London. *Proc. Natl Acad. Sci. USA* **108**, 18 238–18 243. (doi:10.1073/pnas.1103002108)

7. Wu JT *et al.* 2014 Inferring influenza infection attack rate from seroprevalence data. *PLoS Pathog.* **10**, e1004054. (doi:10.1371/journal.ppat.1004054)

8. Ali ST, Kadi AS, Ferguson NM. 2013 Transmission dynamics of the 2009 influenza A (H1N1) pandemic in India: the impact of holiday-related school closure. *Epidemics* **5**, 157–163. (doi:10.1016/j.epidem.2013.08.001)

9. Yu H *et al.* 2012 Transmission dynamics, border entry screening, and school holidays during the 2009 influenza A (H1N1) pandemic, China. *Emerg. Infect. Dis.* **18**, 758–766. (doi:10.3201/eid1805.110356)

10. Cauchemez S, Valleron AJ, Boelle PY, Flahault A, Ferguson NM. 2008 Estimating the impact of school closure on influenza transmission from Sentinel data. *Nature* **452**, 750–754. (doi:10.1038/nature06732)

11. Ali ST, Cowling BJ, Lau EHY, Fang VJ, Leung GM. 2018 Mitigation of influenza B epidemic with school closures, Hong Kong, 2018. *Emerg. Infect. Dis.* **24**, 2071–2073. (doi:10.3201/eid2411.180612)

12. Ryu S, Ali ST, Cowling BJ, Lau EHY. 2020 Effects of school holidays on seasonal influenza in South Korea, 2014–2016. *J. Infect. Dis.* **222**, 832–835. (doi:10.1093/infdis/jiaa179)

13. Bi Q *et al.* 2020 Epidemiology and transmission of COVID-19 in 391 cases and 1286 of their close contacts in Shenzhen, China: a retrospective cohort study. *Lancet Infect. Dis.* **20**, 911–919. (doi:10.1016/S1473-3099(20)30287-5)

14. Davies NG *et al.* 2020 Age-dependent effects in the transmission and control of COVID-19 epidemics. *Nat. Med.* **26**, 1205–1211. (doi:10.1038/s41591-020-0962-9)

15. Jing QL *et al.* 2020 Household secondary attack rate of COVID-19 and associated determinants in Guangzhou, China: a retrospective cohort study. *Lancet Infect. Dis.* **20**, 1141–1150. (doi:10.1016/S1473-3099(20)30471-0)

16. Zhang J *et al.* 2020 Changes in contact patterns shape the dynamics of the COVID-19 outbreak in China. *Science* **368**, 1481–1486. (doi:10.1126/science.abb8001)

17. Pollán M *et al.* 2020 Prevalence of SARS-CoV-2 in Spain (ENE-COVID): a nationwide, population-based seroepidemiological study. *Lancet* **396**, 535–544. (doi:10.1016/S0140-6736(20)31483-5)

18. Stringhini S *et al.* 2020 Seroprevalence of anti-SARS-CoV-2 IgG antibodies in Geneva, Switzerland (SEROCoV-POP): a population-based study. *Lancet* **396**, 313–319. (doi:10.1016/S0140-6736(20)31304-0)

19. Wu JT, Leung K, Bushman M, Kishore N, Niehus R, de Salazar PM, Cowling BJ, Lipsitch M, Leung GM. 2020 Estimating clinical severity of COVID-19 from the transmission dynamics in Wuhan, China. *Nat. Med.* **26**, 506–510. (doi:10.1038/s41591-020-0822-7)

20. Prem K, van Zandvoort K, Klepac P, Eggo RM, Davies NG, Cook AR, Jit M. 2020 Projecting contact matrices in 177 geographical regions: an update and comparison with empirical data for the COVID-19 era. *medRxiv*, 2020.2007.2022.20159772. (doi:10.1101/2020.07.22.20159772)

21. Lauer SA, Grantz KH, Bi Q, Jones FK, Zheng Q, Meredith HR, Azman AS, Reich NG, Lessler J. 2020 The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: estimation and application. *Ann. Intern. Med.* **172**, 577–582. (doi:10.7326/M20-0504)

22. Zheng X *et al.* 2020 Asymptomatic patients and asymptomatic phases of Coronavirus Disease 2019 (COVID-19): a population-based surveillance study. *Natl Sci. Rev.* **7**, 1527–1539.

23. Sugimoto JD, Borse NN, Ta ML, Stockman LJ, Fischer GE, Yang Y, Halloran ME, Longini IM, Duchin JS. 2011 The effect of age on transmission of 2009 pandemic influenza A (H1N1) in a camp and associated households. *Epidemiology* **22**, 180–187. (doi:10.1097/EDE.0b013e3182060ca5)

24. Cauchemez S, Donnelly CA, Reed C, Ghani AC, Fraser C, Kent CK, Finelli L, Ferguson NM. 2009 Household transmission of 2009 pandemic influenza A (H1N1) virus in the United States. *N Engl. J. Med.* **361**, 2619–2627. (doi:10.1056/NEJMoa0905498)

25. Verity R *et al.* 2020 Estimates of the severity of coronavirus disease 2019: a model-based analysis. *Lancet Infect. Dis.* **20**, 669–677. (doi:10.1016/S1473-3099(20)30243-7)

26. Mayurasakorn K, Pinsawas B, Mongkolsucharitkul P, Sranacharoenpong K, Damapong SN. 2020 School closure, COVID-19 and lunch programme: unprecedented undernutrition crisis in low-middle income countries. *J. Paediatr. Child Health* **56**, 1013–1017. (doi:10.1111/jpc.15018)

27. Thomas EY, Anurudran A, Robb K, Burke TF. 2020 Spotlight on child abuse and neglect response in the time of COVID-19. *Lancet Public Health* **5**, e371. (doi:10.1016/S2468-2667(20)30143-2)

28. Loades ME *et al.* 2020 Rapid systematic review: the impact of social isolation and loneliness on the mental health of children and adolescents in the context of COVID-19. *J. Am. Acad. Child Adolesc. Psychiatry* **59**, 1218–1239. (doi:10.1016/j.jaac.2020.05.009)

29. SS Teo S, Griffiths G. 2020 Child protection in the time of COVID-19. *J. Paediatr. Child Health* **56**, 838–840. (doi:10.1111/jpc.14916)

30. Green P. 2020 Risks to children and young people during covid-19 pandemic. *BMJ* **369**, m1669. (doi:10.1136/bmj.m1669)

31. Bayham J, Fenichel EP. 2020 Impact of school closures for COVID-19 on the US health-care workforce and net mortality: a modelling study. *Lancet Public Health* **5**, e271–e278. (doi:10.1016/S2468-2667(20)30082-7)

32. Li F *et al.* 2021 Household transmission of SARS-CoV-2 and risk factors for susceptibility and infectivity in Wuhan: a retrospective observational study. *Lancet Infect. Dis.* **21**, 617–628.

33. Le Vu S *et al.* 2021 Prevalence of SARS-CoV-2 antibodies in France: results from nationwide serological surveillance. *Nat. Commun.* **12**, 1–7.

34. Reicher S *et al.* 2021 Nationwide seroprevalence of antibodies against SARS-CoV-2 in Israel. *Eur. J. Epidemiol.* **36**, 727–734. (doi:10.1007/s10654-021-00749-1)

35. Dattner I, Goldberg Y, Katriel G, Yaari R, Gal N, Miron Y, Ziv A, Hamo Y, Huppert A. 2020 The role of children in the spread of COVID-19: using household data from Bnei Brak, Israel, to estimate the relative susceptibility and infectivity of children. *medRxiv*, 2020.2006.2003.20121145. (doi:10.1101/2020.06.03.20121145)

36. Madewell ZJ, Yang Y, Longini IM, Halloran ME, Dean NE. 2020 Household transmission of SARS-CoV-2: a systematic review and meta-analysis of secondary attack rate. *medRxiv*, 2020.2007.2029.20164590. (doi:10.1101/2020.07.29.20164590)

37. Lyngse FP, Kirkeby CT, Halasa T, Andreasen V, Skov RL, Møller FT, Krause TG, Mølbak K. 2020 COVID-19 transmission within Danish Households: a nationwide study from lockdown to reopening. *medRxiv*, 2020.2009.2009.20191239. (doi:10.1101/2020.09.09.20191239)

38. Jones TC *et al.* 2020 An analysis of SARS-CoV-2 viral load by patient age. *medRxiv*, 2020.2006.2008.20125484. (doi:10.1101/2020.06.08.20125484)

39. Bunyavanich S, Do A, Vicencio A. 2020 Nasal gene expression of angiotensin-converting enzyme 2 in children and adults. *JAMA* **323**, 2427–2429. (doi:10.1001/jama.2020.8707)

40. Koo JR, Cook AR, Park M, Sun Y, Sun H, Lim JT, Tam C, Dickens BL. 2020 Interventions to mitigate early spread of SARS-CoV-2 in Singapore: a modelling study. *Lancet Infect. Dis.* **20**, 678–688.

41. Hilton J, Keeling MJ. 2020 Estimation of country-level basic reproductive ratios for novel Coronavirus (SARS-CoV-2/COVID-19) using synthetic contact matrices. *PLoS Comput. Biol.* **16**, e1008031. (doi:10.1371/journal.pcbi.1008031)

42. Mossong J *et al.* 2008 Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS Med.* **5**, e74. (doi:10.1371/journal.pmed.0050074)

43. Béraud G *et al.* 2015 The French connection: the first large population-based contact survey in France relevant for the spread of infectious diseases. *PLoS ONE* **10**, e0133203. (doi:10.1371/journal.pone.0133203)

44. Wu JT *et al.* 2021 A global assessment of the impact of school closure in reducing COVID-19 spread. Figshare.