

# 1 **Within and between classroom transmission patterns of** 2 **seasonal influenza among primary school students in** 3 **Matsumoto city, Japan**

4 Akira Endo<sup>1,2,3,4,\*</sup>, Mitsuo Uchida<sup>5</sup>, Naoki Hayashi<sup>6,7</sup>, Yang Liu<sup>1,2</sup>, Katherine E. Atkins<sup>1,2,8</sup>, Adam J.  
5 Kucharski<sup>1,2</sup>, Sebastian Funk<sup>1,2</sup>

6 1. Department of Infectious Disease Epidemiology, London School of Hygiene & Tropical  
7 Medicine, London, WC1E 7HT, United Kingdom

8 2. The Centre for Mathematical Modelling of Infectious Diseases, London School of Hygiene &  
9 Tropical Medicine, WC1E 7HT, United Kingdom

10 3. The Alan Turing Institute, London, NW1 2DB, United Kingdom

11 4. School of Tropical Medicine and Global Health, Nagasaki University, Nagasaki, 852-8523, Japan

12 5. Graduate School of Medicine, Gunma University, Gunma, 371-8511, Japan

13 6. Simulation & Mining Division, NTT DATA Mathematical Systems Inc., Tokyo, 160-0016, Japan

14 7. Department of Mathematical and Computing Science, School of Computing, Tokyo Institute of  
15 Technology, Tokyo, 152-8552, Japan

16 8. Centre for Global Health Research, Usher Institute, University of Edinburgh, Edinburgh, EH16  
17 4UX, United Kingdom

18 \* Correspondence to: [akira.endo@lshtm.ac.uk](mailto:akira.endo@lshtm.ac.uk)

## 19 **Significance**

20 Empirical evidence on detailed transmission patterns of influenza among students within and between  
21 classes and grades and how they are shaped by school population structure (e.g. class and school  
22 sizes) has been limited to date. We analysed a detailed dataset of seasonal influenza incidence in 29

23 **NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.**  
primary schools in Japan and found that the reproduction number at school did not show any clear

24 association with the size or the number of classes. Our findings suggest that the interventions that only  
25 focus on reducing the number of students in class at any moment in time (e.g. reduced class sizes and  
26 staggered attendance) may not be as effective as measures that aim to reduce within-class risk (e.g.  
27 mask-wearing and vaccines).

## 28 **Abstract**

29 Schools play a central role in the transmission of many respiratory infections. Heterogeneous social  
30 contact patterns associated with the social structures of schools (i.e. classes/grades) are likely to  
31 influence the within-school transmission dynamics, but data-driven evidence on fine-scale  
32 transmission patterns between students has been limited. Using a mathematical model, we analysed a  
33 large-scale dataset of seasonal influenza outbreaks in Matsumoto city, Japan to infer social  
34 interactions within and between classes/grades from observed transmission patterns. While the  
35 relative contribution of within-class and within-grade transmissions to the reproduction number varied  
36 with the number of classes per grade, the overall within-school reproduction number, which  
37 determines the initial growth of cases and the risk of sustained transmission, was only minimally  
38 associated with class sizes and the number of classes per grade. This finding suggests that  
39 interventions that change the size and number of classes, e.g. splitting classes and staggered  
40 attendance, may have limited effect on the control of school outbreaks. We also found that  
41 vaccination and mask-wearing of students were associated with reduced susceptibility (vaccination  
42 and mask-wearing) and infectiousness (mask-wearing) and hand washing with increased  
43 susceptibility. Our results show how analysis of fine-grained transmission patterns between students  
44 can improve understanding of within-school disease dynamics and provide insights into the relative  
45 impact of different approaches to outbreak control.

## 46 **Background**

47           Influenza virus and other directly transmitted pathogens typically spread over social contact  
48 networks involving frequent conversational or physical contacts (1–4). There is evidence that schools  
49 are important social environments that can facilitate the transmission of influenza via close contacts  
50 between students (5–9). Previous studies have collected contact data between students using  
51 questionnaires and wearable sensor devices and found strong assortativity of contact rates within  
52 classes and grades (10–14), which is likely relevant to the within-school transmission dynamics of  
53 respiratory infections and the effectiveness of school-based interventions. However, such insights  
54 from contact data also need to be validated with real-world outbreak data because contacts as  
55 measured in those studies may not necessarily be fully representative of the types of contacts that lead  
56 to transmission (e.g. with regards to proximity and duration). In this light, the differential transmission  
57 rates of influenza associated with classes and grades have also been estimated from empirical  
58 outbreak data in a few studies (6, 15, 16). However, those studies are limited to the analysis of only  
59 one or two schools and included a relatively small number of cases (< 300). Therefore, robust findings  
60 across schools with different structures that capture the full range of heterogeneity in within-school  
61 transmission dynamics have remained a crucial knowledge gap.

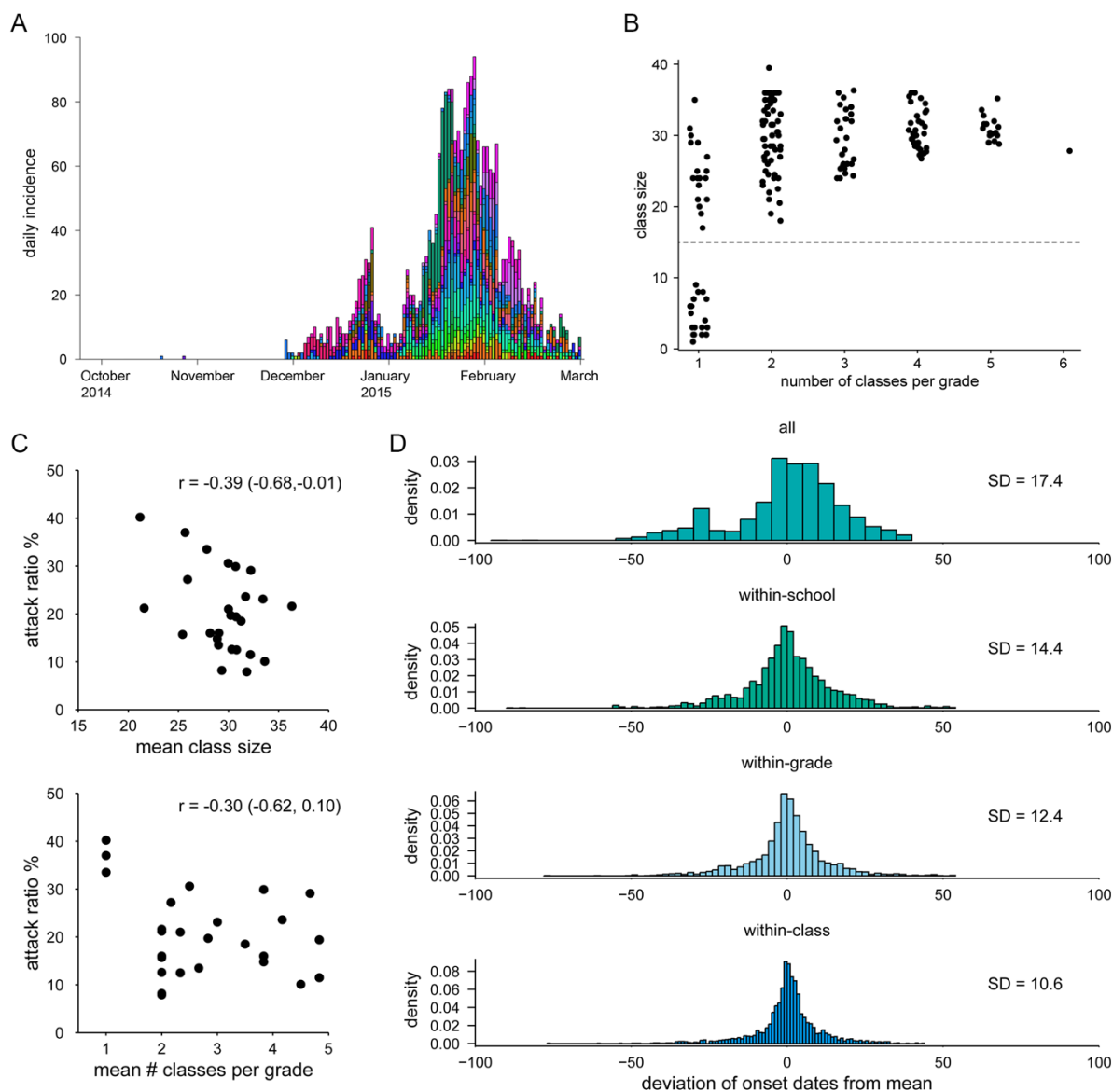
62           Understanding how school population structures (e.g. class and school sizes) shape  
63 transmission dynamics is key to making predictions about outbreak dynamics and interventions in  
64 these settings. Modelling studies of school outbreaks often require a choice between the ‘density-  
65 dependent mixing’ and ‘frequency-dependent mixing’ assumptions (17). The density-dependent  
66 mixing assumes that the transmission rate between a pair of students is constant regardless of the  
67 class/school sizes, while the frequency-dependent mixing assumes an inverse proportionality between  
68 them. As a result, the reproduction number is expected to increase with class/school size with the  
69 density-dependent mixing assumption and remain stable with the frequency-dependent mixing  
70 assumption. Whether the transmission is best characterised by the density-dependent mixing,  
71 frequency-dependent mixing or any other alternative assumption may vary between different modes  
72 of transmission and exposure settings (18–22). However, choices between the assumptions made by  
73 existing studies of school outbreaks vary widely and are not based on a clear empirical consensus (9,

74 23–26). These makes it challenging to interpret simulation studies evaluating school-based  
75 interventions (e.g. reduced class sizes) because the estimated effect sizes can heavily rely on the  
76 assumed mixing patterns (27–31).

77 To fill this knowledge gap in heterogeneous transmission dynamics at school, we applied a  
78 mathematical model of influenza virus transmission to a large-scale dataset from the 2014-15 season  
79 in Matsumoto city, Japan, which included diagnosed influenza reports among 10,923 primary school  
80 students and their household members. The model accounted for within-school transmissions as well  
81 as introductions to and from households and risk from the general community, which constitute key  
82 social layers of transmission (32–34). Using this model, we estimated fine-scale heterogeneous  
83 transmission patterns among students within and between classes and grades, as well as determinants  
84 of transmission rates including school structures and precautionary measures.

## 85 **Results**

86 We analysed citywide survey data of 10,923 primary school students (5–12 years old) in  
87 Matsumoto city, Japan in 2014/15, which included 2,548 diagnosed influenza episodes among  
88 students (Figure 1A). The dataset was obtained from 29 schools with a range of class structures (sizes  
89 and the number of classes per grade), allowing for detailed analysis of within and between class  
90 transmission patterns (Figure 1B). The attack ratio (i.e. the cumulative proportion diseased) in each  
91 school (excluding three distinctively small schools with fewer than 15 students per class) showed  
92 weak to null negative correlations with the mean class size and the mean number of classes per grade  
93 (Figure 1C). The onset dates of students showed a temporal clustering pattern associated with school  
94 structure (Figure 1D). When the students were partitioned into different levels of groupings (i.e. by  
95 class, grade, school and overall), the deviation of onset dates from the within-group mean tended to be  
96 smaller with finer groupings.



97

98 Figure 1. Transmission dynamics of seasonal influenza in primary schools in Matsumoto city, Japan and estimated effects of  
 99 interventions for SARS-CoV-2. (A) Epidemic curve of seasonal influenza by illness onset in primary schools in Matsumoto  
 100 city, 2014/15. Colours represent different schools. Month names denote the 1<sup>st</sup> day of the month. (B) Scatterplot of the class  
 101 sizes and the number of classes per grade in the dataset. Each dot represents a class in the dataset. Dots are jittered along the  
 102 x-axis. Three schools had classes of fewer than 15 students (denoted by dotted horizontal line) and were excluded from the  
 103 model fitting. (C) The scatterplots of the school attack ratio (%) against the mean class size and the mean number of classes  
 104 per grade. The correlation indices ( $r$ ) and the 95% confidence intervals are also shown. (D) Temporal clustering patterns of  
 105 students' onset dates with different levels of groupings reproduced from the school transmission model. The distributions of  
 106 the deviance of each student's onset from the group mean are displayed at overall, school, grade and class levels. The  
 107 standard deviation (SD) of each distribution is also shown.

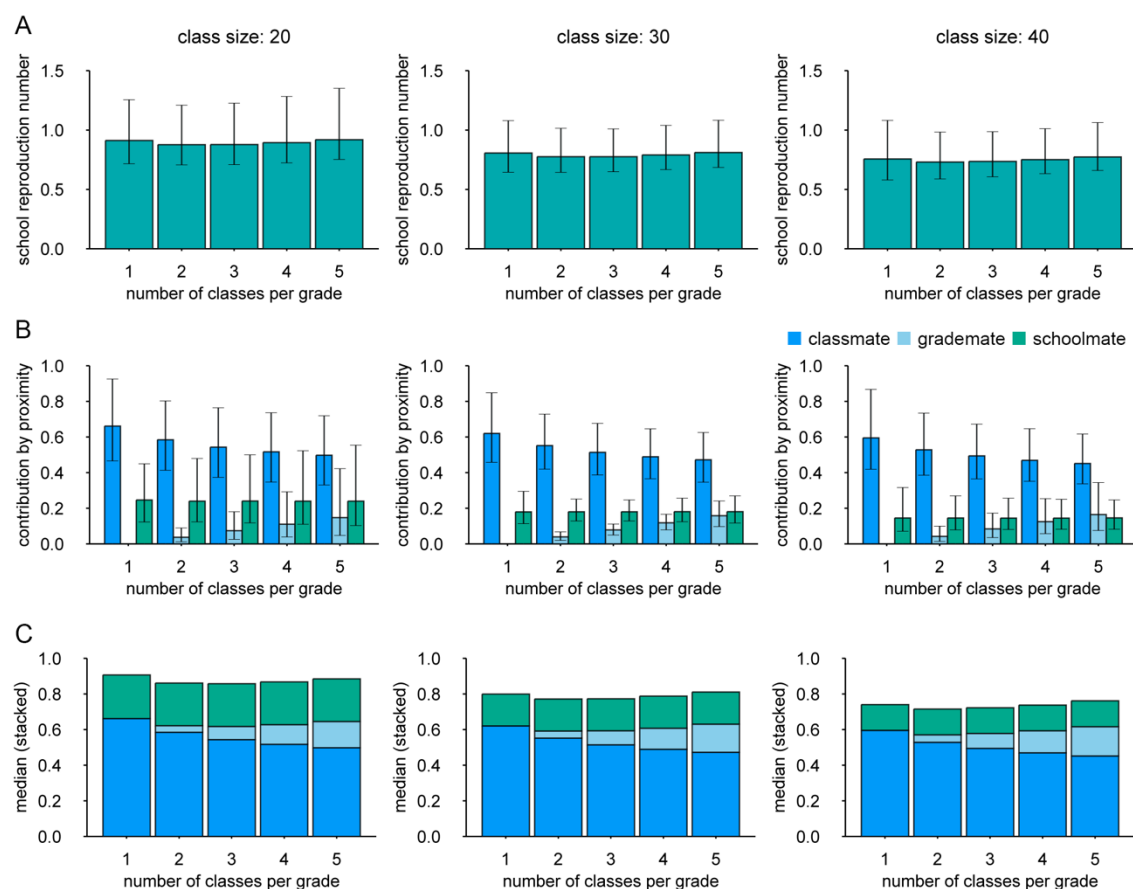
108

109           The temporal clustering shown in Figure 1D supports the hypothesis that the transmission is  
110 more likely within-class, followed by within-grade and within-school. We explored this further by  
111 estimating reproduction numbers within school. Using a mathematical model that accounts for  
112 different levels of interaction within and between classrooms and grades as well as introductions from  
113 households and community, we estimated the within-school effective reproduction number  $R_S$  of  
114 seasonal influenza in primary schools along with the breakdown of transmission risks associated with  
115 class/grade relationships (Figure 2). The relationship between any pair of students in the same school  
116 was classified as either “classmates”, “grademates” (in the same grade but not classmates) or  
117 “schoolmates” (not in the same grade). The estimated  $R_S$  was broken down as a sum of the  
118 contributions from these students, where the class size ( $n$ ) and the number of classes per grade ( $m$ )  
119 were assumed to affect the risk of transmission. The reconstructed overall  $R_S$  in a 6-year primary  
120 school was estimated to be around 0.7–0.9 and was not significantly associated with  $n$  or  $m$  (Figure  
121 2A). Namely, an infected student was suggested to generate a similar number of secondary cases  
122 irrespective of the class structure; although our estimates of  $R_S$  were about 15% smaller for the class  
123 size of 40 than 20<sup>1</sup>, the posterior p-value did not suggest a statistical significance ( $p \sim 0.15$  or above).  
124 As  $R_S$  was likely below 1 across class structures, school outbreaks may not have been sustained  
125 without continuous introductions from households and community. Transmission to classmates  
126 accounted for about two-thirds of  $R_S$  when each grade has only one class and was partially replaced  
127 by transmission to grademates as the number of classes per grade increases, while the sum of within-  
128 grade transmission (i.e. transmission to either classmates or grademates) remained stable (Figures 2B  
129 and 2C). Around 20–30% of overall  $R_S$  was explained by transmission to schoolmates throughout. We  
130 also obtained qualitatively similar results throughout our sensitivity analysis (Figure S3). In a 6-year  
131 school with 3 classes of 30 students, the risk of transmission was estimated to be 1.8% (95% credible  
132 interval (CrI): 1.4–2.5) from a given infected classmate of the same sex, 1.6% (1.2–2.1) the opposite  
133 sex, 0.13% (0.08–0.19) from a given infected grademate and 0.040% (0.029–0.055) from a given

---

<sup>1</sup> For example, the estimated relative reduction was 17% (95% credible interval: -16%–40%) for  $m = 3$ .

134 infected schoolmate (Table S2). The cumulative risk of infection from the community was estimated  
 135 to be 2.2% (1.7–2.7) over the season.



136  
 137 Figure 2. The estimated within-school transmission patterns of seasonal influenza among primary school students in  
 138 Matsumoto city, Japan. (A) The overall school reproduction number ( $R_S$ ) under different class structures. Whiskers represent  
 139 the 95% credible intervals (B) The breakdown of  $R_S$  corresponding to each type of within-school relationships. Whiskers  
 140 represent the 95% credible intervals. Bottom panels: stacked graph of  $R_S$  based on the median estimates.

141  
 142 We incorporated a log-linear regression (35) into the above estimation of  $R_S$  to account for  
 143 covariates that may affect the susceptibility or infectiousness of students. The results suggested that  
 144 vaccines were associated with reduced susceptibility while mask wearing was associated with both  
 145 reduced susceptibility and infectiousness (Table 1). Conversely, hand washing was associated with  
 146 increased susceptibility. Reduced chance of transmission during the winter break (27 December  
 147 2014–7 January 2015) was captured as a 75% estimated decline in the infectiousness of cases whose

148 onset dates were during the break. School grade, which serves as a proxy of students' age, did not  
 149 show a significant association with either susceptibility (relative value 1.10; CrI: 0.88–1.38) or  
 150 infectiousness (relative value 0.78; CrI: 0.59–1.04).

151 Table 1. Covariates and effects estimated in the log-linear regression

Covariate	Frequency in data	Relative susceptibility	Relative infectiousness
School grade (1 year increase)	—	1.10 (0.88–1.38)	0.78 (0.59–1.04)
Vaccine	47.7%	0.88* (0.81–0.96)	0.97 (0.82–1.18)
Mask wearing	51.4%	0.77* (0.71–0.85)	0.65* (0.55–0.78)
Hand washing	80.1%	1.54* (1.36–1.76)	1.25 (0.96–1.65)
Onset in winter break	5.9% (of cases)	—	0.25* (0.15–0.39)

152 Values are median estimates and 95% credible intervals.

153 \* Estimates with 95% credible intervals not crossing 1.

154

155 We estimated the breakdown of the source of infection for student cases based on the  
 156 conditional probability predicted by the model and parameter estimates. The epidemic curve stratified  
 157 by the estimated source of infection suggested that within-school transmission accounted for the  
 158 majority of student cases while schools were open and that the within-household transmission was  
 159 responsible for most of the cases reported during the winter break and shortly after (Figure 3A). The  
 160 aggregated relative contribution suggested that 54.6% (CrI: 53.5–55.7), 38.7% (CrI: 37.9–39.5) and  
 161 6.7% (CrI: 5.9–7.5) of the student cases were acquired from school, household and community,  
 162 respectively (Figure 3B). We estimated the possible relative effects of interventions altering the school  
 163 population structure on the school reproduction number  $R_s$ . We assumed that the estimated relative  
 164 contributions of class/grade relationship to the transmission risk reflect the contact patterns between  
 165 students which may also be relevant to the dynamics of another influenza outbreak at school (and  
 166 potentially those of directly-transmitted disease outbreaks in general) and that the responses to

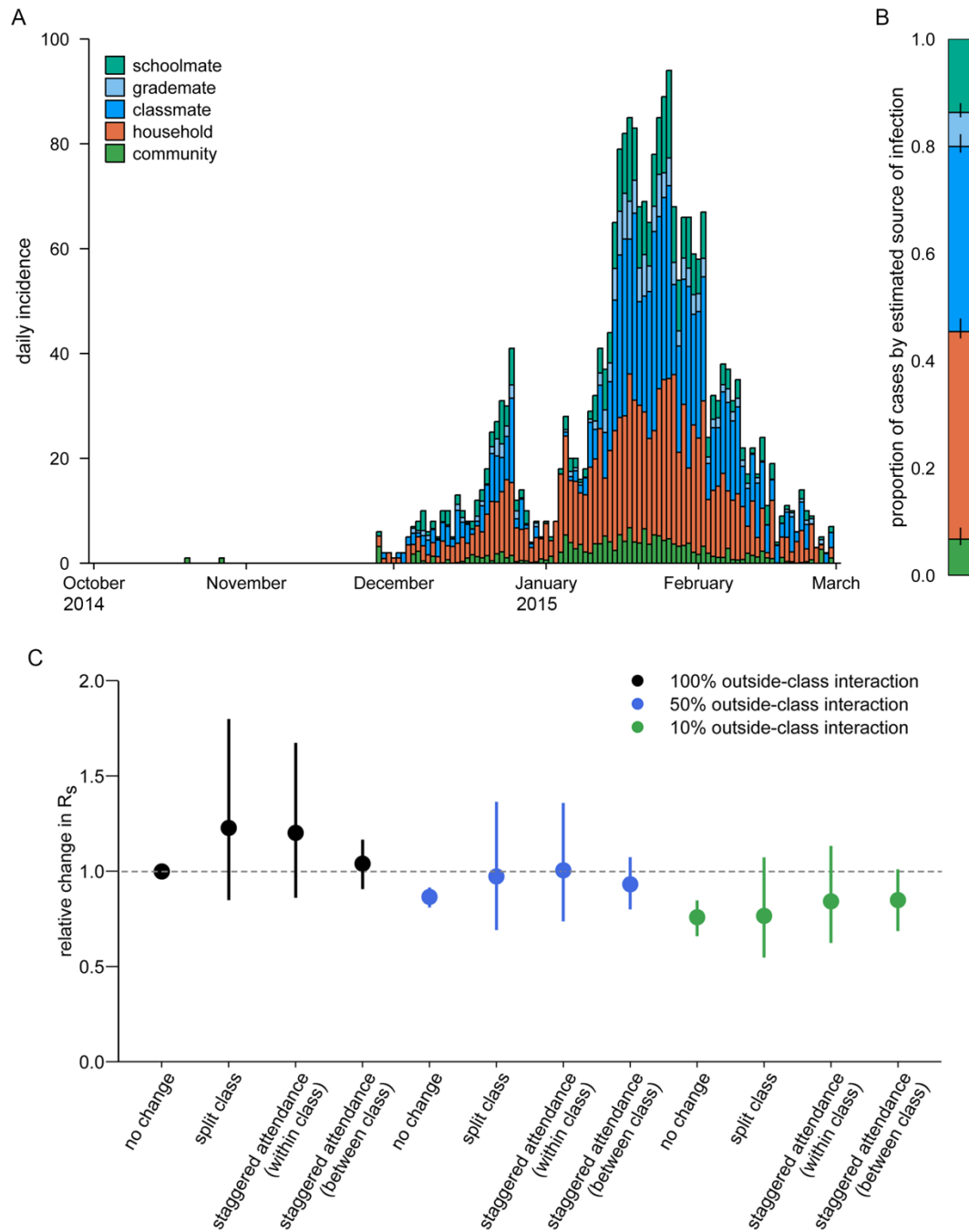


167 interventions can be captured by the estimated relationship between  $R_S$  and the changes in the  
168 variables  $n$  and  $m$  according to each intervention (Table 2). Specifically, in the ‘split class’ scenario,  
169 each class was assumed to be split in half and taught simultaneously in separate classrooms, while in  
170 the ‘staggered attendance’ scenarios only half of the students attend school at the same time by  
171 introducing two different time schedules, e.g. morning and evening classes. The estimated relative  
172 effects of school-based interventions on  $R_S$  in a hypothetical setting of 6-year school with 2 classes  
173 per grade (40 students each) showed that splitting classes or staggered attendance alone was unlikely  
174 to reduce  $R_S$  (or may even be counteractive) (Figure 1D), which is consistent with the aforementioned  
175 estimates of  $R_S$  minimally associated with class sizes and the number of classes. By reducing  
176 interactions between students from different classes (so-called ‘bubbling’ or ‘cohorting’) by 90%,  $R_S$   
177 could be reduced by up to around 20%. Combining split classes/staggered attendance with reduced  
178 interactions outside classes did not suggest incremental benefit in reducing  $R_S$ . Given that these  
179 interventions typically require additional resources including staff and classrooms, the overall benefit  
180 to changing class structures for influenza control may be limited.

181 Table 2. Summary of interventions that changes the size/number of classes

Interventions	Class size ( $n$ )	# classes per grade ( $m$ )	Assumption
Baseline (‘no change’)	40	2	Students contact within and between classes and grades proportionally to the estimated transmission patterns in Figure 2.
Split class	20	4	Each class is split into two and taught simultaneously in separate classrooms. Students may contact each other between classes.
Staggered attendance (within class)	20	2	Each class is split into two and taught separately in two different time slots (e.g. morning and evening). Students in different time slots do not contact each other and thus $R_S$ is calculated for students in one slot.
Staggered attendance (between class)	40	1	Each class is allocated (as a whole) to either of the two different time slots and taught separately. Students in different time slots do not contact each other and thus $R_S$ is calculated for students in one slot.

182



183

184 Figure 3. Reconstruction of students' source of infection. (A) Epidemic curve stratified by the reconstructed source of  
 185 infection. The conditional probability of infection from different sources was computed for each student and aggregated by  
 186 date of illness onset. (B) Breakdown of the reconstructed source of infection. For each student, the source of infection was  
 187 sampled based on the conditional probability to provide the proportion of students infected from each source. Bars denote  
 188 posterior median and whiskers 95% credible intervals. (C) Expected relative changes in the school reproduction number  
 189 under school-based interventions changing the structure of classes. Dots represent medians and whiskers 95% credible

190 intervals. Reduced outside-class transmissions (i.e. from grademates or schoolmates) were also considered (50% reduction:  
191 blue; 90% reduction: green).

192

## 193 **Discussion**

194 We used a mathematical model that stratified transmission within and between classes/grades  
195 to understand the dynamics of influenza transmission among primary school students. The inferred  
196 transmission dynamics of seasonal influenza in Matsumoto city, Japan, 2014-15 season suggested that  
197 the within-school reproduction number  $R_S$  stayed relatively constant regardless of the size or the  
198 number of classes (suggesting ‘frequency-dependent mixing’ (36)), in contrast to common modelling  
199 assumptions. The estimated  $R_S$  of 0.8–0.9, more than half of which was attributable to within-class  
200 transmissions, is consistent with a previous study in the United States (15). This value is also in line  
201 with the reported  $R_0$  of 1.2–1.3 for seasonal influenza (37) because our previous study estimated that  
202 the students in this dataset had infected 0.3–0.4 household members on average during this 2014-15  
203 season (note that  $R_0$  corresponds to the overall number of secondary transmissions per student,  
204 including at school and household) (18). The value of  $R_S$  below 1 suggests that an outbreak cannot  
205 sustain itself within a school alone and that interactions through importing and exporting infections  
206 between households and the general community is likely to play a crucial role in the overall  
207 transmission dynamics. We estimated that school, household and community accounted for 55%, 39%  
208 and 7% of the source of infection for student cases, respectively. The attributable proportion was  
209 lower for schools and higher for households than the previous study (15), which may be explained by  
210 different scales of outbreaks in schools and households. In the Matsumoto city dataset, the overall  
211 attack ratio at school was lower (19%), students had larger households (average size 5.5) and there  
212 were more household cases than student cases (3996 vs 2548), as opposed to 35%, size of 3.4 and 141  
213 vs 129 cases in (15).

214 The estimated breakdown of  $R_S$  revealed a number of notable patterns. As the number of  
215 classes per grade increased, the contribution of within-class transmission risk declined and was

216 replaced by within-grade transmission. Combined with the almost constant overall  $R_s$ , this might  
217 indicate that contact behaviour between students that contributed to transmission was only minimally  
218 affected by the student population density. That is, students may have had a certain number of ‘close  
219 friends’ with whom they had more intimate interactions that could facilitate transmission. In a school  
220 with more classes per grade, some of such friendship may have come from grademates instead of  
221 classmates, but the total number of close friends may have remained similar. This interpretation is in  
222 line with published evidence of influenza spreading predominantly in close proximity (38) and is  
223 likely to influence the expected effect of interventions not only for influenza but also other respiratory  
224 infectious diseases including COVID-19, which share similar routes and range of transmission (39,  
225 40). Further disease-specific studies could elucidate the generalisability of these associations in more  
226 detail.

227         Our results suggested that interventions such as reducing class sizes or the number of students  
228 present (staggered attendance) may not be effective in contrast to what would be expected under the  
229 density-dependent mixing assumption (27–31). If interventions altering class structures are not  
230 accompanied by additional precaution measures and students try to resume their ‘natural’ behaviours  
231 (i.e. the same contact patterns as those in school with the resulting class structures) through so-called  
232 social contact ‘rewiring’ (41), the effect of such interventions can diminish or even reverse. For  
233 example, if other classes are absent due to staggered attendance, students may increase their  
234 interactions with classmates instead of their previous close friends in other classes. Our results are  
235 also consistent with a recent study of interventions against COVID-19 in US schools that did not find  
236 a significant risk reduction associated with reducing class sizes (42). Given the additional logistical  
237 resources required to implement these interventions, we propose that reducing the class sizes or the  
238 number of attending students should be considered only if they enable effective implementation of  
239 precaution measures such as physical distancing, environmental cleaning or forming social bubbles.

240         Using a log-linear regression analysis combined with a transmission model, we identified  
241 several precautionary measures associated with the susceptibility or infectiousness of students.  
242 Vaccines were associated with reduced susceptibility and masks with a reduction in both

243 susceptibility and infectiousness. Influenza vaccine effectiveness in the 2014-15 season was suggested  
244 to be particularly low in Japan due to vaccine mismatch and estimated to be 26% (95% CrI: 7-41%)  
245 for primary-school-age children (6-12 years old) (43). Our estimate of a relative susceptibility of 0.88  
246 (CrI: 0.81-0.96) in vaccinated students, which translates into a vaccine effectiveness of 12% (CrI: 9-  
247 19%), is broadly consistent with this prior estimate. While existing evidence for the effectiveness of  
248 mask policies for the control of respiratory infections is still limited (44, 45), our estimates of small  
249 protective effects acting on the relative susceptibility (0.77; CrI: 0.71-0.85) and infectiousness (0.65;  
250 CrI: 0.55-0.78) lie within a plausible range based on evidence available to date (44, 46-48). Increased  
251 susceptibility associated with hand washing in our analysis, however, does not align with existing  
252 findings (49, 50). Although the underlying cause for this is unclear, the original report on the  
253 Matsumoto city dataset also reported a higher odds ratio (1.4; CrI: 1.27-1.64; unadjusted for  
254 differential exposure) and attributed it to the possible congregation of students washing hands in  
255 communal settings at school (51).

256         Several limitations of this study should be noted. First, the transmission patterns within  
257 schools were estimated from a single dataset of seasonal influenza in primary schools (aged 5-12  
258 years) in Matsumoto city, Japan, and it is unclear to what extent the results can be extrapolated to  
259 other settings, e.g. secondary schools or schools in other countries. Some features of our results may  
260 still be relevant to transmission dynamics in different types of schools if they reflect general social  
261 contact behaviours of schoolchildren; however, the relative contribution of within-class/within-grade  
262 interactions may become smaller for older students (52). The data points used in the inference mostly  
263 consisted of classes of size 20-40 (those with a size smaller than 10 were excluded as they might be  
264 operated differently) and most schools had no more than 5 classes per grade. The scope of the  
265 estimated effect of the school-based interventions was also limited to within this range for internal  
266 consistency and thus may not necessarily be applicable to class structures outside this range (e.g.  
267 splitting a class of 20 students into two). Extrapolating the estimated transmission patterns to other  
268 respiratory infectious diseases also warrants caution because their epidemiological characteristics may  
269 not be identical, although we believe that such an approach may still be useful for diseases sharing

270 similar modes of transmission. Modelling studies using social contact data often assume  
271 proportionality between contacts and the transmission of directly-transmitted diseases (e.g. measles,  
272 influenza and COVID-19) and have many successful applications (7, 33, 53–57). Using the estimated  
273 transmission patterns of influenza as a proxy for other diseases essentially rests on the same  
274 assumption, which nonetheless has limitations and should eventually be validated by disease-specific  
275 studies. Second, some aspect of the outbreaks may have been missing from the dataset. Since the  
276 illness data of teachers were not available, they were not considered throughout the analysis.  
277 However, their role in seasonal influenza transmission may have been minor given a large number of  
278 student cases and the smaller risk in adults (58, 59). Although our student incidence data likely had  
279 good case ascertainment given encouraged medical attendance and confirmation by rapid diagnostic  
280 kits (18), a certain proportion of infections (e.g. asymptomatic or very mild) may have been missing.  
281 We believe that students feeling unwell due to influenza mostly attended medical institutions and  
282 received a test as it was encouraged by schools. Nonetheless, it should be noted that this could have  
283 been a source of bias in the estimated transmission patterns. Students with very mild symptoms (e.g.  
284 only slight sore throat) may visit a medical institution only if they know of other classmates also  
285 diagnosed with influenza. If such cases were common, the contribution of within-class transmissions  
286 in our results might have been an overestimate. Third, since the dataset was obtained from an  
287 observational study, the identified determinants of transmission may not be causal and should not be  
288 viewed as conclusive evidence. The results of our log-linear regression were mostly in line with  
289 existing findings, however, our dataset may still be biased due to unmeasured confounders such as  
290 health awareness. Our estimates of the relative effect of school-based interventions were based on the  
291 assumption that students' behaviours follow the fixed patterns according to the school structure even  
292 under interventions. That is, when the class size or the number of classes were changed by an  
293 intervention, students were assumed to change their behaviour according to the new school structure  
294 (as if it were the original structure) by e.g. rewiring close contacts in a timely manner. This is a  
295 hypothetical expectation that may not exactly be observed in actual interventional settings; for  
296 example, it may take time for students to resume close contacts after the class is split, which can bring  
297  $R_s$  lower than our prediction at least temporarily. We have also neglected the possible effect of the

298 interventions on the transmission outside the school. The actual effects of these interventions should  
299 ideally be validated by empirical data, as in (42).

300 Our analysis disentangled the transmission dynamics of seasonal influenza among primary  
301 school students and highlighted the relative importance of within-class and within-grade transmission.  
302 Since class and school sizes were minimally associated with the within-school reproduction number,  
303 school-based interventions that change classroom structures, e.g. reduced class sizes and staggered  
304 attendance, may have limited effectiveness. Empirical evidence on fine-grained heterogeneous  
305 transmission patterns at school as was obtained from this study would inform public health planning  
306 for future outbreaks of influenza and, potentially, other directly transmitted infectious diseases that  
307 thrive in schools.

## 308 **Materials and methods**

### 309 **Data**

310 We analysed a citywide school-based influenza survey data from the 2014/15 season. The  
311 survey was conducted in Matsumoto city (population size: 242,000 (60)), Japan, enrolling 13,217  
312 students from all 29 public primary schools in the city. During the survey period (from October 2014  
313 to February 2015), the participants were asked to fill out a questionnaire when they were back from  
314 the suspension of attendance due to diagnosed influenza (prospective survey). In March, the  
315 participants were asked to respond to another survey on their experience during the study period,  
316 regardless of whether they had contracted influenza (retrospective survey). A total of 2,548 diagnosed  
317 influenza episodes were reported in the prospective survey, which accounted for 96% of the cases  
318 officially recognised by the schools during the study period. Primary schools in Japan often requested  
319 students suspected of influenza to seek diagnosis at a medical institution. All students reporting an  
320 influenza episode in the prospective survey answered that they had received a diagnosis and at least  
321 95% of them were noticed of type A influenza (indicating that they were lab-confirmed). In the  
322 retrospective survey, 11,390 (86%) participants responded, among which 8,375 reported that they did  
323 not have influenza during the study period.

324 We combined those who responded to the prospective survey (“case group”) and those who  
325 reported no influenza experience in the retrospective survey (“control group”) and obtained a dataset  
326 of 10,923 students. Of those, 71 students from 3 schools with less than 15 students per grade were  
327 excluded because they may have different schooling patterns from other schools (e.g. some students  
328 in different grades shared classrooms). We used individual profiles (sex, school, grade, class,  
329 household composition), onset dates, influenza episodes of household members and precaution  
330 measures students engaged in (vaccine, mask, hand washing) in the subsequent analysis. Further  
331 details of the dataset can be found in the original studies (51, 61).

332 The secondary data analysis conducted in the present study was approved by the ethics  
333 committee at the London School of Hygiene & Tropical Medicine (reference number: 14599).

### 334 **Inference model**

335 We modelled within-school transmission considering class structures as follows. We defined  
336 the “school proximity”  $d$  between a pair of students  $i$  and  $j$  attending the same school as

$$d = \begin{cases} 1 & \text{(different grades, same school)} \\ 2 & \text{(different classes, same grade)} \\ 3 & \text{(different sex, same class)} \\ 4 & \text{(same sex, same class)} \end{cases} \quad (1)$$

337 To investigate the potential effect of reduced class sizes and the number of attending students, we  
338 modelled the transmission between students as a function of two variables: the class size  $n$  and the  
339 number of classes per grade  $m$  (i.e. the number of students per grade is  $nm$ ). Namely, we assumed that  
340 in the absence of any individual covariate effects, the cumulative transmission rate between student  $i$   
341 and  $j$  in proximity  $d$  over the infectious period is represented as

$$\beta_{ij} = \beta_d (n_{i,d})^{-\gamma_d} (m_{i,d})^{-\delta_d}, \quad (2)$$

342 where  $\beta_d$ ,  $\gamma_d$ ,  $\delta_d$  are parameters to be estimated. When  $i$  and  $j$  are in the same grade (i.e.  $d = 2, 3, 4$ ),  
343 the average class size and the number of classes in that grade were used as  $n_{i,d}$  and  $m_{i,d}$ . When  $d = 1$ ,  
344 the school average was used as  $n_{i,d}$  and  $m_{i,d}$ . The exponent parameters within the same class were  
345 assumed to be equal:  $\gamma_3 = \gamma_4$  and  $\delta_3 = \delta_4$ .



346 We modelled the daily hazard of incidence for student  $i$  as a renewal process. Let  $h_\tau$  be the  
347 onset-based transmission hazard as a function of serial interval  $s$  (normalised such that  $\sum_{s=1}^{\infty} h_s = 1$ ;  
348  $h_s = 0$  for  $s \leq 0$ ). We used a gamma distribution of a mean of 1.7 and a standard deviation of 1.0 for  
349 influenza, which resulted in a mean serial interval of 2.2 days (62). The daily hazard of disease onset  
350 attributed to school transmission is given as

$$\lambda_i^S(T) = v_i \sum_j w_j \beta_{ij} h_{T-T_j}, \quad (3)$$

351 where  $v_i$  and  $w_i$  represent the relative susceptibility and infectiousness, respectively, which are  
352 specified for each individual by a log-linear regression model to account for covariates (see  
353 Supplementary materials for detailed methods).

354 In addition to the above within-school transmission, we also considered within-household  
355 transmission and general community transmission. Within-household transmission was incorporated  
356 as the Longini-Koopman model (63) with parameters from a previous study on the same cohort of  
357 students (18). General community transmission was modelled as a logistic curve fitted to the total  
358 incidence in the dataset to reflect the overall trend of the epidemic. See Supplementary materials for  
359 further details of the model.

360 We constructed the likelihood function and estimated the parameters by the Markov-chain  
361 Monte Carlo (adaptive mixture Metropolis) method. We obtained 1,000 thinned samples from  
362 100,000 iterations after 100,000 iterations of burn-in, which yielded the effective sample size of at  
363 least 300 for each parameter. Using the posterior samples, we computed the proximity-specific  
364 reproduction number  $R_d$  in a hypothetical 6-year school with given  $n$  and  $m$  (assumed to be constant  
365 schoolwide) as

$$R_d = \begin{cases} 5nm \cdot \beta_1 n^{-\gamma_1} m^{-\delta_1} & (d = 1) \\ n(m-1) \cdot \beta_2 n^{-\gamma_2} m^{-\delta_2} & (d = 2) \\ n \cdot \frac{\beta_3 + \beta_4}{2} n^{-\gamma_3} m^{-\delta_3} & (d = 3, 4) \end{cases} \quad (4)$$

366 and defined the within-school reproduction number  $R_S$  as a sum of them.

367 We predicted the relative reduction in  $R_s$  under intervention measures changing the number of  
368 attending students and class structures by using posterior samples. Interventions were assumed to  
369 change  $n$  and  $m$  as shown in Table 1, and the predictive distribution of the relative change in  $R_s$  was  
370 computed for each intervention. The estimated  $R_s$  represents the value in a hypothetical condition  
371 where an infectious student spends the whole infectious period at school; the effect of absence due to  
372 symptoms or the staggered attendance was not included in this reduction.

373 All analysis was performed in Julia 1.5.2 and R 4.1.0. Replication code is available on  
374 GitHub ([https://github.com/akira-endo/schooldynamics\\_FluMatsumoto14-15](https://github.com/akira-endo/schooldynamics_FluMatsumoto14-15)).

## 375 **Acknowledgement**

376 This research was partially funded by Lnest Grant Taisho Pharmaceutical Award. AE was financially  
377 supported by The Nakajima Foundation and The Alan Turing Institute. Yang Liu is supported by Bill  
378 & Melinda Gates Foundation [INV-003174], National Institute for Health Research [16/137/109],  
379 European Commission [101003688] and UK Medical Research Council [MC\_PC\_19065]. KEA is  
380 supported by European Research Council Starting Grant [757688]. AJK [206250] and SF [210758]  
381 are supported by the Wellcome Trust.

## 382 **Conflict of interest**

383 AE received a research grant from Taisho Pharmaceutical Co., Ltd.

## 384 **Data and code availability**

385 Due to potentially sensitive information included, the original dataset is not made public and is  
386 available from the corresponding author upon reasonable request. A processed dataset with an  
387 increased level of anonymity, which can still qualitatively reproduce the main study finding (i.e.  
388 breakdown of the school reproduction number breakdown by the class/grade relationship without

389 adjustment for covariates) is publicly available along with the accompanying code on a GitHub  
390 repository ([https://github.com/akira-endo/schooldynamics\\_FluMatsumoto14-15](https://github.com/akira-endo/schooldynamics_FluMatsumoto14-15)).

## 391 **Prior publication**

392 Earlier version of this manuscript is available at Research Square [<https://doi.org/10.21203/rs.3.rs-322366/v1>].

## 394 **References**

- 395 1. N. A. Christakis, J. H. Fowler, Social Network Sensors for Early Detection of Contagious  
396 Outbreaks. *PLOS ONE* **5**, e12948 (2010).
- 397 2. O. le Polain de Waroux, *et al.*, *Identifying human encounters that shape the transmission of*  
398 *Streptococcus pneumoniae and other acute respiratory infections* (2018)  
399 <https://doi.org/10.1016/j.epidem.2018.05.008>.
- 400 3. S. Eubank, *et al.*, Modelling disease outbreaks in realistic urban social networks. *Nature* (2004)  
401 <https://doi.org/10.1038/nature02541>.
- 402 4. L. A. Meyers, M. E. J. Newman, M. Martin, S. Schrag, Applying network theory to epidemics:  
403 Control measures for Mycoplasma pneumoniae outbreaks. *Emerging Infectious Diseases* (2003)  
404 <https://doi.org/10.3201/eid0902.020188>.
- 405 5. W. P. Glezen, Emerging Infections: Pandemic Influenza. *Epidemiologic Reviews* **18**, 64–76  
406 (1996).
- 407 6. L. Wang, *et al.*, Transmission Characteristics of Different Students during a School Outbreak of  
408 (H1N1) pdm09 Influenza in China, 2009. *Sci Rep* **4**, 5982 (2014).
- 409 7. K. T. D. Eames, The influence of school holiday timing on epidemic impact. *Epidemiology and*  
410 *Infection* (2014) <https://doi.org/10.1017/S0950268813002884>.
- 411 8. K. T. D. Eames, N. L. Tilston, E. Brooks-Pollock, W. J. Edmunds, Measured Dynamic Social  
412 Contact Patterns Explain the Spread of H1N1v Influenza. *PLOS Computational Biology* **8**,  
413 e1002425 (2012).
- 414 9. S. Cauchemez, A.-J. Valleron, P.-Y. Boëlle, A. Flahault, N. M. Ferguson, Estimating the impact  
415 of school closure on influenza transmission from Sentinel data. *Nature* **452**, 750–754 (2008).
- 416 10. A. J. K. Conlan, *et al.*, Measuring social networks in british primary schools through scientific  
417 engagement. *Proceedings of the Royal Society B: Biological Sciences* (2011)  
418 <https://doi.org/10.1098/rspb.2010.1807>.
- 419 11. M. Leecaster, *et al.*, Estimates of Social Contact in a Middle School Based on Self-Report and  
420 Wireless Sensor Data. *PLOS ONE* (2016) <https://doi.org/10.1371/journal.pone.0153690>.

- 421 12. J. Fournet, A. Barrat, Contact patterns among high school students. *PLoS ONE* (2014)  
422 <https://doi.org/10.1371/journal.pone.0107878>.
- 423 13. H. Guclu, *et al.*, Social Contact Networks and Mixing among Students in K-12 Schools in  
424 Pittsburgh, PA. *PLOS ONE* **11**, e0151139 (2016).
- 425 14. J. Stehlé, *et al.*, High-resolution measurements of face-to-face contact patterns in a primary  
426 school. *PLoS ONE* (2011) <https://doi.org/10.1371/journal.pone.0023176>.
- 427 15. Simon. Cauchemez, *et al.*, Role of social networks in shaping disease transmission during a  
428 community outbreak of 2009 H1N1 pandemic influenza. *Proceedings of the National Academy  
429 of Sciences* **108**, 2825–2830 (2011).
- 430 16. V. Clamer, I. Dorigatti, L. Fumanelli, C. Rizzo, A. Pugliese, Estimating transmission probability  
431 in schools for the 2009 H1N1 influenza pandemic in Italy. *Theoretical Biology and Medical  
432 Modelling* **13**, 19 (2016).
- 433 17. M. Begon, *et al.*, A clarification of transmission terms in host-microparasite models: Numbers,  
434 densities and areas. *Epidemiology and Infection* (2002)  
435 <https://doi.org/10.1017/S0950268802007148>.
- 436 18. A. Endo, M. Uchida, A. J. Kucharski, S. Funk, Fine-scale family structure shapes influenza  
437 transmission risk in households: Insights from primary schools in Matsumoto city, 2014/15.  
438 *PLoS Computational Biology* (2019) <https://doi.org/10.1371/journal.pcbi.1007589>.
- 439 19. P. H. Thrall, A. Biere, M. K. Uyenoyama, Frequency-Dependent Disease Transmission and the  
440 Dynamics of the Silene-Ustilago Host-Pathogen System. *The American Naturalist* **145**, 43–62  
441 (1995).
- 442 20. E. S. Nightingale, O. J. Brady, C. C.-19 working Group, L. Yakob, The importance of saturating  
443 density dependence for predicting SARS-CoV-2 resurgence. *medRxiv*, 2020.08.28.20183921  
444 (2020).
- 445 21. M. J. Smith, *et al.*, Host–pathogen time series data in wildlife support a transmission function  
446 between density and frequency dependence. *PNAS* **106**, 7905–7909 (2009).
- 447 22. B. Borremans, J. Reijnders, N. Hens, H. Leirs, The shape of the contact–density function matters  
448 when modelling parasite transmission in fluctuating populations. *Royal Society Open Science* **4**,  
449 171308.
- 450 23. N. M. Ferguson, *et al.*, Strategies for mitigating an influenza pandemic. *Nature* (2006)  
451 <https://doi.org/10.1038/nature04795>.
- 452 24. N. M. Ferguson, *et al.*, Strategies for containing an emerging influenza pandemic in Southeast  
453 Asia. *Nature* **437**, 209–214 (2005).
- 454 25. L. Fumanelli, M. Ajelli, S. Merler, N. M. Ferguson, S. Cauchemez, Model-Based  
455 Comprehensive Analysis of School Closure Policies for Mitigating Influenza Epidemics and  
456 Pandemics. *PLoS Comput Biol* **12** (2016).
- 457 26. W. M. Getz, C. Carlson, E. Dougherty, T. C. Porco, R. Salter, An agent-based model of school  
458 closing in under-vaccinated communities during measles outbreaks. *SIMULATION* **95**, 385–393  
459 (2019).

- 460 27. J. Panovska-Griffiths, *et al.*, Determining the optimal strategy for reopening schools, the impact  
461 of test and trace interventions, and the risk of occurrence of a second COVID-19 epidemic wave  
462 in the UK: a modelling study. *The Lancet Child & Adolescent Health* **4**, 817–827 (2020).
- 463 28. M. J. Keeling, *et al.*, The impact of school reopening on the spread of COVID-19 in England.  
464 *Philosophical Transactions of the Royal Society B: Biological Sciences* **376**, 20200261 (2021).
- 465 29. B. Phillips, D. T. Browne, M. Anand, C. T. Bauch, Model-based projections for COVID-19  
466 outbreak size and student-days lost to closure in Ontario childcare centres and primary schools.  
467 *Sci Rep* **11**, 6402 (2021).
- 468 30. A. Best, *et al.*, The impact of varying class sizes on epidemic spread in a university population.  
469 *Royal Society Open Science* **8**, 210712.
- 470 31. A. Bilinski, J. A. Salomon, J. Giardina, A. Ciaranello, M. C. Fitzpatrick, Passing the Test: A  
471 Model-Based Analysis of Safe School-Reopening Strategies. *Ann Intern Med* (2021)  
472 <https://doi.org/10.7326/M21-0600> (June 16, 2021).
- 473 32. S. Cauchemez, *et al.*, Household transmission of 2009 pandemic influenza A (H1N1) virus in  
474 the United States. *The New England journal of medicine* (2009)  
475 <https://doi.org/10.1056/NEJMoa0905498>.
- 476 33. J. Mossong, *et al.*, Social Contacts and Mixing Patterns Relevant to the Spread of Infectious  
477 Diseases. *PLoS Medicine* **5**, e74 (2008).
- 478 34. T. Britton, T. Kypraios, P. D. O’neill, Inference for Epidemics with Three Levels of Mixing:  
479 Methodology and Application to a Measles Outbreak. *Scandinavian Journal of Statistics* **38**,  
480 578–599 (2011).
- 481 35. R. Christensen, *Log-Linear Models and Logistic Regression*, 2nd Ed. (Springer-Verlag, 1997)  
482 <https://doi.org/10.1007/b97647> (June 30, 2021).
- 483 36. M. Begon, *et al.*, A clarification of transmission terms in host-microparasite models: Numbers,  
484 densities and areas. *Epidemiology and Infection* (2002)  
485 <https://doi.org/10.1017/S0950268802007148>.
- 486 37. M. Biggerstaff, S. Cauchemez, C. Reed, M. Gambhir, L. Finelli, Estimates of the reproduction  
487 number for seasonal, pandemic, and zoonotic influenza: A systematic review of the literature.  
488 *BMC Infectious Diseases* (2014) <https://doi.org/10.1186/1471-2334-14-480>.
- 489 38. B. Killingley, J. Nguyen-Van-Tam, Routes of influenza transmission. *Influenza and Other  
490 Respiratory Viruses* **7**, 42–51 (2013).
- 491 39. R. K. Banik, A. Ulrich, Evidence of Short-Range Aerosol Transmission of SARS-CoV-2 and  
492 Call for Universal Airborne Precautions for Anesthesiologists During the COVID-19 Pandemic.  
493 *Anesthesia & Analgesia* **131**, e102–e104 (2020).
- 494 40. M. Klompas, M. A. Baker, C. Rhee, Airborne Transmission of SARS-CoV-2. *JAMA* **324**, 441  
495 (2020).
- 496 41. K. Y. Leung, F. Ball, D. Sirl, T. Britton, Individual preventive social distancing during an  
497 epidemic may have negative population-level outcomes. *Journal of The Royal Society Interface*  
498 **15**, 20180296 (2018).

- 499 42. J. Lessler, *et al.*, Household COVID-19 risk and in-person schooling. *Science* **372**, 1092–1097  
500 (2021).
- 501 43. N. Sugaya, *et al.*, Trivalent inactivated influenza vaccine effective against influenza A(H3N2)  
502 variant viruses in children during the 2014/15 season, Japan. *Eurosurveillance* **21**, 30377  
503 (2016).
- 504 44. J. Howard, *et al.*, An evidence review of face masks against COVID-19. *PNAS* **118** (2021).
- 505 45. B. J. Cowling, G. M. Leung, Face masks and COVID-19: don't let perfect be the enemy of  
506 good. *Eurosurveillance* **25**, 2001998 (2020).
- 507 46. K. Chaabna, S. Doraiswamy, R. Mamtani, S. Cheema, Facemask use in community settings to  
508 prevent respiratory infection transmission: A rapid review and meta-analysis. *International*  
509 *Journal of Infectious Diseases* **104**, 198–206 (2021).
- 510 47. B. J. Cowling, Y. Zhou, D. K. M. Ip, G. M. Leung, A. E. Aiello, Face masks to prevent  
511 transmission of influenza virus: a systematic review. *Epidemiology & Infection* **138**, 449–456  
512 (2010).
- 513 48. H. Bundgaard, *et al.*, Effectiveness of Adding a Mask Recommendation to Other Public Health  
514 Measures to Prevent SARS-CoV-2 Infection in Danish Mask Wearers. *Ann Intern Med* **174**,  
515 335–343 (2021).
- 516 49. J. Xiao, *et al.*, Nonpharmaceutical Measures for Pandemic Influenza in Nonhealthcare  
517 Settings—Personal Protective and Environmental Measures - Volume 26, Number 5—May  
518 2020 - Emerging Infectious Diseases journal - CDC <https://doi.org/10.3201/eid2605.190994>  
519 (June 15, 2021).
- 520 50. A. E. Aiello, *et al.*, Mask Use, Hand Hygiene, and Seasonal Influenza-Like Illness among  
521 Young Adults: A Randomized Intervention Trial. *The Journal of Infectious Diseases* (2010)  
522 <https://doi.org/10.1086/650396>.
- 523 51. M. Uchida, *et al.*, Effectiveness of vaccination and wearing masks on seasonal influenza in  
524 Matsumoto City, Japan, in the 2014/2015 season: An observational study among all elementary  
525 schoolchildren. *Preventive Medicine Reports* **5**, 86–91 (2017).
- 526 52. H. Guclu, *et al.*, Social Contact Networks and Mixing among Students in K-12 Schools in  
527 Pittsburgh, PA. *PLOS ONE* **11**, e0151139 (2016).
- 528 53. L. Munasinghe, Y. Asai, H. Nishiura, Quantifying heterogeneous contact patterns in Japan: a  
529 social contact survey. *Theoretical Biology and Medical Modelling* **16**, 6 (2019).
- 530 54. S. Funk, *et al.*, Combining serological and contact data to derive target immunity levels for  
531 achieving and maintaining measles elimination. *BMC Medicine* **17**, 180 (2019).
- 532 55. D. M. Feehan, A. S. Mahmud, Quantifying population contact patterns in the United States  
533 during the COVID-19 pandemic. *Nat Commun* **12**, 893 (2021).
- 534 56. C. I. Jarvis, *et al.*, Quantifying the impact of physical distance measures on the transmission of  
535 COVID-19 in the UK. *BMC Medicine* **18**, 124 (2020).
- 536 57. J. D. Munday, *et al.*, Estimating the impact of reopening schools on the reproduction number of  
537 SARS-CoV-2 in England, using weekly contact survey data. *medRxiv*, 2021.03.06.21252964  
538 (2021).

- 539 58. J. Brugger, C. L. Althaus, Transmission of and susceptibility to seasonal influenza in  
540 Switzerland from 2003 to 2015. *Epidemics* **30**, 100373 (2020).
- 541 59. P. Arevalo, H. Q. McLean, E. A. Belongia, S. Cobey, Earliest infections predict the age  
542 distribution of seasonal influenza A cases. *eLife* **9**, e50060 (2020).
- 543 60. Matsumoto City, Population by year (1990–2017) (June 29, 2021).
- 544 61. M. Uchida, *et al.*, Prospective epidemiological evaluation of seasonal influenza in all elementary  
545 schoolchildren in Matsumoto city, Japan, in 2014/2015. *Japanese Journal of Infectious Diseases*  
546 (2017) <https://doi.org/10.7883/yoken.JJID.2016.037>.
- 547 62. M. A. Vink, M. C. J. Bootsma, J. Wallinga, Serial Intervals of Respiratory Infectious Diseases:  
548 A Systematic Review and Analysis. *American Journal of Epidemiology* **180**, 865–875 (2014).
- 549 63. I. M. Longini, J. S. Koopman, Household and community transmission parameters from final  
550 distributions of infections in households. *Biometrics* **38**, 115–126 (1982).
- 551