Simulating COVID-19 classroom transmission on a university campus

Arvin Hekmati[a,1] , Mitul Luhar[b] , Bhaskar Krishnamachari[a,c], and Maja Matarić[a]

We study the airborne transmission risk associated with holding in-person classes on university campuses for the original strain and a more contagious variant of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). We adopt a model for airborne transmission risk in an enclosed room that considers room properties, mask efficiency, and initial infection probability of the occupants. Additionally, we study the effect of vaccination on the spread of the virus. The presented model has been evaluated in simulations using fall 2019 (prepandemic) and fall 2020 (hybrid instruction) course registration data of a large US university, allowing for assessing the difference in transmission risk between in-person and hybrid programs and the impact of occupancy reduction, mask-wearing, and vaccination. The simulations indicate that without vaccination, moving 90% of the classes online leads to a 17 to 18$\times$ reduction in new cases, and universal mask usage results in an ~2.7 to 3.6$\times$ reduction in new infections through classroom interactions. Furthermore, the results indicate that for the original variant and using vaccines with efficacy greater than 90%, at least 23% (64%) of students need to be vaccinated with (without) mask usage in order to operate the university at full occupancy while preventing an increase in cases due to classroom interactions. For the more contagious variant, even with universal mask usage, at least 93% of the students need to be vaccinated to ensure the same conditions. We show that the model is able to predict trends observed in weekly infection rates for fall 2021.

COVID-19 | airborne transmission | vaccination

The COVID-19 pandemic (1) has had a profound impact on educational institutions around the world. More than 85 colleges and universities across the United States have reported at least 1,000 cases of COVID-19, and over 680 institutions have reported at least 100 cases (2). More than 124,000 public and private schools in the United States closed in April 2020, impacting more than 55 million students (3). Worldwide, similar disruptions affected more than 1.7 billion students (4, 5). Across the US educational system as a whole, there were over 397,000 confirmed cases of COVID-19 and more than 90 deaths as of December 2020.

In response to the initial COVID-19 outbreak in spring 2020, a large number of colleges and universities across the United States decided to cancel in-person instruction and close student housing (5). Many universities and colleges moved instruction online. The transition from in-person classes to online instruction brought many challenges for both students and instructors. For students, the transition to online instruction exacerbated challenges associated with access to technology and led to concerns regarding absenteeism and accommodation of special needs (5). For instructors, the transition to online instruction led to increased concerns regarding student engagement and evaluation (6, 7).

In fall 2020, many institutions of higher education in the United States returned to in-person instruction. However, this led to a significant increase in new infections. Several colleges and universities decided to reopen in the fall 2020 and provide hybrid classes where a portion of the students could attend the classes in person while the others attended online, providing a partial solution to the problems associated with purely online instruction. In addition, colleges and universities put in place rules about physical distancing, face covering usage, and limited social gatherings. Many institutions also put in place extensive population testing, contact tracing, and quarantining measures for on-campus students, staff, and faculty. This combination of measures had some success in curbing the spread of COVID-19 on campuses (2).

More recently, the availability of effective vaccines has enabled a significant reopening of society in countries with high vaccination rates, including a return to in-person instruction at many educational institutions in the United States. However, there is still significant value in studying the impact of vaccinations due to a combination of factors, such as the

Significance

This paper simulates the spread of COVID-19 at universities via airborne transmission in classroom settings. The transmission risk model used for these simulations accounts for student-specific class schedules, classroom sizes and occupancy, and ventilation rates, as well as vaccination rate and efficacy. We show the simulations reproduce trends observed in weekly infection rates at a large US university. We also evaluate the impact of campus operational policies. Model predictions show moving 90% of classes online can reduce new infections by as much as 18$\times$, and universal mask usage can reduce new infections by up to 3.6$\times$. For full-time in-person instruction, high vaccination rates are predicted to curb transmission even for more contagious variants of severe acute respiratory syndrome coronavirus 2.

[1]To whom correspondence may be addressed. Email: hekmati@usc.edu.

limited availability of vaccines in many parts of the world, eligibility restrictions, vaccine hesitancy, variations in vaccination mandates, and the emergence of more contagious variants of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) (8–11).

There have been a few previous studies evaluating the relative effectiveness of different measures (e.g., masking, reduced occupancy, and testing) on COVID-19 spread (12–16). For example, Cashore et al. (12) used simulations to analyze virus spread at Cornell University using the susceptible-exposed-infectious-removed (SEIR) model and predict outcomes for a full return of students, faculty, and staff in the 2020 fall semester over a 16-wk time period. They simulated several interventions, such as regular PCR testing for individuals, contact tracing, and quarantining for infected persons. The simulations generated predictions for the number of people that came into close contact with infectious persons, became infected, and required hospitalization. In a subsequent study, the authors extended their work to better estimate the number of contacts per day among members of the campus community and analyzed the effects of high contact rates and noncompliance with testing requirements on the number of subsequent infections (13). Ying and O'Clery (17) present an agent-based model of customer movements in a supermarket to assess the various mitigation methods that have been used in supermarkets using a synthetic dataset. Foster and Kinzel (18) studied SARS-CoV-2 airborne transmission risk in classroom settings, considering various ventilation and masking conditions by combining Wells–Riley and computational fluid dynamics models. However, they considered transmission risk in one sample classroom, and they did not analyze the cumulative effects of students attending multiple classrooms. Furthermore, their model did not consider the effect of vaccination.

This paper uses a previously established modeling framework (19) that studied the impact of different university policies on the transmission of the original strain of SARS-CoV-2 through classroom interactions at universities by considering parameters such as occupancy, mask usage, and initial infection rates for students and instructors. In this paper, we incorporate the impact of vaccinations into the modeling framework and evaluate how vaccination rates and vaccine efficacy impact the spread of the virus on a university campus. The simulations in this paper also consider the case of a more transmissible variant of SARS-CoV-2, with properties similar to the Delta variant. We compare model predictions with the real-world infection rates at a large US university to evaluate predictive capability.

Results and Discussion

This section presents model predictions for the impact of mask usage, occupancy reduction, and vaccination on the spread of the SARS-CoV-2 virus due to classroom transmission. We used course registration data from a large university for fall 2019 and fall 2020. The fall 2019 dataset represents the normal operation of the university without any online/hybrid scheduling. Fall 2020 represents the hybrid instruction scheduling of the university where 90% of the courses moved online and the rest of the courses were held in a hybrid mode on the university campus.

In order to better visualize the relation between courses in the dataset that we used, we present conflict graphs for the fall 2019 and fall 2020 datasets. In these graphs, nodes are courses, and edges between two courses show that there are at least k students in common between them. Fig. 1 shows conflict graphs for fall 2019 and fall 2020 datasets for the case of $k = 5$. We used $k = 5$ in this figure for simplicity since lower values for k would show similar

features but result in a denser graph that is harder to visualize. For fall 2019, the courses are densely linked with each other, implying that the virus can spread more widely in the student population. On the other hand, for fall 2020 the courses show sparser links, implying that virus spread in the hybrid schedule is likely to be diminished considerably.

The classroom risk model in this study makes use of information specific to each classroom interaction (i.e., room size, ventilation rate, occupancy, lecture duration, and number of students). Further, infection probabilities for each registered student are calculated by considering their specific course schedule. The average infection probabilities for the university as a whole are then calculated by combining the individual infection probabilities. Information regarding classroom physical and operational properties was available from the university (*Dataset*). The model also accounts for variability in virus emission rates from instructors and students due to the differences in vocalization activity (i.e., lecturing vs. listening) and the usage of masks. We further consider variations in vaccination rate and vaccine effectiveness in preventing breakthrough infections, as well as the reduction in virus emissions from vaccinated individuals with breakthrough infections.

Further details regarding the model are provided in *Materials and Methods*. We list a few key assumptions below. The results consider scenarios with both full classroom occupancy ($\alpha = 1$ or 100%) and reduced classroom occupancy ($\alpha = 0.2$ or 20%) for in-person classes. For the fall 2020 dataset, occupancy is set to 20% for all hybrid classes. Instructors in the classroom are assumed to be active (i.e., talking) 90% of the time, while students are assumed to be active 5% of the time. Following previous works, we assume virus emissions are significantly higher during vocalization (20–22). We also assume a 10-fold increase in viral emissions when generating predictions for the more contagious variant, based on numbers reported for the Delta variant of SARS-CoV-2 (23, 24). For cases in which mask usage is considered, we assume the virus filtering efficiency for both inhalation (f) and exhalation ($\hat{f}$) to be $f = \hat{f} = 0.5$ or 50% based on results reported for typical cloth coverings (25). To simplify notation, we assume $f = \hat{f}$ and only use f in the remainder of this paper. Regarding vaccines, we assume an effectiveness of $\beta = 0.9$ (90%) at preventing infection for the original SARS-CoV-2 strain (26) consistent with available data for the messenger RNA (mRNA) BNT162b2 (Pfizer-BioNTech) and mRNA-1273 (Moderna) COVID-19 vaccines. We assume that effectiveness reduces to $\beta = 0.64$ (64%) for the more contagious variant (24, 27). For cases in which a vaccinated individual has a breakthrough infection, we assume that the viral load is reduced by a factor of 4. The vaccination rate in the university population, γ, is varied systematically.

Two different metrics are used to evaluate classroom transmission of COVID-19. The first is the predicted number of new infections after 1 wk of classes, assuming an initial infection probability of 0.01 in the population. Clearly, the absolute number of new infections must be treated with caution since it is dependent on the initial infection probability. However, this metric serves as a useful measure of the relative efficacy of different interventions under both normal and hybrid operation. The second metric is the effective reproduction number R_0^e, defined as the ratio of the infection probability after 1 wk of classes to the initial infection probability (0.01). Since 1 wk is roughly the time scale over which COVID-19 progresses, we see this metric as being analogous to the basic reproduction number (R_0) that is often interpreted as the number of secondary infections caused by an infectious individual.
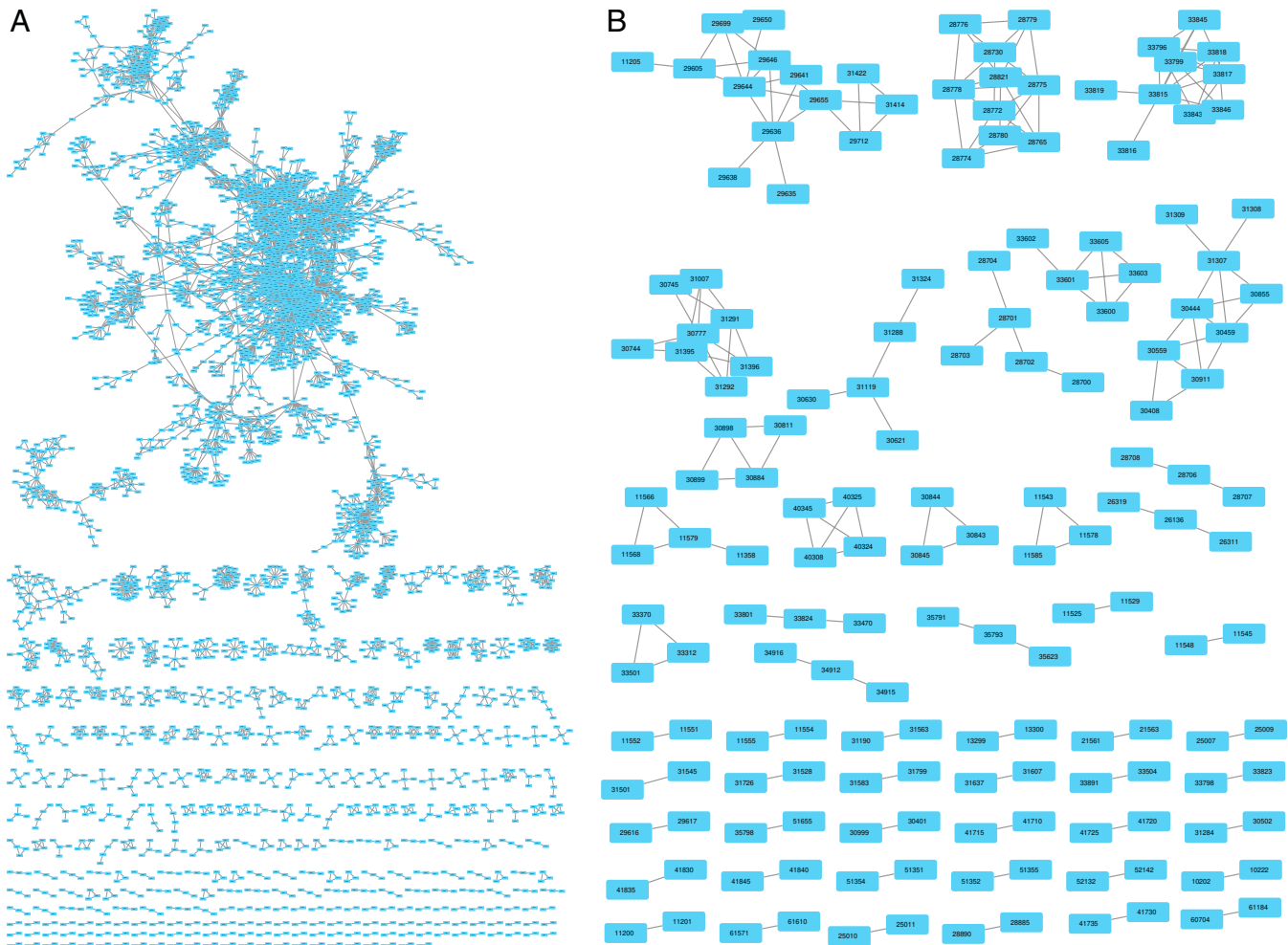
Fig. 1. Dataset visualization: conflict graphs for each semester. (*A*) Fall 2019. (*B*) Fall 2020.

Fig. 2 shows the effective reproduction number specific to individual courses for the fall 2019 dataset, assuming prepandemic operation, i.e., full occupancy, no masking, and no vaccination. For most courses, model predictions indicate that $R_0^e < 5$. As expected, predicted transmission risk is highest ($R_0^e > 20$) in courses characterized by long-duration interactions (e.g., laboratories or studios) in classrooms with low air change rates.

Table 1 provides a summary of new infections and effective reproduction numbers (together with 95% confidence intervals; *Materials and Methods*) for a subset of the no-vaccination scenarios considered. The simulations predict a significant reduction in new infections for fall 2020 compared to fall 2019 across all masking and occupancy conditions. In the case of the original SARS-CoV-2 variant, the results indicate that universal mask usage results in $3.6\times$ reduction in new infections through classroom interactions for both the fall 2019 and fall 2020 datasets. Reducing class occupancy to 20% by switching to hybrid instruction results in 2.2 to $2.3\times$ further reduction in new infections. Importantly, the transition to having 90% of the courses online from fall 2019 to fall 2020, even without mask usage or reduced occupancy, results in an $18\times$ reduction in predicted new cases. It is interesting to note that our simulations predict that mask usage (even at 50% effectiveness) leads to a greater reduction in new cases compared to a reduction in occupancy to 20%.

For the more contagious variant, the results indicate that universal mask usage results in 2.8 to $2.9\times$ reduction in new infections through classroom interactions. A reduction in class

occupancy to 20% results in $2.5\times$ further reduction in new infections. Similar to predictions for the original virus strain, the transition to having 90% of the courses from fall 2019 to fall 2020 alone leads to a $17\times$ reduction in new cases. Importantly, even though the relative reduction in cases due to mask usage, reduced occupancy, and online instruction is similar between the original strain and the contagious variant, the more contagious variant leads to a significantly higher number of new cases in absolute terms.
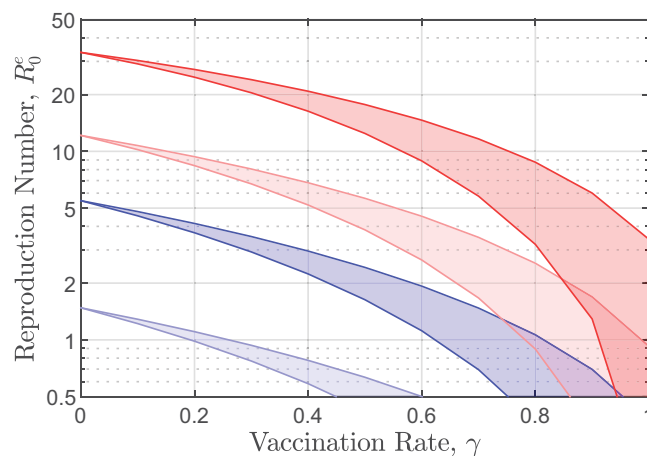
Also shown in Table 1 are estimates for the effective reproduction number (R_0^e) for the different scenarios considered. The relative impact of mask usage and reduced occupancy on R_0^e is similar to the predicted impact on the number of new infections. However, the reduction in R_0^e from the fall 2019 to fall 2020 dataset is less pronounced. For example, the simulations predict a $3.4\times$ reduction in R_0^e compared to an $18\times$ reduction in new cases for the original variant in the absence of masking and reduced occupancy. This is because the R_0^e parameter does not account for the purely online student population; it just represents the ratio of infections for the population attending classes. Estimates for R_0^e are significantly higher for the more contagious variant of the virus.

Next, we consider the effect of vaccines on the effective reproduction number. Predictions from a number of different simulations with varying vaccination rates, vaccine efficacy, and mask usage are summarized in Fig. 3. As expected, an increase in vaccination rates and universal masking leads to a significant
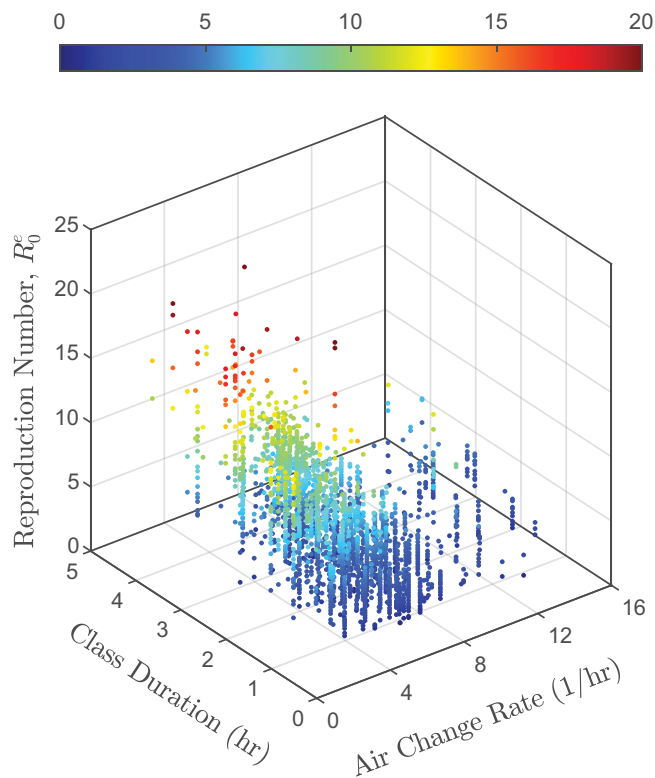
Fig. 2. Course-specific R_0^e plotted against classroom air change rate and class duration. For clarity, marker color also represents R_0^e.

Fig. 3. Estimates for R_0^e plotted against vaccination rates (γ) for the fall 2019 dataset, representing normal university operations. Red patches show the limits for the more contagious Delta variant with (lighter) and without (darker) masks; blue patches show the same information for the original variant. The upper and lower bounds of the patches represent limits where the vaccines are 50 and 90% effective, respectively.

reduction in the risk of transmission on campus and R_0^e. In Table 2, we present the critical vaccination ratio required to prevent an increase in the number of new infections on campus, i.e., to ensure $R_0^e \leq 1$. To arrive at these estimates, we increase the vaccination ratio, γ, in increments of 0.01 and calculate the corresponding value of R_0^e together with 95% confidence intervals. These predictions are used to identify the range of γ values that ensure $R_0^e \leq 1$. For the original strain and with universal masking, the upper-bound prediction for γ in Table 2 indicates that a 23% vaccination rate is sufficient to keep $R_0^e \leq 1$. Without mask usage, the critical vaccination threshold increases to 64%. On the other hand, the increase in viral emissions and reduction in vaccine efficacy for the more contagious, Delta-like variant leads to a much higher critical vaccination threshold of 93% even with universal masking. Without mask usage, even a 100% vaccination rate is not sufficient to limit R_0^e and contain the spread of COVID-19 on

campus. Note that the critical vaccination ratios presented above could be interpreted as the minimum bound for herd immunity due to the fact that we only consider the classroom transmission, while the spread could be higher if out-of-class interactions are taken into account.

Fig. 4 compares model predictions for infection probability at the university against real infection data for fall 2021. For fall 2021, we assume near-normal operation with a similar schedule to fall 2019. We predict infection probabilities for the students using registration information from fall 2019. However, we assume universal mask usage and consider vaccination rates ranging from $\gamma = 0.85$ to $\gamma = 0.95$. The actual vaccination rate is expected to fall within this range. Given the prevalence of the Delta variant for fall 2021, we further assume conditions corresponding to the more contagious variant; i.e., we assume higher airborne virus emissions and reduced vaccine effectiveness, $\beta = 0.64$ (24, 27). Note that the model predictions shown in Fig. 4 make use of real infection rate from the prior week and propagate infection probabilities forward through 1 wk of classroom interactions; i.e., they represent a weekly update. Model predictions closely follow the trends observed in the real infection rate data. Real infection rates fall between model prediction for vaccination rates $\gamma = 0.90$ and 0.95.

Table 1. Number of new infections and effective reproductive number, R_0^e, for the original and more contagious variants of SARS-CoV-2 in the absence of vaccination

	Fall 2019 schedule				Fall 2020 schedule			
	$\alpha = 1$		$\alpha = 0.2$		$\alpha = 1$		$\alpha = 0.2$	
	$f = 0$	$f = 0.5$	$f = 0$	$f = 0.5$	$f = 0$	$f = 0.5$	$f = 0$	$f = 0.5$
Original strain								
New infections over 1 wk	1860 ± 80	500 ± 40	770 ± 50	220 ± 30	100 ± 20	30 ± 10	40 ± 10	10 ± 7
R_0^e	5.5 ± 0.2	1.5 ± 0.1	2.3 ± 0.2	0.63 ± 0.08	1.6 ± 0.3	0.4 ± 0.2	0.7 ± 0.2	0.2 ± 0.1
More contagious variant								
New infections over 1 wk	11400 ± 200	4100 ± 100	4200 ± 100	1630 ± 80	670 ± 50	230 ± 30	230 ± 30	90 ± 20
R_0^e	33.5 ± 0.5	12.1 ± 0.3	12.4 ± 0.3	4.8 ± 0.2	10.5 ± 0.7	3.6 ± 0.4	3.7 ± 0.4	1.5 ± 0.3

f represents mask filtration efficiency, and α represents occupancy.

Table 2. Critical vaccination rate (γ) required to have $R_0^e = 1$ under normal operation (for the fall 2019 course registration dataset)

	Original strain ($\beta = 0.9$)	More contagious variant ($\beta = 0.64$)
No mask, $f = 0$	0.62 ± 0.02	—
With mask, $f = 0.5$	0.19 ± 0.04	0.92 ± 0.01

This assumes a vaccination effectiveness (β) of 90% for the original strain and 64% for the Delta variant.

In summary, our results show that while moving classes online can curb the spread of infection, universal masking and vaccination mandates can also make a significant impact. Further, a comparison with real-world data shows that the model is able to generate reasonable predictions for infection rates. However, it must be emphasized that model predictions for transmission risk presented here are limited to classroom interactions alone. We do not consider extracurricular activities or social interactions. Further, we do not model the impact of any population testing, contact tracing, or quarantine policies implemented at the university.

Materials and Methods

The methods used in this paper involve the development of a suitable model and simulations performed using that model on a dataset. We describe in the following subsections our model and the dataset used.

General Risk Model. Airborne transmission has been shown to be central to the spread of SARS-CoV-2 in enclosed spaces (21, 28–30). Airborne transmission is defined as transmission via inhalation of aerosols or small droplets that can remain suspended for a long time (31). Aerosols also include droplet nuclei produced through the process of rapid desiccation of exhaled respiratory droplets (32). Spatial distancing, face covering, and hand hygiene guidelines are known to be effective at limiting transmission via larger droplets and direct contact. However, even with these precautionary measures, airborne transmission can be occur in enclosed spaces with a large number of occupants and over extended periods of contact, i.e., in classroom settings.

In this paper, we make use of a well-mixed indoor air room model for airborne virus emission and exposure in an enclosed space. The goal is to provide an estimate for airborne virus concentrations and dosage for a known number of occupants and duration of contact. Airborne virus concentrations will depend on the number of infectious persons in the room and whether the occupants are being active or passive, as well as any mitigating factors such as the use of face

coverings; enhanced heating, ventilating, and air-conditioning (HVAC) protocols; and limited occupant density due to spatial distancing. The dosage for exposed individuals will further depend on the duration of contact and the effectiveness of any face coverings. The main assumption in this model is that it considers perfectly mixed conditions in the room, i.e., a uniform concentration of virus particles. The model assumes that any airborne particles are mixed throughout the space quickly and perfectly. This assumption implies a uniform transmission risk for all occupants in the room. Such mixed flow or continuously stirred reactor models are common in indoor air quality modeling (33, 34). In the context of airborne disease transmission, such models are typically referred to as Wells–Riley models after pioneering studies in this field (35–37). Recent efforts have used such models to calculate SARS-CoV-2 airborne transmission risks for specific commercial interactions or case studies (21, 22). A more complete version of the model used in this paper has also been developed and made publicly available by Jimenez (38).

We recognize that the perfectly mixed assumption is a significant oversimplification since virus concentrations are higher closer to the source of emission (i.e., an infectious person) and in the air jet created by talking or vocalizing. Prevailing air currents from HVAC systems and any open doors or windows can also lead to localized hot spots. Consequently, some occupants in the room will be subject to higher exposure than predicted by the model, and others will be subject to lower exposure. Nevertheless, an increase in the exposure predicted by the model should be considered indicative of higher transmission risk overall. Higher-fidelity computational fluid dynamics simulations and/or experiments that capture the intricacies of droplet and aerosol transport from respiratory emissions in the presence of background airflow should be pursued wherever possible (39, 40). However, since running such complex simulations or experiments is not feasible for all potential scenarios, this model provides useful first-order estimates of airborne transmission risk in enclosed spaces.

In this model, we consider an enclosed room of volume v (m^3) with volumetric airflow rate through the HVAC system Q_{hvac} (m^3/s), such that the number of air changes per unit time in the room is $E_{hvac} = Q_{hvac}/v$ (s^{-1}). The total number of people in the room is n. Each person is assumed to inspire and expire with an airflow rate of Q (m^3/s) on average. The probability of a person being initially infected is q_i, and the virion emission rate for an infected person is E_i^a (virions/s) if the person is being active (e.g., lecturing loudly) and E_i^p (virions/s) if the person is being passive (e.g., listening quietly). A virion is a single infectious virus particle. The probability of a person being active is p_a. The probability of a person being passive is therefore $p_p = 1 - p_a$. We further assume the use of face coverings with the filtration efficiency f and $\hat{f}$ for inhalation and exhalation, respectively. If we assume $f = \hat{f}$, we can replace $\hat{f}$ with f in the following equations. Finally, we assume that the occupants remain in the enclosed room for the duration of T. Using these parameters, we can estimate the average airborne virus concentration C (virions/m^3) using a room-scale mass balance as

$$C(n, q_i, p_a) = (1 - \hat{f}) \frac{n}{E_{hvac} v} q_i \left(p_a E_i^a + (1 - p_a) E_i^p \right) \quad \textbf{[1]}$$

and estimate the average virus dose D (virions) for an occupant as

$$D(n, q_i, p_a) = C(n, q_i, p_a)(1 - f)QT$$
$$= (1 - \hat{f})(1 - f) \frac{n}{E_{hvac} v} q_i \left(p_a E_i^a + (1 - p_a) E_i^p \right) QT. \quad \textbf{[2]}$$

Note that the above equations assume steady state conditions. Further, for simplicity, this formulation neglects virus removal due to settling and the decay in the number of viable (or infectious) virus particles over time. The primary sink of virions is assumed to be the airflow through the HVAC system. These assumptions are reasonable given the high air change rates in the classrooms (see *Dataset*
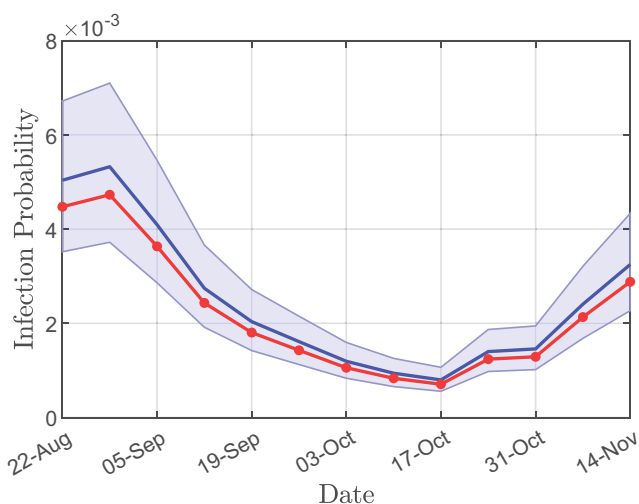
Fig. 4. Model predictions for infection probability compared with real data (red line with markers). The shaded blue region shows predictions for vaccination rates ranging from $\gamma = 0.85$ to $\gamma = 0.95$. The solid blue line shows predictions for $\gamma = 0.90$.

description). Specifically, the average air change rate for the classrooms in the dataset is 6.1 changes per hour, which translates into an air change time scale of $\sim$10 min. Per Stokes' law, the settling velocity for a 5-μm aerosol with density comparable to that of water is expected to be $V_s \approx 8 \times 10^{-4}$ m/s. For a typical classroom of height $h = 2.7$ m ($\sim$9 ft), this yields a settling time scale of $h/V_s \approx$ 3400 s, or about 1 h. There is significant variability in estimates for how long SARS-CoV-2 remains viable in aerosols or droplet nuclei; prior studies suggest a half-life of $\sim$1.1 to 1.2 h (41). In other words, settling and inactivation time scales for airborne SARS-CoV-2 are expected to be almost an order of magnitude higher compared to the time scale associated with air exchange.

The virus emission rates E_i^a and E_i^p can be estimated based on known virus concentrations and aerosol volumes for typical active and passive activities (20, 21, 42). For instance, Stadnytskyi et al. (20) estimated that 25 s of active or loud speaking led to the emission of between 60 and 320 nL of oral fluid. The viral load in the sputum for the original variant was estimated to be $c_v \approx 7 \times 10^6$ RNA copies cm^{-3}, although it may be as high as $O(10^9)$ RNA copies cm^{-3} (43). Based on these estimates, and assuming that the number of virions is similar to the number of RNA copies (44), the virus output for an active, infected person is expected to be $E_i^a \approx 17 - 90$ virions/s for the original variant. Further, the data presented in Buonanno et al. (21) suggest that virus emissions are roughly 40 times higher while speaking when compared to resting conditions. Assuming $E_i^a/E_i^p \approx 40$, virus emissions from passive persons are expected to be $E_i^p \approx 0.4 - 2.3$ virions/s. For the more contagious Delta variant, we assume that the average viral emissions of an infected person are 10$\times$ higher than for the original variant, based on recent reports (23, 24)

Given the average virus dosage, we can calculate the infection probability for one individual in the room after the encounter as

$$P_i = 1 - e^{-\frac{D(n, q_i, p_a)}{D_0}}, \qquad [3]$$

where D_0 is the dose that leads to transmission in roughly 63% of cases (21, 36). The exponential mapping used to translate virus dose into a transmission probability implicitly accounts for the variation in physiological responses to the same exposure as well as the room-scale variation in exposure that the well-mixed model neglects (i.e., arising from concentration hot spots). To our knowledge, the infectious dose for SARS-CoV-2 remains uncertain, but previous estimates for SARS-CoV-1 and Influenza A suggest that 300 to 800 virions are needed to cause infection in 50% of the population (45, 46). If the infectious dose is $D_0 = 1000$ virions, the respiratory emission estimates provided above suggest that active infectious persons with $E_i^a \approx 17 - 90$ virions/s can emit roughly 60 to 320 infectious doses per hour, while passive persons can emit roughly 1.8 to 7 infectious doses per hour. These ranges are consistent with the estimates provided in previous studies (21, 22, 40), suggesting that infectious persons undergoing light activity and talking can generate over 100 quanta per hour, where a quantum is defined as the dose required to cause infection in 63% of susceptible persons (36). We recognize that there is significant variability in our estimates for both virus emissions and infectious dose. As a result, any predictions for absolute infection risk must be treated with caution. Nevertheless, predictions generated using the physics-based model presented in this section should still provide useful estimates for relative risk under different scenarios.

The use of a well-mixed indoor air model is expected to yield robust predictions for classroom interactions in which masking is universal. However, the perfect-mixing assumption can underestimate airborne transmission risk, especially for shorter-range unmasked interactions, such as those expected in some social situations (e.g., dining). This is because masks serve to filter pathogens and also disrupt respiratory jets (47), causing additional mixing. A more complete model that accounts for the effect of respiratory jets, varying aerosol and droplet size distributions, and the relative contributions of various removal mechanisms to quantify indoor transmission risk for COVID-19 can be found in Bazant and Bush (40). Recent studies show that the use of surrogate measurements for respiratory activity (e.g., from carbon dioxide monitors) can also be used to evaluate the risk of airborne SARS-CoV-2 transmission in indoor settings (48).

Classroom Risk Model. Next, we take the model presented above and adapt it to consider classroom interactions. A classroom is assumed to be occupied by instructors (teachers) and students. Instructors are more likely to be active (i.e., lecturing) during a class, while students are more likely to be passive (i.e.,

listening). Therefore, to better model classroom interactions, we assume different activity levels for instructors and students, and we also consider the effect of differing initial infection probabilities for instructors and students. Specifically, we assume instructors and students have activity rates of p_a^t and p_a^s, respectively. Similarly, we assume instructors and students have initial infection probabilities q_i^t and q_i^s, respectively. We further assume that we have 1 instructor and N^s students in a classroom attending the class in person, given by

$$N^s = \alpha n, \qquad [4]$$

where n is the total number of students enrolled in the class and α is the occupancy ratio of the students who attend the class in person. The average viral dose from m infected students in a classroom can be calculated as

$$D_m^s = D(m, 1, p_a^s), \qquad [5]$$

and the average viral dose from one infected instructor can be calculated as

$$D^t = D(1, 1, p_a^t). \qquad [6]$$

The infection probability for a student in a given classroom after one session, for the case that the instructor and m students are initially infected, is given by

$$P_i(1, m) = 1 - e^{-\frac{D^t + D_m^s}{D_0}}. \qquad [7]$$

The infection probability for a student in the case that the instructor is not initially infected but m students are initially infected is given by

$$P_i(0, m) = 1 - e^{-\frac{D_m^s}{D_0}}. \qquad [8]$$

To find the total infection probability for a student in a class session, we have to first compute the probability that m students out of N^s will be infectious given the initial infection probability of 1 student as q_i^s. For this purpose, we use the following binomial probability:

$$p_i^s(m, N^s) = \binom{N^s}{m}(q_i^s)^m (1 - q_i^s)^{N^s - m}. \qquad [9]$$

Finally, to calculate the total infection probability for any one student after one class meeting given the initial infection probability for the instructor, q_i^t, and the students, $p_i^s(m, N^s)$, we have

$$P_{i,class}^s = \sum_{m=0}^{N^s} p_i^s(m, N^s)[q_i^t P_i((1, m)) + (1 - q_i^t)P_i(0, m)]. \qquad [10]$$

Using the above formulation, we now formally define three important metrics for assessing the impact of classroom interactions over 1 wk: 1) $P_{i,week}^s$, the individual infection probability for students[*]; 2) $\hat{P}_{i,week}^s$, the average infection probability; and 3) R_0^e, the effective reproduction number. We define these metrics below.

We present these metrics with reference to the time period of 1 wk in part as a modeling choice–our analysis and simulations could be carried out for any time period. However, 1 wk is a natural time scale to focus on for two reasons. First, class schedules repeat weekly. Second, COVID-19 symptoms appear at most after 2 wk, and on average, symptomatic patients show symptoms after 1 wk (49). We assume individuals would not attend classes once they are symptomatic.

If a particular student j attends k classes with infection probabilities of $p_1, p_2, \ldots, p_k$, the individual infection probability for that particular student after attending 1 wk of classes is

$$P_{i,week}^s(j) = 1 - (1 - p_1)^{n_1}(1 - p_2)^{n_2} \cdots (1 - p_k)^{n_k}, \qquad [11]$$

where n_i is the number of sessions for the class i. Since infection for student j can be treated as a Bernoulli event with probability $P_{i,week}^s(j)$, the expected number of new infected students after holding 1 wk of classes is

$$N_{i,week}^s = \sum_{j=1}^{N^s} P_{i,week}^s(j), \qquad [12]$$

---

[*]For simplicity and ease of exposition, we focus only on students because we assume a high student–faculty ratio at the university.

and the estimated SD in the number of infected students after a week is defined as

$$\sigma_{i,week}^s = \sqrt{\sum_{j=0}^{N_s} P_{i,week}^s(j)[1 - P_{i,week}^s(j)]}.$$ **[13]**

Assuming a normal distribution for the sum, the 95% confidence interval in the number of new infections can therefore be estimated as $N_{i,week}^s \pm \Delta_{i,week}^s$, where

$$\Delta_{i,week}^s = 1.96\, \sigma_{i,week}^s.$$ **[14]**

We then define the average infection probability after 1 wk of classes as

$$\hat{P}_{i,week}^s = \frac{N_{i,week}^s}{N^s}.$$ **[15]**

Since N^s is deterministic, the 95% confidence interval for the average infection probability is $\hat{P}_{i,week}^s \pm (\Delta_{i,week}^s / N^s)$.

A well-known parameter for infection spread in epidemics is R_0, referred to as the reproductive number, which indicates the average number of individuals infected by one initially infected individual in a population. For classroom interactions over the course of a week, we can define a similar ratio of new cases to initial cases by taking the ratio of infection probabilities before and after the week. We thus define the effective reproductive number R_0 from 1 wk of operating classes as

$$R_0^e = \frac{\hat{P}_{i,week}^s}{q_i^s}.$$ **[16]**

In order to calculate the confidence interval for R_0^e, given that we assume the initial infection probability of the students (i.e., q_i^s is deterministic in the model), the 95% confidence interval for R_0^e can be estimated to be $R_0^e \pm \Delta_{i,week}^s / (N^s q_i^s)$.

The predictions in Tables 1 and 2 reflect the 95% confidence interval estimates provided in the preceding two paragraphs. The parameters used to generate model predictions (e.g., initial infection rates, mask filtration efficiency, room properties, pathogen emission rates, and infectious dose) are considered to be deterministic quantities with known values, although we recognize that there are significant uncertainties in some of these assumed parameter values.

Modeling the Effect of Vaccination. To consider the effect of vaccinations, three additional parameters are introduced: 1) the fraction of the population that is vaccinated, γ; 2) vaccine effectiveness in preventing infection, β; and 3) the viral dose from vaccinated individuals with breakthrough infections is assumed to be reduced by a factor η relative to individuals that are not vaccinated. Given these assumptions and parameters, the average viral dosage from the students and the instructor in Eqs. **5** and **6** is updated as follows:

$$D_m^s = D(m, 1, p_a^s)[\gamma(1 - \beta)\eta + (1 - \gamma)],$$ **[17]**

$$D^t = D(1, 1, p_a^t)[\gamma(1 - \beta)\eta + (1 - \gamma)].$$ **[18]**

Following that, the infection probabilities calculated in Eqs. 7 and **8** are modified as follows:

$$P_i(1, m) = \left(1 - e^{-\frac{D^t + D_m^s}{D_0}}\right)\left[\gamma(1 - \beta) + (1 - \gamma)\right],$$ **[19]**

$$P_i(0, m) = \left(1 - e^{-\frac{D_m^s}{D_0}}\right)\left[\gamma(1 - \beta) + (1 - \gamma)\right].$$ **[20]**

Note that we do not model some policies implemented by universities such as regular testing of students and faculty, contact tracing, and quarantining close contacts of infected individuals; these could potentially reduce the R_0^e further.

Dataset. We obtained registration information of all students for a large US university for both fall 2019 and fall 2020 (50). Both datasets include information for each student that registered for classes. For fall 2019, we consider only classes that were held in person (most of them were). However, classes in fall 2020 were either online or hybrid. In hybrid mode, a fraction of students are assumed to attend the class in person, and the rest are assumed to watch the class online. For fall 2019, there were 5,986 courses with 34,042 students on campus. For fall 2020, there were 523 hybrid courses with 6,376 students registered for those classes. The remaining classes were entirely online and are therefore not considered in this study.

We also obtained a dataset of buildings at the same university that contained information about classroom sizes, ventilation rates, and maximum occupancy (or capacity) (50). This dataset was used to estimate the physical parameters (classroom volume v, air change rate E_{hvac}, etc.) appearing in Eqs. **1** and **2**. In this dataset, classroom capacity varies between 5 and 430. The average student enrollment is 35. The air change rate in the rooms varies from 1.8 to 15 h^{-1}, with an average of 6.1 h^{-1}. Most classrooms are in buildings serviced by centralized HVAC systems with air filter ratings ranging from minimum efficiency reporting value (MERV)–14 to MERV-16. Roughly 30% of classrooms are in building serviced by decentralized or rooftop HVAC systems with air filters ratings ranging from MERV-8 to MERV-16. Classroom area varies between 32 and 3,668 ft^2, with an average of 519 ft^2. Finally, course duration varies from 50 min to 5 h with an average of 1 h and 51 min. Note that our simulations used data specific to each classroom and course for calculating infection probabilities.

The University of Southern California's Institutional Review Board (IRB) approval was required for this project. IRB approval (Study ID UP-20-00400) was obtained to access and use the student registration datasets described in this research. All anonymized relevant study data will be deposited on a private repository and will be available to interested researchers upon request.

Data Availability. All anonymized relevant study data has been deposited on the public repository GitHub (https://github.com/ANRGUSC/covid-university-data) (50). The datasets are encrypted, and the decryption key will be available to interested researchers upon request.

1. World Health Organization, WHO Director-General's opening remarks at the media briefing on COVID-19 - 11 March 2020 (2020). https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19–11-march-2020. Accessed 11 May 2022.
2. *Tracking COVID at United States colleges and universities* (https://www.nytimes.com/interactive/2020/us/covid-college-cases-tracker.html) (2020). Accessed 12 October 2020.
3. Map: Coronavirus and school closures in 2019–2020. *Education Week* (2020). https://www.edweek.org/leadership/map-coronavirus-and-school-closures-in-2019-2020/2020/03. Accessed 12 October 2020.
4. V. Duong, J. Luo, P. Pham, T. Yang, Y. Wang, "The ivory tower lost: How college students respond differently than the general public to the COVID-19 pandemic" in *2020 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)* (2020), pp. 126–130.
5. *Impact of the COVID-19 pandemic on education*. Wikipedia (2020). https://en.wikipedia.org/wiki/Impact_of_the_COVID-19_pandemic_on_education. Accessed 12 October 2020.
6. O. B. Adedoyin, E. Soykan, COVID-19 pandemic and online learning: The challenges and opportunities. *Interact. Learn. Environ.* 1–13 (2020).
7. D. C. Barton, Impacts of the COVID-19 pandemic on field instruction and remote teaching alternatives: Results from a survey of instructors. *Ecol. Evol.* **10**, 12499–12507 (2020).
8. C. J. L. Murray, P. Piot, The potential future of the COVID-19 pandemic: Will SARS-CoV-2 become a recurrent seasonal infection? *JAMA* **325**, 1249–1250 (2021).
9. A. Fontanet *et al.*, SARS-CoV-2 variants and ending the COVID-19 pandemic. *Lancet* **397**, 952–954 (2021).
10. S. Kashte, A. Gulbake, S. F. El-Amin III, A. Gupta, COVID-19 vaccines: Rapid development, implications, challenges and future prospects. *Hum. Cell* **34**, 711–733 (2021).
11. L. O. Gostin, D. A. Salmon, H. J. Larson, Mandating COVID-19 vaccines. *JAMA* **325**, 532–533 (2021).
12. J. M. Cashore *et al.*, COVID-19 mathematical modeling for Cornell's fall semester (2020). https://people.orie.cornell.edu/pfrazier/COVID_19_Modeling_Jun15.pdf. Accessed 8 November 2020.
13. J. M. Cashore *et al.*, Addendum: COVID-19 mathematical modeling for Cornell's fall semester (2020). https://covid.cornell.edu/_assets/files/covid_19_modeling_addendum.pdf. Accessed 8 November 2020.
14. R. Zafarnejad, P. M. Griffin, Assessing school-based policy actions for COVID-19: An agent-based analysis of incremental infection risk. *Comput. Biol. Med.* **134**, 104518 (2021).
15. M. Mandal *et al.*, A model based study on the dynamics of COVID-19: Prediction and control. *Chaos Solitons Fractals* **136**, 109889 (2020).
16. D. M. Kennedy, G. J. Zambrano, Y. Wang, O. P. Neto, Modeling the effects of intervention strategies on COVID-19 transmission dynamics. *J. Clin. Virol.* **128**, 104440 (2020).
17. F. Ying, N. O'Clery, Modelling COVID-19 transmission in supermarkets using an agent-based model. *PLoS One* **16**, e0249821 (2021).
18. A. Foster, M. Kinzel, Estimating COVID-19 exposure in a classroom setting: A comparison between mathematical and numerical models. *Phys Fluids (1994)* **33**, 021904 (2021).
19. A. Hekmati, M. Luhar, B. Krishnamachari, M. Matarić, "Simulation-based analysis of COVID-19 spread through classroom transmission on a university campus" in *2021 IEEE International Conference on Communications Workshops (ICC Workshops)* (2021), pp. 1–6.

20. V. Stadnytskyi, C. E. Bax, A. Bax, P. Anfinrud, The airborne lifetime of small speech droplets and their potential importance in SARS-CoV-2 transmission. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 11875–11877 (2020).

21. G. Buonanno, L. Stabile, L. Morawska, Estimation of airborne viral emission: Quanta emission rate of SARS-CoV-2 for infection risk assessment. *Environ. Int.* **141**, 105794 (2020).

22. S. L. Miller *et al.*, Transmission of SARS-CoV-2 by inhalation of respiratory aerosol in the Skagit Valley Chorale superspreading event. *Indoor Air* **31**, 314–323 (2021).

23. T. Avril, A. Nathan, S. Gantz, The delta variant may spread as fast as chickenpox, and other facts about new CDC data. *The Philadelphia Inquirer*, 30 July 2021. https://www.inquirer.com/health/coronavirus/delta-variant-cdc-chickenpox-contagious-data-vaccinated-20210730.html. Accessed 10 May 2022.

24. M McMorrow, Improving communications around vaccine breakthrough and vaccine effectiveness. https://context-cdn.washingtonpost.com/notes/prod/default/documents/8a726408-07bd-46bd-a945-3af0ae2f3c37/note/57c98604-3b54-44f0-8b44-b148d8f75165. Accessed 10 May 2022.

25. A. Konda *et al.*, Aerosol filtration efficiency of common fabrics used in respiratory cloth masks. *ACS Nano* **14**, 6339–6347 (2020).

26. M. G. Thompson *et al.*, Interim estimates of vaccine effectiveness of BNT162b2 and mRNA-1273 COVID-19 vaccines in preventing SARS-CoV-2 infection among health care personnel, first responders, and other essential and frontline workers–Eight U.S. locations, December 2020-March 2021. *MMWR Morb. Mortal. Wkly. Rep.* **70**, 495–500 (2021).

27. Ministry of Health, Israel, "Explanation About the Effectiveness of the Vaccine for Coronavirus in Israel," https://www.gov.il/en/departments/news/06072021-04. Accessed 10 May 2022.

28. L. Morawska *et al.*, How can airborne transmission of COVID-19 indoors be minimised? *Environ. Int.* **142**, 105832 (2020).

29. S. J. Dancer *et al.*, Putting a balance on the aerosolization debate around SARS-CoV-2. *J. Hosp. Infect.* **105**, 569–570 (2020).

30. Y. Li *et al.*, Probable airborne transmission of SARS-CoV-2 in a poorly ventilated restaurant. *Build. Environ.* **196**, 107788 (2021).

31. R. Zhang, Y. Li, A. L. Zhang, Y. Wang, M. J. Molina, Identifying airborne transmission as the dominant route for the spread of COVID-19. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 14857–14863 (2020).

32. R. Tellier, Y. Li, B. J. Cowling, J. W. Tang, Recognition of aerosol transmission of infectious agents: A commentary. *BMC Infect. Dis.* **19**, 101 (2019).

33. W. W. Nazaroff, Indoor particle dynamics. *Indoor Air* **14** (suppl. 7), 175–183 (2004).

34. W. W. Nazaroff, Indoor bioaerosol dynamics. *Indoor Air* **26**, 61–78 (2016).

35. W. Wells, On air-borne infection: Study II. Droplets and droplet nuclei. *Am. J. Epidemiol.* **20**, 611–618 (1934).

36. E. C. Riley, G. Murphy, R. L. Riley, Airborne spread of measles in a suburban elementary school. *Am. J. Epidemiol.* **107**, 421–432 (1978).

37. G. N. Sze To, C. Y. H. Chao, Review and comparison between the Wells-Riley and dose-response approaches to risk assessment of infectious respiratory diseases. *Indoor Air* **20**, 2–16 (2010).

38. J. L. Jimenez, Z. Peng, COVID-19 aerosol transmission estimator. https://tinyurl.com/covid-estimator. Accessed 10 May 2022.

39. V. Vuorinen *et al.*, Modelling aerosol transport and virus exposure with numerical simulations in relation to SARS-CoV-2 transmission by inhalation indoors. *Saf. Sci.* **130**, 104866 (2020).

40. M. Z. Bazant, J. W. M. Bush, A guideline to limit indoor airborne transmission of COVID-19. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2018995118 (2021).

41. N. van Doremalen *et al.*, Aerosol and surface stability of SARS-CoV-2 as compared with SARS-CoV-1. *N. Engl. J. Med.* **382**, 1564–1567 (2020).

42. S. Asadi *et al.*, Aerosol emission and superemission during human speech increase with voice loudness. *Sci. Rep.* **9**, 2348 (2019).

43. R. Wölfel *et al.*, Virological assessment of hospitalized patients with COVID-2019. *Nature* **581**, 465–469 (2020).

44. R. Sender *et al.*, The total number and mass of SARS-CoV-2 virions. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2024815118 (2021).

45. T. Watanabe, T. A. Bartrand, M. H. Weir, T. Omura, C. N. Haas, Development of a dose-response model for SARS coronavirus. *Risk Anal.* **30**, 1129–1138 (2010).

46. I. Schröder, COVID-19: A risk assessment perspective. *J. Chem. Health Saf.* **27**, 308 (2020).

47. J. W. Tang, T. J. Liebner, B. A. Craven, G. S. Settles, A Schlieren optical study of the human cough with and without wearing masks for aerosol infection control. *J. R. Soc. Interface* **6** (suppl. 6), S727–S736 (2009).

48. M. Z. Bazant *et al.*, Monitoring carbon dioxide to quantify the risk of indoor airborne transmission of COVID-19. *Flow* **1**, E10 (2021).

49. J. A. Backer, D. Klinkenberg, J. Wallinga, Incubation period of 2019 novel coronavirus (2019-nCoV) infections among travellers from Wuhan, China, 20-28 January 2020. *Euro Surveill.* **25**, 2000062 (2020).

50. A. Hekmati, M. Luhar, B. Krishnamachari, M. Matarić, University COVID Data. GitHub. https://github.com/ANRGUSC/covid-university-data. Deposited 11 May 2022.