Received: 6 December 2021 | Revised: 21 February 2022 | Accepted: 25 February 2022

DOI: 10.1111/ina.13015

ORIGINAL ARTICLE

WILEY

Beyond well-mixed: A simple probabilistic model of airborne disease transmission in indoor spaces

Sijian Tan[1] | Zhihang Zhang[2] | Kevin Maki[2] | Krzysztof J. Fidkowski[1] | Jesse Capecelatro[1,3]

[1]Department of Aerospace Engineering, University of Michigan, Ann Arbor, Michigan, USA

[2]Department of Naval Architecture and Marine Engineering, University of Michigan, Ann Arbor, Michigan, USA

[3]Department of Mechanical Engineering, University of Michigan, Ann Arbor, Michigan, USA

Correspondence
Sijian Tan, Department of Aerospace Engineering, University of Michigan, Ann Arbor, MI, USA.
Email: sijian@umich.edu

Abstract

We develop a simple model for assessing risk of airborne disease transmission that accounts for non-uniform mixing in indoor spaces and is compatible with existing epidemiological models. A database containing 174 high-resolution simulations of airflow in classrooms, lecture halls, and buses is generated and used to quantify the spatial distribution of expiratory droplet nuclei for a wide range of ventilation rates, exposure times, and room configurations. Imperfect mixing due to obstructions, buoyancy, and turbulent dispersion results in concentration fields with significant variance. The spatial non-uniformity is found to be accurately described by a shifted lognormal distribution. A well-mixed mass balance model is used to predict the mean, and the standard deviation is parameterized based on ventilation rate and room geometry. When employed in a dose–response function risk model, infection probability can be estimated considering spatial heterogeneity that contributes to both short- and long-range transmission.

KEYWORDS
aerosol, CFD, disease transmission, probabilistic model, risk, well-mixed

1 | INTRODUCTION

Understanding and predicting the transmission of infectious diseases are critical for assessing risk to individuals and informing policy changes. Airborne diseases, such as influenza, tuberculosis, measles, and SARS-CoV-2, most effectively spread in indoor settings, especially in spaces with poor ventilation and overcrowding. For example, it is now recognized that airborne transmission is the dominant transmission mode of SARS-CoV-2 during the COVID-19 pandemic,[1,2] with the majority of super-spreading events occurring indoors.[1,3-5,]

The likelihood of infection depends on the exposure dose—the number of viral particles inhaled by a susceptible individual—which is a consequence of both short- and long-range transmission routes.[2,6-8] Short-range transmission is controlled by the initial advection of exhaled respiratory droplets, gravitational settling of larger droplets, and evaporation into droplet nuclei.[9] Long-range transmission of droplet nuclei from a host to a susceptible person is governed by complex turbulent airflow, mechanical and passive ventilation, thermal buoyancy effects, and air filtration.

Due to their small size, expiratory droplet nuclei move passively with background air currents, and their concentration field, $C(\boldsymbol{x}, t)$, can be modeled by an advection-diffusion equation

$$\frac{\partial C}{\partial t} + \nabla \cdot (\boldsymbol{u}C) - \nabla \cdot (\mathcal{D}\nabla C) = S, \tag{1}$$

where $\boldsymbol{u}$ *is* the background fluid velocity, $\mathcal{D}$ is the diffusion coefficient, and S is a source term that accounts for pathogen inactivation and other contributions to changes in the concentration field. A direct solution to Equation (1) is generally computationally expensive, particularly in the Navier–Stokes calculation of the background velocity. Thus, risk analyses that require a large number of iterations must rely on a simplified, or coarse-grained, representation.

Integrating Equation (1) over the entire space results in an ordinary differential equation—the so-called well-mixed mass balance model—whose solution after an exposure time t with initially zero concentration is[15]

$$\overline{C}(t) = \frac{q}{\lambda V}(1 - e^{-\lambda t}), \tag{2}$$

where q is the pathogen generation rate (by one or more infected sources) due to breathing, speaking, etc., λ is a loss coefficients that accounts for ventilation, deposition, etc., and V is the volume of the room. In this expression, $\overline{C}(t)$ represents the average concentration field assuming a homogeneous environment in which expelled particles are uniformly distributed within the enclosed space. Under these assumptions, the total number of infectious particles inhaled by a susceptible individual (dose) is related to the concentration field according to $d(t) = \int_0^t p\overline{C}(\tau)\,\mathrm{d}\tau$,[10] where p is the pulmonary ventilation rate (breathing rate) of an individual. When combined with an exponential dose–response function, the probability of infection, defined as the ratio of number of infections, N_I, to number of susceptible individuals, N_S, can be estimated according to

$$\mathcal{P} \equiv \frac{N_I}{N_S} = 1 - e^{-d/k}, \tag{3}$$

where k is a constant that depends on the infectiousness of the virus.[11] Alternatively, the probability can be expressed in terms of the quantum of infection (or quanta) introduced by Wells,[12] which gives rise to the well-known Wells–Riley model.[13] The aforementioned approaches act as the basis of most epidemiological models used today for assessing the risk of airborne diseases. For example, such models have been used to predict the risk of infection associated with measles,[13] tuberculosis,[14-16] and more recently SARS-CoV-2.[3,4,17,18]

A key drawback of the Wells–Riley model and dose–response functions based on a well-mixed mass balance is that the concentration of infectious material is assumed to be evenly distributed throughout the enclosed space at each instant in time. It is well established that this is rarely true, even in spaces with high ventilation rates and strong mixing.[19-21] Consequently, infection risk due to short- and long-range exposure cannot be distinguished, and all susceptible individuals are treated as equally vulnerable. As pointed out in a review by Sze To and Chao,[22] "[the] spatial distribution of airborne pathogens is one of the most important factors in infection risk assessment of respiratory disease."

A variety of approaches have been proposed to address the shortcomings of well-mixed models in recent years. Multi-zone (or zonal ventilation) models divide the space into sub-volumes (zones), each of which are assumed well-mixed with homogeneous composition.[19,23,24] Susceptible individuals located in different zones are exposed to a different dose and thus experience different levels of infection risk. This approach is typically used in settings with multiple rooms or spaces with partitions, such as in hospitals,[25-27] commercial airliners,[28,29] and apartment buildings.[30] Several extensions to the multi-zone model have been proposed. Nicas[31] developed a probabilistic model based on Markov chains, where each zone is treated as well-mixed but the concentration from one zone to another is described probabilistically, allowing for some variability in infection risk to be captured. Noakes and Sleigh[19] proposed a stochastic version of the Wells–Riley model, where mixing between zones is limited. While multi-zone models have shown significant improvement over previous epidemiological models, the treatment of each zone as well-mixed precludes them from identifying infection to susceptible people in close proximity to the infectious source, unless detailed information regarding transport between interconnected spaces is known a-priori.[32]

Sun and Zhai[33] generalized the Wells–Riley model by introducing a distance index and a ventilation index to quantify the impact of social distance and ventilation effectiveness on the probability of infection. Guo et al.[34] showed that this model yields reasonable accuracy against available data. However, they point out that it does not account for the location of infectors and fails to capture the spatial distribution of infection probability. Accordingly, Guo et al.[34] introduced a so-called spatial flow impact factor into the Wells–Riley model and demonstrated success in predicting the spatial distribution of infection probability, which was used to identify optimal placement of individuals and facilities (e.g., air purifiers) in a simulated hospital ward. However, such an approach requires detailed flow measurements via experiments or computational analysis of the specific confined space prior to its use.

Computational fluid dynamics (CFD) has gained popularity in recent years for quantifying the spatial distribution of pathogens in confined spaces. Due to the wide range of length and time scales present in turbulent flows, obtaining the local velocity field, $\boldsymbol{u}$, in Equation (1) via a direct solution to the Navier–Stokes equations is often not tractable except for early stage, short-range propagation

Practical implications

- Spatial variation in aerosol concentration is significant in indoor settings, with standard deviations comparable to the mean, and should be accounted for during risk assessment.
- The classical well-mixed mass balance provides an accurate estimate of the mean concentration, but fails to capture spatial variation that distinguishes short- and long-range transmission.
- The concentration distribution was found to be nearly lognormal for all of the indoor settings considered.
- The model developed in this work can be integrated within existing epidemiological models to capture non-uniform mixing for general indoor settings.
- Treating the exposure dose as a probability density function results in a distribution of infection probabilities (as opposed to a single value obtained from classical approaches), that informs impact of social distancing among other safety guidelines.

of expiratory flows such as coughs or sneezes.[35-37] Instead, the mean flow field is often obtained via the Reynolds-averaged Navier-Stokes (RANS) equations, for example, using a $k-\varepsilon$ turbulence model.[38-43] The concentration field obtained from CFD can then be employed in a dose–response function or used to inform multi-zone models.[21,39,44] Due to its high computational cost, CFD is traditionally used to study specific scenarios under a limited number of permutations. Thus, its use for general risk assessment remains limited.

In this work, a general model is developed that accounts for spatial non-uniformity in indoor spaces. A large database of CFD results is generated for a wide range of representative spaces (see Figure 1). The location of infected individuals, room geometry, ventilation rate, and exposure duration are varied to obtain a statistically significant representation of disease transmission in confined spaces. Details on the numerical simulations followed by a statistical analysis of the concentration fields are provided in Section 2. In Section 3, a model is proposed for its probability density function and is employed in a dose–response function to quantify the effect of non-uniform mixing on risk of infection.

2 | SIMULATION DETAILS

This section provides details on the numerical simulations used to generate spatio-temporal concentration fields that will be used to inform the probabilistic model proposed in Section 3. We consider several representative indoor spaces: an urban bus; a 30-student classroom; a large college-style lecture hall; and a music rehearsal space, as depicted in Figure 1. Room dimensions and ventilation specifications are modeled after spaces at the University of Michigan. Details on the computational mesh and boundary conditions used in each case are summarized in Appendix A.

2.1 | Governing equations

The airflow is assumed to be incompressible and turbulent, governed by

$$\frac{\partial \mathbf{u}}{\partial t}+\boldsymbol{u}\cdot\nabla\boldsymbol{u}=-\nabla p+\nabla\cdot\left[\nu_t\left(\nabla\boldsymbol{u}+\nabla\boldsymbol{u}^T\right)\right]-\boldsymbol{g}\cdot\boldsymbol{x}\nabla\left(\frac{\rho}{\rho_0}\right), \tag{4}$$

where p is the kinematic pressure, $\boldsymbol{g}$ is the acceleration due to gravity, and ν_t is the effective viscosity that accounts for both molecular and turbulent diffusion. The turbulent viscosity is obtained from the $k-\varepsilon$ turbulence model.[45] Due to the small density variations of the air, the Boussinesq approximation is employed, so that density variations only appear in the gravitational term. The nominal air density is $\rho_0 = 1.2\,\mathrm{kg/m^3}$ and the local density, ρ, varies based on the temperature field according to the ideal gas law. The equation governing temperature is

$$\frac{\partial T}{\partial t}+\nabla\cdot(\boldsymbol{u}T)-\nabla\cdot(\alpha_t\nabla T)=0, \tag{5}$$

where $\alpha_t=\nu_t/\mathrm{Pr}_t+\nu/\mathrm{Pr}$ is the effective thermal diffusivity, and $\mathrm{Pr}_t=0.9$ and $\mathrm{Pr}=0.71$ are the turbulent and laminar Prandtl numbers, respectively. The concentration field is solved according to Equation (1), where the diffusion coefficient is given by $D=\nu_t/\mathrm{Sc}_t+\nu/\mathrm{Sc}$, where $\mathrm{Sc}_t=\mathrm{Sc}=1$ are the turbulent and laminar Schmidt numbers, respectively.

For each case, a precursor simulation is performed to generate a fully developed turbulent flow field inside the domain. The resulting flow field is used as the initial condition at which point the infected individual begins shedding aerosols. Occupants are represented by human-size manikins (standing and sitting) with uniform temperature of 32°C. The mouth and nose are modeled as an integrated round patch with a diameter of 4 cm and temperature of 34°C (see Figure 1B). A turbulence intensity of 10% and a mixing length of 7.5 mm are enforced at the vicinity of mouth. In each simulation, it is assumed that only one infected person is present. Multiple simulations are run for each case, varying the location of the infector. A constant virus shedding rate of $q=50\,\mathrm{s^{-1}}$ is enforced as a boundary condition at the nose/mouth of the infected individual with a breathing rate of $p=6$ l/min, corresponding to a highly contagious person speaking continuously and loudly.[18,46]

2.2 | Flow field and concentration distribution

In total, 174 simulations are performed: 163 of the 30-student classroom; eight of the lecture hall; one of the music hall; and two of the urban bus. Each indoor space is characterized by its volume V, height H, aspect ratio $\sqrt{A}/H$, where $A=V/H$ is the floor area, and

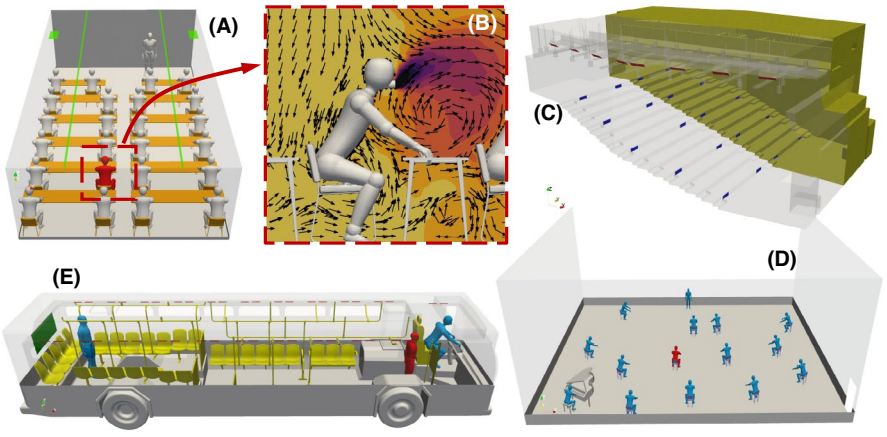

FIGURE 1 Schematic of the various geometries used in the numerical simulations. (A) Classroom. (B) Zoom-in of the infected individual (color shows concentration, arrows depict fluid streamlines). (C) Lecture hall. (D) Rehearsal hall. (E) Urban bus

ventilation rate denoted by the number of air changes per hour (ACH). The volume, ACH, and duration contribute to the mean concentration, $\overline{C}$, per Equation (2), where the loss coefficient is related to ACH according to $\lambda = ACH/3600$. The variation in concentration field is found to depend on the room geometry ($\sqrt{A}/H$) and relative position of the infector to the location of the return and supply vents (see Appendix B). Thus, for the analysis presented herein, averaging is performed over the volume of the indoor space $\Omega \in V$ and over different realizations by varying the location of the infector, according to

$$\overline{C} = \frac{1}{N}\sum_{i=1}^{N}\left[\frac{1}{V}\int_{\Omega} C(\mathbf{x},t)\mathrm{d}V\right], \tag{6}$$

and

$$C_{\mathrm{std}}^2 = \frac{1}{N}\sum_{i=1}^{N}\left[\frac{1}{V}\int_{\Omega}\left(C(\mathbf{x},t)-\overline{C}\right)^2\mathrm{d}V\right], \tag{7}$$

where $\overline{C}$ and C_{std} are the mean and standard deviation of the concentration field, respectively, $C(\mathbf{x},t)$ is the local concentration at time t, and N is the number of realizations. It should be noted that in the region near the infector's mouth, the concentration is extremely high and quickly drops before reaching susceptible individuals. Thus, computational cells containing the largest 0.1% values near the infector are removed.

The mean and standard deviation of the concentration field obtained from each case are summarized in Appendix B. Several important observations can immediately be made. First, the well-mixed mass balance model (2) yields an accurate prediction of the mean concentration when compared to the CFD results for all of the cases considered. Second, the variation about the mean, characterized by C_{std}, is large, with values comparable to or larger than the mean. This has important consequences in risk assessment since the infection dose of an individual may vary significantly about the mean, depending on the relative proximity to the infector or supply and return vents. In addition, the variation in concentration is seen to be correlated with the room geometry: C_{std} tends to increase with increasing aspect ratio.

Instantaneous snapshots of the airflow and concentration field in two of the classroom configurations are shown in Figure 2. Turbulent mixing and recirculation are driven by the return and supply vents. The concentration field is diffused and convected by the air currents away from the infector to neighboring susceptible individuals. Because each case considers one infector, increasing the aspect ratio from $\sqrt{A}/H = 2$ in Figure 2B to $\sqrt{A}/H = 3$ in Figure 2C yields larger variation in the concentration field. This can be attributed to inadequate diffusion over the duration t droplet nuclei are being emitted. In the absence of any background airflow, transport of C is entirely controlled by diffusion. Because of the diffusion time $A/\mathcal{D} \gg \mathrm{d}t$, a larger surface area results in greater non-uniformity.

The probability density function (PDF) of concentration throughout the space in Figure 2D shows (I) significant regions with concentration below the mean, corresponding to locations far from the infector; (II) high probability events of concentration near the mean value; and (III) a long tail with values significantly above the mean corresponding to regions near the infector. Thus, the PDF captures effects of both short-range (high C) and long-range (low and intermediate C) transmission.

Larger volume spaces, namely the lecture hall and music rehearsal hall considered here, exhibit vastly different behavior compared to the smaller classrooms. As shown in Figure 3, the large height of the lecture hall results in significant vertical displacement of the concentration field. We attribute this behavior to buoyancy effects. The combined effect of convection by ventilation (or wind if windows are open) and buoyancy can be characterized by the densimetric Froude number according to $\mathrm{Fr} = U/\sqrt{g'H}$,[47] where U is a velocity scale associated with the supply vents (proportional to the ACH), and $g' = g\beta\Delta T$, where $\beta = 1/293\ \mathrm{K}^{-1}$ is the thermal expansion coefficient of air at room temperature, and $\Delta T \approx 14$ K is the temperature difference between the infector and ambient air. If the ACH is sufficiently high to mix the space, the spatial variation in the concentration field will decrease. However, if there exists a significant temperature difference across the height of the room, the air can become stratified, trapping droplet nuclei near the ceiling (as shown in Figure 3). For the cases considered herein, $\mathrm{Fr} = 1$. Thus, in spaces with large H (relative to the height of a person), the concentration distribution is expected to be controlled by buoyancy.

3 | PROBABILISTIC MODEL

3.1 | Model development

To capture the spatial non-uniformity in aerosol concentration for risk assessment, we propose to model the PDF based on a small number of input parameters that characterize the indoor space. We adopt a presumed-shape PDF approach, where the concentration distribution is taken to be a shifted and scaled lognormal distribution. The two parameters that govern a standard lognormal distribution of a random variable X are μ and σ, corresponding to the mean and standard deviation of $\ln(X)$, respectively. The probability density function of X is

$$P(X) = \frac{1}{X\sigma\sqrt{2\pi}}e^{-\frac{1}{2\sigma^2}(\ln X-\mu)^2}, \tag{8}$$

and the mean and standard deviation of X are given by $\overline{X} = e^{\mu+\sigma^2/2}$ and $X_{\mathrm{std}} = \sqrt{(e^{\sigma^2}-1)e^{2\mu+\sigma^2}}$, respectively. To model the concentration distribution, we scale and shift X to match the mean and standard deviation of the concentration field, C, according to

$$C = \frac{X-\overline{X}}{X_{\mathrm{std}}}C_{\mathrm{std}} + \overline{C}. \tag{9}$$

Solving for X in terms of C and substituting into Equation (8) yields the probability density of C.

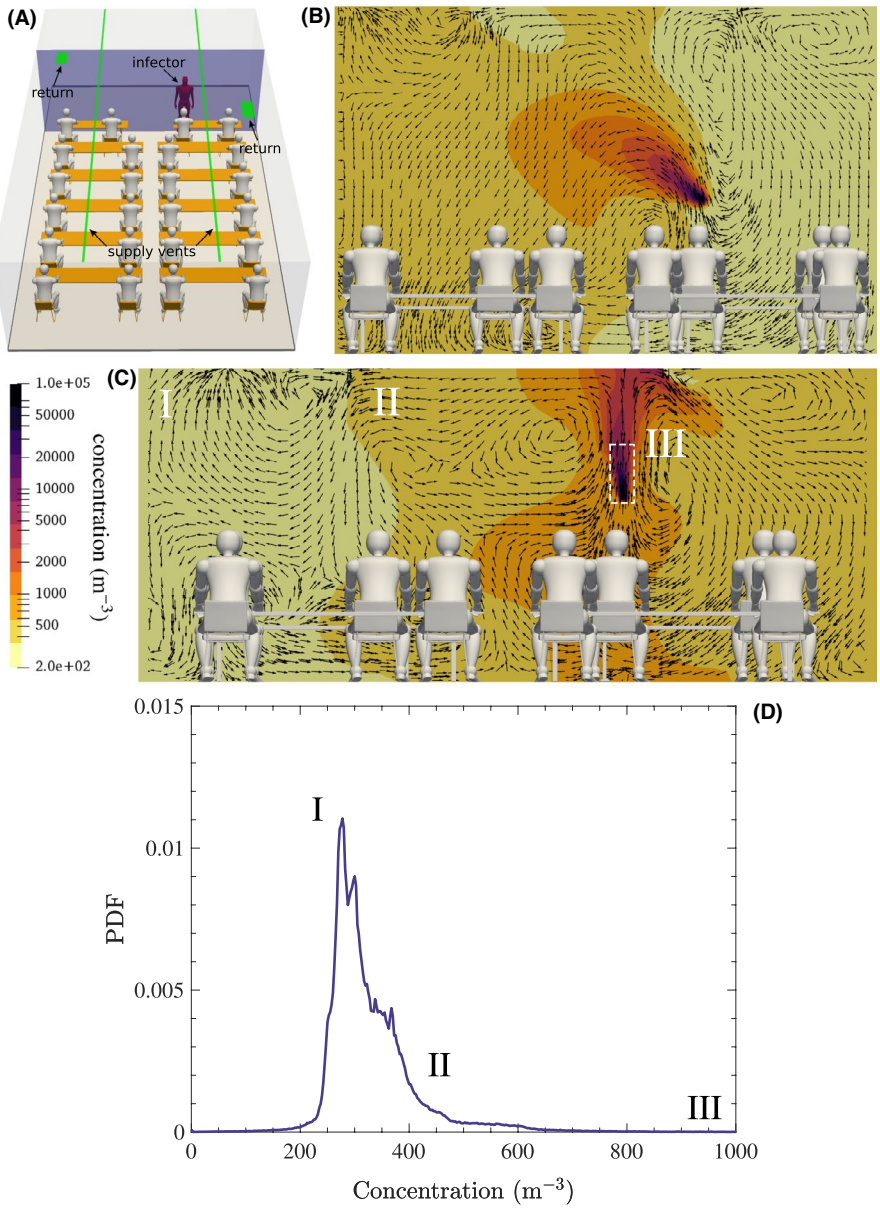

FIGURE 2 Concentration C contours and streamlines inside classrooms with ACH = 3 at different aspect ratios $\sqrt{A}/H$. (A) Schematic of the classroom indicating location of the infector, supply vents, and return. (B) $\sqrt{A}/H = 2$. (C) $\sqrt{A}/H = 3$. (D) PDF of concentration from (C) highlighting high probability events of moderate concentration (I), moderate probability of high concentration (II), and low probability of maximum concentration (III)

The parameters μ and σ govern the shape of the template lognormal distribution for the nondimensional variable, X. Based on fitting experiments with the CFD simulation dataset, we have found that $\mu = 0$ and $\sigma = 0.9$ perform well in minimizing the error in the probability distributions. The PDF of C then only depends on the concentration mean, $\overline{C}$, and the concentration standard deviation, C_{std}. The well-mixed model per Equation (2) is used for the former, and the latter is parameterized using the room geometry and ventilation parameters.

Linear regression of the CFD data of the classrooms, lecture hall, and rehearsal hall (see Table B1) shows that a reasonable model for the standard deviation is given by

$$\frac{C_{std}}{\overline{C}_{WM}} = \beta_0 + \beta_1 \frac{ACH}{ACH_0} + \beta_2 \frac{\sqrt{A}}{H} + \beta_3 \frac{H}{H_0}, \tag{10}$$

where $\overline{C}_{WM}$ is the mean concentration obtained from the well-mixed mass balance (2), $ACH_0 = 1$, $H_0 = 1$ m, and the coefficients are $\beta_0 = -1.01$, $\beta_1 = 0.11$, $\beta_2 = 0.33$, and $\beta_3 = 0.10$. As described above, ACH captures the effects of turbulent mixing on the spatial distribution of the concentration field, while $\sqrt{A}/H$ and H account for inadequate diffusion from the infector to the surroundings and buoyancy, respectively. We have found that a two-parameter model using ACH and $\sqrt{A}/H$ is sufficient for the classrooms, but inclusion of H is needed to capture the spatial non-uniformity in the lecture hall and music hall. The model given in Equation (10) results in an R^2 value of 0.97 over all of the cases listed in Table B1. The model is developed for $0 \le ACH \le 6$ h^{-1}, $2 \le \sqrt{A}/H \le 3$, and $2.45 \le H \le 8.26$ m, and therefore, caution should be exercised when using values outside of this range.

As shown in Figure 4, the normalized standard deviation increases linearly with aspect ratio. As noted earlier, ACH increases mixing in the room, resulting in smaller values of C_{std} in general. However, ACH also acts to dilute the air, resulting in a reduction of the mean concentration–recall ACH appears in the loss coefficient λ in Equation (2). Consequently, C_{std}/C_{WM} is observed to increase

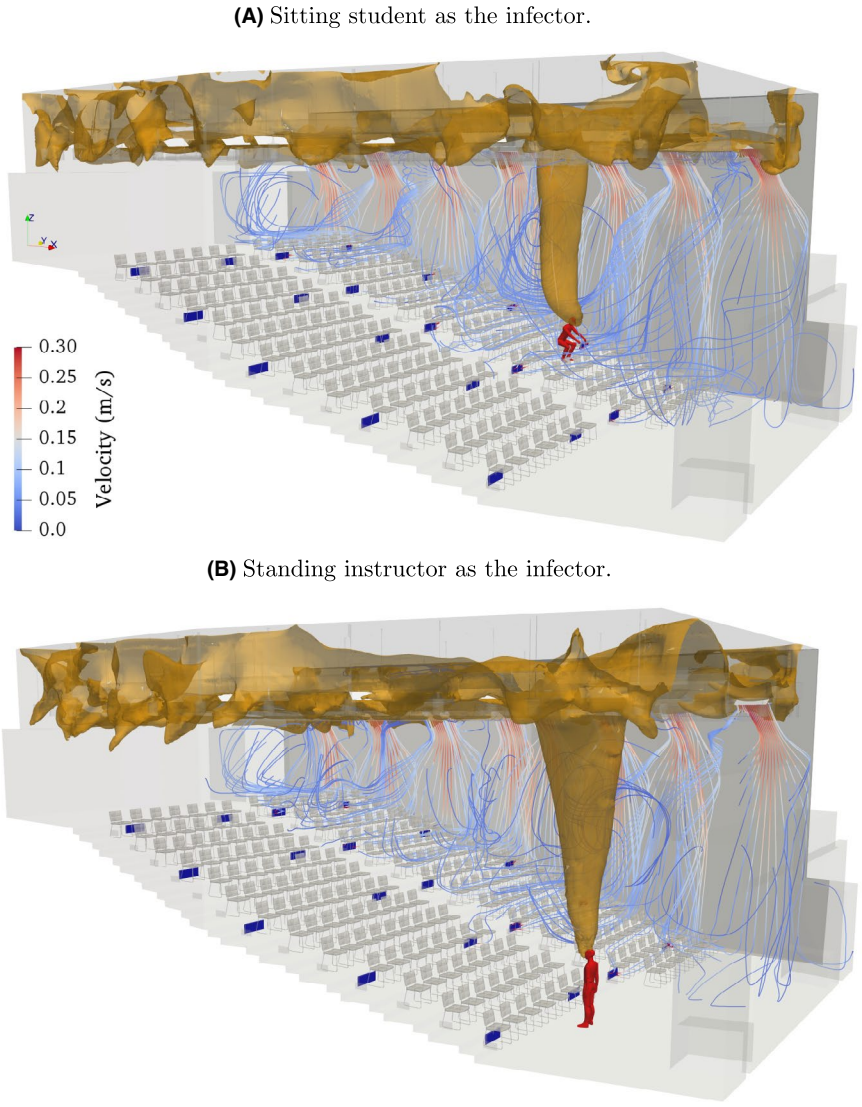

FIGURE 3 Instantaneous airflow streamlines and contour of concentration ($C = 120\ \mathrm{m}^{-3}$) inside the lecture hall at $t = 5400$ s

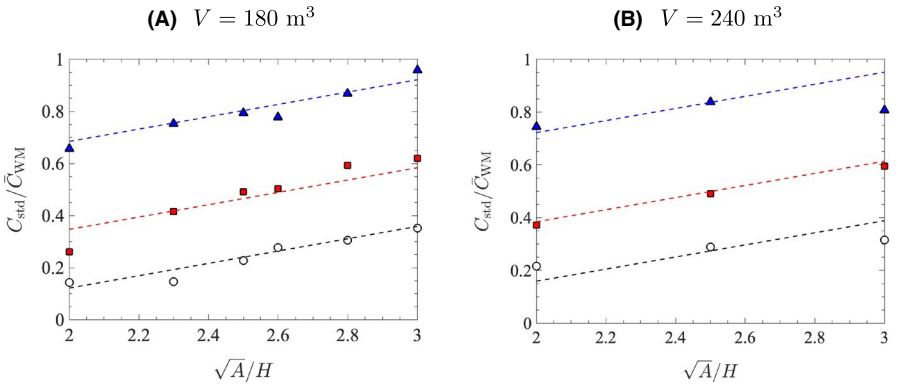

FIGURE 4 Standard deviation of the concentration field inside a classroom normalized by the mean obtained from a well-mixed mass balance. CFD results (symbols), model given by Equation (10) (dashed lines) for ACH = 1 (°), ACH = 3 ((■), and ACH = 6 ((▲)

with ACH. Further comparisons between the standard deviation obtained from CFD with Equation (10) are given in Table 1. Overall good agreement is observed over all of the indoor spaces considered. The largest discrepancy can be observed with the urban bus operated at ACH=16. As the regression was performed using $\mathrm{ACH} \in [0, 6]$, this value falls outside of the training set. It should also be noted that the buses were simulated using a clean air delivery rate of 20% (see Appendix A4), and thus, the ventilation rate used in (10) is 5 times larger than the ACH used in predicting the mean.

Examples of the fully integrated PDF model, combining Equations (8–10) with $\mu = 0$ and $\sigma = 0.9$, are shown in Figure 5. For each case, the general trend observed in the CFD results is captured exceptionally well. The heavy-tailed nature of the distribution yields significant probability of concentration greater than the

mean, corresponding to locations near the infector. Meanwhile, regions with local concentration below the mean are most probable, representative of long-range exposure. In the following section, the proposed model will be employed within a dose–response function to assess its impact on infection probability.

3.2 | Demonstration

Dose–response modeling is commonly used to estimate individual-level probability of infection based on the concentration of infectious aerosols.[11,22,48] We consider the exponential dose–response function given by Equation (3), where the dose represents the total number of infectious particles an individual would inhale according to $d = pCt$.[10] The pulmonary ventilation rate is taken to be $p = 1 \times 10^{-4}$ m^3/s (6 l/min), the concentration C is sampled from the proposed PDF model (9), and t is the duration. The parameter k typically varies between 75 and 500.[11,48] Watanabe et al.[11] found $k = 410$ provides a good fit for SARS coronavirus (SARS-CoV-1) based on data sets of infected mice, and this value is used in the present study.

The shifted lognormal concentration PDF is sampled using a discrete inverse cumulative distribution function (CDF) method. $N = 10^4$ random numbers u_i are generated from a uniform distribution between 0 and 1, and these are mapped to concentration values via $C_i = CDF^{-1}(u_i)$, where the discrete CDF is evaluated by cumulative sums of the PDF values, and its inverse is computed by linear interpolation of point data. Specifically, to obtain the CDF, which is the integral of the PDF, the shifted lognormal PDF is sampled at 1000 uniformly spaced concentration values and the CDF at each of these samples is computed by summing and normalizing the PDF values at all lower concentrations. This discrete representation of the CDF is then linearly interpolated to allow for the calculation of the CDF and its inverse at an arbitrary concentration value.

Predictions of infection probability are shown in Figure 6. In each case, a single infected person is presumed to reside within the space. The relative placement of the infector to susceptible individuals and to return and supply vents is accounted for in the modeled concentration PDF from which the samples are drawn. It can be seen that the probability of infection based on a dose obtained from the well-mixed mass balance is representative of the mean value obtained from sampling the concentration PDF. As expected, increasing ACH in the classroom leads to lower infection risk. Due to the large size of the music and lecture halls, infection probability is lower compared to the classroom despite having relatively low ACH.

TABLE 1 Predictions of C_{std} from CFD and the model given by Equation (10)

Description	ACH	C_{std} (CFD)	C_{std} (model)
Classroom	1	335.26	329.56
Classroom	3	255.7	251.38
Classroom	6	233.56	211.58
Lecture hall	1	52.8	45.75
Lecture hall	6	20.24	19.44
Rehearsal hall	0	43.48	51.22
Urban bus	1.6	469.33	595.75
Urban bus	16	74.53	1903.6

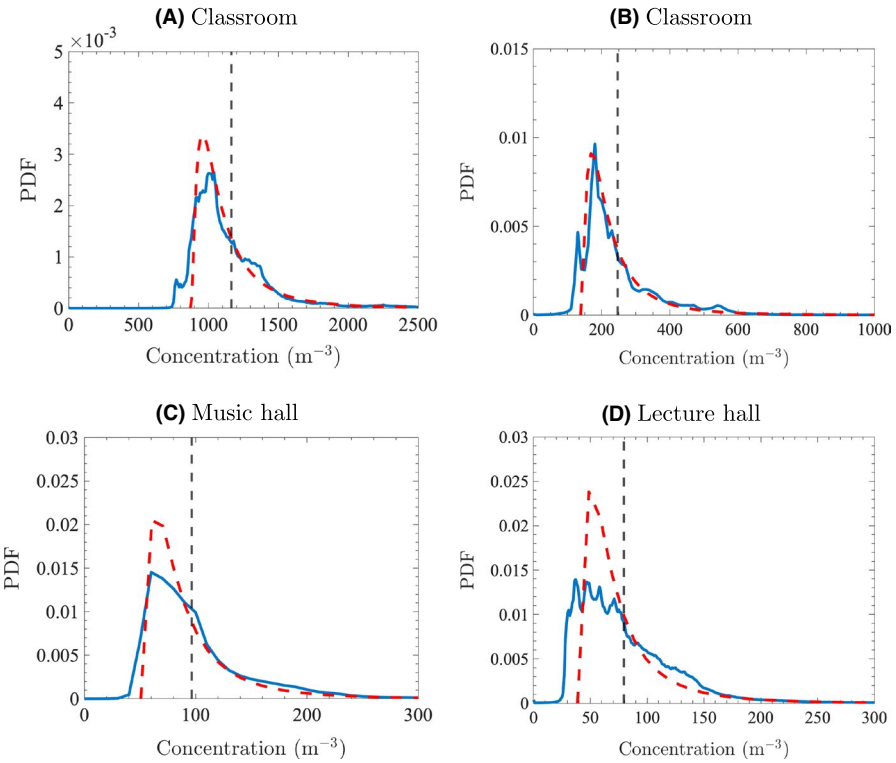

FIGURE 5 Concentration PDF obtained from CFD (——) and model using Equations (8)–(10) with $\mu = 0$ and $\sigma = 0.9$ (– –). Mean obtained from the well-mixed mass balance (– –). (A) Classroom with $V = 120$ m^3 and ACH = 1 averaged over $N = 6$ infector positions. (B) Classroom with $V = 240$ m^3 and ACH = 3, averaged over $N = 3$ infector positions. (C) Rehearsal hall with ACH = 0 and $N = 1$. (D) Lecture hall with ACH = 1, averaged over $N = 4$ infector positions

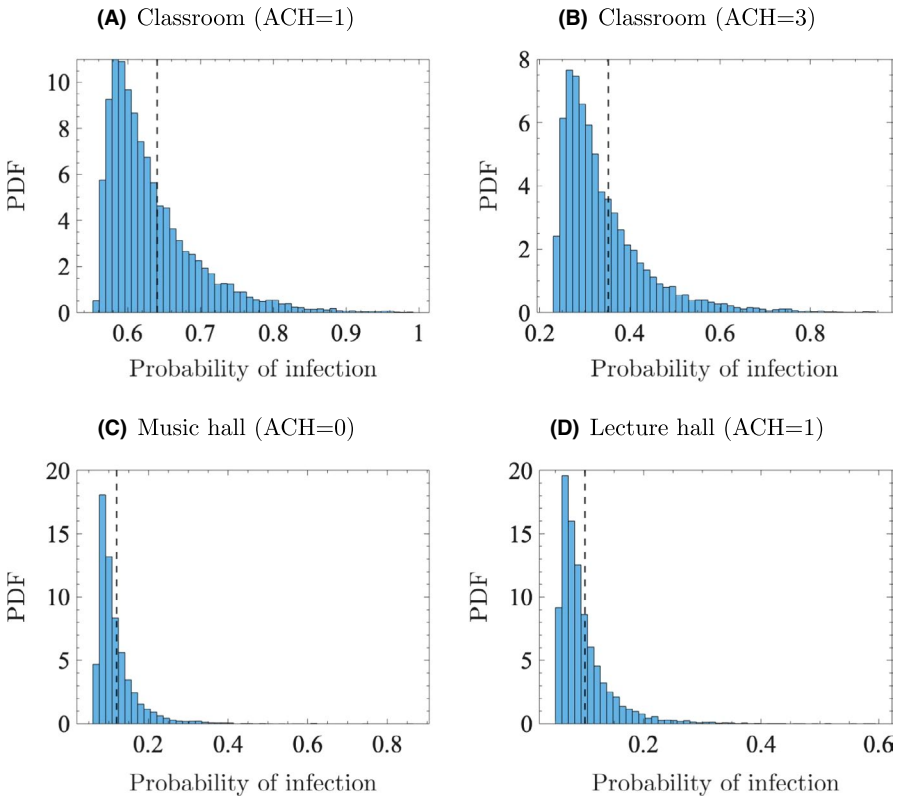

FIGURE 6 Histograms of the probability of infection in different scenarios. Each case consists of 10^4 uniform random samples of the concentration PDF obtained from the proposed model. The vertical line denotes probability of infection using a dose obtained from the well-mixed mass balance. The classrooms have $V = 180\ \mathrm{m}^3$ and $\sqrt{A}/H = 2.5$

Importantly, the probability of infection sampled from the proposed PDF model provides critical information that cannot be gleaned from a well-mixed assumption alone. Namely, the distribution of infection probability about the mean is large. At two standard deviations above and below the mean, probability of infection can be more than 2 times larger or 50 times smaller than what a well-mixed assumption would predict, as a consequence of the relative position of the susceptible individual. Therefore, upper and lower bounds of infection probability, or ideally the entire distribution, should be taken into account when determining safety guidelines for mitigating airborne transmission.

4 | CONCLUSIONS

This work presents a simple probabilistic model that captures spatial heterogeneity of airborne pathogens. Dispersion of expiratory droplet nuclei in classrooms, lecture halls, and buses was simulated using the three-dimensional unsteady Reynolds-averaged Navier–Stokes equations. A database containing 174 simulations was generated and used to quantify the spatial distribution of droplet nuclei for a wide range of ventilation rates, exposure times, and room configurations.

The classical well-mixed mass balance was found to provide an accurate estimate of the mean concentration. However, the concentration field was found to be highly non-uniform, with standard deviations comparable to the mean. We attribute this spatial variation to three mechanisms: (i) turbulent mixing characterized by the ventilation rate; (ii) inadequate diffusion characterized by the lateral area of the room; and (iii) vertical displacement due to buoyancy in spaces with high ceilings.

To date, well-mixed models are primarily used to make quantitative risk assessment for decision-making. We show quantitatively how much the concentration distribution (mean and standard deviation) varies and its implications on risk of infection. It was found that the concentration field can be modeled using a shifted and scaled lognormal distribution. A simple expression for the standard deviation was proposed based on easy-to-measure parameters (ventilation rate and room geometry). By sampling the exposure dose from the modeled concentration distribution in a dose–response function, events corresponding to short- and long-range transmission are captured. This was found to yield an order of magnitude change in probability of infection compared to what a well-mixed assumption predicts depending on the relative position of the susceptible individual.

Finally, we note the proposed model is complementary to many existing epidemiological models. For example, additional effects can be included in the well-mixed mass balance, such as deactivation due to ultraviolet lights or air filters.[10] The concentration PDF can also be employed within multi-zone models to incorporate spatial heterogeneity within each zone. In addition, the effect of masks can be incorporated by adjusting the mean in the well-mixed mass balance and the breathing rate used in estimating the dose. Due to its low computational cost, the proposed model can be easily integrated into risk analyses for determining safety guidelines for mitigating airborne transmission, which to date have relied on well-mixed assumptions.[17,48]

ACKNOWLEDGEMENTS

This work was supported by the College of Engineering at the University of Michigan. The computing resources and assistance provided by the staff of the Advanced Research Computing at the University of Michigan are greatly appreciated.

CONFLICT OF INTEREST

No conflict of interest declared.

AUTHOR CONTRIBUTIONS

ST and ZZ conducted simulations, performed grid refinement, processed data, and wrote portions of the paper. KM contributed to the CFD code and guided the analysis. KJF assisted with the probabilistic model and Monte Carlo sampling of the dose–response function. JC contributed to the data analysis, model development, and writing of the manuscript.

ORCID

Sijian Tan https://orcid.org/0000-0002-1649-5883
Jesse Capecelatro https://orcid.org/0000-0002-1684-4967

REFERENCES

1. Zhang R, Li Y, Zhang A, Wang Y, Molina M. Identifying airborne transmission as the dominant route for the spread of COVID-19. *Proc Natl Acad Sci USA*. 2020;117:14857-14863.
2. Delikhoon M, Guzman M, Nabizadeh R, Baghani A. Modes of transmission of severe acute respiratory syndrome-Coronavirus-2 (SARSCoV-2) and factors influencing on the airborne transmission: a review. *Int J Environ Res Public Health*. 2021;18:395.
3. Miller S, Nazaroff W, Jimenez J, et al. Transmission of SARS-CoV-2 by inhalation of respiratory aerosol in the Skagit Valley Chorale superspreading event. *Indoor Air*. 2021;31:314-323.
4. Morawska L, Tang J, Bahneth W, et al. How can airborne transmission of COVID-19 indoors be minimised? *Environ Int*. 2020;142:105832.
5. Shen Y, Li C, Dong H, et al. Community outbreak investigation of SARS-CoV-2 transmission among bus riders in eastern China. *JAMA Intern Med*. 2020;180(12):1665.
6. Stilianakis N, Drossinos Y. Dynamics of infectious disease transmission by inhalable respiratory droplets. *J R Soc Interface*. 2010;7:1355-1366.
7. Liu L, Li Y, Nielsen P, Wei J, Jensen R. Short-range airborne transmission of expiratory droplets between two people. *Indoor Air*. 2017;27:452-462.
8. Wei J, Li Y. Airborne spread of infectious agents in the indoor environment. *Am J Infect Control*. 2016;44:S102-S108.
9. Balachandar S, Zaleski S, Soldati A, Ahmadi G, Bourouiba L. Host-to-host airborne transmission as a multiphase flow problem for science-based social distance guidelines. *Int J Multiph Flow*. 2020;132:103439.
10. Evans M. Avoiding COVID-19: aerosol guidelines. arXiv preprint arXiv:2005.10988. 2020.
11. Watanabe T, Bartrand T, Weir M, Omura T, Haas C. Development of a dose-response model for SARS coronavirus. *Risk Anal*. 2010;30:1129-1138.
12. Wells W. Airborne contagion and air hygiene. An ecological study of droplet infections. *JAMA*. 1955;159:90.
13. Riley E, Murphy G, Riley R. Airborne spread of measles in a suburban elementary school. *Am J Epidemiol*. 1978;107:421-432.
14. Nicas M. An analytical framework for relating dose, risk, and incidence: an application to occupational tuberculosis infection. *Risk Anal*. 1996;16:527-538.
15. Gammaitoni L, Nucci M. Using a mathematical model to evaluate the efficacy of TB control measures. *Emerg Infect Dis*. 1997;3:335.
16. Beggs C, Noakes C, Sleigh P, Fletcher L, Siddiqi K. The transmission of tuberculosis in confined spaces: an analytical review of alternative epidemiological models. *Int J Tuberc Lung Dis*. 2003;7:1015-1026.
17. Bazant M, Bush J. A guideline to limit indoor airborne transmission of COVID-19. *Proc Natl Acad Sci USA*. 2021;118(17):e2018995118.
18. Vuorinen V, Aarnio M, Alava M, et al. Modelling aerosol transport and virus exposure with numerical simulations in relation to SARS-CoV-2 transmission by inhalation indoors. *Saf Sci*. 2020;130:104866.
19. Noakes C, Sleigh P. Mathematical models for assessing the role of airflow on the risk of airborne infection in hospital wards. *J R Soc Interface*. 2009;6:S791-S800.
20. Anchordoqui L, Chudnovsky E. A physicist view of COVID-19 airborne infection through convective airflow in indoor spaces. *SciMed J*. 2020;2:68-72.
21. Zhang Z, Capecelatro J, Maki K. On the utility of a well-mixed model for predicting disease transmission on an urban bus. *AIP Adv*. 2021a;11:085229.
22. Sze To G, Chao C. Review and comparison between the Wells-Riley and dose-response approaches to risk assessment of infectious respiratory diseases. *Indoor Air*. 2010;20:2-16.
23. Axley J. Multi-zone dispersal analysis by element assembly. *Build Environ*. 1989;24:113-130.
24. Bouia H, Dalicieux P. Simplified modeling of air movements inside dwelling room. In: *Proceedings of Building Simulation'91 Conference, IBPSA (The International Building Performance Simulation Association)*. 1991;106-110.
25. Nicas M, Miller S. A multi-zone model evaluation of the efficacy of upper-room air ultraviolet germicidal irradiation. *Appl Occup Environ Hyg*. 1999;14:317-328.
26. Ko G, Burge H, Nardell E, Thompson K. Estimation of tuberculosis risk and incidence under upper room ultraviolet germicidal irradiation in a waiting room in a hypothetical scenario. *Risk Anal*. 2001;21:657-674.
27. Noakes C, Khan M, Gilkeson C. Modeling infection risk and energy use of upper-room ultraviolet germicidal irradiation systems in multi-room environments. *Sci Technol Built Environ*. 2015;21:99-111.
28. Ko G, Thompson K, Nardell E. Estimation of tuberculosis risk on a commercial airliner. *Risk Anal*. 2004;24:379-388.
29. Jones R, Masago Y, Bartrand T, Haas C, Nicas M, Rose J. Characterizing the risk of infection from *Mycobacterium tuberculosis* in commercial passenger aircraft using quantitative microbial risk assessment. *Risk Anal*. 2009;29:355-365.
30. Li Y, Duan S, Yu I, Wong T. Multi-zone modeling of probable SARS virus transmission by airflow between flats in Block E, Amoy Gardens. *Indoor Air*. 2005;15:96-111.
31. Nicas M. Markov modeling of contaminant concentrations in indoor air. *AIHAJ*. 2000;61:484-491.
32. Mora L, Gadgil A, Wurtz E. Comparing zonal and CFD model predictions of isothermal indoor airflows to experimental data. *Indoor Air*. 2003;13:77-85.
33. Sun C, Zhai Z. The efficacy of social distance and ventilation effectiveness in preventing COVID-19 transmission. *Sustain Cities Soc*. 2020;62:102390.
34. Guo Y, Qian H, Sun Z, et al. Assessing and controlling infection risk with Wells–Riley model and spatial flow impact factor (SFIF). *Sustain Cities Soc*. 2021;67:102719.
35. Fabregat A, Gisbert F, Vernet A, Dutta S, Mittal K, Pallarès J. Direct numerical simulation of the turbulent flow generated during a violent expiratory event. *Phys Fluids*. 2021;33:035122.
36. Monroe K, Yao Y, Lattanzi A, Raghav V, Capecelatro J. Role of pulsatility on particle dispersion in expiratory flows. *Phys Fluids*. 2021;33:043311.

37. Chong K, Ng C, Hori N, Yang R, Verzicco R, Lohse D. Extended lifetime of respiratory droplets in a turbulent vapor puff and its implications on airborne disease transmission. *Phys Rev Lett*. 2021;126:034502.
38. Noakes C, Sleigh P, Escombe A, Beggs C. Use of CFD analysis in modifying a TB ward in Lima, Peru. *Indoor Built Environ*. 2006;15:41-47.
39. Qian H, Li Y, Nielsen P, Huang X. Spatial distribution of infection risk of SARS transmission in a hospital ward. *Build Environ*. 2009;44:1651-1658.
40. He Q, Niu J, Gao N, Zhu T, Wu J. CFD study of exhaled droplet transmission between occupants under different ventilation strategies in a typical office room. *Build Environ*. 2011;46:397-408.
41. Abuhegazy M, Talaat K, Anderoglu O, Poroseva S. Numerical investigation of aerosol transport in a classroom with relevance to COVID-19. *Phys Fluids*. 2020;32:103311.
42. Li Y, Qian H, Hang J, et al. Probable airborne transmission of SARSCoV-2 in a poorly ventilated restaurant. *Build Environ*. 2021;196:107788.
43. Zhang Z, Han T, Yoo K, Capecelatro J, Boehman A, Maki K. Disease transmission through expiratory aerosols on an urban bus. *Phys Fluids*. 2021b;33:015116.
44. Gao N, Niu J, Perino M, Heiselberg P. The airborne transmission of infection between flats in high-rise residential buildings: tracer gas simulation. *Build Environ*. 2008;43:1805-1817.
45. Launder B, Spalding D. The numerical computation of turbulent flows. In: *Computer Methods in Applied Mechanics and Engineering*. 1974;3:269-289.
46. Abkarian M, Mendez S, Xue N, Yang F, Stone H. Speech can produce jet-like transport relevant to asymptomatic spreading of virus. *Proc Natl Acad Sci USA*. 2020;117:25237-25245.
47. Wykes MD, Chahour E, Linden P. The effect of an indoor-outdoor temperature difference on transient cross-ventilation. *Build Environ*. 2020;168:106447.
48. Swanson T, Guikema S, Bagian J, Schemanske C, Payne C. COVID-19 aerosol transmission simulation-based risk analysis for in-person learning. *medRxiv*. 2021.

How to cite this article: Tan S, Zhang Z, Maki K, Fidkowski KJ, Capecelatro J. Beyond well-mixed: A simple probabilistic model of airborne disease transmission in indoor spaces. *Indoor Air*. 2022;32:e13015. doi:10.1111/ina.13015

APPENDIX A

ADDITIONAL SIMULATION DETAILS

Details on the numerical simulations performed in this study are provided in Tables A1 and A2 below. In each case, the computational domain is divided into a set of non-overlapping hexahedral cells. Local grid refinement is applied to small features, such as the return and supply vents, occupants, interior walls, tables, and seats. A grid refinement study of the present configuration can be found in Zhang et al.[43].

A.1 | Classroom

For the classroom simulations, three different volumes are considered: 120 m^3, 180 m^3, and 240 m^3. For the 120 m^3 cases, fixed dimensions (8.95 m × 5.46 m × 2.45 m) are used to analyze the effect of different positions of infectors and different ACH values. For the 180 m^3 and 240 m^3 cases, to analyze the effect of the geometry, multiple dimensions with different aspect ratio values are used (shown in Table B1).

For all the classroom cases, similar HVAC systems are applied. The width of the inlet vents is 5 cm and the length of the inlet vents equals to the length of the classroom (shown in Figure 1A). The size of the outlet vents is the same for each case, which is 0.3 m × 0.6 m × 2 (two outlets). The maximum ACH value of this HVAC is 6, and different ACH values (1,3,6) are used for simulations.

The infector is placed at six different locations within the simulations, both close and far relative to the supply/return vents within the room (see Figure A1).

A.2 | Lecture hall

A stereolithography (STL) file is created using computer-aided design (CAD) software to represent the geometry of the Chrysler Lecture Hall at the University of Michigan (shown in Figure 1C), which has a large volume of 1753 m^3. Manikins of 46 sitting students and a standing instructor, seats, and instruments are placed inside the hall. Four different infector positions are chosen (one for each simulation). The entire simulation period is 90 min, where the sick person emits in the whole period. Different ACH values (1,6) are used for simulations.

A.3 | Music rehearsal hall

A rectangular box is adopted to represent the simplified geometry of a music rehearsal hall, which has a large volume of 2488 m^3 with dimensions of 19.82 m × 15.24 m × 8.24 m ($L \times W \times H$). Manikins of 14 sitting students and a standing conductor, seats, and instruments are placed inside the hall. The entire simulation period consists of two 40-min classes with a 20-min break in between, where the sick person only emits during classes. No HVAC system is present in this hall, and the airflow is mainly driven by pulmonary ventilation of the occupants and temperature gradients. A door is set to balance the pressure inside the domain.

A.4 | Urban bus

For the bus simulations, an urban bus that is used on the campus of the University of Michigan is studied. The geometry including the interior of the cabin, windows, doors, seats, handrails, ventilation supply and return are determined from a laser scanner and used for generating the computational grid of the fluid domain. A rendering of the bus is shown in Figure 1E. The bus dimensions are 12.1 m × 2.58 m × 2.95 m ($L \times W \times H$) with a total interior volume of 52 m^3. A total of 42 supply vents are located along both sides of the bus ceiling and have an orientation such that air exits vertically downward. The single ventilation fan draws air from the passenger compartment through a return vent in the back of the bus and adds 20% fresh air from outside before returning the air to the cabin through supply

TABLE A1 Mesh resolution and the range of total number of cells

Scenario	Mesh resolution [mm]					No. of cells [million]
	Background	Supply vents	HVAC return	Mouth	Other surfaces	
Classroom	100	6.2	12.5	3.1	6.2	4.07–4.24
Lecture hall	100	12.5	12.5	3.1	25	6.47
Rehearsal	125	–	–	3.9	7.8	3.17
Urban Bus	125	2	31.2	3.9	7.8	5.86

TABLE A2 Boundary conditions and material properties

Boundary name	Boundary conditions
HVAC supply vents	Velocity inlet with uniform flow, 25°C, 10% turbulence intensity, and turbulence length scale is 5×10^{-3} m
HVAC return	Pressure outlet
Mouth	Constant exhalation, 6 l/min, 34°C, 2.5% turbulence intensity, turbulence length scale is 7.5×10^{-3} m and $q = 50\ s^{-1}$
Manikins	No-slip and 32°C
Other surfaces	No-slip and adiabatic walls
Material parameters	$\nu = 1.5 \times 10^{-5}\ m^2/s$, $T_0 = 20°C$, $\beta = 3 \times 10^{-3}\ K^{-1}$, $Pr = 0.71$, $Pr_t = 0.9$, $g = 9.81\ m/s^2$, $Sc = 1$, $Sc_t = 1$

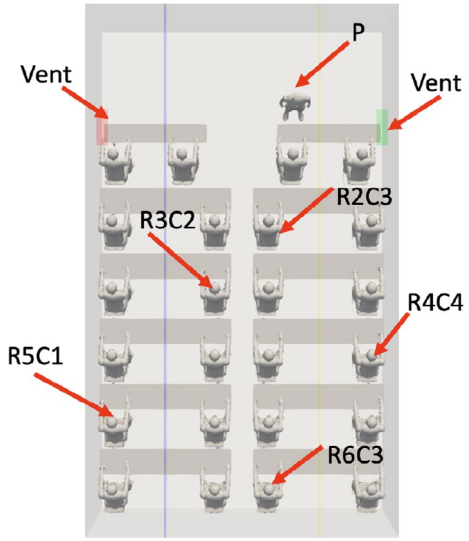

FIGURE A1 Six positions of the infector used in the classroom simulations.

vents. Each supply vent has a dimension of 9 in by 1 in (0.229 m by 0.0254 m), and the single return vent is 4 ft by 1.5 ft (1.22 m by 0.457 m). This HVAC system can provide a maximum flow rate of 2,500 ft^3 /min (70.8 m^3/min), which is equivalent to an ACH of 16 considering the fresh air rate of 20%. With such ventilation system, the airflow moves up and down in transverse direction and the net flow is rearward through the compartment. Manikins are placed at different locations inside the bus: a driver sitting behind the wheel and standing passengers. Further details can be found in Zhang et al.[43].

APPENDIX B

Simulation results

TABLE B1 Mean and standard deviation of the concentration field obtained from CFD. Mean obtained from the well-mixed mass balance (2) shown for reference

Description	Volume [m^3]	$\frac{\sqrt{A}}{H}$ [–]	ACH [h^{-1}]	Duration [min]	$\overline{C}$ (Well-mixed) [m^{-3}]	$\overline{C}$ [m^{-3}]	C_{std} [m^{-3}]
Classroom	120	2.85	1	90	1166	1165.7	335.2
Classroom	120	2.85	3	90	495	498	255.7
Classroom	120	2.85	6	90	250.2	259	233.6
Classroom	180	2	1	90	776.9	780.6	111.2
Classroom	180	2	3	90	329.6	332.7	86.3
Classroom	180	2	6	90	166.6	178.7	109.4
Classroom	180	2.3	1	90	776.9	759.5	113.7
Classroom	180	2.3	3	90	329.6	356.7	136.9
Classroom	180	2.3	6	90	166.6	176.3	125.6
Classroom	180	2.5	1	90	776.9	790.7	177.2
Classroom	180	2.5	3	90	329.6	361.2	162.6
Classroom	180	2.5	6	90	166.6	169.8	132.3
Classroom	180	2.6	1	90	776.9	790.7	215.5
Classroom	180	2.6	3	90	329.6	359.2	165.7
Classroom	180	2.6	6	90	166.6	170.7	129
Classroom	180	2.8	1	90	776.9	779.6	236.7
Classroom	180	2.8	3	90	329.6	347.3	195.6
Classroom	180	2.8	6	90	166.6	166.3	144.3
Classroom	180	3	1	90	776.9	786.2	273.6
Classroom	180	3	3	90	329.6	342.3	204.5
Classroom	180	3	6	90	166.6	161.3	155.9
Classroom	240	2	1	90	582.6	600.5	126
Classroom	240	2	3	90	247.2	271.9	91.8
Classroom	240	2	6	90	125.0	132.7	93
Classroom	240	2.5	1	90	582.6	592.6	167.9
Classroom	240	2.5	3	90	247.2	263	121.4
Classroom	240	2.5	6	90	125	130	105
Classroom	240	3	1	90	582.6	581.7	183
Classroom	240	3	3	90	247.2	264.3	147
Classroom	240	3	6	90	125	128.4	101
Lecture Hall	1753	2.45	1	90	79.8	83.9	52.8
Lecture Hall	1753	2.45	6	90	17.1	21.6	20.2
Music Hall	2488	2.1	0	80	96.5	96.5	122.1
Bus	52	2.2	16	15	213	183.7	74.5
Bus	52	2.2	1.6	15	713.7	680	469.3