Analysis of Fall 2021 Classroom COVID-19 Transmission

Cornell Covid Modeling Team
October 2021

This document presents an analysis of the risk of transmission of COVID-19 during classes in fall 2021 incorporating the effect of the following mitigations:

- masking in class,
- distancing,
- ventilation,

and for the following groups of individuals:

- undergraduate students,
- faculty instructors,
- graduate student teaching assistants (TA).

This document uses a theoretical model and data published in the scientific literature to estimate the probability faced by an individual of becoming infected with COVID due to classroom interaction over the course of the entire Fall 2021 semester. For example, the nominal estimate of this probability for vaccinated faculty instructors is 0.017%, or about 1 in 6000.

It updates an analysis done over the summer to incorporate new information available from the literature and Cornell's Fall 2021 semester data about vaccine effectiveness and prevalence in the Cornell student population. Estimates of vaccine effectiveness were revised downward in the scientific literature through a sequence of papers and public reports described herein, and prevalence at the start of the semester was higher than anticipated, leading to an increase in estimated risks compared to those calculated in the summer.

The risk estimated in this document depends on parameters, about which we have uncertainty. For this reason, we produce *distributions* of risk estimates, rather than single values, and report a few key statistics of these distributions (Table 1). In particular, we report the median, which can be interpreted as our best estimate of the risk faced, and we refer to it as our nominal estimate. The estimated median per-person risk of classroom infection over the entire course of the fall semester is 0.404% for undergraduate students (Section 2.1), 0.035% and 0.017% for unvaccinated and vaccinated faculty instructors respectively (Section 2.2), and 0.005% and 0.002% for unvaccinated and vaccinated graduate student TAs respectively (Section 2.2).

We also report the 95% quantile of the risk distribution, which is a risk that would be faced if the parameters governing reality fall on the more pessimistic end of the modeled range, and the 5% quantile, which is a risk faced if the parameters fall on the more optimistic end. Among a

number of assumptions we especially emphasize an assumption of limited student talking in the assumed classroom environment.

Group	Nominal (median) estimates of the per-person risk of infection	Optimistic (5% quantile) estimates of the per-person risk of infection	Pessimistic (95% quantile) estimates of the per-person risk of infection
Undergraduate students	0.404%	0.088%	1.092%
Unvaccinated faculty instructors	0.035%	0.012%	0.080%
Vaccinated faculty instructors	0.017%	0.006%	0.047%
Unvaccinated graduate student TAs	0.005%	0.002%	0.011%
Vaccinated graduate student TAs	0.002%	0.0006%	0.006%

Table 1: Nominal, optimistic and pessimistic estimates of per-person infection risk over the course of the fall semester for different groups due to lecture transmission over the Fall 2021 semester.

While the risk faced by any given individual is quite low, it is not zero. Moreover, the number of students in classrooms at Cornell is large and the estimates above predict that some infections will occur.

With this in mind, we have been closely following instances when multiple students test positive in the same classroom. The rate at which this occurs is consistent with either no in-class transmission or rare in-class transmission of a similar order of magnitude to our analysis in this document. Large classes have many students and so coincidental positives are possible. Also, certain classes are more popular with individuals who have other risk factors (e.g., membership in active Greek-life social organizations), which makes it extremely likely for there to be multiple positives in the same class, even in the absence of classroom transmission. Individual investigation of cases as well as adaptive testing of classrooms with positive cases has not revealed any instances of classroom transmission, which is consistent with the finding here that classroom transmission is likely to be rare.

We are collecting seating assignment data from Bailey Hall, a lecture hall on campus, in an attempt to either find classroom transmission or establish an empirical upper bound on how rare it is, though this analysis is still in its preliminary stages.

Instructor positives have been rare. Detailed investigations of the handful of instructor cases that have occurred suggest that they were very unlikely to have been created by a classroom transmission, e.g., because the instructor hadn't taught recently in person at the time of the positive test result, or there were no student positives in their class. Similarly, we have not observed any instances of likely transmission between a student and a TA. The risk estimates

suggest it is plausible, and even likely, that we would not see an instructor or TA infected by classroom transmission during the fall semester.

In contemplating the risks, it is important to emphasize that vaccination confers a great deal of protection against serious symptoms once infected. Our Fall 2021 Modeling Report synthesizes the literature to estimate the risk that a vaccinated infected person faces of experiencing symptoms severe enough to require hospitalization. For a few age cohorts, it finds the risk to be:

- 0.036% - 0.072% for those aged 18-30 years old
- 0.22% - 0.32% for those aged 40-50
- 1.8% - 3.6% for those aged 65-75

The risk to other age ranges is between the ones given above: the risk to 30 to 40-year-olds lies between that of 18 to 30 and 40 to 50-year-olds; the risk to 50 to 65-year-olds lies between that of 40 to 50 and 65 to 76-year-olds.

In all of this modeling, as in all modeling, we emphasize limitations due to modeling assumptions. Some of the key assumptions include:

- Breathing is the primary activity of students in class rather than talking and there is no singing or yelling.
- Compliance with masking on behalf of all individuals in the classroom is 100%.
- Social distancing is maintained throughout the lecture experience, including any questions posed by students to faculty or TAs after class.
- Classroom ventilation is within certain standards discussed herein.
- We do not include any additional risks incurred during face-to-face office hours or in other student/instructor/TA interaction outside the classroom.

1. General Modeling Approach for Transmission Risk Estimation

Our estimation of transmission risk takes two steps.

- We simulate the expected number of secondary infections per hour in a classroom with one infectious individual under different *activity conditions* and *intervention measures*. In particular, the activity conditions are:
    - vaccination level among the attendees
    - type of respiratory activity.

  The intervention measures are:
    - enforcement of masking
    - seating density
    - ventilation
    - indoors vs. outdoors.
- Based on the simulation results, we then calculate the risk of infection due to lecture transmission in the fall 2021 semester for different groups of individuals, taking into

account the duration of lecture attendance throughout the semester and the prevalence among students.

We first briefly state the assumptions made about the different conditions, intervention measures, and student prevalence. Detailed justification for these assumptions is given in Section 3.1.

1.1 Activity Conditions

Condition 1: Vaccination level

The level of vaccination among the population is a key variable in the simulation because vaccination (of either the source or the susceptible individual) reduces the risk of transmission, resulting in fewer secondary infections.

We focus our calculation on fully vaccinated individuals and unvaccinated individuals. Those who are partially vaccinated, or those that received the second dose less than 14 days ago, are not explicitly considered. The risk of these population groups lies between the risk for fully vaccinated and unvaccinated individuals. Hereafter, the word "vaccinated" is meant exclusively for fully vaccinated individuals.

The vaccination level in the Cornell population is high. As of mid August, 94% of Cornell undergraduates have been fully vaccinated, and an updated statement from the Provost states that we anticipate 99% of the students to be soon vaccinated. We conservatively use 90% as the nominal vaccination level in the Cornell population (i.e., the population in Cornell-managed surveillance or exempt students and employees). The model allows us to adjust this vaccination level for specific instances to get additional results that pertain to special events, for example.

Our modeling includes parameters for vaccine efficacy. Vaccine efficacy has two aspects. On the one hand, vaccination may reduce the viral load of a source case, thereby reducing their chance of infecting others; on the other hand, vaccination offers protection against infection for susceptible individuals. We focus on the Delta variant, which already makes up 83% of US cases as of July 20, 2021 [CNN]. We assume that the Delta variant had dominated all infections by the time the semester started. This may be slightly conservative, yet not unrealistic given Delta's fitness and high transmissibility. We assume the Delta variant is **2.4** times more transmissible than the non-variant strain (details in Section 3.1).

We have uncertainty about vaccine efficacy parameters because there is significant disagreement in the literature about these parameters. To quantify this uncertainty, we place Bayesian prior probability distributions on these parameters instead of using fixed values. We detail these prior distribution here:

- Prior on vaccine efficacy in reducing the viral load of a source case:

    - The literature reports varying results, ranging from a 2.8-4.5 fold reduction in viral load [Levine-Tiefenbrun et al. 2021] (on average scaling viral load by a factor of 0.29, i.e., reducing transmission by 71%), 40%-50% reduction of transmission risk [Harris et al. 2021] to no reduction in viral load [Brown et al. 2021] or peak viral load [Pouwels et al. 2021].
    - Given the significant uncertainty around this parameter, we set the prior to be **uniformly distributed over three values: 0%, 50% and 71%,** corresponding respectively to the pessimistic, middle, and optimistic ends of the ranges reported in the literature.
- Prior on vaccine efficacy in protecting against infection for susceptible individuals:
    - We collect estimates from the literature, including **42%** [Puranik et al. 2021], **79%** [Sheikh et al. 2021], **88%** [Bernal et al. 2021], **40%** [Pouwels et al. 2021] for Pfizer and **76%** [Puranik et al. 2021], **66%** [Fawlkes et al. 2021] for Moderna.
    - We set the prior to be **uniformly distributed over these six values**.
    - This prior is optimistic in the sense that some of these results are measured shortly after vaccination, while Pouwels et al. 2021 observed that the protection provided by vaccination decays over time.
    - This prior is conservative (i.e., pessimistic) in that the studies above were performed on general populations. Cornell has a larger fraction of the younger population, and vaccine efficacy was observed to be higher for younger people and lower for older [Pouwels et al. 2021]. However, there is not sufficient evidence in the literature to support further investigation of age-stratified vaccine efficacy.
    - Limitations of this approach include: (1) it does not explicitly consider the statistical error reported in each of the individual point estimates in the papers above, though these errors are small compared to the difference in estimates across papers, and (2) it does not differentiate between different types of vaccines.

Condition 2: Type of respiratory activity

We model COVID-19 transmission via *droplets* and *aerosols*. Droplets are large particles that fall to the ground within 6 ft of emission; aerosols are small particles that suspend in the air and spread to the entire room. The mechanism is described in detail in Section 3.3.

Activities at different respiratory intensities (such as breathing, speaking, and singing) are associated with different levels of droplet and aerosol emission. We model a normal lecture as dominated by "breathing" because students mostly sit and listen. In a lecture where a handful of students respond to scattered questions, the additional risk is negligible.

In our calculations we assume that infectious students do not cough or sneeze during class. Thus, the focus of this analysis is on protection against asymptomatic or pre-symptomatic students attending class. Other measures (surveillance, a general awareness of COVID

symptoms, and access to testing for students with COVID-like symptoms) are thus important for ensuring that symptomatic students do not attend class.

1.2 Intervention Measures

Interventions in classrooms can be used to reduce the risk of transmission in those spaces. We next describe how they are modeled in our simulation. Table 2 presents an overview of the qualitative modeled effectiveness of different intervention measures in reducing droplet and aerosol transmission.

Intervention measures	**Droplets**	**Aerosols**
Distancing	reduces droplet risk	/
Masking	works equally for both, reduces exhalation and inhalation	
Ventilation	/	reduces aerosol risk
Outdoors	/	makes aerosol risk negligible

Table 2: Summary of qualitative modeled effects of intervention measures on droplet and aerosol transmission risk. "/" means no effect.

We now describe the modeling for each intervention in detail.

Intervention 1: Distancing

- We consider three distancing options:
    - **1'** distancing: fully dense seating where the pre-pandemic seating capacity is fully used; this is our default assumption, though we consider others in case of the need for additional interventions;
    - **6'** distancing: socially distanced seating used during the pandemic;
    - **3’** distancing: a density that is in between the two extremes.
- In all cases, we pessimistically assume that all students enrolled in a class attend every lecture and that classroom assignments are such that all seats in the classroom are needed for students in the class. In reality, the number of students in a classroom on any given day is smaller than this pessimistic assumption. In addition, students who have behaviors that put them at higher risk of catching COVID may also have lower attendance rates than the general student population, which would lower risk compared to our assumption that all students attend every class.
- In arranging the unvaccinated students among the available seats, we simulate “clump” seating, which pessimistically assumes that unvaccinated students tend to sit in nearby available seats (because the same demographic factors that cause students to be unvaccinated upon arrival to Ithaca may also create social connections).
- Examples of seating arrangements at different classroom densities are given below in Figure 1.

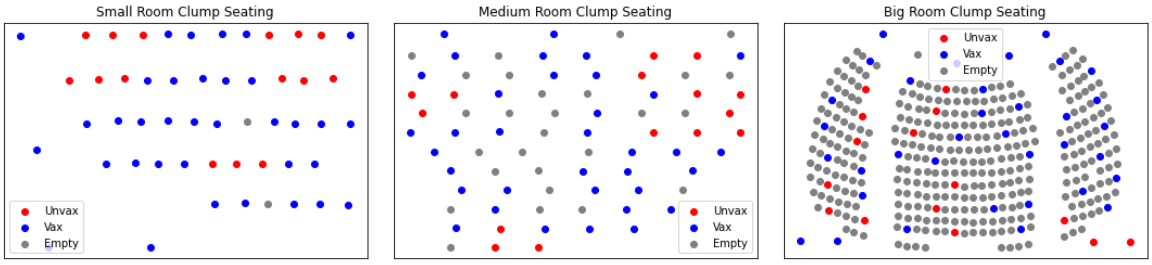

Figure 1: Examples of seating arrangements of 50 students, including vaccinated ("Vax") and unvaccinated ("Unvax"), under three classroom densities (small room, i.e., fully dense seating; medium room, i.e., 3' social distancing; and big room, i.e., AY20-21 6' socially distanced seating). Under clump seating, unvaccinated students tend to sit near each other.

Intervention 2: Masking

The effectiveness of masking is subject to various sources of uncertainty. The material of a mask and whether it is properly worn both significantly affect the effectiveness of masking. Based on experimental and observational studies in the literature [Wang et al. 2020, Kumar et al. 2020, Konda et al. 2020, Doung-ngern et al. 2020, Howard et al. 2020], we assume a pessimistic estimate of **50%** and an optimistic estimate of **80%** for one-way masking effectiveness. That is, if an infectious individual is masked, *or* if a susceptible individual is masked, the risk of transmission is reduced by 50% - 80%. If both the infectious and susceptible individuals are masked, the risk of transmission is thus scaled by a factor of **4% - 25%**.

Effective July 30, 2021, Cornell is requiring all individuals, including fully vaccinated ones, to wear masks indoors [Announcement]. It is therefore reasonable to assume that all individuals in the classroom are masked. For two-way masking, we set the prior on the risk reduction factor to be **Normal(0.145, 0.0536^2)** (derived from assuming a normal prior with a symmetric 95% credible interval of [0.04, 0.25]), **truncated below at 0**.

The smaller the risk reduction factor, the more protection masking offers. Equivalently, one can think of two-way masking effectiveness as normally distributed with mean 1-0.145 = 85.5% and the same variance.

Intervention 3: Ventilation

Ventilation facilitates air exchanges in the room and reduces the time that virus-containing aerosols suspend in the air. We simulate ventilation by considering different numbers of air exchanges per hour (ACH).

Rooms on Cornell campus have different ventilation conditions. Some are equipped with more frequent mechanical ventilation while others use natural ventilation through windows and doors. According to communication with Facilities and Campus Services, most rooms with mechanical ventilation have between 1 and 2 ACH, with some having significantly higher ACH. We simulate three scenarios:

- no ventilation: the amount of aerosols is not reduced. This is a pessimistic estimate of the ventilation in naturally ventilated rooms.
- 1 ACH: reduces the amount of aerosols to ½ of their value without ventilation. This is a pessimistic estimate of the amount of ventilation in a typical classroom on campus.
- 3 ACH: reduces the amount of aerosols to ¼ of their value without ventilation. This models the amount of ventilation in a room on campus that has above average ventilation.

For rooms relying on natural ventilation, it is very difficult to measure their air exchange rates exactly. However, they roughly lie somewhere between the "no ventilation" scenario and the 1 ACH scenario, and thus the "no ventilation" scenario provides a conservative estimate.

1.3 Prevalence among students

Prevalence (more precisely, prevalence in classrooms) is defined as the percentage of students that are infectious with COVID and attending class (i.e., not in isolation because of a positive test result or quarantine because of lack of vaccination or symptoms following close contact, and not missing class due to symptoms or other reasons) at a given time. It determines the probability that a randomly selected classroom hour contains a positive student.

As prevalence among students is subject to significant uncertainty, we place a prior distribution on possible values it could take. The calculation is outlined below, with more details in Section 3.1.

- We start by modeling the total number of student infections in the fall 2021 semester.
    - We assume a **lognormal(6.872, 0.372)** distribution on the total number of student infections in the semester. The mode of this distribution (i.e., the point with the largest probability density) is **840** and the 97.5% quantile is 2000.
    - The lognormal distribution is heavy-tailed, so it accounts for the possibility of having large numbers of infections.
    - We truncate the distribution below at 488, the number of cases observed at the time parameters were fit for this analysis (Sep. 19) since the start of the fall semester (Aug. 26).
- Then, we estimate prevalence from the total number of infections in the semester as follows:
    - We assume each case is free and infectious for half a week before isolation due to surveillance testing.
    - We assume for simplicity that prevalence is constant over the semester, though the calculation could be expanded to multiple stages with different prevalence. Moving to a calculation with multiple stages would have very little effect on the average risk over the course of the semester.
    - We assume recovered cases return to being equally susceptible as the rest of the population. This makes our calculation slightly conservative, as it does not take into account natural immunity due to past infections.

- Under these assumptions, the prevalence at any time is estimated by evenly distributing the cases and the time infectious in classrooms that they create over the entire semester into ½-week periods across the entire undergraduate population. This gives:
  Prevalence = (total # cases over the semester * ½ week infectious) / (undergraduate population * 14 weeks per semester) = (total # cases over the semester) / (undergraduate population * 28).
- For example, 840 cases in the entire semester corresponds to 840/(15000*28) = **0.2%** prevalence at any time, assuming an undergraduate population of 15000 [University Facts].

2. Results

We report estimated transmission risk in the following scenarios:

- infection risk for students in classrooms
- infection risk for both faculty and graduate student TAs in classrooms

2.1 Infection risk for students in classrooms

We consider a "nominal classroom setting" to be one where students are densely seated (1' distancing with no empty seats), all masked, and sitting and listening (i.e., breathing).

Our simulation assumes a 90% vaccination level in the classroom and outputs a distribution over the number of secondary infections per classroom hour in a dense all-masked classroom given one infectious student in the classroom. This distribution depends on the prior distributions on the vaccine efficacy parameters, as specified in Section 1. 90% is much lower than the current student vaccination level, and is also lower than the vaccination level at the start of the semester when some students without a medical or religious exemption had not yet been vaccinated. Thus, our analysis is somewhat conservative in this regard.

Next, we extrapolate this per-hour, conditional risk to estimate the risk of classroom transmission for students over the entire semester. This depends on the prevalence p in the student population, i.e., the percent of students that are positive and not in isolation at a given time. Below, we first explain how to calculate the risk of classroom transmission for any prevalence level and then present concrete estimates.

In our calculation, we assume that all positives that occur in classrooms occur when everyone else in the classroom is susceptible, i.e., not infected or infectious. The fact that a few positives occur in the same classrooms at the same time makes reality slightly more optimistic than our estimates --- in our estimates, everyone else in the classroom is susceptible while in reality, the other infectious individual cannot be infected again; in addition, the increase in the risk of infection created by adding a second positive is smaller than the increase created by adding a

first positive[1] (see Section 3.3 for a brief discussion). The level of optimism introduced by this fact is really extremely small, however, and our assumption produces nearly the same estimate as one that allowed multiple positives to be in the same classroom.

Our calculation takes the following steps:

- By the definition of prevalence, the number of students that are positive and attending classes at any time is the product of prevalence p and the size of the student population N. (Recall that p is the fraction of students attending classes who are infectious at any moment in time.)
    - We choose N to be 15K, corresponding to the undergraduate population [University Facts].
- Then, the number of classroom hours with an infectious student over the entire semester is $p * N$ times the average number of hours that a student spends in the classroom over the entire semester T.
    - We consider an average student enrolled for 15 credits with 45 hours of course-related work per week, including both lecture, non-lecture time in classrooms (e.g., recitation), and time spent outside of class on homework and other coursework.
    - We assume ½ of the 45 hours is spent in the classroom.
    - Over a 14-week semester, a student spends 45 / 2 * 14 = 315 hours in the classroom. We choose this to be the value of T.
    - This estimate is slightly optimistic in that non-lecture classroom time, such as recitations, may involve more talking than lectures.
- Finally, we multiply $p * N * T$ by the number of secondary infections per classroom hour from one infectious individual to obtain the expected number of student secondary infections due to classroom transmission over the semester.
- To get the per-person risk of infection, we further divide the expected number of student secondary infections over the semester by N.

We now present the estimated distribution of per-student infection risk due to classroom transmission, where model parameters (prevalence, vaccine efficacy, and masking effectiveness) follow the prior distributions specified in Section 1. The median risk is **0.404%** (corresponding to 13 cases) and the (5%, 95%) quantile values are **(0.088%, 1.092%)** (corresponding to 61 and 164 cases).

---

[1] This follows when the dose resulting from each positive is independent and identically distributed and from concavity of the probability of infection as a function of the dose, given in Section 3.3. In particular, let V_1 and V_2 be the (strictly positive) dose to a given susceptible person associated with the first and second positives in a classroom, so that the dose is V_1 if there is one positive in the classroom and $V_1 + V_2$ if there are two. We assume that V_1 and V_2 are independent and identically distributed after marginalizing over the random locations of the two positive individuals. Then let P(v) be the probability of infection given a dose v, as given in Section 3.3. Since P is concave, and also using that P(0)=0, then for strictly positive V_1 and V_2 are positive, then $P(V_1 + V_2) - P(V_1) <= P(V_2) - P(0)$, which implies that $P(V_1 + V_2) <= P(V_1) + P(V_2)$. So then because V_1 and V_2 are identically distributed, $E[p(V_1 + V_2) - p(V_1)] <= E[p(V_1) + p(V_2) - p(V_1)] = E[p(V_2)] = E[p(V_2) - p(0)]$. The left-hand side $E[p(V_1 + V_2) - p(V_1)]$ is the increase in the risk of infection created by adding a second positive, and the right-hand side $E[p(V_2) - p(0)]$ is the increase in risk from adding the first positive.

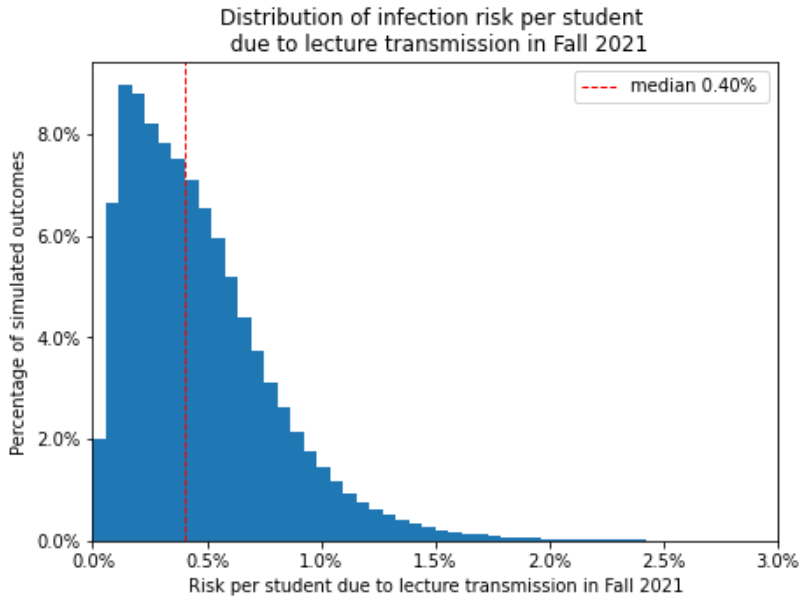

Figure 2: Distribution of simulated per-student infection risk due to classroom transmission in Fall 2021.

2.2 Infection risk for instructors and TAs

Next, we estimate the risk of classroom transmission for a faculty instructor or graduate student TA. We assume the instructor / TA is masked and sufficiently distanced (>6 feet) from the students that the instructor / TA is not subject to risk of droplet transmission. Lectures often entail some one-on-one discussion between the instructor / TA and individual students at the beginning and end of the lecture. We assume that social distancing is maintained during these interactions and that they do not add significantly to the total interaction time between instructors / TAs and students.

Our simulation assumes a 90% vaccination rate among students and produces probability distributions over the risk of infection for a masked instructor / TA per classroom hour from one infectious student. We consider the cases where the instructor / TA is unvaccinated or vaccinated separately (as the vaccine efficacy against infection is a parameter subject to uncertainty).

Next, using the same approach as in Section 2.1, we extrapolate this per-hour, conditional risk to estimate the risk for instructors / TAs over the entire semester. We first outline the steps of the calculation for faculty instructors:

- As in the previous section, we assume an average undergraduate student spends 22.5 hours per week, or $T = 315$ hours per semester, in the classroom. We further assume

that ⅔ of this time (210 hours) is spent in lectures with a faculty instructor, while the rest ⅓ (105 hours) is spent in recitation sections with a graduate student TA.

- Recall that we let p denote the prevalence among undergraduate students and N the undergraduate population. With the same reasoning as above, the total number of classroom hours with a positive student spent by all undergraduate-teaching faculty members over the semester is $p * N * T * 2/3$.
- We then multiply this with the simulation output (risk per hour from one infectious student) to get the expected number of classroom infections among faculty instructors over the semester.
- Finally, to obtain per-person risk, we divide this by N_f, the number of faculty members that teach undergraduate classes.
    - The Ithaca campus has roughly 1700 faculty members in Fall 2020 [University Facts].
    - We assume half of them teach undergraduate classes, i.e., $N_f = 850$.

The risk for graduate student TAs can be calculated similarly with the following differences in parameter values:

- The number of classroom hours that graduate student TAs spend with a positive undergraduate student over the semester is half of that for faculty instructors, i.e., $p * N * T * 1/3 = 7875$.
- We estimate the population of graduate student TAs N_g as follows:
    - In fall 2020, there were 6239 graduate students [University Facts].
    - Assuming half of them work as TAs for classes or recitations, we get $N_g = 3120$.

      Note that some of these TAs may not interact face-to-face with students, e.g., if their primary responsibility is grading, and that the risk calculated is the average risk across all TAs, including these individuals.
- This estimate for graduate student TAs assumes a classroom-like environment with limited student talking. The risk may be higher for graduate students that lead recitations with significant talking among the students.

Table 3 and Figure 3 present the estimated distribution of per-instructor/TA infection risk due to classroom transmission for unvaccinated and vaccinated faculty and graduate students respectively, where model parameters follow the prior distributions specified in Section 1. The faculty and graduate students are overwhelmingly vaccinated, so the risk for vaccinated individuals may be a close approximation to the overall risk to the population.

Group	**Median (# cases)**	**5% quantile (# cases)**	**95% quantile (# cases)**
Unvaccinated faculty instructors	0.035%	0.012%	0.080%
Vaccinated faculty instructors	0.017%	0.006%	0.047%

Unvaccinated graduate student TAs	0.005%	0.002%	0.011%
Vaccinated graduate student TAs	0.002%	0.0006%	0.006%

Table 3: Median and (5%, 95%) quantiles of per-person infection risk for different instructor and TA groups due to lecture transmission over the Fall 2021 semester.

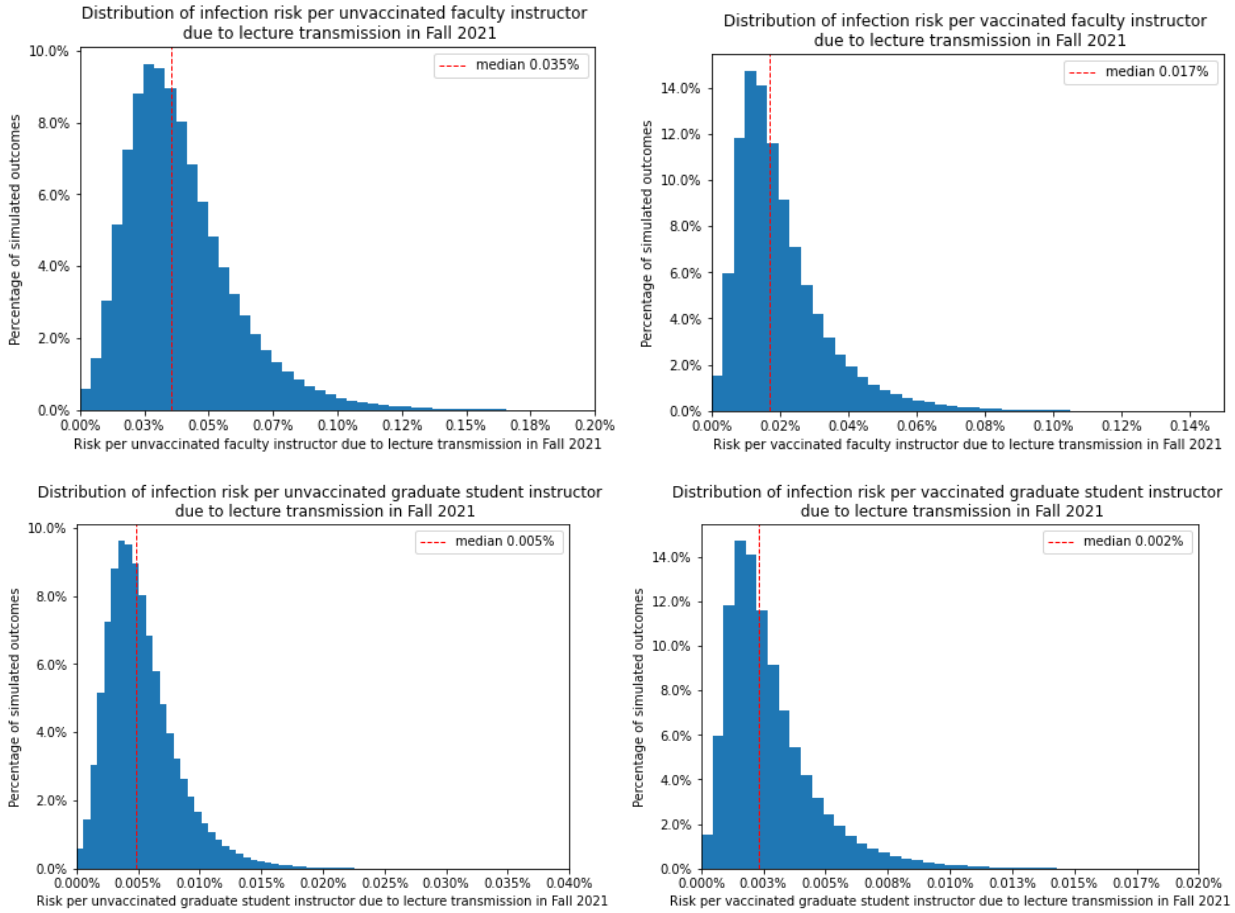

Figure 3: Distribution of simulated per-instructor/TA infection risk due to classroom transmission in Fall 2021. “Graduate student instructor” in titles actually refers to graduate student teaching assistants.

3. Detailed Methods

In this section, we describe our detailed simulation methodology for simulating the number of secondary cases in a classroom with one positive individual. We first state the assumptions and parameter estimates in different components of the simulation. We then present the simulation tool, which is implemented in Python. Finally, we introduce the mathematical model used for estimating transmission risk under different conditions.

3.1 Assumptions and Parameters

The Delta variant

- We assume the Delta variant had dominated all infections by the time the semester started.
- The Delta variant is 2.4 times more transmissible than the non-variant strain.
    - The Alpha variant was approximately 50% more transmissible than the original SARS-CoV-2, and Delta is approximately 60% more transmissible than Alpha [Callaway 2021]. This gives a multiplicative increase of 1.5 * 1.6 = 2.4.

Level of vaccination in the population

- Vaccination rate for undergraduates was approximately 94% at the time data was pulled for these simulations were performed. (It has since risen close to 100%.)
- We use 90% as a conservative estimate for the Cornell population in general.

Prevalence in the population

- In the Cornell population, we assumed the prevalence follows lognormal(6.872, 0.372). (Recall that prevalence reflects only those who are free and infectious; it excludes those in isolation.)
- This is derived by setting the mode to be 840 and the 97.5% quantile to be 2000.
    - At the time parameters were fit for this analysis (Sep. 19), we have seen 488 student cases since the start of the semester and 32 student cases in the past week. Assuming the same rate of 32 cases/week for the next 11 weeks till the end of the semester (Dec. 7), we expect to see 488 + 32 * 11 = 840 cases in total in the fall semester. Thus, we set 840 to be the mode, i.e., the point with the highest probability density. At the time of completing this written explanation of our analysis results (mid-October), the assumed mode of 840 student cases seems conservative as the number of new student cases per week has fallen below 32.

Vaccine efficacy

- We assume vaccination reduces the viral load of a source case by 0, 50%, and 71% with equal probability [Levine-Tiefenbrun et al. 2021, Harris et al. 2021, Brown et al. 2021, Pouwels et al. 2021].
- We assume vaccination reduces the risk of infection for susceptible individuals by 40%, 42%, 66%, 76%, 79%, and 88% with equal probability [Puranik et al. 2021, Sheikh et al. 2021, Bernal et al. 2021, Pouwels et al. 2021, Puranik et al. 2021, Fawlkes et al. 2021]

Masking effectiveness

- An optimistic estimate for one-way masking effectiveness is 80%.
    - Masking of infectious individuals: Kumar et al. 2020 used fluid dynamics simulation and estimated that 12% of the airflow (carrying virus particles) leaks around the side of a mask. Wang et al. 2020 observed that masking by the

primary case and family contacts before the primary case developed symptoms was 79% effective in reducing transmission.
    - Masking of susceptible individuals: Konda et al. 2020 experimentally measured the filtration efficiency of masks made from different materials and found that many materials could block particles > 0.3 μm with at least 96% filtration efficiency. (Morawska et al. 2009 observed that most human expiratory activities generate droplets/aerosols with size larger than this.) Doung-ngern et al. 2020 reported that people wearing a mask all the time during contact with a COVID-19 patient are 84% less at risk of infection. Howard et al. 2020 noted that wearing masks provides the additional protection by preventing touching the nose and mouth, which is another vector of transmission.
- A conservative estimate for one-way masking effectiveness is 50%.
    - Konda et al. 2020 observed that improper mask-wearing (e.g., having a gap between the face and the mask) can result in a large decrease in filtration efficiency. Unmasking to eat or drink would also reduce the protection. Due to these considerations, we choose a conservative estimate of 50%.
- Together, these imply that if masking is enforced, where the source and susceptible are both masked, transmission risk is reduced from the no-masking scenario by a multiplicative constant of (1-80%) * (1-80%) = 0.04 to (1-50%) * (1-50%) = 0.25.
- Effective July 30, 2021, Cornell is requiring all individuals, including fully vaccinated ones, to wear masks indoors. For two-way masking, we set the prior on the risk reduction factor to be the Normal(0.145, 0.0536), truncated below at 0.

Type of respiratory activity

- Our simulation takes into account both droplet and aerosol transmission, which have been identified as two major modes of transmission of COVID-19 [CDC, 2021]. Droplets are large in size and typically deposit within a short distance and time after being released; aerosols are fine particles that suspend in the air for a long time and may spread across the entire room.
- We assume breathing is the dominant type of respiratory activity for students attending lectures and that the effect of occasional speaking (e.g., asking and answering questions) is negligible.
- We present the details in Section 3.3.

3.2 Simulation tool

We implement a simulation tool in Python to simulate gatherings under different scenarios.

Classroom seating

We simulate seating at different density levels by assigning a fixed number of individuals to classrooms of different sizes. We assume that an average class contains N = 50 individuals. From university floor plans, we identify three representative rooms, namely Hollister 206, Gates

G01, and Rockefeller 201, that correspond to roughly 1', 3', and 6' distancing respectively for 50 students. The corresponding seating capacities are presented in Table 4.

Room	**Pre-COVID max capacity (1' distancing)**	**COVID capacity (6' distancing)**
Hollister 206	52	12
Gates G01	156	22
Rockefeller 201	383	56

Table 4: Seating capacity for Hollister 206, Gates G01, and Rockefeller 201.

Using the seating plan tool developed by David Shmoys' CTRO team, we identify seats in these classrooms and assign maximally distanced seats such that approximately 50 individuals can fit in each room. The generated seating plans are displayed in Figure 4.

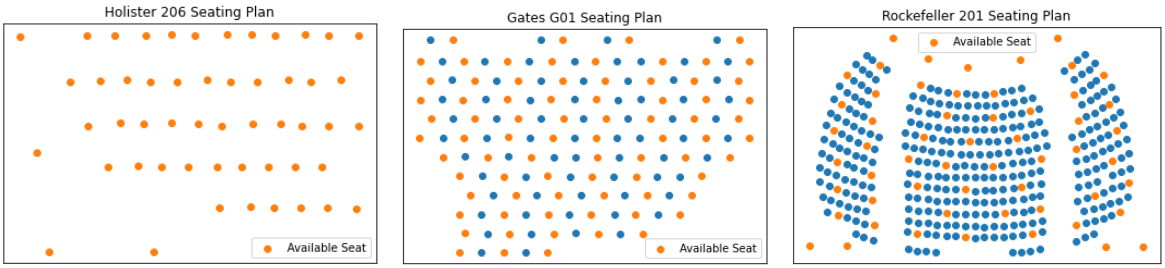

Figure 4: Seating plans generated for Hollister 206, Gates G01, and Rockefeller 201. They correspond to 1', 3', and 6' social distancing, respectively.

The orange dots represent available seats (i.e., seats that are allowed to be occupied) and the blue dots are considered unavailable to our simulation tool. The rooms used have extra space above and beyond what is required to accommodate the social distancing requirements we have assumed. This extra room does not offer additional benefit in our simulations as it is not used.

We assign students to available seats assuming that unvaccinated students tend to sit together (the same demographic factors that cause students to be unvaccinated upon arrival to Ithaca may also create social connections). This is more pessimistic than assigning students randomly to seats regardless of their vaccination statuses. Figure 5 shows examples of seating assignments at different distancing levels. The red dots represent unvaccinated students, the blue dots represent vaccinated students, and the grey dots represent unavailable seats due to the distancing requirement.

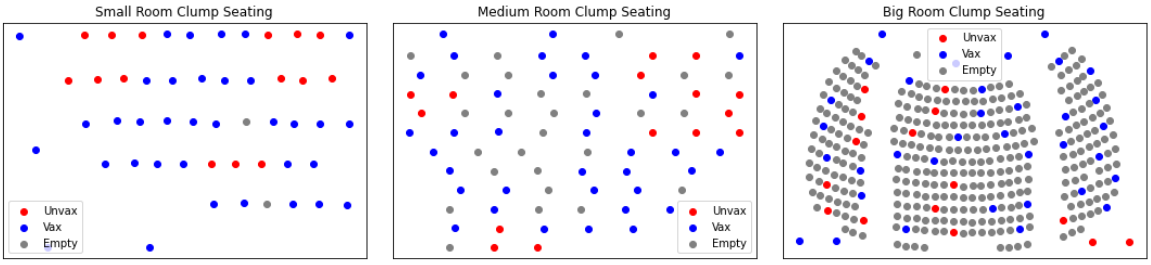

Figure 5: Example seating assignments of 50 students at different levels of social distancing. "Small room" refers to Hollister 206 with 1' distancing, "medium room" refers to Gates G01 with 3' distancing, and "big room" refers to Rockefeller 201 with 6' distancing. Under clump seating, unvaccinated students tend to sit near each other.

Simulation Model

We simulate the number of secondary infections that occur in a 1 hour class with 50 students, where initially a single student is infected. We vary the percentage of students vaccinated and test the effect of masking, decreasing room density, changing seating arrangements, and class types on the number of secondary infections.

When generating an initial source infection in our simulation, we flip a biased coin to decide whether the source is vaccinated or unvaccinated. By the assumption that vaccination reduces the risk for a vaccinated susceptible person by a factor of one minus vaccine efficacy (VE) , we set $P(inf \mid vax) = (1 - VE) * P(inf \mid unvax)$. The probability that the source case is vaccinated can then be calculated using Bayes Rule:

$$P(vax \mid inf) = \frac{P(inf \mid vax)\,P(vax)}{P(inf \mid vax)\,P(vax) + P(inf \mid unvax)\,P(unvax)} = \frac{(1-VE) * P(vax)}{1-VE * P(vax)}$$

We average over 500 trials to obtain our final results.

3.3 Modeling the risk of droplet and aerosol transmission

Droplet and aerosol transmission are modeled separately and we describe them below.

Assumptions and parameters

Both the droplet and the aerosol model are based on the *exponential dose-response model* (see Equation 4 in Buonanno et al. 2020 and Watanabe et al. 2010). A dose-response model calculates the transmission probability as a function of *dose*, the amount of virus particles a susceptible person is exposed to. According to the exponential dose-response model, the transmission probability given dose D takes the following form,

$$P(transmission) = 1 - exp(-c \cdot D),$$

where c is a positive constant. Observe that P(transmission) is concave in the dose D, a fact used in Section 2.1 to observe that the increase in risk created by adding a second positive in the classroom to be smaller than the increase in risk created by the first positive.

Droplet transmission

In an earlier analysis, we developed a model that estimates the risk of droplet transmission (over short distances) as a function of the relative location of the infectious and susceptible persons and their duration of interaction. In particular, the probability of infection for a susceptible person at distance and angle (r, θ) away from a source case for a duration of interaction T is given by

$$P_{droplet}(transmission, r, \theta, T; \alpha) = 1 - exp(- c_3 \cdot 1\{\theta \; in \; cone; \alpha\} \cdot \frac{\phi(r)}{r} \cdot T).$$

where, based on Sun and Zhai 2020,

$$\phi(r) = - 0.1819 * ln(r) + 0.43276,$$

and $1\{\theta \; in \; cone; \alpha\}$ is an indicator function for whether the susceptible person is seated within a degree-α "cone" $[-\alpha, 180 + \alpha]$ in which the source case is emitting droplets (we assume that students are in their seats, and thus emit droplets in a cone whose center faces forward; see Figure 1 of the document).

Using maximum likelihood estimation on a dataset of transmission on high-speed trains [Hu et al. 2021], we find $(c_3^*, \alpha^*) = (0.0135, 15)$.

Note that these parameters apply to the original SARS-CoV-2 strains that were predominant early in the pandemic. As stated above, we multiply the transmission risk by 2.4 to account for the Delta variant at the end of our analysis and as described at the bottom of this document.

Aerosol transmission

We complement our droplet analysis with the results from Schijven et al. 2021, who developed a model for predicting the transmission risk in an enclosed space due to aerosols only, under the assumption that emitted aerosols are dispersed across the entire room. The estimated risk depends on the aerosol-emitting activity (such as breathing, speaking, and singing), the level of ventilation of the room, the duration of interaction, and the virus concentration of the source case (measured in the number of virus particles per unit volume). One key property is that the risk due to aerosol transmission is uniform across all locations of the room.

The risk of aerosol transmission over time T is estimated by the exponential dose-response model:

$$P_{aerosol}(transmission,\ T)\ =\ 1\ -\ exp(-\ D(T)/1440),$$

where the D denotes the "dose", i.e., the amount of virus particles that a susceptible person receives from the infectious person over time. This depends on the number of virus particles emitted over time T, which we denote $N(T)$, in the following way

$$D(T)\ =\ N(T)\ *\ \frac{inhalation\ rate\ of\ susceptible\ (volume\ /\ time)}{volume\ of\ the\ room}$$

where $N(T)$ depends on the type of aerosol-emitting activity, the viral load of the infectious person, the ventilation condition of the room, and the duration of interaction T. It is estimated that breathing emits aerosols containing 3300 virus RNA copies per hour (assuming a nominal viral load 10^8 cp/ml). We refer the readers to Equations (4) - (14) in Schijven et al. 2021 for details of the calculation.

We implement a few additional calculations when deploying this model for the classroom simulation:

1. We average the risk over the distribution of the source viral load, which Schijven et al. 2021 estimated to be log-normal.
2. We take into account the effect of vaccination and masking, as described in the previous assumptions.
3. We implement different ventilation conditions, namely no ventilation (a conservative estimate of the amount of ventilation in naturally-ventilated spaces), 1 air exchange per hour, and 3 air exchanges per hour. The amount of aerosols present in the room per hour is reduced to a factor of ½ and ¼ under the latter two conditions.

Combining the risk from droplets and aerosols

In our simulation, the risk for a susceptible individual is

$$max(P_{droplet}(transmission),\ P_{aerosol}(transmission)),$$

i.e., the larger of the predicted risk due to droplet transmission and the risk due to aerosol transmission. This is justified by the fact that we inferred the parameters for the droplet model assuming all infections in the dataset result from droplet transmission; as such, the inferred model implicitly accounts for aerosol transmission in the dataset. (In practice, the risk due to droplets dominates at short distances and approaches 0 at large distances; the risk due to aerosol is uniform across all locations.)

Finally, we multiply the calculated risk by 2.4 to account for the increased transmissibility of the Delta variant.