Infectious Disease Modelling 6 (2021) 1025–1045

Contents lists available at ScienceDirect

Infectious Disease Modelling

journal homepage: www.keaipublishing.com/idm

Mathematical modelling of the spread of COVID-19 on a university campus

Kaitlyn Muller*, Peter A. Muller

800 E. Lancaster Avenue, Department of Mathematics and Statistics, Villanova University, Villanova, PA, USA

ARTICLE INFO

Article history:
Received 20 April 2021
Received in revised form 29 July 2021
Accepted 8 August 2021
Available online 14 August 2021
Handling editor: Dr. Lou Yijun

Keywords:
Coronavirus
COVID-19
University/college campus
Epidemiological models
SEIR
Contact tracing
Surveillance testing 2010 MSC: 92-10

ABSTRACT

In this paper we present a deterministic transmission dynamic compartmental model for the spread of the novel coronavirus on a college campus for the purpose of analyzing strategies to mitigate an outbreak. The goal of this project is to determine and compare the utility of certain containment strategies including gateway testing, surveillance testing, and contact tracing as well as individual level control measures such as mask wearing and social distancing. We modify a standard SEIR-type model to reflect what is currently known about COVID-19. We also modify the model to reflect the population present on a college campus, separating it into students and faculty. This is done in order to capture the expected different contact rates between groups as well as the expected difference in outcomes based on age known for COVID-19. We aim to provide insight into which strategies are most effective, rather than predict exact numbers of infections. We analyze effectiveness by looking at relative changes in the total number of cases as well as the effect a measure has on the estimated basic reproductive number. We find that the total number of infections is most sensitive to parameters relating to student behaviors. We also find that contact tracing can be an effective control strategy when surveillance testing is unavailable. Lastly, we validate the model using data from Villanova University's online COVID-19 Dashboard from Fall 2020 and find good agreement between model and data when superspreader events are incorporated in the model as shocks to the number of infected individuals approximately two weeks after each superspreader event.

1. Introduction

The majority of us have felt the impact of COVID-19 on our daily lives and many understand how important it is for public health and our own individual safety to take part in social distancing, mask wearing, and other strategies for mitigating the spread of the virus. College campuses, in particular, provide an ideal breeding ground for coronavirus as we have group living situations and lots of population mixing within class settings and other social events. In addition, the younger population present on campuses tend to have more social contacts as well as an increased likelihood of asymptomatic infection (Cashore et al., 2020b; Poletti et al., 2020). This all leads to an increased chance of community spread. Campuses across the country

* Corresponding author.
E-mail addresses: k.muller@villanova.edu (K. Muller), peter.muller@villanova.edu (P.A. Muller).
Peer review under responsibility of KeAi Communications Co., Ltd.

https://doi.org/10.1016/j.idm.2021.08.004

were forced to develop public health strategies in order to attempt to host students on campus in the past year. There has been a lot of work on modelling the spread of COVID-19 in a campus setting (Paltiel et al., 2020; Losina et al., 2020; Cashore et al., 2020b; Lopman et al., 2020; Gressman & Peck, 2020; Bahl et al., 2020) aimed at providing insights to university administrators on how best to prepare for on-campus semesters. These results, in particular, highlight the need for expansive (and expensive) surveillance testing programs and other strategies.

In this work we develop and investigate a model of the spread of coronavirus on a college campus. It is our goal to highlight the relative importance of individual- and institution-level mitigation measures. We focus on a medium-sized university that hosts primarily undergraduate students on campus. In addition, we assume that students have returned to campus and take a mix of in-person, online, and hybrid courses as has been the norm across many campuses since Fall 2020 (The Chronicle of Higher Education and Davidson College's College Crisis Initiative, 2020). We assume students, faculty, and staff are expected to wear masks, use hand sanitizer, and take part in other well-known public health measures while on campus. In-person classes will be assumed to be socially distant.

We use an epidemiological SEIR-type model to describe the spread of the disease with the addition of different population types for students and faculty. In addition we separate asymptomatic/pre-symptomatic and symptomatic individuals and include classes for quarantined and isolated individuals. This model is not intended for predictive purposes. Instead, we aim to give a quantitative measure of the potential effectiveness of individual measures such as mask wearing and social distancing (measured by individual model parameters) as well as institution-level measures: gateway testing, surveillance testing, and contact tracing. We find that the scale of on-campus outbreaks is largely determined by student behaviors. Our results indicate that a gateway test in conjunction with effective contact tracing or robust surveillance testing, are necessary in order to keep the number of infections at a more reasonable level and make it more feasible to continue to house and teach students on campus. We find that contact tracing, on its own, is a rather robust and powerful mitigation technique when done efficiently. In addition, we find that surveillance testing on its own must be done often (at least weekly) and to the entire population in order to be effective as a sole mitigation measure as was found in similar studies (Paltiel et al., 2020).

We also investigate the effects of surveillance testing done on a smaller scale. Surveillance testing can be a difficult task for universities without access to medical schools or other lab facilities and is extremely expensive. We aim to determine if a small scale surveillance testing program is useful in maintaining effective control of the spread of the virus. We find that small scale surveillance strategies are not sufficient on their own but can reduce infection rates when used in conjunction with contact tracing. However, it is imperative to have a robust contact tracing program when a robust surveillance strategy is not a possibility. Finally, we validate this model by fitting to estimated active cases per day data from the Villanova University COVID-19 Dashboard (Villanova University, 2020). We find good agreement between our predicted active cases and the data once we modify our model to include the possibility of superspreader events modelled as shocks to the system approximately two weeks post event.

Similar studies have been conducted to determine which strategies would be effective for re-opening of college campuses. Much of this work was completed prior to the Fall 2020 semester. In (Paltiel et al., 2020) the focus was on determining an effective surveillance testing strategy and also the cost effectiveness of each strategy. Our work differs in that we use a continuous model as opposed to their discrete dynamical system model as well as our inclusion of separate classes for students and faculty and the inclusion of contact tracing as a distinct strategy. Our results from considering a surveillance testing only scenario are similar. In addition (Paltiel et al., 2020), incorporates random shocks or influxes of new positive cases which they use to model infections due to contact with the broader community and/or superspreader events. We neglect the effect of the broader community but we also use shocks at certain points in time to study the effect of superspreader events like parties. For example, the data from Villanova University (Villanova University, 2020) shows a sizable increase in cases approximately two weeks after Halloween. Another, more recent, study (Losina et al., 2020) has similar to results to (Paltiel et al., 2020).

The work by (Cashore et al., 2020b,a), has been used to make decisions for re-opening strategies at Cornell University. This work differs from our own as it considers a stochastic compartmental model. Other key differences include their assumptions that surveillance testing of the entire population is achievable every 5 or 7 days. In addition, they do not make a distinction between the contact patterns of students and faculty. Finally, they assume contact-traced individuals are only identified in the exposed class. This assumption assumes a very robust contact tracing system that identifies positive cases and their contacts in such a timely manner that the contacts have not yet become infectious, i.e. within 2–3 days of exposure. We allow for the possibility that due to delays in reporting of symptoms as well as our assumed surveillance testing being possibly less often that there may be infected individuals caught by contact tracing in the presymptomatic/asymptomatic class as well. The most interesting result of (Cashore et al., 2020b) is the conclusion that re-opening results in fewer infections than the campus being entirely virtual. We do not make any comparisons to an entirely virtual semester in this paper.

The most similar study to our own is (Lopman et al., 2020), which also uses a continuous-time dynamical system for a model with similar groups delineated within the campus population. They go further to distinguish between on-campus and off-campus students which is likely necessary since they are focused on campuses of much larger size. The most significant distinction between our work and theirs is that we allow for surveillance testing (screening in (Lopman et al., 2020)) to potentially test less than the entire population. They investigate testing the entire population in intervals varying from weekly to once a semester while we allow for a variable percentage of the population to be tested from daily to biweekly frequencies. We hope to shed some light on the effectiveness of surveillance testing in this perhaps more financially feasible fashion. We also investigate the effects of student behaviors such as adherence to mask wearing and social distancing policies by

modelling the contact rate to depend on number of contacts daily (assumed to be related to social distancing) and probability of transmission (assumed to be related to mask wearing). We investigate scenarios with different levels of adherence to each behavior to determine their effect on total infections over the semester.

We also note there have been agent-based models developed in (Bahl et al., 2020; Gressman & Peck, 2020). These differ significantly from our approach as they focus on individuals and their behaviors within the population instead of considering collective behavior as we do. With their focus on individuals, (Gressman & Peck, 2020) is able to examine the effect of class sizes more easily and concludes that large class sizes would be a main driver of outbreaks on campuses. We note that this brief introduction does not cover the quickly changing landscape of COVID-19 literature, but covers the most relevant literature to modelling the spread on college campuses currently available.

The paper will be organized as follows. In Section 2 we will discuss our methods which will include a description of our model and the assumptions made. We will also describe our method of solution and the various simulations performed. In Section 3 we will discuss the results of our simulations. In particular, we will first discuss a series of hypothetical situations, §3.1, highlighting the difference between three levels of adherence to social distancing and mask wearing policies. In addition we will investigate hypothetical scenarios as we vary levels of surveillance testing and contact tracing. We will also perform sensitivity analysis, §3.2, to determine which parameters the model is most sensitive to concerning the total number of infections over the course of a semester as well as the estimated basic reproduction number. Finally, in §3.3, we will provide a model fit to data obtained from Villanova University's COVID dashboard data from Fall 2020 (Villanova University, 2020).

2. Methods

2.1. Assumptions

As noted before, we develop a SEIR-type epidemiological model modified to incorporate our knowledge of the characteristics of COVID-19 as well as those of a college campus community. We note that by using an SEIR-type model we are implicitly assuming a homogeneously mixed population. For example, it is thought that a population like that of a cruise ship or a heavily populated area like New York City may be well modelled in this way (Bansal et al., 2007). While this is not entirely realistic as people always act as individuals, we believe campus populations are significantly more mixed than a standard population of say an average town or small city. We do allow for some heterogeneity, however, by separating each class into faculty and students as we expect students to have much higher contact rates with each other than faculty. We note that there

Table 1
Definitions of variables and parameters. $i = 1$ represents quantities related to the student population and $i = 2$ the faculty.

Variable/parameter	Description
t	time
$S_i(t)$	susceptible population at time t
$E_i(t)$	exposed population at time t
$A_i(t)$	infectious asymptomatic/presymptomatic population at time t
$I_i(t)$	infectious symptomatic population at time t
$Q_i(t)$	quarantined population at time t
$L_i(t)$	isolated population at time t
$R_i(t)$	recovered population at time t
$D_i(t)$	deceased population at time t
c_i	contacts per case per day
p_c	probability of transmission per contact
f_{ji}	proportion of contacts of infectious type j with type i
τ	relative infectiousness of asymptomatic individuals (in comparison to symptomatic)
N_i	total population on campus
ρ	average time from surveillance test to quarantine/isolation
P_T	proportion of population being surveillance tested
S_p	specificity of test
S_e	sensitivity of test
ψ	average time to quarantine/isolate contacts from contact tracing
T	proportion of contacts caught by tracing
p_e, p_a	probability that case caught is in the exposed, asymptomatic classes
μ	average time in quarantine
σ	average time to become asymptomatic/presymptomatic (i.e. infectious) after exposure
φ	average time to transition from presymptomatic to symptomatic
η_0	average time for asymptomatic to recover
η_i	average time for symptomatic population to recover
ξ	average time to self-isolate after becoming symptomatic
r	proportion of symptomatic cases that self-isolate
ζ_i	average time to death after becoming symptomatic
d_i	probability of death for symptomatic population
P_i	proportion of population infections that are symptomatic
d	number of days that are contact traced

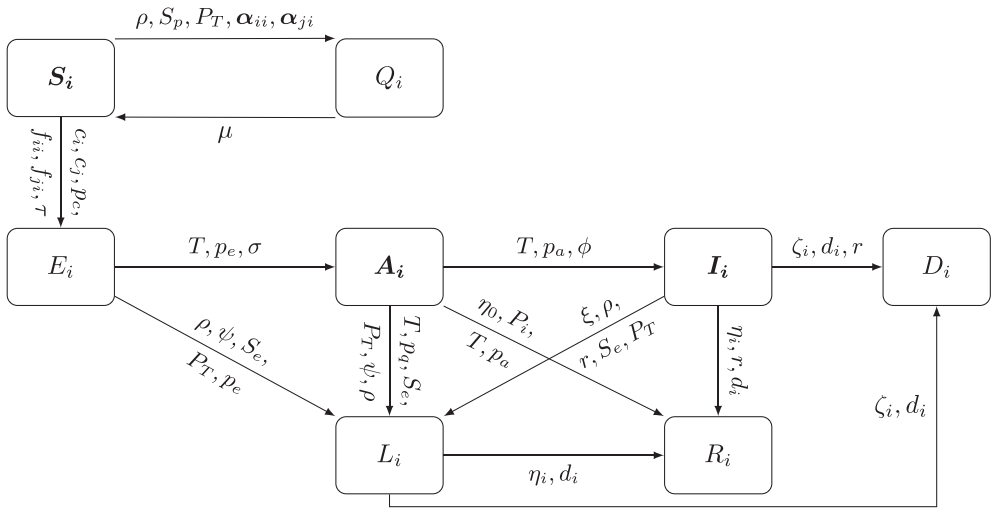

Fig. 1. Flow diagram of a college campus COVID-19 outbreak model. Bold compartments indicate those with possible transmission interactions between students and faculty.

has been some work on more heterogeneously mixed models (Bahl et al., 2020; Gressman & Peck, 2020) that have different takeaways than our work. No results from this study should be used to make predictions about number of infections, deaths, etc.

To build our model, we make the following assumptions.

Table 2
List of chosen model parameter values and ranges for parameters found in literature.

Parameter	Range of Values	DefaultValue	Units	References
c_1	[7, 17]	12	days^{-1}	Cashore et al. (2020a)
c_2	[1, 4]	2	days^{-1}	Cashore et al. (2020a) & Assumption
p_c	[0.6, 3.7]%	2.15%	dimensionless	Luo et al. (2020); Bi et al. (2020); Jing et al. (2020); Liang et al. (2020); Davies et al. (2021)
f_{ij}	[0, 1]	$f_{11} = 0.7$ $f_{22} = 0.3$	dimensionless	Assumption
τ	[0.5, 1]	0.75	dimensionless	Centers for Disease Control and Prevention (2020); He et al. (2020); Luo et al. (2020)
N_1	6000	6000	number of people	Assumption
N_2	1000	1000	number of people	Assumption
ρ	7–14	7	days	Assumption
P_T	[0, 1]	0.25	dimensionless	Assumption
S_p	[0.918, 1]	0.959	dimensionless	Bisoffi et al. (2020)
S_e	[0.624, 0.9412]	0.7826	dimensionless	Bisoffi et al. (2020)
ψ	1	1	days	Assumption
T	[0.32, 0.79]	0.555	dimensionless	Spencer et al. (2021)
p_e, p_a	[0, 1]	0.15	dimensionless	Assumption
μ	14	14	days	Assumption
σ	[2, 4]	3	days	Lauer et al. (2020); Centers for Disease Control and Prevention (2020); Tindale et al. (2020)
φ	[2, 3]	2.5	days	Tindale et al. (2020)
η_0	14	14	days	Voinsky et al. (2020)
η_1	14	14	days	Voinsky et al. (2020)
η_2	14	14	days	Voinsky et al. (2020)
ξ	[1,6]	3.5	days	Centers for Disease Control and Prevention (2020)
r	[0.15, 0.65]	0.4	dimensionless	Reese et al. (2020)
ζ_1	16	16	days	Centers for Disease Control and Prevention (2020)
ζ_2	16	16	days	Centers for Disease Control and Prevention (2020)
d_1	0.000359	0.000359	dimensionless	Centers for Disease Control and Prevention (2021a) & Assumption
d_2	0.0105	0.0105	dimensionless	Centers for Disease Control and Prevention (2021a); McChesney & Bichsel (2020); Penn State University Budget Office (2017) & Assumption
P_1	[0.2, 0.6]	0.4	dimensionless	Byambasuren et al. (2020); Poletti et al. (2020); Centers for Disease Control and Prevention (2020)
P_2	[0.3, 0.8]	0.55	dimensionless	Byambasuren et al. (2020); Mizumoto et al. (2020); Poletti et al. (2020)
d	2	2	days	Assumption

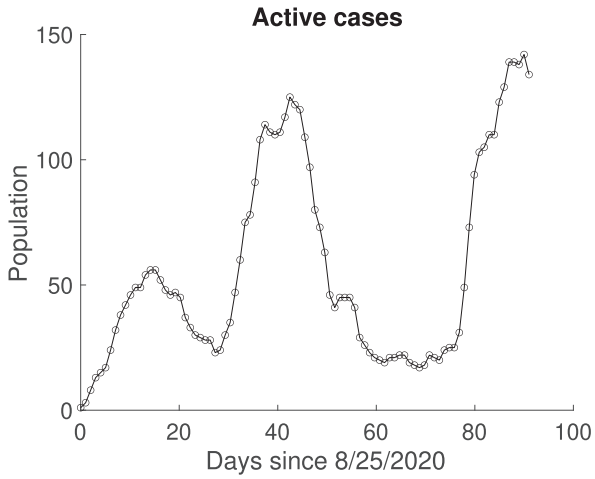

Fig. 2. Estimated Active Cases Per Day, Villanova Fall 2020, using new cases per day from (Villanova University, 2020) and assuming they remain isolated for 14 days.

Table 3
List of initial condition parameters and ranges.

Parameter	Range of Values	Units	References
Initial percent infected	[0.25, 4]	dimensionless	(Cashore et al., 2020b) & e-mail from Villanova's President
Initial Proportion of Self-isolation of Symptomatic Students	[0.15, 0.65]	dimensionless	(Reese et al., 2020Reese et al., 2020)
Initial Proportion of Self-isolation of Symptomatic Faculty	[0.5, 0.75]	dimensionless	Assumption
Gateway Test Sensitivity	[0.624, 0.9412]	dimensionless	(Bisoffi et al., 2020 et al., 2020)
Gateway Test Specificity	[0.9502, 100]	dimensionless	(Bisoffi et al., 2020 et al., 2020)

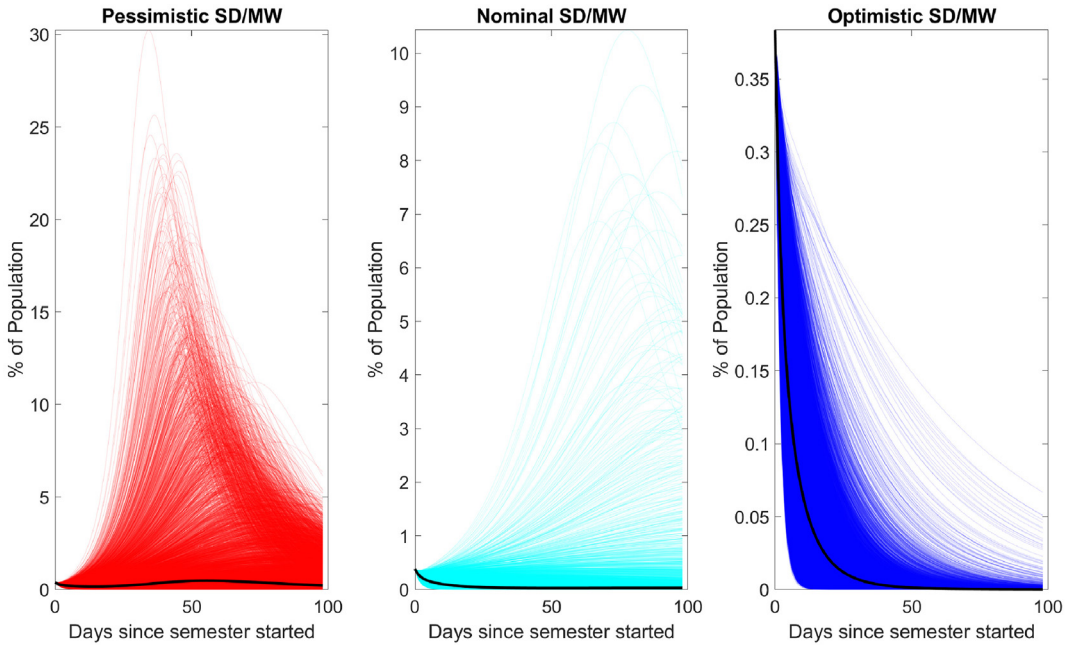

Fig. 3. Student Percent Infected over time for various levels of social distancing and mask wearing adherence. Left: Pessimistic adherence. Middle: Nominal adherence. Right: Optimistic adherence. Bold, black curves represent the mean of all curves in each subfigure. Each family of curves represents the variability due to unknown/random parameters in the model.

Table 4
Mean Percent of Students and Faculty Infected over the course of the semester assuming **pessimistic social distancing** and pessimistic (left), nominal (middle), and optimistic (right) mask wearing adherence.

	Pessimistic MW		Nominal MW		Optimistic MW	
Scenario	Students	Faculty	Students	Faculty	Students	Faculty
1	96.27%	79.45%	80.90%	56.55%	26.62%	15.02%
2	21.03%	12.01%	10.50%	6.12%	4.17%	3.00%
3	13.51%	8.24%	8.89%	3.89%	2.62%	2.32%
4	87.65%	64.88%	49.85%	30.51%	6.33%	4.03%
5	9.14%	5.36%	4.05%	2.93%	2.44%	2.24%

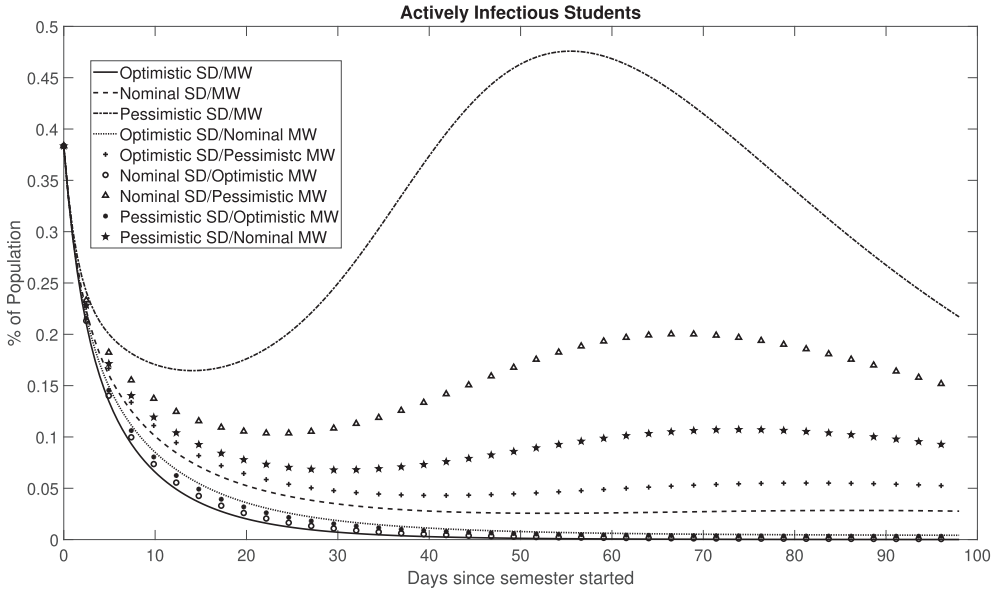

Fig. 4. Mean student percent actively infected for various levels of social distancing and mask wearing adherence.

As1. We assume the campus population may be divided into two groups with different collective behavior: students and faculty. This assumption is based on the fact that we expect students to have significantly more contacts with each other and perhaps others on campus, which increases their chance of catching and transmitting disease. In addition, the average age of undergraduates is much lower than of the faculty and thus students will be more likely to be asymptomatic and have a much smaller case fatality rate. We neglect the effect of other groups on campus such as staff and graduate students. There is a greater variability in the amount of contact these groups have with other people on campus which makes modelling more challenging in their case. In addition faculty will likely be the largest group of employees who are interacting with students often and hence are the secondary group we take into account along with the students.
As2. We assume the total population is constant (including deaths) and that there are no student transfers nor faculty hiring during the semester as well as no deaths due to circumstances other than COVID-19. It is therefore assumed that the campus is a closed community during the semester, i.e. no one interacts with any person not a part of the campus community. This is obviously not the case but greatly simplifies the model. We therefore neglect the effect of student and faculty interactions off campus or visitors to the campus. We assume that there are no breaks during the semester and that students are discouraged from traveling during the semester.

Table 5
Mean Percent of Students and Faculty Infected over the course of the semester assuming **nominal social distancing** and pessimistic (left), nominal (middle), and optimistic (right) mask wearing adherence.

	Pessimistic MW		Nominal MW		Optimistic MW	
Scenario	Students	Faculty	Students	Faculty	Students	Faculty
1	90.32%	68.25%	63.05%	39.75%	14.48%	8.04%
2	14.21%	8.15%	7.04%	4.35%	3.49%	2.70%
3	8.15%	5.19%	3.72%	2.84%	2.42%	2.24%
4	70.04%	46.31%	25.07%	14.11%	3.56%	2.73%
5	5.41%	3.59%	2.96%	2.46%	2.34%	2.20%

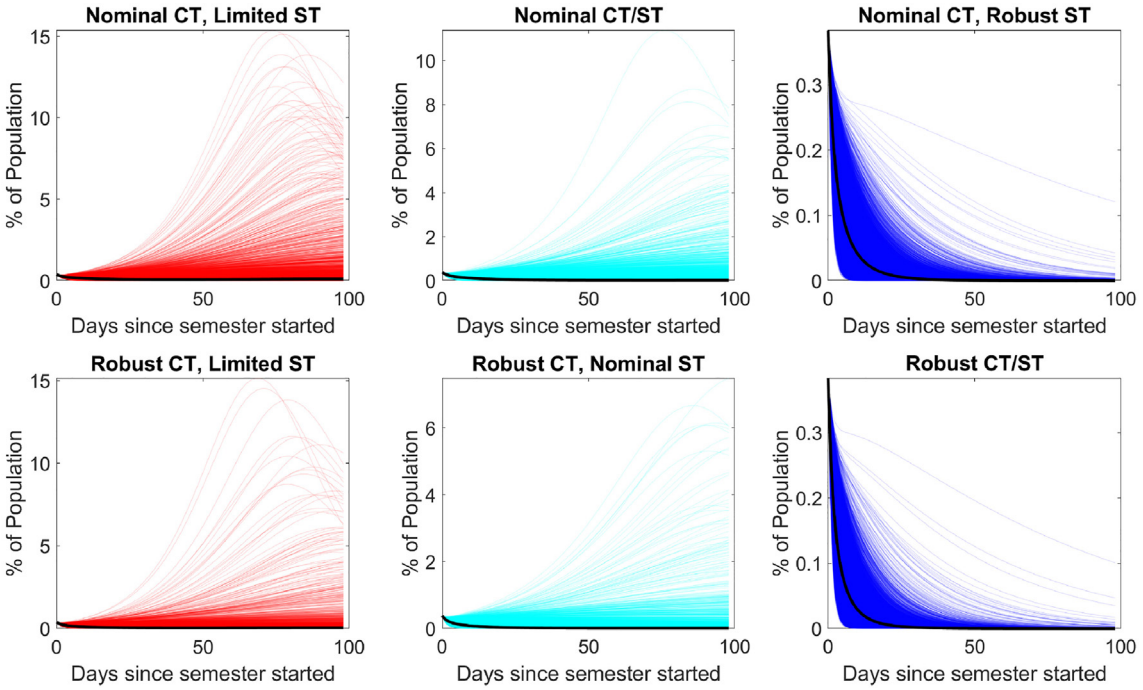

Fig. 5. Student Percent Infected over time for various levels of contact tracing and surveillance testing. Top row: Nominal contact tracing. Bottom row: Robust contact tracing. Left: Limited surveillance testing. Middle: Nominal surveillance testing. Right: Robust surveillance testing. Bold black curves are the mean of all curves in each subfigure. Each family of curves represents the variability due to unknown/random parameters in the model.

As3. We break down the overall student and faculty populations into the following classes: susceptible (S_i), exposed (not infectious, E_i), asymptomatic/presymptomatic (infectious, A_i), symptomatic (infectious, I_i), quarantined (Q_i), isolated (L_i), recovered (R_i), and deceased (D_i). The quarantined class includes anyone identified by a false positive test or those identified via contact tracing that do not in fact have the disease. This group will return to the susceptible class after completing an assumed required 14 day quarantine.

As4. The driver of new infections is assumed to be contact between susceptible and infectious individuals. In addition, we assume different contact rates for students and faculty. Contact rate is broken down into three parameters in our model. The first factor is the average number of contacts per person per day (c_i) multiplied by the probability of transmission given contact (p_c). We then include a factor f_{ji} which indicates which proportion of those contacts per day are made between an infected individual of type j with an individual of type i. For example f_{12} is the proportion of the contacts an infected student makes per day with faculty and f_{11} is the proportion of contacts infected students make with other students, therefore $f_{12} + f_{11} = 1$. Similarly, $f_{21} + f_{22} = 1$. This enables us to distinguish between the different expected contact patterns of students versus faculty.

As5. We assume that contact rates with asymptomatic/presymptomatic and symptomatic individuals may be different. There is still emerging research in this area but we allow the relative infectiousness for the asymptomatic class to be less than or equal to that of the symptomatic class.

As6. We assume recovered individuals obtain at least a temporary immunity that will last the length of the semester and hence recovered individuals do not become susceptible again. Studies are still ongoing regarding the duration, if any, of temporary immunity (Wu et al., 2021).

Table 6
Mean Percent of Students and Faculty Infected over the course of the semester assuming **optimistic social distancing** and pessimistic (left), nominal (middle), and optimistic (right) mask wearing adherence.

	Pessimistic MW		Nominal MW		Optimistic MW	
Scenario	Students	Faculty	Students	Faculty	Students	Faculty
1	72.91%	48.32%	35.27%	19.85%	6.80%	4.19%
2	8.57%	5.15%	4.72%	3.25%	3.00%	2.49%
3	4.50%	3.24%	2.76%	2.39%	2.30	2.19%
4	35.41%	20.44%	8.05%	4.83%	2.65%	2.34%
5	3.35%	2.64%	2.53%	2.28%	2.26%	2.17%

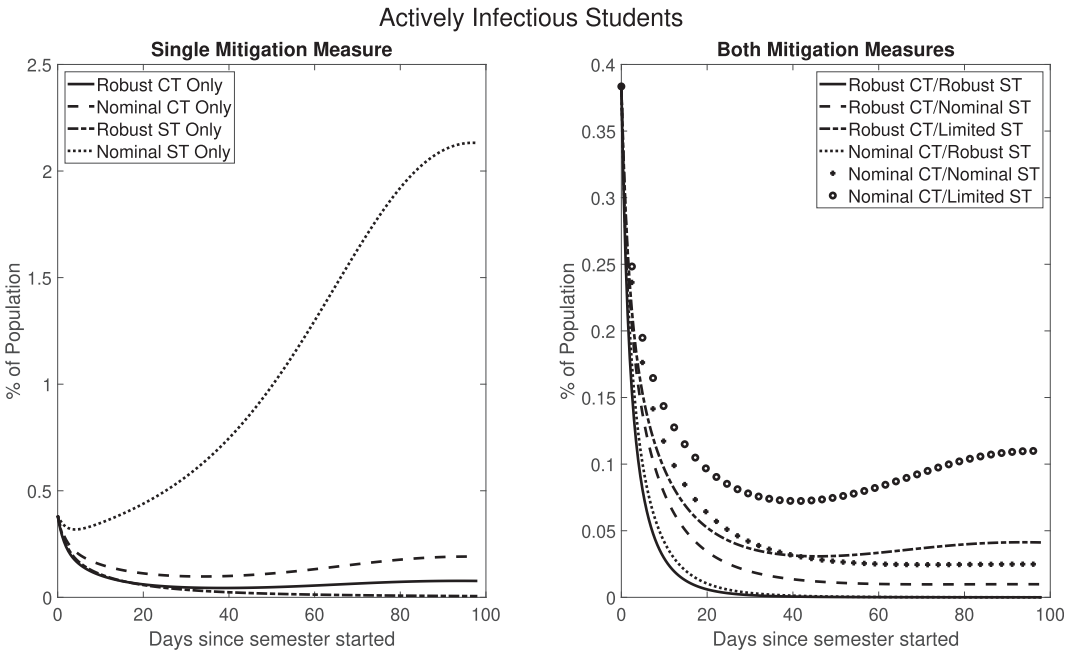

Fig. 6. Student Percent Infected over time for various levels of contact tracing and surveillance testings, each used on its own. Here again we only show mean curves.

As7. We assume the effect of social distancing, mask usage, hand washing, etc. is modelled by a change in the contact rate between susceptible and infected people. We anticipate these measures will reduce both the number of contacts per day and the probability of transmission. We vary the effectiveness randomly as it is not known precisely how much each of these measures affects contact and transmission.

As8. We assume average recovery times for symptomatic and asymptomatic students and faculty to be the same, but given the age disparity between the two groups, we use different death rates for each.

As9. In the case where no contact tracing or surveillance testing is used, we eliminate the possibility of an asymptomatic person being isolated. In the case with contact tracing and/or surveillance testing we assume that both exposed and asymptomatic/presymptomatic individuals may be isolated as a result. We then assign a certain proportion of those cases that are caught by tracing (T) and an average number of days (ψ) it takes for the tracing process to isolate an individual.

As10. In the model with a gateway test and/or surveillance testing, we assume that this test will not catch every case. We prescribe a set sensitivity and specificity of the test used from within a range found in (Bisoffi et al., 2020).

As11. It is assumed that some proportion of symptomatic people are placed into isolation in a set number of days after presenting symptoms. We assume there may be some instances of people whose symptoms are so mild that they do not self report and/or those who hide symptoms or do not abide by isolation rules. Therefore it is possible for symptomatic people to interact with susceptible populations.

As12. The model does not account for long-term illness or hospitalization. These are important considerations and possible, but do not impact the analysis of the effectiveness of individual and institutional mitigation measures, which aim to reduce the number of infections rather than the severity of the disease, and therefore are not considered here.

Table 7
Mean Percent of Students and Faculty Infected over the course of the semester comparing nominal and robust contact tracing.

Scenario	% Students	% Students	% Faculty	% Faculty
	Nominal CT	Robust CT	Nominal CT	Robust CT
1	66.09%	66.09%	40.93%	40.93%
2	8.80%	5.94%	5.18%	3.82%
3	4.27%	3.24%	3.08%	2.61%
4 with Limited ST	37.49%	37.49%	20.93%	20.93%
5 with Limited ST	3.74%	2.99%	2.82%	2.48%
4 with Nominal ST	17.27%	17.27%	9.40%	9.40%
5 with Nominal ST	2.99%	2.66%	2.47%	2.33%
4 with Robust ST	2.80%	2.80%	2.40%	2.40%
5 with Robust ST	2.39%	2.34%	2.22%	2.20%

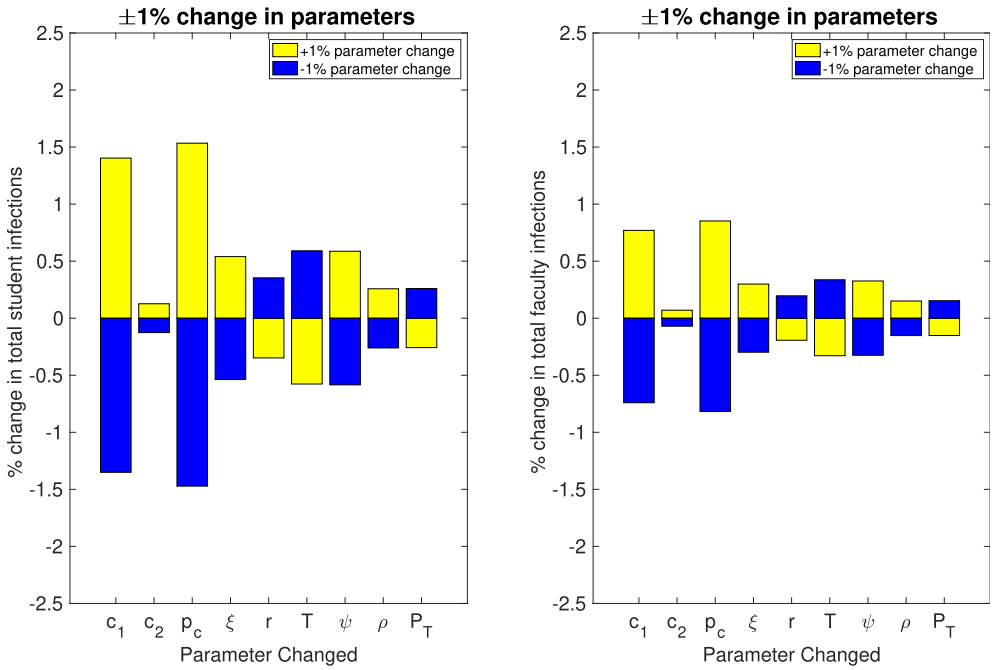

Fig. 7. Effect of a ± 1% perturbation in parameter value on the total number of students (left) and faculty (right) infected in a semester.

As13. This model does not address the effects of underlying medical conditions or the disparity between different racial/ethnic groups. We recognize that some groups are more likely to be impacted by a COVID-19 outbreak, even if the model does not explicitly show this.

2.2. Model explanation

Our model, given by Equation (1), is a SEIR-type system modified to include the two population types of students and faculty. There are eight compartments into which each population type could pass through: susceptible (S_i), exposed (E_i), asymptomatic/presymptomatic (A_i), symptomatic (I_i), quarantined (Q_i), isolated (L_i), recovered (R_i), and deceased (D_i). In Equation (1), an index of 1 represents variables and parameters related to the student population and an index of 2 represents faculty. All variable and parameter definitions are listed in Table 1.

Table 8
Estimated Parameter Values and Error for different fits.

	τ, S_e, S_p fixed to default value		τ, S_e, S_p fit	
Parameter	Jumpspost 14 days	Halloween jumppost 10 days	Jumpspost 14 days	Halloween jumppost 10 days
p_c	0.0095	0.0061	0.0219	0.0276
c_1	10.9953	9.3031	8.2020	10.9833
c_2	2	1.7071	1.9997	1.9992
ξ	5.5393	3.0453	3.9962	1.7224
P_1	0.43	0.2779	0.5221	0.5626
P_2	0.7	0.5468	0.7024	0.7444
r	0.5994	0.6044	0.6495	0.2714
T	0.7429	0.5470	0.3320	0.4663
p_e	0.9418	0.3076	0.8670	0.7204
p_a	0.2569	0.0763	0.5057	0.2393
jump size 1	34.8053	54.5155	22.9506	38.0779
jump size 2	133.1988	207.7258	124.1705	133.8524
jump size 3	232.0961	316.7951	225.0505	215.6279
f_{11}	0.9223	0.3759	0.7879	0.4142
f_{22}	0.1928	0.3744	0.3362	0.2674
τ	0.75	0.75	0.8614	0.8552
S_e	0.7826	0.7826	0.8223	0.8552
S_p	0.959	0.959	0.9597	0.9540
Error	220.8991	123.7509	220.14	125.9956

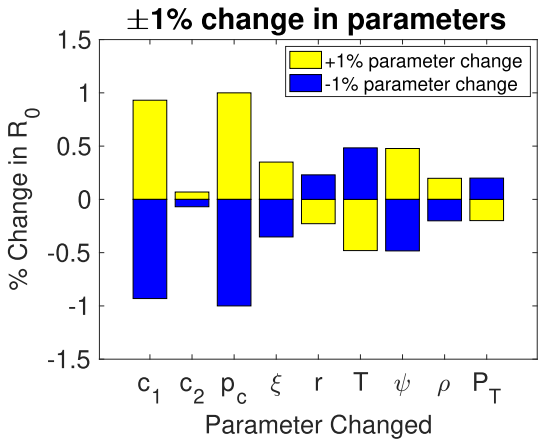

Fig. 8. Effect of a ± 1% perturbation in parameter value on R_0.

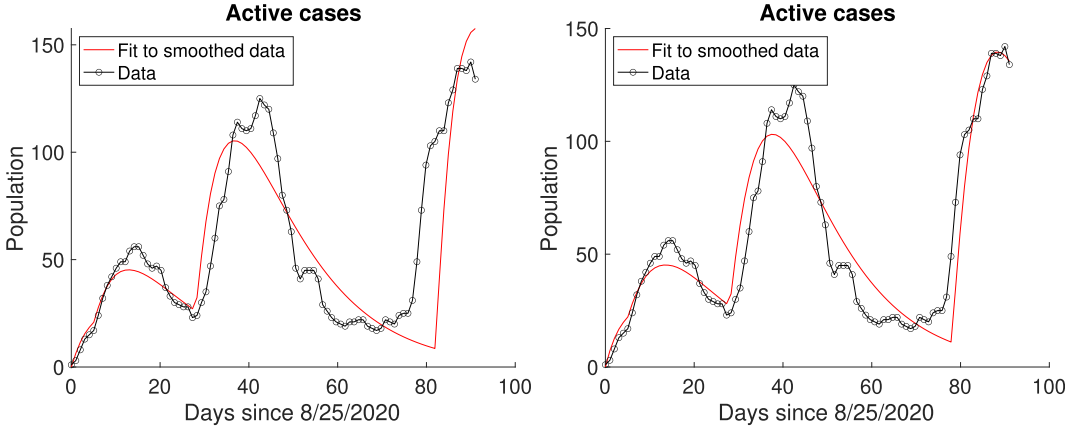

Fig. 9. Model Fit to Estimated Data, assuming default fixed parameters, all jumps 14 days post event (left), Halloween jump moved to 10 days post event (right).

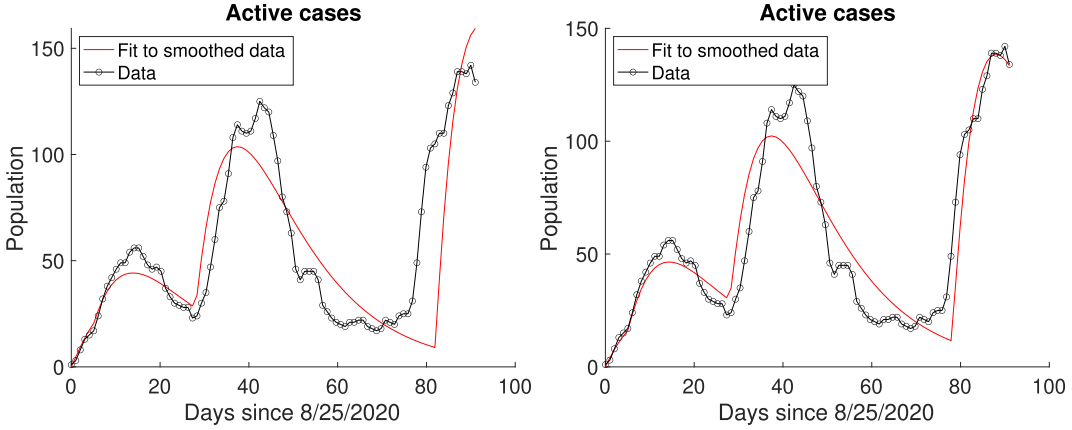

Fig. 10. Model Fit to Estimated Data, Fitting τ, S_e, S_p additionally, assuming default fixed parameters, all jumps 14 days post event (left), Halloween jump moved to 10 days post event (right).

$$
\begin{aligned}
\frac{dS_i}{dt} &= -\frac{c_i p_c f_{ii}(I_i+\tau A_i)}{I_i+S_i+A_i+R_i}S_i - \frac{c_j p_c f_{ji}(I_j+\tau A_j)}{I_j+S_j+A_j+R_j}S_i - \frac{1}{\rho}(1-S_p)P_T S_i - \alpha_{ii}S_i - \alpha_{ji}S_i + \frac{1}{\mu}Q_i \\
\frac{dE_i}{dt} &= \frac{c_i p_c f_{ii}(I_i+\tau A_i)}{I_i+S_i+A_i+R_i}S_i + \frac{c_j p_c f_{ji}(I_j+\tau A_j)}{I_j+S_j+A_j+R_j}S_i - \left(\frac{(1-Tp_e)}{\sigma} + \frac{Tp_e}{\psi} + \frac{1}{\rho}S_e P_T(1-Tp_e)\right)E_i \\
\frac{dA_i}{dt} &= \frac{(1-Tp_e)}{\sigma}E_i - \frac{(1-Tp_a)}{\varphi}P_i A_i - \frac{1}{\eta_0}(1-P_i)(1-Tp_a)A_i - \frac{Tp_a}{\psi}A_i - \frac{1}{\rho}S_e P_T(1-Tp_a)A_i \\
\frac{dI_i}{dt} &= \frac{(1-Tp_a)}{\varphi}P_i A_i - \frac{1}{\xi}rI_i - \frac{1}{\eta_1}(1-r)(1-d_i)I_i - \frac{1}{\zeta_i}(1-r)d_i I_i - \frac{1}{\rho}S_e P_T(1-r)I_i \\
\frac{dQ_i}{dt} &= \frac{1}{\rho}(1-S_p)P_T S_i + \alpha_{ii}S_i + \alpha_{ji}S_i - \frac{1}{\mu}Q_i \\
\frac{dL_i}{dt} &= \frac{1}{\xi}rI_i + \frac{Tp_e}{\psi}E_i + \frac{Tp_a}{\psi}A_i + \frac{1}{\rho}S_e P_T((1-Tp_a)A_i + (1-r)I_i + (1-Tp_e)E_i) \\
&\quad -\frac{1}{\eta_i}(1-d_i)L_i - \frac{1}{\zeta_i}d_i L_i \\
\frac{dR_i}{dt} &= \frac{1}{\eta_0}(1-P_i)(1-Tp_a)A_i + \frac{1}{\eta_i}(1-r)(1-d_i)I_i + \frac{1}{\eta_1}(1-d_i)L_i \\
\frac{dD_i}{dt} &= \frac{1}{\zeta_1}(1-r)d_i I_i + \frac{1}{\zeta_i}d_i L_i
\end{aligned}
\tag{1}
$$

for $i = 1, 2$ and $j \neq i$ above and where

$$
\alpha_{ji} = \frac{\frac{c_j d}{5}(1-p_c)Tf_{ji}\left(\frac{rI_j}{\psi+\xi} + \frac{S_e P_T}{\psi+\rho}\left[(1-Tp_a)A_j + (1-r)I_j + (1-Tp_e)E_j\right] + \frac{(1-S_p)}{\psi+\rho}P_T S_j\right)}{I_i+S_i+A_i+R_i}
$$

for $j = i$ and $j \neq i$.

In our model the susceptible class loses members when individuals become infected due to contact with infectious individuals, receive a false positive test result, or are contact traced. This class will also regain quarantined individuals after they complete an assumed required 14 day quarantine. Once a susceptible person has been exposed to the virus in such a way that they will eventually become infectious they are moved to the exposed class. They then leave the exposed class either due to contact tracing/surveillance testing, which moves them to isolation, or they have become infectious and move to the asymptomatic/presymptomatic class. Asymptomatic/presymptomatic people then can also be isolated due to contact tracing/surveillance testing or, if they are not detected, they will eventually either move to the symptomatic class or the recovered class. Any individual who is caught by contact tracing or surveillance testing who is not actually infected will be moved to the quarantine class until they complete the required 14 day quarantine and then return to the susceptible class. Symptomatic people may be isolated, recover, or die. Isolated individuals may recover or die. Recovered people will remain recovered as we assume (6) at least short-term immunity during the semester. The flow of this process is depicted in Fig. 1.

We note that beyond the additional classes the main changes from the standard SEIR model are the inclusion of contact tracing and surveillance testing. Our method of modelling contact tracing is somewhat different than those considered in other COVID-19 university models (Paltiel et al., 2020; Cashore et al., 2020b; Gressman & Peck, 2020; Lopman et al., 2020). We assume that some proportion of the exposed and asymptomatic/presymptomatic classes are removed in some average amount of time. We note this proportion depends on the proportion of contacts reached, T, as well as the probability a contact is traced while they are in the exposed or asymptomatic classes, p_e or p_a respectively. This reflects our assumption that all interactions occur on campus. This type of contact tracing term has been used before in other generic SEIR-type models, for example in (Feng, 2007). We also assume some proportion of exposed and asymptomatic classes is removed due to surveillance testing. These terms depend on the sensitivity of the test given as well as how often the testing is completed and the proportion of the total population that is tested in that time frame. This term is very similar to that used in (Paltiel et al., 2020) however we include the possibility that the proportion of the total population tested is less than 100%. In addition, the terms modelling those caught by contact tracing who are not actually infected are slightly different than other models. We aim to estimate the number of people who have tested positive daily and the number of contacts those individuals had with the susceptible class in order to determine how many susceptible individuals are removed by contact tracing. We note this term has a similar structure to the standard SEIR term modelling contact between those infected with susceptible individuals. It is modified to include only those infected that the university is aware of, i.e. $\left(\frac{rI_2}{\psi+\xi} + \frac{S_e P_T}{\psi+\rho}((1-Tp_a)A_2 + (1-r)I_2 + (1-Tp_e)E_2) + \frac{(1-S_p)}{\psi+\rho}P_T S_2\right)$. This term includes false positives from the surveillance testing program as these individuals will also have their contacts traced. For simplicity, we assume that contact-traced individuals do not then have their contacts traced. Our model accounts for individuals sharing d days worth of contacts. Finally, we account for the fact that not all contacts are caught by

the contact tracing program by considering only a proportion, T, of an individual's contacts are traced. We also add in an additional factor of 1/5 in the term of those in the susceptible class who are contact traced. Using data from their local health department (Cashore et al., 2020b) found that about 1/5 of contacts actually met the standard for quarantine as set by the CDC. We assume everyone who is traced who is not actually ill will move to the quarantined class and is required to fulfill a full 14 day quarantine regardless of a subsequent negative test result. We acknowledge there are varying accepted lengths for quarantine and isolation and we choose to have both be a full 14 days.

We also carefully model the initial conditions for all population classes since gateway testing inevitably impacts these. To investigate hypothetical scenarios in §3.1 and §3.2, we assume that there are no individuals initially in the exposed, recovered, or deceased classes. For the infected classes we assume a proportion, P_i of the infectious population is symptomatic and $1 - P_i$ of the infectious population is asymptomatic with $i = 1, 2$ for students and faculty, respectively. For scenarios with no gateway testing (1, 2), the quarantine and isolation classes also begin empty. For scenarios with a gateway test (3, 4, 5), some of the infectious population is instead placed in isolation with the amount based on the sensitivity of the test used. Additionally, we allow the possibility that some people will self-isolate based on symptoms. We also account for the fact that some susceptible individuals will receive false positive tests based on the test specificity and add this number of false positives into the initial quarantine class. For data fitting in §3.3, we use the data to obtain the initial condition for the isolated and quarantined classes and estimate the remaining initial conditions.

2.3. Parameter choices

Equation (1) depends on a significant number of parameters, most of which are not known exactly. There are some that have been estimated from data but many of these estimates include uncertainty and some are not known at this time due to limited publicly available data from college campuses. We attempt to make reasonable estimates based on the literature whenever possible which usually gives a range of values. We model any parameters given with a range of values as uniform random variables over those ranges to reflect our uncertainty. Future work may involve better estimation of parameter-specific probability distributions. We will now describe the ranges or estimates used in our simulations of hypothetical outbreaks in §3.1. The possible ranges of parameter values and their corresponding default values used in this paper are listed in Table 2.

The contact rate, which has units days^{-1} depends on the number of contacts per infected case per day (c_i) as well as the probability of transmission per contact (p_c). We assume that students and faculty have a different average number of contacts. For students we assume 7–17 contacts per day based on the discussion in (Cashore et al., 2020a). We assume faculty have fewer daily contacts than students, 1–4 in this paper, based on the fact that it was found in (Cashore et al., 2020b,a) that the majority of faculty close contacts will be off campus which we do not model. In §3.1, we will break the c_1 range into three subranges to account for various levels of adherence to social distancing among students.

The probability for transmission due to a close contact, p_c, was found to be 2.4% in (Jing et al., 2020) and 3.7% in (Luo et al., 2020). These values were estimated from contact tracing data in China at the beginning of the pandemic and do not take into account mask wearing and social distancing. It has been estimated that masks may reduce transmission by as much as 80% (Liang et al., 2020). We therefore assume that the low end of our range is 20% of the average of the two estimates from literature and consider the range 0.6–3.7%. In §3.1, we will break this range up into three subranges to account for various levels of adherence to mask wearing from both faculty and students.

To further model the differences in contact between students and faculty we also introduce estimates for the proportion of total contacts an infected student/faculty has with susceptible students/faculty. At this time, these values are educated guesses due to a lack of data. We assume the proportion of contacts an infected student makes with other students (f_{11}) is 70% and therefore the proportion of contacts an infected student makes with faculty (f_{12}) is 30%. We assume faculty primarily interact with students on campus therefore f_{21} is set at 70% and f_{22} is 30%.

One final factor for the contact rate is the possible reduction in transmission due to contact with an asymptomatic individual compared to that with a symptomatic one, represented by τ. The estimates of this difference in transmissibility vary but most agree that asymptomatic individuals are less likely to spread the virus (Centers for Disease Control and Prevention, 2020; Luo et al., 2020). In (He et al., 2020) it was found that there was no statistically significant difference in the estimates for asymptomatic versus symptomatic transmission. We therefore consider $\tau \in [0.5, 1]$.

The number of days between surveillance test results (ρ) is a measure of how long it takes on average for the tests to be performed and results returned so that positive cases are then isolated or quarantined. This value is university-protocol-dependent, so we consider frequencies of testing from daily to biweekly to account for various protocols. In addition, the proportion of the population that is tested (P_T) is also policy-dependent. Since most surveillance testing protocols randomly sample from the total campus population, we assume that if P_T of the total population is tested, then the same proportion P_T of each of the susceptible, exposed, asymptomatic, and symptomatic classes is also tested. We allow P_T to vary from 5 to 100% to account for various protocols. Finally, we take the ranges for the specificity (S_p) and sensitivity (S_e) of tests to be 95.02–100% and 62.4–94.12%, respectively, based on a study of available tests used in an ER setting (Bisoffi et al., 2020).

Parameters related to contact tracing are also campus-protocol-dependent. However (Spencer et al., 2021) looked at contact tracing programs in more than 100 health departments and found that between 32% and 79% of close contacts (identified by the individual) were reached within the first 24 hours. We use these percentages for the proportion of contacts reached (T) in §3.1. We also assume in most cases that the average amount of time between identifying a positive case and

reaching their contacts is $\psi = 1$ day. The probability that a case caught by contact tracing is in the exposed (p_e) or asymptomatic (p_a) class is not estimated in the literature to our knowledge and therefore we let them range from 0 to 1. The parameter d represents the number of days of contact the tracers will consider. The CDC recommends tracing beginning two days prior to symptoms (Centers for Disease Control and Prevention, 2021b). We assume that most people limit contact post symptoms and therefore use $d = 2$ in all simulations. We also assume, as it is the policy at most institutions, that traced contacts stay in quarantine for $\mu = 14$ days.

The time from exposure to symptom onset has been estimated to be around 5–6 days with one estimate as high as 7 (Lauer et al., 2020; Centers for Disease Control and Prevention, 2020; Tindale et al., 2020). Our model treats this value as the sum of σ, the number of days between exposure and joining the presymptomatic/asymptomatic class where individuals are deemed infectious, and φ, the number of days between becoming infectious and developing symptoms. In (Tindale et al., 2020) it was estimated that an infectious person infected another approximately 2–3 days before symptom onset. Therefore we set φ to be between 2 and 3 days and σ to be between 2 and 4 days.

There are three separate recovery times listed in our parameter set: η_0 for those who never develop symptoms, η_1 for symptomatic students and η_2 for symptomatic faculty. We set all three to be 14 days (Voinsky et al., 2020) as there is not much said in the literature on the difference between recovery times for different age groups and levels of symptoms. For the sake of our model, we are mainly concerned with the period of infectiousness when these individuals may cause new infections due to contact with susceptible individuals. For future generalizability, we keep these as three separate parameters in case more data becomes available on the differences between asymptomatic and symptomatic periods of infectiousness.

The next parameter to consider is ξ which represents the amount of time on average before a symptomatic person self-reports (or self-isolates). This value is challenging to estimate, however, in (Centers for Disease Control and Prevention, 2020) they estimate the median number of days from symptom onset to getting a test among positive patients as 1–6 days with a base value of 3 days. We assume that obtaining a test is equivalent to self-reporting and therefore being isolated and use this range for ξ. It was estimated in (Reese et al., 2020) that those aged 18–49 with symptoms self report 15–65% of the time (with an average of 34%) which we use as our range for r.

Our model has a few parameters related to death. First, ζ_1 and ζ_2 represent the average number of days between symptom onset and death for students and faculty, respectively. The current best estimates by (Centers for Disease Control and Prevention, 2020) are relatively the same for all age groups so we set both to 16 days. Again, we keep these parameters separate in case more information is made available in the future. Second, we require a parameter for the probability of death for each group (d_i). We use (Centers for Disease Control and Prevention, 2021a) to estimate d_i. To do this, we require an estimate of the age distribution of students and faculty. For students, since we model a primarily undergraduate institution, we assume the majority fall into the age range 18–29 years and estimate the probability of death to be the number of deaths due to COVID-19 divided by the total number of infections which gives us $d_1 = 0.000359$. We note that since we are unable to estimate precisely how many of these reported infections are asymptomatic this estimate is most likely not accurate but given the smaller size of this probability we do not expect it to have a substantial effect on results. We also note that this neglects the possibility of nontraditional students since we assume these students to make up a small percentage of the overall student body and are also less likely to live on campus. Therefore they will not impact the disease spread as much as other students. For faculty the age range varies much more. From (McChesney & Bichsel, 2020) we estimate that 13% of faculty are 65 and older. We will use the CDC data for ages 65–74 for this group, since there is no further breakdown of this group in (McChesney & Bichsel, 2020). To estimate the proportion of faculty in other age groups, we refer to the age distribution of faculty at Penn State University in 2017 (Penn State University Budget Office, 2017). Based on this, 2.4% of faculty are in the 18–29 group, 24.7% are in 30–39, 28.1% are in 40–49. The remaining age group, of 50–64 did not align with the breakdown in (Penn State University Budget Office, 2017), but based on the rest of the age groups, we estimate that 31.8% of faculty are in 50–64. Based on (Centers for Disease Control and Prevention, 2021a), we estimate the probability of death in each of these groups as we did for the student population. We then compute a weighted average of the probabilities of death for these age groups to estimate $d_2 = 0.0105$.

Next we consider the parameters that represent the proportion of infections that are asymptomatic versus symptomatic. It is believed that younger populations will more likely be asymptomatic. We therefore define two parameters for the proportion of symptomatic infections in students (P_1) and faculty (P_2). These estimates vary widely, for example the CDC (Centers for Disease Control and Prevention, 2020) has a range from 10 to 70% over all age groups. In (Byambasuren et al., 2020) they found studies with estimates for the proportion of asymptomatic infections ranging from 4 to 40%. In (Mizumoto et al., 2020) it was found that 17.9% of infected passengers on the Diamond Princess cruise ship were asymptomatic. In addition (Poletti et al., 2020) finds much lower estimates of symptomatic infections for example finding 18.09% of 0–19 years display symptoms. We therefore define our P_1 range for students as 20–60% where the low end comes from an average of the 0–19 and 20–39 age group estimates in (Poletti et al., 2020) and the high end comes from (Centers for Disease Control and Prevention, 2020; Byambasuren et al., 2020). For faculty, P_2 ranges from 30 to 80% symptomatic infections where the low end comes from the 30–49 age group estimate in (Poletti et al., 2020) and the high end from (Byambasuren et al., 2020; Mizumoto et al., 2020).

Lastly we discuss parameters related to the initial conditions listed in Table 3. In (Cashore et al., 2020b) they estimate that the proportion of the student population that is infected at the beginning of the semester ranges from 0.5 to 4%. This estimate was formed before any gateway testing had been completed. Since (Cashore et al., 2020b) an e-mail from the Office of the President at Villanova University reported that 0.41% of gateway tests from Villanova students and faculty returning to campus were positive (Personal Communication, 2020). We therefore include the possibility of a smaller percent infected and

modify the range to be 0.25–4% to include the possibility of lower prevalence. We assume all who test positive begin the semester in isolation. Within this group there will be some false positives who begin in the quarantine class who can then return to the susceptible class after 14 days. We also assume there are some who may not test positive or who self-isolate due to symptoms. We assume the proportion of students who self-isolate is equal to r discussed above. We also assume faculty may be more likely to self-isolate and increase this range to 50–75% for the initial conditions only since classes haven't started yet. For those who are not isolated but are infected we divide these into the asymptomatic/symptomatic classes using the proportions P_1 and P_2 defined above for students and faculty, respectively. For simplicity we assume at the start of the semester there is no one in the recovered class though there are likely some in the population who had the disease and still possess some form of immunity.

2.4. Shocks and data fitting

We will validate our model using data from Villanova University's COVID Dashboard for the fall 2020 semester (Villanova University, 2020). We note that the dashboard data only provides new positive tests per day and only includes positive cases of which the university was made aware. We will assume that each positive case will remain active (i.e. still appear as a case in the data) and isolated for 14 days in order to estimate the number of isolated individuals at any given time. This estimated active cases (or isolated cases) per day is shown in Fig. 2. We will fit our model to this data using the MATLAB function fminsearchbnd from (D'Errico, 2012) using the sum of $L_1(t)$ and $L_2(t)$ to fit to our estimated data. We will use the default value of parameter values we are more confident in from literature and estimate the remaining using the data. The parameters estimated are p_c, c_1, c_2, ξ, P_1, P_2, r, T, p_e, p_a.

It is clear from the Villanova data and other state, national, etc. trends that there are spikes in cases that will not appear by simply modelling SEIR-type dynamics. We assume that there are certain days/events on campus that act as superspreader events causing these spikes. For the Villanova data we hypothesize there are spikes 2 weeks after the following events: the first weekend on campus (prior to class starting), Labor Day weekend, and Halloween. We model these spikes by having a sudden increase or shock in the number of infections approximately 14 days after these expected events. In simulations we simply start and stop the simulation at these dates and add in an expected number of new members to the asymptomatic and symptomatic classes (proportioned using P_1) to the new initial conditions and remove these people from the susceptible classes. We assume that shocks affect only students. Thus, the shock for superspreader event k occurring at time t_k is modelled as

$$\begin{aligned} \Delta S_1 &= -\text{jump size } k, \\ \Delta A_1 &= (1-P_1)(\text{jump size } k), \\ \Delta I_1 &= P_1(\text{jump size } k) \end{aligned} \tag{2}$$

for $t = t_k + 14$. In the numerical work, these shocks re-initialize Equation (1) at $t = t_k + 14$. We note that we also estimate the size of these shocks (jump size k) in the data fitting procedure in §3.3 while the dates on which they happen, t_k, are fixed.

3. Results and discussion

We use MATLAB's built-in `ode45` solver to solve our system of differential equations and provide results for the number of people in each population on each day of a 98 day semester. We present three different sets of results in order to portray the possible scenarios predicted by our model as well as to provide analysis and validation of our model. We first discuss a series of hypothetical epidemics. Here we will consider various ranges of the model's parameters to see the breadth of possible outcomes. We include these results to reflect the uncertainty in parameter estimates as well as to demonstrate the different results depending on institutional policies and individual choices. To further understand the effects of certain policies and behaviors we perform sensitivity analysis on several parameters related to mitigation measures. Finally, we will fit our model to estimated data from Villanova University's fall 2020 semester COVID dashboard, (Villanova University, 2020), to validate our model and provide better estimates of university-specific parameters.

3.1. Hypothetical epidemics

We begin by considering the vast range of possible outcomes due to our uncertainty in some parameter values as well as the variability in containment strategies used across the country on different campuses. We will measure success of different strategies by determining the change in the percent of the total population infected over the course of the semester. We will mainly discuss scenarios that highlight the effects of institution-level practices, i.e. contact tracing and surveillance testing, as well as individual practices, i.e. social distancing and mask wearing. We consider five scenarios of campus-level mitigations measures:

1. No mitigation measures are implemented.
2. Contact tracing is implemented, but there is no gateway nor surveillance testing.
3. Gateway testing and contact tracing are implemented, but there is no surveillance testing.
4. Gateway and surveillance testing are implemented, but there is no contact tracing.

5. Gateway and surveillance testing and contact tracing are all implemented.

First we consider the effects of individual practices such as social distancing and mask wearing. We therefore set parameters related to university measures, as well as those related to the initial conditions, to the default values in Tables 2 and 3 to ensure we are seeing the effects of the individual level practices. With regards to contact tracing and surveillance testing we set $T = 0.555$, $\psi = 1$, $P_T = 0.25$, and $\rho = 7$. For the gateway test and initial condition we assume the initial percent of the population that is infected is 2.1%, the proportion of students/faculty that self-isolate at the start are 0.4/0.624 respectively, the sensitivity of the gateway test is 78.26%, and the proportion of students/faculty that are symptomatic are 0.4/0.55 respectively. We randomly sample c_1 and p_c from the ranges given in Table 2 for 250,000 simulations. We then consider three cases for social distancing and mask wearing parameters. We assume social distancing is modelled by the number of contacts per individual. We assume that in the case of social distancing the student contact numbers are as follows: pessimistic $c_1 \in [13.667, 17]$, nominal $c_1 \in [10.333, 13.667)$, and optimistic $c_1 \in [7, 10.333)$. For mask wearing we assume that p_c, the probability of transmission, is proportional to the level of mask wearing adherence. We assume the following three cases: pessimistic $p_c \in [2.667, 3.7]$, nominal $p_c \in [1.6333, 2.667)$, and optimistic $p_c \in [0.6, 1.6333)$.

Figs. 3 and 4 focus on Scenario 5 when all institution-level strategies are in place. Fig. 3 shows the daily percent of students that are actively infected (i.e. the sum of the asymptomatic and symptomatic student classes) for different levels of adherence to social distancing (SD) and mask wearing (MW) policies from pessimistic SD/MW to optimistic SD/MW from left to right. The mean curve of each level of adherence is displayed as a thick black line. In all plots we assume gateway testing, surveillance testing, and contact tracing have been performed with the parameters specified above. We note that as we increase adherence the peak levels of infections reduce from approximately 30%–10% to 0.35%. In particular we note that the peak level of infections actually occurs at the beginning of the semester in the optimistic scenario, hence there is no outbreak. Fig. 4 shows the mean curves for different combinations of mask wearing and social distancing adherence. We see again that the only case when there is a peak value larger than the initial percent infected is in the case when we assume both pessimistic adherence in social distancing and mask wearing. In addition, we see monotonically decreasing infection levels whenever mask wearing is optimistic or nominal, except in the case when social distancing is pessimistic. In Tables 4–6 we display the mean total percent of students and faculty infected over the course of the semester for all five institution-level scenarios. To highlight when individual measures are the only means of mitigation, we consider Scenario 1 across Tables 4–6. We find that if we hold mask wearing at the pessimistic level (left column) and increase adherence to social distancing (increase table number) this leads to a drop from approximately 96%–90% and finally 72% total percent infected students as we reach optimistic social distancing. On the other hand, if we hold pessimistic social distancing levels, Table 4 only, and look at varying mask wearing levels, we see the total percent of students infected decreases from 96% to 80% and finally 26% as we reach optimistic mask wearing. Clearly, mask wearing adherence is a critical part of any mitigation strategy. We note that while we do not explicitly show the faculty curves they are similar to those of the students with the exception of having much lower peak values as demonstrated in Tables 4–6.

Next we consider comparing contact tracing (CT) and surveillance testing (ST) as strategies. We again fix certain parameters to focus on comparing only the university-level mitigation measures. We assume $c_1 = 12$, $c_2 = 2$, and $p_c = 2.15\%$. We use the same parameters affecting the initial conditions as specified above. We randomly sample T, P_T, and ρ from the ranges given in Table 2 for 300,000 simulations. We then define nominal contact tracing to be $T \in [0.32, 0.555)$ and robust contact tracing to be $T \in [0.555, 0.79]$ with $\psi = 1$ in both cases. For surveillance testing, we define limited testing to be $P_T \in [5, 25)\%$ and $\rho \in [7, 14]$ days, nominal testing to be $P_T \in [25, 75]\%$ and $\rho \in [7, 14]$ days, and robust testing to be $P_T \in [75, 100]\%$ and $\rho \in [1, 7]$ days.

Fig. 5 shows the daily percent of students that are actively infected assuming constant levels of social distancing and mask wearing defined above. Mean curves for each case are displayed as thick black lines. The results of the six cases of various levels of university-level practices show the full breadth of possibilities. From left to right, we see that increasing levels of surveillance testing while holding constant levels of contact tracing causes peak infection levels to decrease dramatically. Also, in the case of limited surveillance testing (left) increasing contact tracing from nominal (top) to robust (bottom) levels does not reduce the maximum peak values but does decrease the number of curves we see that hit those peak levels. In the case of nominal surveillance testing (middle) we see that increasing contact tracing levels decreases peak values by almost half. Lastly, when we have robust surveillance testing (right) the increased level of contact tracing decreases the number of infectious cases earlier in the semester for more simulations.

Table 7 shows the mean total percent of the student and faculty populations that are infected over the course of the semester for the various scenarios. We first note the utility of gateway testing, which when added to contact tracing (moving from Scenario 2 to 3), reduces the percent infected by approximately half. In addition, the total percent of students infected in Scenario 3 is 4.27% and 3.24% for nominal and robust contact tracing programs, respectively. For Scenario 4 to reach a similar level of infection we require surveillance testing to be performed at robust levels (75–100% of the population at least weekly). Lastly, Fig. 6 shows the mean curves for a single mitigation measure (along with a gateway test) on the left and dual mitigation measures on the right. On the right we note that in all cases we do not see peak values reaching initial infection levels again during the semester, indicating successful strategies. However, on the left we see that nominal surveillance testing only does not control the outbreak as the other single measure cases do. We therefore conclude that surveillance testing must be robust if used on its own as a mitigation measure. We also note that if a contact tracing strategy is already in place, further adding in surveillance testing lowers the percent infected by about 1–2% for different level strategies as seen in Table 7. We therefore

find that if a robust surveillance testing strategy is not available then it serves an institution most to develop a robust and efficient contact tracing program.

3.2. Sensitivity analysis

In this section, we investigate the effect of perturbing parameters related to disease mitigation measures at both the university- and individual-level. These parameters are c_1, c_2, p_c, ξ, r, T, ψ, ρ, and P_T. In addition to quantifying the effect perturbations of these parameters have on the total number of student and faculty that are infected over the course of a semester, we also quantify the effect these perturbations have on R_0.

We calculate the basic reproduction number using the next generation matrix method. While in cases with contact tracing and surveillance testing this model yields an infinite number of disease free equilibria (DFE) we choose to linearize about the DFE where all individuals remain in an uninfected classes, in our case in S_1, S_2, Q_1, or Q_2, to compare with the standard DFE where all individuals remain susceptible. There are more details on the calculation of R_0 in A. In the case of no contact tracing or surveillance testing, our R_0 estimates range from as low as 0.2456 (in cases with excellent social distancing and mask wearing) to as high as 5.1909 (in cases with poor distancing/mask wearing and assuming very few symptomatic individuals isolating). The mean R_0 is 1.7478, which is lower than many estimates for R_0 but most previous estimates do not take into account mask wearing and social distancing.

To perform the sensitivity analysis, each parameter was initialized as the default value found in Table 2. Then, each mitigation measure parameter was perturbed positively and negatively by a certain percentage while keeping all other parameters fixed. The resulting percent change in total infected population and R_0 value were computed. Fig. 7 shows the percent change in total number of students and faculty infected over a semester caused by a $\pm$ 1% perturbation in each mitigation measure parameter. Similarly, Fig. 8 shows the percent change in R_0 by these perturbations. As can be seen in these Figures, the total number of students and faculty infected and R_0 are most sensitive to the c_1 and p_c parameters, which relate to the number of daily contacts students have and the probability of transmission per contact, respectively. Interestingly, a $\pm$ 1% change in any of the parameters mostly caused no more than a $\pm$ 1% change in the outcomes except for the effect of c_1 and p_c on the total student population infected. This highlights that maintaining a safe in-person semester can be the most affected by student behaviors. We note that the model is much less sensitive to the rate and speed at which individuals self-isolate upon becoming symptomatic. This is likely due to the assumption that contact tracing and surveillance testing are both taking place which lessens the burden of responsibility on individuals. Looking at Figs. 7 and 8, we see that outcomes and R_0 values were next most sensitive to changes in T and ψ after student behaviors, which are related to the level of contact tracing performed on campus. This further confirms that contact tracing is a powerful measure that can provide excellent control when implemented robustly and efficiently. We, however, stress these results be viewed through the lens of our assumption that all contacts occur on campus. It is likely that campus contact tracing programs will be affected by contacts made outside of campus which would be harder to trace.

3.3. Data fitting

As mentioned in Section 2 we fit our model to the estimated data from the Villanova University COVID Dashboard for the Fall 2020 semester (Villanova University, 2020). We note as before that we see from the data three distinct peaks. It is clear from our hypothetical results that our model, on its own, will not capture this type of behavior. We hypothesize that the independent peaks are due to days/periods of increased social activity on campus. Specifically, we assume there is an influx of new cases due to the return to campus (that would not be reflected by Equation (1)), Labor Day weekend (which is the only time students had a day off from classes during that semester), and Halloween. Halloween is always a time of increased social activity and was towards the end of the semester so we hypothesize students were growing weary of social distancing protocols. In order to model these expected spikes or superspreader events we use Equation (2) to add in a number of new infections coming from the susceptible class, spread across the asymptomatic and symptomatic classes, approximately 14 days after the event. We show results for assuming all 3 spikes occur 14 days post event as well as the case when the last spike is 10 days post Halloween weekend (for which we apply Equation (2) at $t = t_{\text{Halloween}} + 10$). It is expected that there will be some variation in how quickly spikes occur after known superspreader events. We estimate the following parameters: p_c, c_1, c_2, ξ, P_1, P_2, r, T, p_e, p_a, f_{11}, and f_{22} as these parameters are dependent on the specific population on campus. In addition we estimate the sizes of the three jumps (jump size k) as well as the initial conditions for the exposed, asymptomatic, symptomatic, and recovered classes. We assume from dashboard data that the initial quarantined, isolated, and deceased classes are all zero. We set the remaining parameters to be their default values as listed in Table 2. These fits are shown in Fig. 9. We also show cases when, in addition, we estimate τ, S_e, and S_p in Fig. 10. τ in particular is not precisely known, in addition our knowledge of S_e and S_p is limited due to multiple tests being used on campus as well as natural variations due to collection and processing error. Note that all parameters to be fitted have their initial value chosen from a uniform random distribution in the ranges set in Table 2.

Overall in both cases there is generally a good agreement between the model and the data. In particular we are able to capture the third peak well when we assume the jump associated with Halloween occurs 10 days after instead of 14. This leads to a substantial decrease in error as seen in Table 8. The second peak is the hardest to capture and therefore the size of the error in all fits is still substantial. We see that p_c is found to be significantly higher when we don't use the default values for

τ, S_e, and S_p. Our estimated contact rates fall into the ranges in Table 2 and in (Cashore et al., 2020a). In particular they fall in our nominal contact rate range. There is even more variability in our estimates for P_1 and P_2 though in all cases P_1 is lower than P_2 as expected since students are on average much younger than faculty and hence more likely to have an asymptomatic infection. We also note that p_e is always larger than p_a which indicates it was more likely for contact tracing to find contacts in the exposed state rather than the presymptomatic/asymptomatic state which is expected and desired. This variation in estimates is not surprising given the novelty of the virus and the level of uncertainty in parameter values at this time. We expect that as more information is gathered these parameters may not need to be estimated and the fit to data will likely change.

3.4. Model limitations

We conclude our discussion by mentioning briefly the limitations of our model. As with any model not every aspect of a phenomenon is captured and the results should be viewed through the lens of this knowledge. We first note that this model is one that assumes homogeneous mixing, likely overestimating transmission rates, as individuals all have different behavior and contact patterns. We see in (Gressman & Peck, 2020) much lower infection rates than in compartmental models. However we note that in (Bansal et al., 2007) it is noted that error due to the assumption of homogeneous mixing is less when considering heavily mixed populations like cities. We assume college campuses fall into this category. We also note that we have not incorporated the effects of staff and graduate students on campus. We ignored these groups due to an assumption that they have less contact with students who are the primary drivers of infection on campus. On certain campuses, however, graduate students, in particular, have a lot of contact with undergraduate students and therefore should be considered in that case. Staff likely primarily have contacts off campus. Depending on the level of transmission in the community-at-large these contacts may or may not have an impact on the spread on campus. We have neglected to include the effect of contacts off campus. In most other models, for example (Lopman et al., 2020), the effect of off-campus interactions is modelled as a constant rate of new infections per day occurring on campus. Whether or not this has a large impact on overall infections on campus is dependent on the level of community spread as well as the amount of time students spend off campus. In this model, we assume students are asked to limit their time off campus however it is likely many students do spend time off campus, in particular on weekends or other days off. This is likely the assumption that limits our model results the most and we intend to introduce these effects into the model during future study.

It is well known that SEIR-type models do not predict the multiple surges or seasonality seen in many infectious disease epidemics. In this model we have addressed this by including shocks to the number of infected individuals after expected superspreader events. There are likely some dates which are known at the beginning of the semester to likely have superspreader activities, like spring break. However it is impossible to predict all possible superspreader events prior to the start of the school year. Therefore, predictions using these types of models will likely miss the effects of these types of events entirely. It is also possible that there is additional time variation occurring in parameters like contact rates dependent, perhaps, on season that have not been captured by this model. More needs to be known about this specific virus before we can accurately model time dependence in parameter values. Lastly, most parameter values in this work are assumed to be drawn from a uniform distribution defined by a range of possible values found in literature. It is possible to better estimate probability density functions for these parameters which would likely lead to less variation in our set of hypothetical epidemics. We leave this as future work as well.

4. Conclusion

In this work we have developed an SEIR-type model for the spread of COVID-19 on a medium-sized university campus with a primarily residential undergraduate student population. We have found that student behaviors, in particular mask wearing and social distancing, largely determine the ultimate size of an on-campus outbreak. In contrast faculty largely do not contribute to the spread of the virus. We find that for those universities who are unable to develop a robust surveillance testing program contact tracing can provide excellent control of the spread. Smaller scale surveillance testing can provide further reduction in total infections but that reduction is not substantial. We find that our model compares well with collected data from Villanova University's fall 2020 COVID-19 dashboard (Villanova University, 2020) when we further include superspreader events into the model. This does point to the difficulty of truly predicting total infections when it is unknown what events will lead to such large sudden influxes of new infections but there are common events such as fall/spring breaks which may be included when using this type of model for predictive purposes.

CRediT author contributions

Kaitlyn Muller: Data Curation, Writing - Original Draft, Supervision, Project Administration, Formal Analysis. **Peter Muller:** Software, Writing - Review & Editing, Visualization. All other roles were shared equally by the authors.

Appendix A. R_0 calculation

In order to estimate R_0 we utilize the standard next generation matrix method (Diekmann et al., 2010). This method requires one to calculate the Jacobian of the sub-system of differential equations containing only equations that pertain to infected compartments, i.e. including E_i, A_i, I_i, L_i and then evaluate the Jacobian at a disease-free equilibrium (DFE). For this

system there are an infinite number of DFEs and we choose the one that most resembles the standard unique DFE from SIR/SEIR models, i.e. the equilibrium where there are no infections. In our model, this corresponds to everyone remaining in the susceptible class or the quarantine class.

Since at the DFE there is no one in a disease state, we set $E_i, A_i, I_i, L_i = 0$ for $i = 1, 2$. After plugging these values into Equation (1) we are left with the following equations

$$\begin{aligned}
\frac{dS_1}{dt} &= -\frac{(1-S_p)P_T S_1}{\rho} - c_1\frac{d}{5}(1-p_c)Tf_{11}\frac{(1-S_p)P_T S_1}{\rho+\psi}\frac{S_1}{S_1+R_1} \\
&\quad -c_2\frac{d}{5}(1-p_c)Tf_{21}\frac{(1-S_p)P_T S_2}{\rho+\psi}\frac{S_1}{S_1+R_1} + \frac{Q_1}{\mu} = 0 \\
\frac{dS_2}{dt} &= -\frac{(1-S_p)P_T S_2}{\rho} - c_1\frac{d}{5}(1-p_c)Tf_{12}\frac{(1-S_p)P_T S_1}{\rho+\psi}\frac{S_2}{S_2+R_2} \\
&\quad -c_2\frac{d}{5}(1-p_c)Tf_{22}\frac{(1-S_p)P_T S_2}{\rho+\psi}\frac{S_2}{S_2+R_2} + \frac{Q_2}{\mu} = 0
\end{aligned} \tag{A.1}$$

along with the Q_i equations, which are the negatives of (A.1), and the constraints that $S_1 + Q_1 + R_1 + D_1 = N_1$ and $S_2 + Q_2 + R_2 + D_2 = N_2$. We assume that the DFE of interest is when $R_i = 0$ and $D_i = 0$ for $i = 1, 2$. Therefore to find the DFE we seek to solve this system of equations for S_i and Q_i for $i = 1, 2$. However, when we look at the non-zero entries of the Jacobian we find that the only dependence on the population values is in the form of the ratio of S_1 to S_2. Therefore we need only to solve for this quantity.

In the case when there are no measures taken by the university, i.e. when $T = 0$ and $P_T = 0$ the equations above simplify to

$$\begin{aligned}
\frac{dS_1}{dt} &= \frac{Q_1}{\mu} = 0 \\
\frac{dS_2}{dt} &= \frac{Q_2}{\mu} = 0
\end{aligned} \tag{A.2}$$

and therefore $Q_i = 0$ and $S_i = N_i$. In the case when there is no contact tracing ($T = 0$) but surveillance testing is conducted we reduce to the following equations

$$\begin{aligned}
\frac{dS_1}{dt} &= -\frac{(1-S_p)P_T S_1}{\rho} + \frac{Q_1}{\mu} = 0 \\
\frac{dS_2}{dt} &= -\frac{(1-S_p)P_T S_2}{\rho} + \frac{Q_2}{\mu} = 0
\end{aligned} \tag{A.3}$$

and therefore $Q_i = \frac{(1-S_p)P_T\mu}{\rho}S_i$. Plugging this into the constraints $S_1 + Q_1 = N_1$ and $S_2 + Q_2 = N_2$ it may be shown that $S_1/S_2 = N_1/N_2$ as in the previous case. In the case when $P_T = 0$, i.e. no surveillance testing but contact tracing is conducted we obtain the same equations as in the no measures cases. Finally in the case of both contact tracing and surveillance testing we have the following system of linear equations to solve ($S_1' = 0$, $S_2' = 0$):

$$\begin{aligned}
\left(-\frac{(1-S_p)P_T}{\rho} - c_1\frac{d}{5}(1-p_c)Tf_{11}\frac{(1-S_p)P_T}{\rho+\psi}\right)S_1 - c_2\frac{d}{5}(1-p_c)Tf_{21}\frac{(1-S_p)P_T S_2}{\rho+\psi} + \frac{Q_1}{\mu} &= 0 \\
\left(-\frac{(1-S_p)P_T}{\rho} - c_2\frac{d}{5}(1-p_c)Tf_{22}\frac{(1-S_p)P_T}{\rho+\psi}\right)S_2 - c_1\frac{d}{5}(1-p_c)Tf_{12}\frac{(1-S_p)P_T S_1}{\rho+\psi} + \frac{Q_2}{\mu} &= 0
\end{aligned} \tag{A.4}$$

along with the constraint equations. If we define the coefficient of S_1 to be A and of S_2 to be B in the first equation of (A.4) and the coefficient of S_1 to be C and of S_2 to be D in the second equation of (A.4) we obtain the following formula for S_1/S_2 at the equilibrium point using Maple:

$$\left(\frac{S_1}{S_2}\right)^* = \frac{(-BN_2\mu + CN_1\mu - N_1)}{(AN_2\mu - DN1\ \mu - N_2)} \tag{A.5}$$

We may then plug the values for this ratio into the Jacobian for each case of measures taken. To calculate R_0 we then separate the Jacobian into the matrix F, which contains terms related to new infections, and the matrix V, which contains terms related to moving between compartments after initial exposure. We write out the non-zero entries of F and V below assuming both measures are taken. The first two rows of F are given in equation (A.6) since the third through eighth rows are all zeros. Likewise, the non-zero entries of V are given by (A.7). If at least one measure is removed then the ratio in (A.5)

simplifies to N_1/N_2. In the case of one or no mitigation measures, we simply set $T = 0$ and/or $P_T = 0$ to remove the effect of contact tracing and surveillance testing, respectively.

$$F_{1:2,1:8} = \begin{bmatrix} 0 & 0 & c_1 p_c \frac{d}{5} f_{11}\tau & c_2 p_c f_{21}\left(\frac{S_1}{S_2}\right)^* \tau & c_1 p_c f_{11} & c_2 p_c f_{21}\left(\frac{S_1}{S_2}\right)^* & 0 & 0 \\ 0 & 0 & c_1 p_c f_{12}\left(\frac{S_2}{S_1}\right)^* \tau & c_2 p_c f_{22}\tau & c_1 p_c f_{12}\left(\frac{S_2}{S_1}\right)^* & c_2 p_c f_{22} & 0 & 0 \end{bmatrix} \tag{A.6}$$

$$\begin{aligned}
V_{11} &= -\frac{1-Tp_e}{\sigma} - \frac{Tp_e}{\psi} - \frac{S_e P_T}{\rho}, \\
V_{22} &= -\frac{1-Tp_e}{\sigma} - \frac{Tp_e}{\psi} - \frac{S_e P_T}{\rho}, \\
V_{31} &= \frac{1-Tp_e}{\sigma}, \\
V_{33} &= -\frac{(1-Tp_a)P_1}{\varphi} - \frac{(1-P_1)(1-Tp_a)}{\eta_0} - \frac{Tp_a}{\psi} - \frac{S_e P_T}{\rho}, \\
V_{42} &= \frac{1-Tp_e}{\sigma}, \\
V_{44} &= -\frac{(1-Tp_a)P_2}{\varphi} - \frac{(1-P_2)(1-Tp_a)}{\eta_0} - \frac{Tp_a}{\psi} - \frac{S_e P_T}{\rho}, \\
V_{53} &= \frac{(1-Tp_a)P_1}{\varphi}, \\
V_{55} &= -\frac{r}{\xi} - \frac{(1-r)(1-d_1)}{\eta_1} - \frac{(1-r)d_1}{\zeta_1} - \frac{S_e P_T}{\rho} \\
V_{64} &= \frac{(1-Tp_a)P_2}{\varphi}, \\
V_{66} &= -\frac{r}{\xi} - \frac{(1-r)(1-d_2)}{\eta_2} - \frac{(1-r)d_2}{\zeta_1} - \frac{S_e P_T}{\rho}, \\
V_{71} &= \frac{Tp_e}{\psi} + \frac{S_e P_T}{\rho}, \\
V_{73} &= \frac{Tp_a}{\psi} + \frac{S_e P_T}{\rho}, \\
V_{75} &= \frac{r}{\xi} + \frac{S_e P_T}{\rho}, \\
V_{77} &= -\frac{1-d_1}{\eta_1} - \frac{d_1}{\zeta_1}, \\
V_{82} &= \frac{Tp_e}{\psi} + \frac{S_e P_T}{\rho}, \\
V_{84} &= \frac{Tp_a}{\psi} + \frac{S_e P_T}{\rho}, \\
V_{86} &= \frac{r}{\xi} + \frac{S_e P_T}{\rho}, \\
V_{88} &= -\frac{1-d_2}{\eta_2} - \frac{d_2}{\zeta_2}, \\
V_{ij} &= 0, \text{ for all other } i,j.
\end{aligned} \tag{A.7}$$

We then have that R_0 is the largest eigenvalue of the next generation matrix $-FV^{-1}$. Given the complexity of these expressions they are evaluated during simulations in MATLAB.

Declaration of competing interest

The authors declare no conflicts of interest regarding this paper.

Acknowledgments

The authors would like to thank Dr. Angela DiBenedetto and the rest of the COVID-19 Science Advisory Committee at Villanova University for their valuable discussions on the spread of COVID-19. In addition we thank Dr. Elise Pasles and Dr. Jesse Frey for their input on the statistical aspects of sampling a population for surveillance testing. We would also like to thank the reviewers and editor for their comments. This work received funding from Villanova University's Falvey Memorial Library Scholarship Open Access Reserve (SOAR) Fund.

References

Bahl, R., Eikmeier, N., Fraser, A., Junge, M., Keesing, F., Nakahata, K., & Wang, L. Z. (2020). *Modeling COVID-19 spread in small colleges*.

Bansal, S., Grenfell, B. T., & Meyers, L. A. (2007). When individual behaviour matters: Homogeneous and network models in epidemiology. *Journal of The Royal Society Interface, 4*, 879–891.

Bisoffi, Z., Pomari, E., Deiana, M., Piubelli, C., Ronzoni, N., Beltrame, A., Bertoli, G., Riccardi, N., Perandin, F., Formenti, F., Gobbi, F., Buonfrate, D., & Silva, R. (2020). Sensitivity, specificity and predictive values of molecular and serological tests for COVID-19: A longitudinal study in emergency room. *Diagnostics, 10*, 669.

Bi, Q., Wu, Y., Mei, S., Ye, C., Zou, X., Zhang, Z., Liu, X., Wei, L., Truelove, S. A., Zhang, T., Gao, W., Cheng, C., Tang, X., Wu, X., Wu, Y., Sun, B., Huang, S., Sun, Y., Zhang, J., … Feng, T. (2020). Epidemiology and transmission of COVID-19 in 391 cases and 1286 of their close contacts in Shenzhen, China: A retrospective cohort study. *The Lancet Infectious Diseases, 20*, 911–919.

Byambasuren, O., Cardona, M., Bell, K., Clark, J., McLaws, M.-L., & Glasziou, P. (2020). Estimating the extent of asymptomatic COVID-19 and its potential for community transmission: Systematic review and meta-analysis. *Official Journal of the Association of Medical Microbiology and Infectious Disease Canada, 5*, 223–234.

Cashore, J. M., Duan, N., Janmohamed, A., Wan, J., Zhang, Y., Henderson, S., Shmoys, D., & Frazier, P. (2020a). *Addendum: COVID-19 mathematical modeling for cornell's fall semester.* https://cpb-_us-_w2.wpmucdn.com/sites.coecis.cornell.edu/dist/3/341/files/2020/08/Addendum_COVID_19_Modeling_Jul_17-_VD.pdf Accessed: 2021-03-20.

Cashore, J. M., Duan, N., Janmohamed, A., Wan, J., Zhang, Y., Henderson, S., Shmoys, D., & Frazier, P. (2020b). *COVID-19 mathematical modeling for cornell's fall semester.* https://cpb-_us-_w2.wpmucdn.com/sites.coecis.cornell.edu/dist/3/341/files/2020/10/COVID_19_Modeling_Jun15-_VD.pdf Accessed: 2021-03-20.

Centers for Disease Control and Prevention. (2020). *Covid-19 pandemic planning scenarios.* https://www.cdc.gov/coronavirus/2019-_ncov/hcp/planning-_scenarios.html Accessed: 2021-02-11.

Centers for Disease Control and Prevention. (2021a). *Demographic trends of COVID-19 cases and deaths in the US reported to CDC.* https://covid.cdc.gov/covid-_data-_tracker/#demographics Accessed: 2021-02-11.

Centers for Disease Control and Prevention. (2021b). *Operational considerations for adapting a contact tracing program to respond to the COVID-19 pandemic in non-US settings.* https://www.cdc.gov/coronavirus/2019-_ncov/global-_covid-_19/operational-_considerations-_contact-_tracing.html Accessed: 2021-03-26.

Davies, N. G., Abbott, S., Barnard, R. C., Jarvis, C. I., Kucharski, A. J., Munday, J. D., Pearson, C. A., Russell, T. W., Tully, D. C., Washburne, A. D., Wenseleers, T., Gimma, A., Waites, W., Wong, K. L., van Zandvoort, K., Silverman, J. D., CMMID COVID-19 Working Group, Diaz-Ordaz, K., Keogh, R., … Edmunds, W. J. (2021). Estimated transmissibility and impact of SARS-CoV-2 lineage B. 1.1. 7 in England. *Science*, 372.

D'Errico, J. (2012). *fminsearchbnd, fminsearchcon.* https://www.mathworks.com/matlabcentral/fileexchange/8277-_fminsearchbnd-_fminsearchcon Accessed: 2021-02-01.

Diekmann, O., Heesterbeek, J., & Roberts, M. G. (2010). The construction of next-generation matrices for compartmental epidemic models. *Journal of The Royal Society Interface, 7*, 873–885.

Feng, Z. (2007). Final and peak epidemic sizes for SEIR models with quarantine and isolation. *Mathematical Biosciences and Engineering, 4*, 675.

Gressman, P. T., & Peck, J. R. (2020). Simulating COVID-19 in a university environment. *Mathematical Biosciences, 328*, 108436.

He, D., Zhao, S., Lin, Q., Zhuang, Z., Cao, P., Wang, M. H., & Yang, L. (2020). The relative transmissibility of asymptomatic COVID-19 infections among close contacts. *International Journal of Infectious Diseases, 94*, 145–147.

Jing, Q.-L., Liu, M.-J., Zhang, Z.-B., Fang, L.-Q., Yuan, J., Zhang, A.-R., Dean, N. E., Luo, L., Ma, M.-M., Longini, I., Kenah, E., Lu, Y., Ma, Y., Jalali, N., Yang, Z.-C., & Yang, Y. (2020). Household secondary attack rate of COVID-19 and associated determinants in Guangzhou, China: A retrospective cohort study. *The Lancet Infectious Diseases, 20*, 1141–1150.

Lauer, S. A., Grantz, K. H., Bi, Q., Jones, F. K., Zheng, Q., Meredith, H. R., Azman, A. S., Reich, N. G., & Lessler, J. (2020). The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: Estimation and application. *Annals of Internal Medicine, 172*, 577–582.

Liang, M., Gao, L., Cheng, C., Zhou, Q., Uy, J. P., Heiner, K., & Sun, C. (2020). Efficacy of face mask in preventing respiratory virus transmission: A systematic review and meta-analysis. *Travel Medicine and Infectious Disease, 36*, 101751.

Lopman, B., Liu, C., Le Guillou, A., Handel, A., Lash, T. L., Isakov, A., & Jenness, S. (2020). *A model of COVID-19 transmission and control on university campuses.* medRxiv:2020.06.23.

Losina, E., Leifer, V., Millham, L., Panella, C., Hyle, E. P., Mohareb, A. M., Neilan, A. M., Ciaranello, A. L., Kazemian, P., & Freedberg, K. A. (2020). College campuses and COVID-19 mitigation: Clinical and economic value. *Annals of Internal Medicine*, (1–13).

Luo, L., Liu, D., Liao, X., Wu, X., Jing, Q., Zheng, J., Liu, F., Yang, S., Bi, H., Li, Z., Liu, J., Song, W., Zhu, W., Wang, Z., Zhang, X., Huang, Q., Chen, P., Liu, H., Cheng, X., … Mao, C. (2020). Contact settings and risk for transmission in 3410 close contacts of patients with COVID-19 in Guangzhou, China: A prospective cohort study. *Annals of Internal Medicine, 173*, 879–887.

McChesney, J., & Bichsel, J. (2020). *The aging of tenure-track faculty in higher education: Implications for succession and diversity.* ERIC. ED603016 https://files.eric.ed.gov/fulltext/ED603016.pdf.

Mizumoto, K., Kagaya, K., Zarebski, A., & Chowell, G. (2020). Estimating the asymptomatic proportion of coronavirus disease 2019 (COVID-19) cases on board the Diamond Princess cruise ship, yokohama, Japan, 2020. *Euro Surveillance, 25*, 2000180.

Paltiel, A. D., Zheng, A., & Walensky, R. P. (2020). Assessment of SARS-CoV-2 screening strategies to permit the safe reopening of college campuses in the United States. *JAMA network open, 3*, Article e2016818.

Penn State University Budget Office. (2017). Faculty distribution by age: Fall 2017. https://budget.psu.edu/Factbook/hrdynamic/FacultyDistributionByAgePSULaw.aspx?ReportCode=Age&YearCode=2017humors&FBPlusIndc=N.%20Accessed:%202021-02-11.

Poletti, P., Tirani, M., Cereda, D., Trentini, F., Guzzetta, G., Sabatino, G., Marziano, V., Castrofino, A., Grosso, F., Castillo, G. D., Piccarreta, R., Force, A. L. C.-. T., Andreassi, A., Melegaro, A., Gramegna, M., Ajelli, M., & Merler, S. (2020). *Probability of symptoms and critical disease after SARS-CoV-2 infection.*

Reese, H., Iuliano, A. D., Patel, N. N., Garg, S., Kim, L., Silk, B. J., Hall, A. J., Fry, A., & Reed, C. (2020). *Estimated incidence of COVID-19 illness and hospitalization—United States, February–September, 2020* (p. ciaa1780). Clinical infectious diseases: an official publication of the Infectious Diseases Society of America.

Spencer, K. D., Chung, C. L., Stargel, A., Shultz, A., Thorpe, P. G., Carter, M. W., Taylor, M. M., McFarlane, M., Rose, D., Honein, M. A., & Walke, H. (2021). COVID-19 case investigation and contact tracing efforts from health departments—United States, June 25–July 24, 2020. *Morbidity and Mortality Weekly Report, 70*, 83.

The Chronicle of Higher Education and Davidson College's College Crisis Initiative. (2020). *Here's our list of colleges' reopening models*. https://www.chronicle.com/article/heres-_a-_list-_of-_colleges-_plans-_for-_reopening-_in-_the-_fall/?bc_nonce=m8iw713xyabkthloqwve&cid=reg_wall_signup Accessed: 2020-10-01.

Tindale, L. C., Stockdale, J. E., Coombe, M., Garlock, E. S., Lau, W. Y. V., Saraswat, M., Zhang, L., Chen, D., Wallinga, J., & Colijn, C. (2020). Evidence for transmission of COVID-19 prior to symptom onset. *Elife, 9*, Article e57149.

Villanova University. (2020). Fall 2020 COVID-19 dashboard (Aug 17–Dec 6). https://app.powerbi.com/view?r=eyJrIjoiYmUzZjJkMzQtNmQ3Mi00MTBhLWFkNWYtYjQwN2NhZTc0ZDZhIiwidCI6Ijc2NWE4ZGU1LWNmOTQtNDRmMC05Y2FmLWFlNWJmOGNmYTM2NiIsImMiOjF9&pageName=ReportSection.%20Accessed:%202021-03-03.

Voinsky, I., Baristaite, G., & Gurwitz, D. (2020). Effects of age and sex on recovery from COVID-19: Analysis of 5769 Israeli patients. *Journal of Infection, 81*, e102–e103.

Wu, J., Liang, B., Chen, C., Wang, H., Fang, Y., Shen, S., Yang, X., Wang, B., Chen, L., Chen, Q., Wu, Y., Liu, J., Yang, X., Li, W., Zhu, B., Zhou, W., Wang, H., Li, S., Lu, S., … Zheng, X. (2021). SARS-CoV-2 infection induces sustained humoral immune responses in convalescent patients following symptomatic COVID-19. *Nature Communications, 12*, 1813.