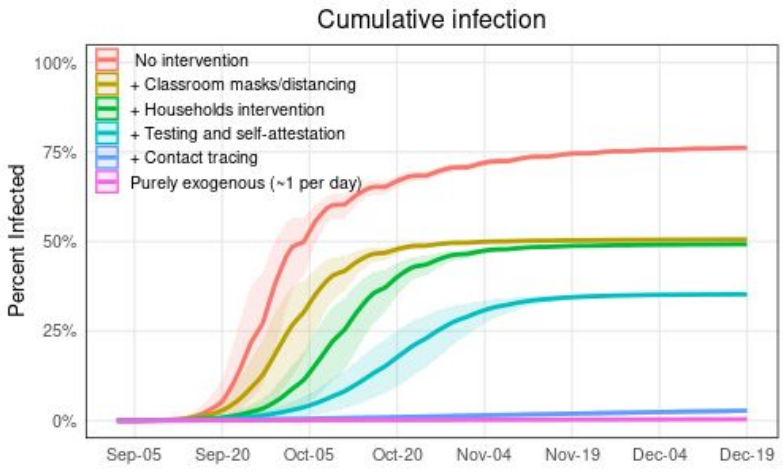
BU COVID-19 Modeling: Insight on the Relative Impact of Proposed Interventions for a Large University in an Urban Environment

Katia Bulekova[1], Brian Gregor[1], Eric D. Kolaczyk[2,3], Wenrui Li[3], and Laura White[4]
August 3, 2020

Overview

As with most institutions of higher learning world-wide, Boston University's goal of reopening at some appropriate level in the Fall of 2020 is impacted by many factors. Throughout the early summer of 2020, we conducted a detailed modeling exercise to provide insight into the expected relative efficacy of potential interventions taken around a Fall 2020 BU re-opening. Our approach was to do so in a manner informed as much as reasonably possible by BU data, with a focus primarily on the dynamics around (i) classroom instruction, and (ii) residential housing. In turn, we are making available our software and model parameters, to facilitate other institutions conducting similar analyses.

We used agent-based network modeling to simulate outcomes for the student and instructional portion of the BU population this fall during the ongoing COVID-19 pandemic. A susceptible - exposed - infected - removed (SEIR) model dictates the progression of disease for individuals. Networks determine interactions between individuals that can cause infection. It is assumed the population is "screened and cleaned" prior to the start of the semester. Infections are introduced from exogenous sources (spontaneous infection, from the BU perspective). Testing, quarantine and isolation strategies are implemented. The models are stochastic -- multiple trials are run, with average trends reported and uncertainty quantified.



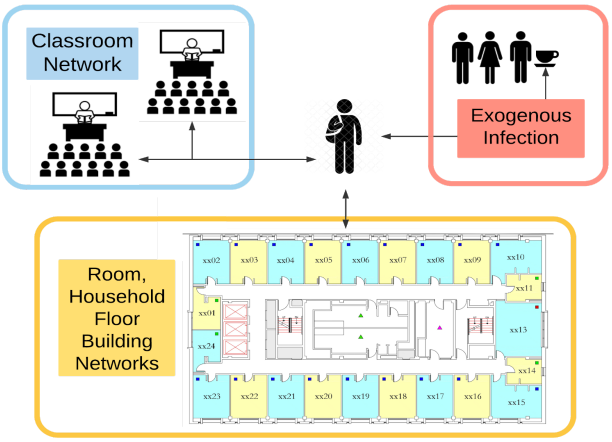
The figure to the left illustrates the type of insight that can be gained from our modeling approach. For an urban institution, rather than zero infections, a more realistic goal is to *keep the BU rate of infections proportional to a constant background rate* (or linear cumulative rate) in the Boston area as a whole. The results of our modeling support that it is a feasible goal that BU keep Covid-19 infections near the ideal "linear regime" using the full

[1] Research Computing Services, IS&T, Boston University
[2] Hariri Institute for Computing (Director), Boston University
[3] Department of Mathematics & Statistics, Boston University
[4] Department of Biostatistics, Boston University

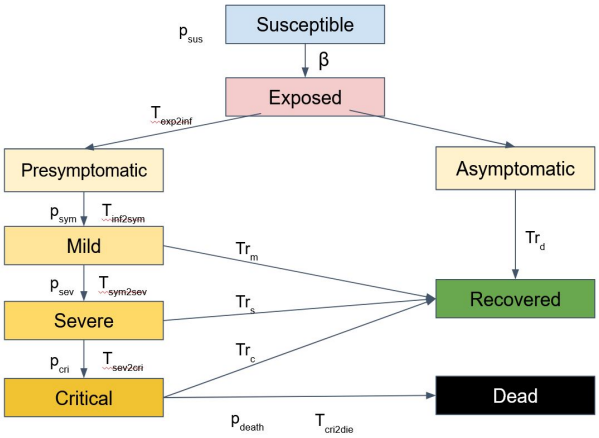
combination of interventions and measures, starting with a screened/cleaned population, assuming sufficiently low and constant background and adequate compliance by the population. Additional simulations (not shown) suggest that there is some reasonable robustness to background infection levels in the greater Boston area. However, in general, compliance around interventions is critical.

Modeling Details

Our models are network-based, with nodes representing individuals in the BU population and links representing non-trivial contacts that may admit transmission. Our choice of networks and their construction reflects what are felt to be primary drivers of disease transmission on campus, and are informed in turn by available institutional data. Classroom contacts reflect transmission through proximity, where classrooms are treated as square units of varying size, with students (and instructor) distributed uniformly throughout, and contacts captured through a grid of 10 neighbors each (i.e., 8 points of the compass, plus 2). De-densification is captured by reducing the assumed transmission rate. Housing contacts are captured at the level of roommates, households, floors, and buildings, where households are approximated by units of at most 12 in proximity to a common bathroom. These contacts are assumed to have differential transmission rates, with roommates highest, then households, and then floors.



Our model without interventions allows for transmission along any of the above contact links. A baseline transmission rate is modulated as a function of age. Upon exposure, individuals may transition to a symptomatic state, or remain asymptomatic. Those showing symptoms (termed 'presymptomatic' prior to this) may proceed from mild to severe to critical stages, with death a possibility for some portion of those reaching a critical stage. Distributions for the various waiting times between states, and the relevant parameters (i.e., transition probabilities and distributional parameters) have been specified in a manner informed as best possible by the large and still-evolving literature around COVID-19.



In the context of this model, interventions are introduced through some combination of deletion of contact links or individuals, and/or modulation of the baseline infection rate. Informed by our understanding of intended and/or proposed interventions, we developed a sequence of models with the following characteristics: (i) in-person classroom attendance fractured into cohorts dictated by social distancing (e.g., MWF classes attending in thirds, TR classes attending in halves, etc.) and assuming masks worn by all occupants; (ii) shared bathroom usage in residential housing extended no further than households of at most 12 in size; (iii) self-attestation, with a one-day lag for those displaying symptoms; (iv) testing of everyone on a regular basis, at currently understood RT-PCR false positive rates and one-day reporting lag, with isolation for those testing positive; and (v) contact tracing, with 14-day quarantine for those exposed to an infected individual (who, in turn, are tested and isolated if themselves found to be infected).

Software Details

Simulations were run using the workflow described below. The primary software tool used in our modeling was the python-based Covasim package[5], suitably modified to incorporate data on the BU population and proposed intervention strategies. Input to and output from this package was managed using scripts written in the R statistical programming language. All of these materials, as well as illustrative input and output data sets, have been made available to the open source community in a github repository.[6]

The workflow for working with our simulation code is as follows. Once the code has been checked out from the github repository the provided sample data can be used to run the simulation end to end. The required input data consists of comma-separated-value (CSV) files that describe the demographics of the population, course registrations and schedules, and information on student resident housing. A set of R scripts are provided that convert the input data into networks and then save them as GraphML formatted files. The simulations are then run using Python and our *bu_covid* library in conjunction with the Covasim python package. Multiple simulations can be run in parallel to generate a range of probable outcomes. We use separate Python scripts to set up and test different intervention and testing strategies. The Python code produces a set of CSV files containing all of the generated data. Finally, an additional set of R scripts are used to plot the results. On a modern computer (Intel Xeon Gold 6132) it takes approximately 15 seconds to run one simulation for a population of approximately 30,000 people over a 109 day semester. Collecting the results for 1000 such simulations requires about 200 GB of RAM.

---

[5] The original Covasim package is developed and maintained by the Institute for Disease Modeling, which specializes in mathematical modelling of infectious disease, as part of the Bill & Melinda Gates Foundation's Global Health Division.

[6] https://github.com/bu-rcs/BU-COVID

Appendix: Representative Parameterization

Our models require that a variety of distributions and parameters be specified. We list those here, along with our baseline setting for the initial model and variants we have considered in some of our sensitivity analyses. For many of these choices, we use the Covasim default settings, which are well documented and referenced.[7,8] For the others, we supply references supporting our choice when possible. Note that all times below are in days.

Table A. Disease specific parameters. These parameters will not vary for different intervention scenarios, but might be varied as more is learned about COVID-19.

Parameter	Definition	Baseline	Variants	Sources
s	Length of time after exposure before a person is infectious	log-normal μ = 4.6 σ = 4.8		Covasim default
i	Length of time after person is infectious before person has symptoms	log-normal μ = 1.0 σ = 0.9		Covasim default
r_d	Recovery time for asymptomatic cases	log-normal μ = 8 σ = 2		Covasim default
r_m	Recovery time for mild cases	log-normal μ = 8 σ = 2		Covasim default
r_s	Recovery time for severe cases	log-normal μ =14 σ = 2.4		Covasim default
r_c	Recovery time for critical cases	log-normal μ = 14 σ = 2.4		Covasim default
p_{sus}	Odds ratio of susceptibility	age-varying (0.34 – 1.47)		Covasim default

[7] "InstituteforDiseaseModeling/covasim: COVID-19 ... - GitHub." https://github.com/InstituteforDiseaseModeling/covasim. Accessed 28 Jul. 2020.
[8] "an agent-based model of COVID-19 dynamics and ... - medRxiv." 15 May. 2020, https://www.medrxiv.org/content/10.1101/2020.05.10.20097469v1.full.pdf. Accessed 28 Jul. 2020.

Parameter	Definition	Baseline	Variants	Sources
p_{sym}	Prob. of devel. symptoms	age-varying (0.5 – 0.9)		Covasim default
p_{sev}	Prob. of severe symptoms	age-varying (0.0005 – 0.2457)		Covasim default
p_{cri}	Prob. of critical symptoms	age-varying (0.00003 – 0.1742)		Covasim default
p_{death}	Prob. of death	age-varying (0.00002 – 0.093)		Covasim default
Asymptomatic factor	Multiplier of β for asymptomatic cases	0.5		CDC planning scenarios[9] (Note: this value was selected based on version 1.0 of CDC planning scenarios; new numbers indicate this might actually be lower)
sym2sev	Duration from symptomatic to severe symptoms	log-normal μ = 6.6 σ = 4.9		Covasim default
sev2crit	Duration from severe symptoms to critical symptoms	log-normal μ =3.0 σ = 7.4		Covasim default
crit2die	Duration from critical symptoms to death	log-normal μ = 6.2 σ = 1.7		Covasim default
ili_prev	Prevalence of influenza-like-illness symptoms in the population	N/A		Not included in the current model.

[9] "COVID-19 Pandemic Planning Scenarios | CDC." 10 Jul. 2020, https://www.cdc.gov/coronavirus/2019-ncov/hcp/planning-scenarios.html. Accessed 28 Jul. 2020.

Table B. Graph specific parameters. These parameters might vary from graph to graph. For instance we might set these parameters differently for housing versus classroom networks. We calculate transmission probabilities, β, using a transmission calculator based on the Wells-Riley equation modified to calculate transmission probabilities for COVID-19 in classroom settings.[10]

Parameter	Definition	Baseline	Variants	Sources
Classroom transmission ($\beta_{classroom}$)	Probability of transmission in a classroom	0.02		Transmission calculator (500 sq ft room, 25 students per room, no masks)
Housing (roommates) ($\beta_{roommate}$)	Probability of transmission between roommates	0.7		Transmission calculator (600 minutes of exposure in 144 sq ft room; no masks)
Housing (households /share bathroom) ($\beta_{household}$)	Probability of transmission between individuals in household (e.g. sharing bathroom)	0.08		Transmission calculator (60 minutes of exposure in 144 sq ft room; no masks)
Housing (same floor) (β_{floor})	Probability of transmission between individuals on the same floor	0.03		Transmission calculator (30 minutes of exposure in 144 sq ft room; no masks)
Viral load	Multiplier of β to account for varying viral load through infectious period	log-normal μ = 0.84 σ = 0.3		Covasim default
n_imports	Average number of daily exogenous infections	Poisson (1)	Poisson(3)	Based on Suffolk county daily case incidence in June/July. Variant is based on Suffolk county case incidence in March/April. In June/July there were approximately 7/100,000 cases daily. In March/April this was approximately 18.7/100,000 cases daily.[11] Our BU population is approximately 30,000, so the daily exogenous rate of infections is approximately 5-6 in March/April and 2 in June/July. Our baseline scenario represents the scenario where things continue to improve and the variant represents things getting worse once classes resume, but not reaching the levels observed in the spring.

[10] "COVID-19 Airborne Transmission Tool Available | CIRES." 25 Jun. 2020, https://cires.colorado.edu/news/covid-19-airborne-transmission-tool-available. Accessed 28 Jul. 2020.
[11] "COVID-19 pandemic in Massachusetts - Wikipedia." https://en.wikipedia.org/wiki/COVID-19_pandemic_in_Massachusetts. Accessed 28 Jul. 2020.

Table C. Intervention parameters. These are parameters that can vary for different testing or distancing scenarios.

Parameter	Definition	Baseline	Variants	Sources
Classroom intervention	Reduction of classroom β due to de-densification and mask usage	0.5		Transmission calculator (10 students in 500 sq ft classroom, 100% mask usage with 50% efficacy)
Quarantine factor	Multiplier of β to reflect impact on quarantined contacts	0.0 for all networks		BU protocols eliminate potential transmission since students are removed from housing and classes
Isolation factor	Multiplier of β for isolated diagnosed cases	0.0 for all networks		BU protocols eliminate potential transmission since students are removed from housing and classes
Quarantine period	Number of days to quarantine	14 days		
test_symp_prob	Probability of self-attestation of symptoms of infected, unquarantined individual	0.75	0.65, 0.85	Hypothetical information
test_asymp_prob	Probability of self attestation of symptoms for asymptomatic person	0		ditto
test_symp_quar_prob	Probability of self attestation of symptoms in a symptomatic quarantined person	0.75	0.65, 0.85	ditto

Parameter	Definition	Baseline	Variants	Sources
test_asymp_quar_prob	Probability of self attestation of symptoms in an asymptomatic quarantined person	0		ditto
test_sensitivity	Probability of a true positive	0.9	0.8	Woloshin et al,[12] Padhye[13]
test_specificity	Probability of a true negative	1		Padhye, Kucirka et al,[14] FDA data[15]
test_loss_prob	Probability that people will not receive their results	0		Covasim default; consistent with BU protocols
test_delay	Waiting time for testing results	24hrs	48 hours, 72 hours	BU specific protocols
trace_prob	Probability that a contact can be traced	1 for roommates 0 for all others		BU protocol specifies that all roommates of infected individuals will be quarantined. Other types of contacts are handled using trace_sens and trace_spec parameters.
trace_sens	Sensitivity of contact tracing (i.e. probability infected contact is identified from contact tracing)	0.3		Best guess. Contacts from classrooms and housing that are infected are assumed to be identified with this probability.

[12] "False Negative Tests for SARS-CoV-2 Infection — Challenges ...." https://www.nejm.org/doi/full/10.1056/NEJMp2015897. Accessed 28 Jul. 2020.
[13] "Reconstructed diagnostic sensitivity and specificity ... - medRxiv." 29 Apr. 2020, https://www.medrxiv.org/content/10.1101/2020.04.24.20078949v1. Accessed 28 Jul. 2020.
[14] "Variation in False-Negative Rate of Reverse Transcriptase ...." 13 May. 2020, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7240870/. Accessed 28 Jul. 2020.
[15] "Quest SARS-CoV-2 rRT-PCR test - FDA." https://www.fda.gov/media/136231/download. Accessed 28 Jul. 2020.

Parameter	Definition	Baseline	Variants	Sources
trace_spec	Specificity of contact tracing (i.e. probability uninfected contact is not identified from contact tracing)	0.9		Best guess. Contacts from classrooms and housing that are not infected are assumed to not be identified with this probability.
trace_time	Length of time taken to identify and notify contacts	24hrs	48 hours (2 days)	BU student health services