Advances in Biology Laboratory Education
Publication of the Association for Biology Laboratory Education
Volume 42, Article 13, 2022

Simulating Parasite Transmission and Epidemics Using Contact Networks

Karena H. Nguyen[1], Evelyn C. Rynkiewicz[2], and David J. Civitello[1]

[1]Emory University, Department of Biology, 1510 Clifton Road NE, Atlanta, GA 30322, USA
[2]Fashion Institute of Technology, Department of Science and Mathematics, 227 West 27th Street, New York City, NY 10001, USA
(**karena.ha.nguyen@emory.edu; evelyn_rynkiewicz@fitnyc.edu**; **david.james.civitello@emory.edu**)

Undergraduate students have an intuitive understanding that contact between an infectious individual and a susceptible individual can cause parasite transmission or infection. While undergraduate students can describe factors that facilitate parasite transmission, such as differences in immunity or contact patterns among hosts, it can be challenging to teach the mathematical representations that connect these processes to population-level outcomes that are of interest to public and environmental health experts. Here, we present an inquiry-based laboratory exercise where students construct a contact network based on shared courses, extracurricular activities, and residential housing. We first contrast this approach with the "classic" compartment or SIR-type (Susceptible-Infected-Recovered) model that ignores variation in contact patterns. Students then simulate epidemics through their self-generated network using dice to illustrate the role of demographic stochasticity (i.e., luck or chance) in the progression of epidemic clusters. Once familiar with infection spread in the network, students make predictions about the size of epidemics that begin at different nodes of the network with and without public health interventions (e.g., masking, social distancing). These exercises have been adapted for non-major and upper-level biology undergraduates and have been implemented successfully in-person and in remote teaching modalities.

Keywords: disease ecology, epidemiology, mathematical modeling, contact network, inquiry-based learning

Introduction

Parasite transmission, defined as the movement of an infectious agent from one host to another, is a key aspect of epidemics. Transmission can occur, for instance, via dermal contact, inhalation, consumption, vectors, or by sexual activity. These different modes of transmission are not difficult for students to describe. However, many students struggle with the concept that parasite transmission can be quantified mathematically and that the rates or probabilities at which individuals transition between susceptible, infected, and recovered categories can be defined by different mathematical representations. Additionally, it can be difficult for students to compare the strengths and weaknesses of different frameworks and how they relate to biological realities (Bansal et al. 2007, Bansal et al. 2010).

The now classic "SIR" (Susceptible-Infected-Recovered) models were developed from early mathematical approaches to modeling infectious disease outbreaks. SIR models use differential equations to track how hosts move from one category to another as a parasite spreads through a population and hosts clear infection or develop immunity (Kermack and McKendrick 1927). SIR models are invaluable tools in public health, wildlife, and livestock disease modeling. Additionally, they can be used to predict the magnitude of epidemics based on the traits of individual hosts and parasites, as was observed first-hand by the general public during the Covid-19 pandemic. However, one of the major drawbacks of the classic SIR approach is that within each category, SIR models treat all individuals as identical. The assumption of homogeneity ignores important sources of variation tied to the contact behavior and immune status of individual hosts that could drive the dynamics of infectious diseases in populations (e.g., superspreaders). Disease researchers are increasingly turning to contact network models of outbreaks to incorporate these aspects of variability among hosts and to gain novel insights into disease dynamics.

To introduce students to network models of parasite transmission and control, we developed an activity where students are asked to evaluate their connectedness to each other based on shared courses, extracurricular activities, and residential housing. Based on their connections, students generate an adjacency matrix, which is a representation of links or “edges” (direct links between individuals that could enable transmission) among “nodes” (students/hosts), and use these data to produce a contact network for their class. For the next activity, students choose different individuals (i.e., nodes) within the network to simulate the start of epidemics and visualize how chance impacts the spread of an infectious agent. Students then predict how the connectedness of the individual initiating the epidemic in the network impacts the likelihood of infection spreading through the network. Lastly, students are asked to hypothesize how different public intervention scenarios, such as masking or social distancing, would change the contact network and influence epidemics.

This activity can be adapted for a variety of settings. For remote courses or in-person courses with low connectedness among students, we provide mock student profiles that mimic a moderately connected, on-campus environment. For non-majors, this activity has been adapted to ask more qualitative questions about the contact network and the potential impact of interventions. For upper-level biology majors, we have scaffolded the activity to ask more quantitative questions and include programming steps that allow students to manipulate specific parameters in the model.

Student Outline

Objectives

Build a contact network based on personal or pre-assigned courses, extracurricular activities, and residential housing
Simulate the direct transmission of a parasite in a static contact network
Hypothesize how network position (i.e., connectedness) influences parasite spread
Predict how public health interventions influence network connections and epidemic size

Introduction

Parasites, used here to include all infectious agents (e.g., viruses, bacteria, and eukaryotes), are a critical concern for public and environmental health. Parasites exert powerful effects on wildlife, crops, and livestock, thereby altering the composition of natural communities, the function of natural and managed ecosystems, and the productivity and sustainability of agriculture. Thus, understanding how parasites jump from one species of host to another, spread within host populations, and cause harm to host and non-host populations are critical objectives for human epidemiologists, veterinarians, disease ecologists, microbiologists, parasitologists, geneticists, and scientists from many other fields.

Parasite transmission is a fundamental process in epidemiology and is defined as the successful colonization of a new, susceptible host by parasites produced by a currently infected host. For a parasite, establishing in a new host is analogous to reproduction for a nonparasitic organism because it increases the parasite population size. Transmission is also critical for parasite persistence. Without transmitting to new hosts, parasite populations simply go extinct if their current host defeats their infection or dies.

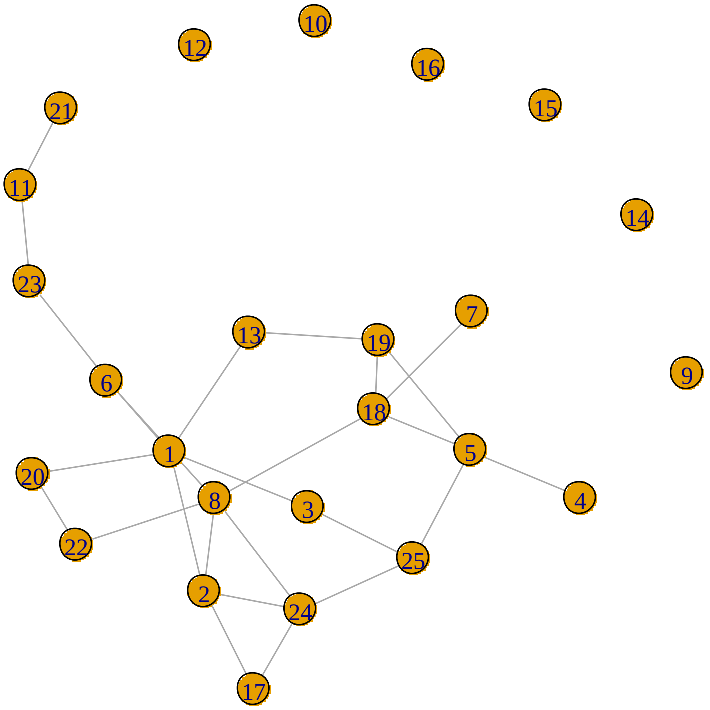
Classic epidemiological theory is based on differential equation models borrowed from physics and chemistry that treat individuals as identical particles that perfectly follow the population-averaged rules of contact, transmission, recovery, etc. in an infinitely large environment (Kermack and McKendrick 1927). These models, often called “SIR” models in reference to their categories of individuals – Susceptible, Infected, Recovered – form the foundation of our understanding of epidemics because (1) they are relatively easy to formulate, understand, and communicate, and (2) they can often be “solved” with (relatively) simple algebra and calculus.

Despite their utility, a drawback of the classic SIR approach is that it ignores variation. Individual variation can change epidemic dynamics in powerful ways. Biologists typically focus on two types of variation: consistent differences among individuals (called “demographic heterogeneity”) and differences in event-to-event outcomes among individuals even when we expect that they are similar (called “demographic stochasticity”).

In this activity, you are going to examine a network epidemic model that incorporates heterogeneity in contacts. That is, some hosts contact many others while some hosts engage in few contacts. We will assume that network is static, meaning that that the structure of the contact network does not change on the time scale of the epidemic. You will then start epidemics at randomly selected nodes in the network, which incorporates stochasticity in transmission and recovery into the model. Transmission will depend on a probability that a contact between an infected host and an uninfected one causes a new infection. Similarly, individuals will vary in their duration of infectiousness.

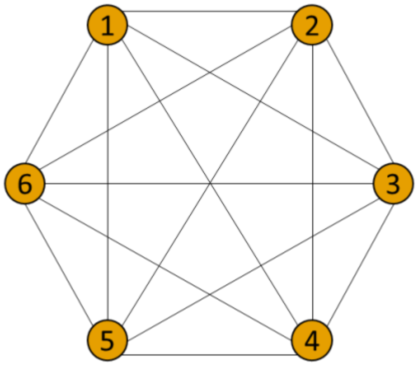
What are contact networks?

Contact networks are one way to construct and visualize the transmission dynamics we just described. Networks are graphical tools that display individual hosts (“**nodes**”) as circles and the contacts between hosts as lines (“**edges**”) that connect them. The **degree** of a node describes how many edges it has (i.e., the degree is equal to the number of other nodes a focal node is connected to). This tool can help visualize what nodes in the network are the most connected to others and subsequently, which nodes have the highest probability of becoming **superspreaders** (i.e., individuals who infect a greater than average number of people). An example network can be seen below:



This network has 25 hosts ("**nodes**") and illustrates that there is variation in the connections between nodes. For example, clusters of connections occur between nodes, but there are also nodes that have no connection to any others in the network. Thus, the potential for the parasite to spread in the network will be related to the contact structure (i.e., how many edges a node has) and where that node is in the network when the parasite is introduced.

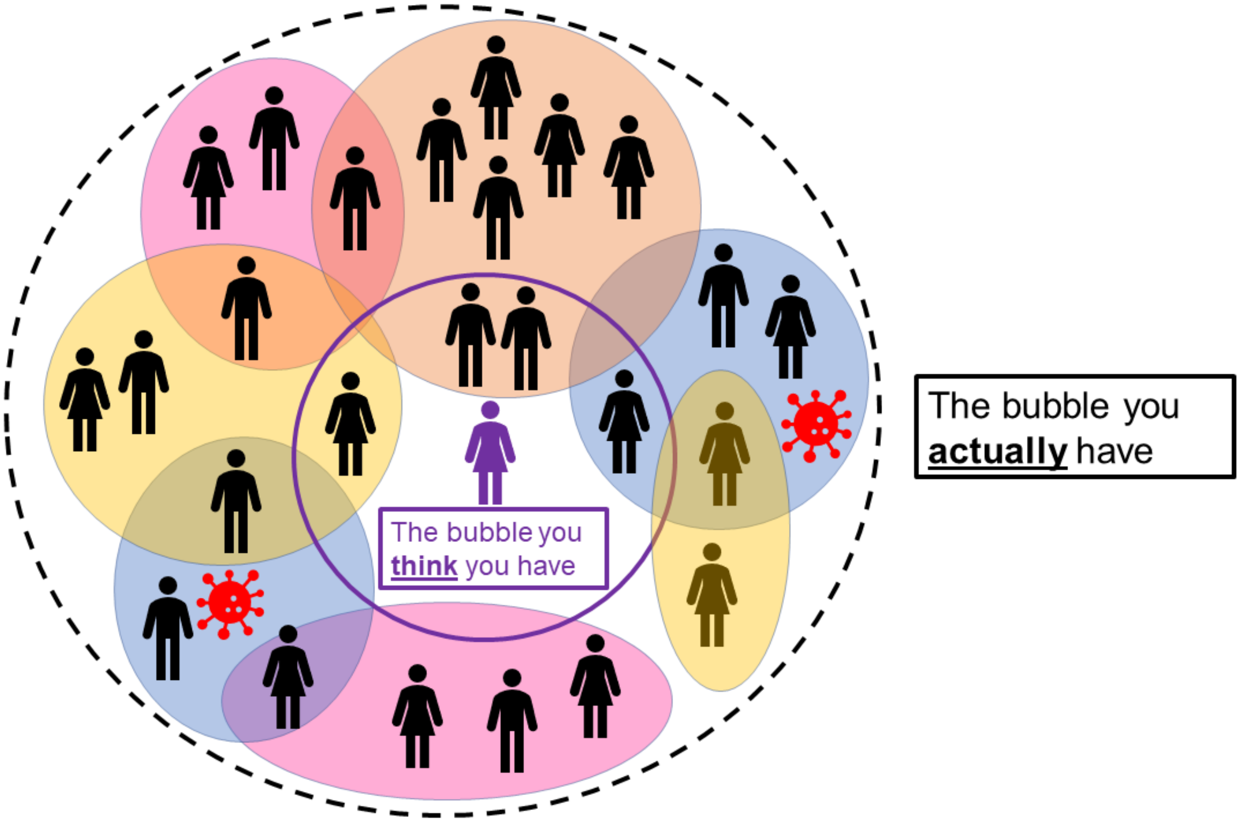
We can also visualize how the previous contact network differs from the classic SIR approach, which assumes that all individuals can contact one another and that there are no differences in the number of contacts among hosts. These assumptions would create a fully saturated contact network, which we illustrate below with a 6-person network for visual clarity. Note that under these assumptions, 1) all individuals are directly connected and 2) all individuals have the same connections to other hosts.



Now that we have introduced contact networks, the classic SIR approach, and some of their defining features, **use the figure below to brainstorm when contact networks might provide tools to improve our understanding of disease dynamics:**

- What message is this image conveying?
- How does it relate to the biology of parasite transmission?
- How are these aspects of biology incorporated into contact networks?

In the following activity, you will construct a potential contact network based on personal or pre-assigned shared courses, extracurricular activities, and residential housing with students in the course. You will first record these “links” using Google Sheets in a data format called an **adjacency matrix**, which labels rows and columns with the names of each node (i.e., host) and records the presence or absence of a link between two nodes using a 1 or 0, respectively, in the corresponding cell. The adjacency matrix contains all the information of the contact network in a format that is easy to record, but hard for us to immediately visualize. We will import this information into R, a programming language used for statistical analyses and scientific computing, to visualize these contact networks and simulate epidemics of a directly transmitted parasite. Lastly, you will hypothesize and discuss how public health interventions (e.g., masking, social distancing) might impact disease spread through a contact network.



Methods and Data Collection

Part A: Downloading R and R Studio (optional)

If your instructor has chosen to include the simulation portions of the exercise, you will need to download R and R Studio.

1. R is freely available for download at https://www.r-project.org/.
    a. Follow the directions under “Getting Started” and choose a CRAN mirror that is closest to your physical location.
2. Once you have downloaded R, navigate to https://www.rstudio.com/products/rstudio/ to download R Studio for your desktop.
    a. Follow the appropriate directions to download the correct version for your OS.

   i. One strong suggestion for Windows users: Do not install into C:/Program Files/R. Instead, we recommend, C:/R. Windows' default antivirus settings will interfere with installing R packages if you ignore this advice.

Part B: Building a Contact Network

The first step in building a contact network is to determine how individual hosts are connected. This information is turned into an adjacency matrix, where a 1 signifies that two hosts are connected and a 0 signifies that the hosts are not connected. Note that this representation is binary, meaning that it signifies whether hosts are connected or not. For the purposes of this activity, even if two nodes have multiple connections, we still use a "1" to represent that a connection of any type exists. All connections are assumed to be equal with respect to parasite transmission and we assume that transmission cannot occur directly between two individuals that are not connected.

1. To build an adjacency matrix, you begin by recording the name of each student in the class as a row and column. An example is shown below.

$$Adjacency\ matrix = \begin{array}{cccc} & Jake & Marta & Jen \\ Jake & & & \\ Marta & & & \\ Jen & & & \end{array}$$

2. Next, place a 1 on the diagonal of the matrix (this is somewhat of a data artifact/historical convention and it does not mean anything for a person to "contact themselves").

$$Adjacency\ matrix = \begin{array}{cccc} & Jake & Marta & Jen \\ Jake & 1 & & \\ Marta & & 1 & \\ Jen & & & 1 \end{array}$$

3. For each student pair possible in the matrix, input a 0 if there is no contact between them and a 1 if students share a course, extracurricular activity, or residential housing. For example, Jake and Marta share a course together. Marta plays rugby with Jen. Jake and Jen do not share any connections.

$$Adjacency\ matrix = \begin{array}{cccc} & Jake & Marta & Jen \\ Jake & 1 & 1 & 0 \\ Marta & & 1 & 1 \\ Jen & & & 1 \end{array}$$

Note, this matrix will be symmetric, with the top right mirroring the bottom left (i.e., if you folded the matrix along the major diagonal, which extends from the top left to the bottom right of the matrix). Your final adjacency matrix used in the simulation will look like the one below.

$$Adjacency\ matrix = \begin{array}{cccc} & Jake & Marta & Jen \\ Jake & 1 & 1 & 0 \\ Marta & 1 & 1 & 1 \\ Jen & 0 & 1 & 1 \end{array}$$

4. Go through this process for your class! (Your instructor may break up the class into smaller groups depending on class size.) The instructor will provide a blank template with the names of class attendees along the top row and first column as a Google Sheets file. Work together to build an adjacency matrix for the students attending this class. Discuss your shared courses, clubs/activities, and residential housing buildings, and fill in 1s for shared links and 0s to denote the absence of a link. As a reminder, **do not count sharing this course**, as this would result in connections among everyone in our network. Additionally, even if you have multiple connections with another student, still mark that connection in the adjacency matrix as a “1”. There are matrices that account for different types of connections, but they are not specifically explored here.

5. Once this class-specific matrix is filled in completely, export it from Google Sheets as a “.csv” file named “AdjMatrix.csv” in your preferred working directory for R.

6. Open “R Studio” and the R script provided (named “ContactNetwork.r”).
    a. Load the necessary R packages, network and igraph, by running the following lines
        i. library(network)
        ii. library(igraph)

7. Import the data into R Studio using the following code:
    a. setwd([my_directory]) # replace [mydirectory] with the folder location of AdjMatrix.csv
    b. AM <- read.csv(“AdjMatrix.csv”, header=TRUE, row.names=1, check.names=FALSE)
        i. This line of code will store the matrix as a variable called AM (for *adjacency matrix*).

8. Convert the adjacency matrix into an “edge list”, which is another data format for networks that is needed for the functions we will use to simulate epidemics. Run the following code:
    a. EL = as.edgelist(network(as.matrix(AM), directed=F, matrix.type="adjacency"))

9. Visualize our contact network by plotting the “edge list”. The first line creates the graph object, the second and third lines change some visual characteristics, and the fourth line plots the graph in a circular layout:
    a. gr = graph_from_edgelist(EL, directed=FALSE)
    b. V(gr)$altcolor = "white"’
    c. V(gr)$altcolor[1]<-"white"
    d. plot(gr, layout=layout_in_circle, vortex.color=V(gr)$altcolor)

Answer the following questions after you have constructed your adjacency matrix and contact network:

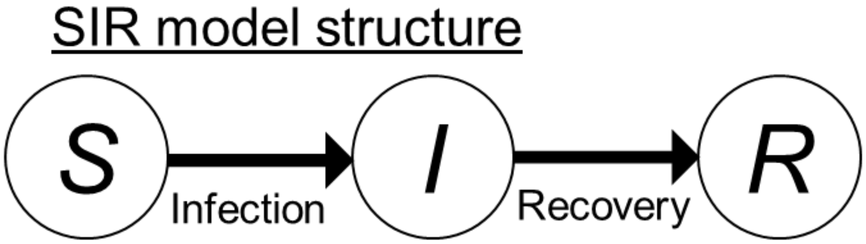
1. Describe the adjacency matrix.
    a. Are there some people with many connections and some with few? Or does everyone have a similar number of contacts?
    b. Are there certain factors that are more common or that lead to more links between people?
2. Describe the contact network.
    a. Do all **nodes** have the same number of **edges**?
3. Which node is the most connected?
    a. Use the “Edgelist.Degrees()” function provided in the R script to calculate the degree of each node in your network.
    b. To do this, you must first run the function definition, which begins on Line 38. Next, you can use the function by providing our edgelist, EL, as the argument.
4. Write a hypothesis about how the **degree** of the host first infected by a new infectious agent will influence the size of the epidemic, meaning how many people will eventually become infected. We will calculate this as the proportion of the class that becomes infected. Explain your rationale for this hypothesis.

For upper-level students, complete this additional step:

5. Make a histogram to illustrate the variation in degree among the individual nodes in the network. Explain whether you think there is a little or a lot of variation among individuals in their degree.

Part C: Simulating an Epidemic

Our contact network represents heterogeneous contacts among our hosts, meaning that not all hosts have the same number of contacts. We will use a stochastic version of a classic SIR model in this network, which tracks susceptible (S), infected (I), and recovered (R) individuals, to simulate epidemics of a directly transmitted parasite. The flow diagram below indicates how individuals transition from infection to recovery:



1. Manually simulate epidemics in the network using dice.
    a. Draw or copy the contact network produced in Part A in your lab notebook or on a sheet of paper.
    b. Start with yourself as the only infected individual (commonly known as "Patient Zero").
    c. Now, roll a six-sided die or use a six-sided dice simulator for every node connected to you.
        i. If you roll a 5 or 6, that node becomes infected. Fill the node to indicate "infected".
        ii. If you roll 1-4, the node does not become infected and remains "susceptible". Leave the node blank to indicate "susceptible".
    d. Next, roll/simulate a six-sided die to determine your own status.
        i. If you get a 5 or 6, you recover from or clear the infection. Mark an X to indicate "recovered".
        ii. If you roll 1-4, you do not recover and remain infected. Fill in your node to indicated "infected".
    e. Repeat this process (in a round) for each infected node until one of the following conditions are satisfied: (1) all individuals have been infected or (2) no infected nodes remain (in both cases no additional hosts remain to be infected).
    f. Note: If an individual remains infected, you must return to that individual in the next round and roll to see if they recover or remain infected.
2. As a class, construct a graph that visualizes epidemic size (as a total number or as a proportion of hosts to ever get infected) vs. your degree of connectedness in the network.

Next, we are going to simulate epidemics in the network using the model in R, which mimics the manual simulation we just completed.

1. Load the "epinet" package, by running the line below:
    a. library(epinet)

2. Use the "summarize.SIR()" function to obtain a final summary of a single simulated outbreak. The result will provide the identity (but not the degree) of "Patient Zero" and the proportion of hosts in the network that became infected before the outbreak ended.
    a. Again, you must first run the function definition, which starts on Line 60.
    b. Next, you can use this function, by supplying our edgelist, EL, and a value for beta, the transmission parameter (we provide a default value of 0.15. Values between 0.15 - 0.5 typically work well).

3. Use the Spread.by.Degree function to simulate multiple, independent outbreaks on our network. Each time, the epidemic will begin with a randomly selected individual and progress until either (1) all individuals have been infected or (2) no infected nodes remain (in which case the parasite has gone extinct). This time, the function will return the degree of Patient Zero, but not their identity, and the proportion of the network that was infected by the end of the outbreak. We will need to simulate many epidemics in order to glean patterns between node degree and epidemic size.

   a. Again, you must run the function definition for Spread.by.Degree, which starts on Line 69
   b. Next, you can use this function to simulate 10,000 outbreaks with beta = 0.15 and store the result in the object “Sim_baseline”, by running the following line:
      i. Sim_baseline <- Spread.by.Degree(edgelist=EL, beta=0.15, iterations=10000)

4. Make a figure that effectively illustrates the relationship between degree and the ultimate size of the outbreak.
   a. *For non-majors*, the instructor can use the given code (see Lines 85-89) to create a box plot that relates the proportion infected individuals to the degree of Patient Zero. The instructor will then walk the class through how to interpret the results as they are presented there and lead a discussion about how the graph summarizes the raw data that was seen in the first graph. Instructors can ask the students to brainstorm or sketch out other ways that the data could be displayed to show the relationship between a node’s degree and the proportion of the network infected.
   b. *For majors*, either code or draw out a graph that would effectively summarize the raw data seen in the previous graph. How would different summary statistics (e.g., mean, median, or range) show or not show details of the simulation (e.g., the mean may show the trend in the relationship between degree and proportion infected but does not show that chance events can lead to any node starting an epidemic, for instance)?

Part D: Predicting the Effects of Public Health Interventions

Now that we have simulated epidemics in the network, **discuss and simulate how public health interventions may impact parasite transmission in the context of a contact network.**

Consider the following questions and discuss within your group:

1. Consider masking and social distancing. How might these interventions impact the contact network?
   a. Note: View the “Code block for Part C”, beginning on Line 97, for two new simulations and plots.
      i. Sim_masks re-runs our first simulation, but represents universal mask wearing that reduces transmission by 90%.
      ii. Sim_SD reruns our first simulation, but represents social distancing by assuming that individuals, on average, cut out 25% of their contacts.
   b. Note: Do not forget to run the intermediate steps for this process on Lines 107 and 109.
2. How do these interventions differ in the way they attempt to disrupt transmission?
3. How could you use this data to plan control strategies during an epidemic? Are there any other pieces of information that you think would improve the dataset?
4. Did building a dataset like this help you understand how people study disease spread? How does this relate to Covid-19 pandemic control strategies?

Answer the following questions after you have simulated epidemics in the contact network:

5. What is the relationship between degree of “Patient Zero” and the proportion of the network that becomes infected? Does this support your previous prediction? Explain why or why not.

Discussion

For this laboratory exercise, you were asked to build a contact network based on your connectedness with classmates, simulate an epidemic based on the contact network, and discuss the implications of public health interventions on the contact network. These exercises should have illustrated that the connectivity of hosts, the likelihood of parasite transmission, and the effect of public health interventions are key factors that influence the size and duration of epidemics and that understanding these dynamics can help us solve real-world epidemics. For example, researchers were able to trace the origin of an antibiotic resistant Klebsiella pneumoniae hospital outbreak by comparing whole-genome sequence isolates from infected patients and building a network of possible transmission links between patients (Snitkin et al. 2012). Similar analyses are currently underway for the Covid-19 pandemic (Nande et al. 2021), which further exemplifies the utility of contact networks in elucidating drivers of disease. Moving forward, you can use this laboratory exercise to explore, test, and discuss how other environmental and social factors, such as directionality (e.g., sexually transmitted infections), the intensity of host interactions, and temporary host and/or parasite interactions, could impact disease transmission.

Cited References

Bansal S, Grenfell BT, and Meyers LA. 2007. When individual behaviour matters: homogeneous and network models in epidemiology. Journal of the Royal Society Interface 4: 879-91.

Bansal S, Read J, Pourbohloul B, and Meyers LA. 2010. The dynamic nature of contact networks in infectious disease epidemiology. Journal of Biological Dynamics 4: 478-89.

Kermack WO and McKendrick AG. 1927. A contribution to the mathematical theory of epidemics. Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character 115: 700-721.

Nande A, Adlam B, Sheen J, Levy MZ, and Hill AL. 2021 Dynamics of COVID-19 under social distancing measures are driven by transmission network structure. PLOS Computational Biology 17(2): e1008684.

Snitkin ES, Zelazny AM, Thomas PJ, Stock F, NISC Comparative Sequencing Program Group, Henderson DK, Palmore TN, Segre JA. 2012. Tracking a hospital outbreak of carbapenem-resistant Klebsiella pneumoniae with whole-genome sequencing. Science Translational Medicine 4: 148ra116.

Materials

Students require a laptop or computer with Google Sheets (freely available through Google Drive), R (freely available for download from **https://www.r-project.org/**), and "R Studio" (freely available for download from **https://www.rstudio.com/products/rstudio/**). At least one 6-sided die should be purchased for each student for in-person classes. If possible, we recommend that more than one die per student should be purchased so that the manual dice-rolling portion of the laboratory can proceed at a quicker pace. If additional dice cannot be purchased, a dice roll simulator could be used instead (search for "roll dice" on Google to use their simulator).

Notes for the Instructor

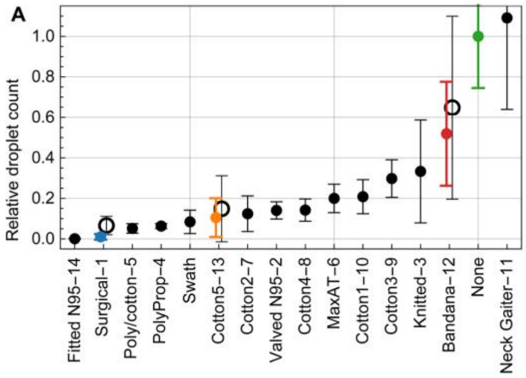
This activity takes students from building a dataset, making and testing hypotheses and predictions, to simulating disease spread in multiple ways. We have included variations of this activity that can be used for different levels of students (e.g., non-science majors, biology majors). An additional strength of this activity is the high level of interaction between students while constructing their contact network. Building a dataset from self-generated, personal data gives students an opportunity to translate observations into usable data and increases personal investment in interpreting the outcome of the activity. This is especially beneficial with remote teaching, as it creates connections between classmates when they are not on campus.

Networks of around 15-25 are ideal. For instance, splitting a class of 40 or more students into two groups makes the networks more manageable to construct and provides two replicate networks to observe that may have different structures and therefore different outcomes. Using Google Docs/Sheets is a useful tool for remote and in-person learning for recording their contact networks, and it allows the instructor to "check in" on groups without being in the group at the time. Additionally, we realize that there may be a variety of factors that lead to low connectedness among students. If this is the case for your class, we have created a selection of mock student profiles that could be used instead of self-generated connections (Appendix A). Additionally, for instructors who would simply like to show the results of the mock student contact network and the subsequent simulations, we have also included these figures (Appendix A).

The most challenging part of this activity is that students struggle with the underlying mechanics of the outbreak computer simulations. The manual dice-rolling component is meant to mitigate this confusion by having students compute different epidemic scenarios and sharing their results with the class or group members. Emphasizing the element of randomness, with some known parameters, such as transmission or recovery probabilities, helps students visualize what is going on "behind the scenes" in disease modeling.

In practice, this activity is best completed over two class periods, especially if a background lecture and introduction to R/R Studio needs to be given first. Splitting the activity up into two class periods also allows the instructor to verify the student's adjacency matrices and scaffold the R script according to the skill level of the class.

This activity also allows instructors and students to explore and simulate patterns and points from outside readings. There are a wealth of news and primary source articles about the Covid-19 pandemic that provide data from which to modify the code in this activity so the students can have a "hands-on" experience with modeling disease spread. For instance, the figure below from Fischer et al. (2020) shows the relative reduction in airborne particle spread using different types of homemade or readily available facemasks (e.g., bandana, cotton facemask, N95) compared to wearing no mask. An interesting result of this study was that neck "gaiters" made from polyester and spandex increased respiratory droplet spread compared to no mask because it reduced particle size and did not trap airborne saliva droplets efficiently. Thus, these data could be used to reduce or increase the transmission term, beta, by a given proportion. Students could then run the simulation again to see how effective different masks could be at slowing or stopping an epidemic.



Depending on the level of the students, articles from newspapers and magazines can also be great resources for information on real-life disease examples. Suggested readings are included below and show examples of superspreader events during the Covid-19 pandemic and vaccination/herd immunity issues during recent measles outbreaks. Ask students to either apply what they read to the actual simulation or discuss how the issues in their readings would relate to the use of contact networks or disease simulations in the control of a disease epidemic. Having students draw or sketch out different network scenarios or changing the parameters of either the code or the dice-roll simulation are also two great ways to help students make connections between what they are reading about and this class activity.

Lastly, for instructors unfamiliar with R/R Studio, we have provided a short list of resources to get started:

https://www.datacamp.com/community/blog/how-to-learn-r

R for Data Science by Hadley Wickham & Garrett Grolemund (https://r4ds.had.co.nz/)

https://www.youtube.com/c/RProgramming101

Suggested readings to assign to students

Cheng, Y, Ma N, Witt C, Rapp S, Wild PS, Andreae MO, Pöschl U, and Su H. 2021. Face masks effectively limit the probability of SARS-CoV-2 transmission. Science: eabg6296.

Fischer EP, Fischer MC, Grass D, Henrion I, Warren WS, Westman E. 2020. Low-cost measurement of face mask efficacy for filtering expelled droplets during speech. Science Advances 6: eabd3083.

Kolata , Gina. 2020, August 13. On Native American Land, Contact Tracing Is Saving Lives. The New York Times. Retrieved from: https://www.nytimes.com/2020/08/13/health/coronavirus-contact-tracing-apaches.html

Otterman, Sharon. 2019, January 17. New York Confronts Its Worst Measles Outbreak in Decades. The New York Times. Retrieved from: https://www.nytimes.com/2019/01/17/nyregion/measles-outbreak-jews-nyc.html

Zimmer, Carl. 2020, June 30. Most People With Coronavirus Won't Spread It. Why Do a Few Infect Many?. The New York Times. Retrieved from: https://www.nytimes.com/2020/06/30/science/how-coronavirus-spreads.html

Cited References

Fischer EP, Fischer MC, Grass D, Henrion I, Warren WS, Westman E. 2020. Low-cost measurement of face mask efficacy for filtering expelled droplets during speech. 2020. Science Advances 6: eabd3083.

Acknowledgments

We want to thank students in BIOL347 at Emory University and SC203 at the Fashion Institute of Technology who provided feedback to improve this activity.

About the Authors

K.H. Nguyen is a postdoctoral fellow in the Department of Biology at Emory University and teaches at Morehouse College. She is supported by an Institutional Research and Academic Career Development Awards (IRACDA) (NIH/NIGMS 2K12GM000680). E.C. Rynkiewicz is an Assistant Professor of Ecology in the Department of Science and Mathematics at the Fashion Institute of Technology. Rynkiewicz teaches non-science majors courses in ecology, biology, disease ecology, and biodesign. D.J. Civitello is an Assistant Professor in the Department of Biology at Emory University and teaches disease ecology and coastal ecology for biology majors.

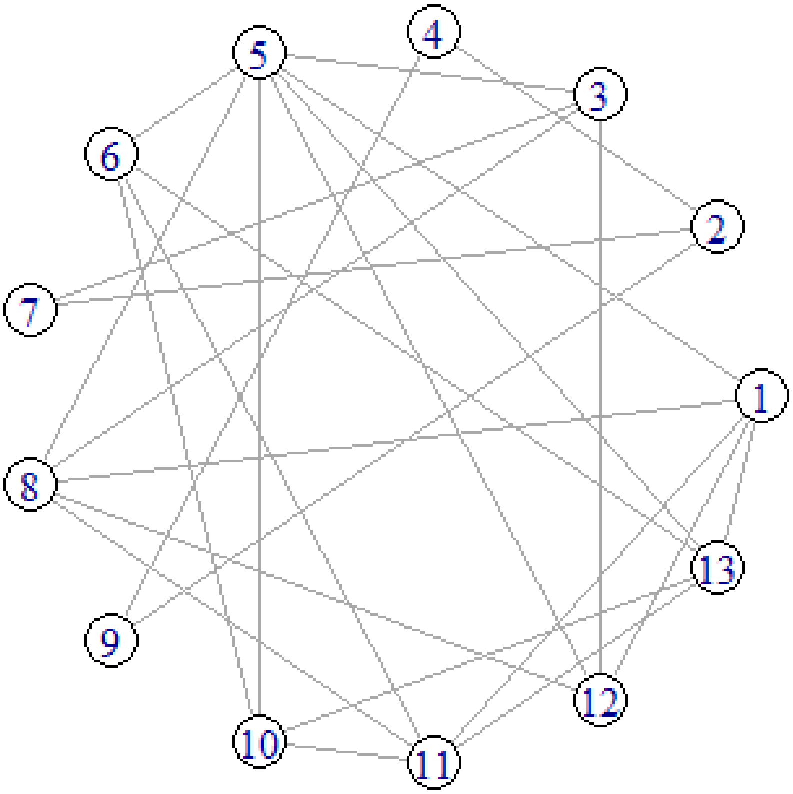
Appendix A

Mock student profiles

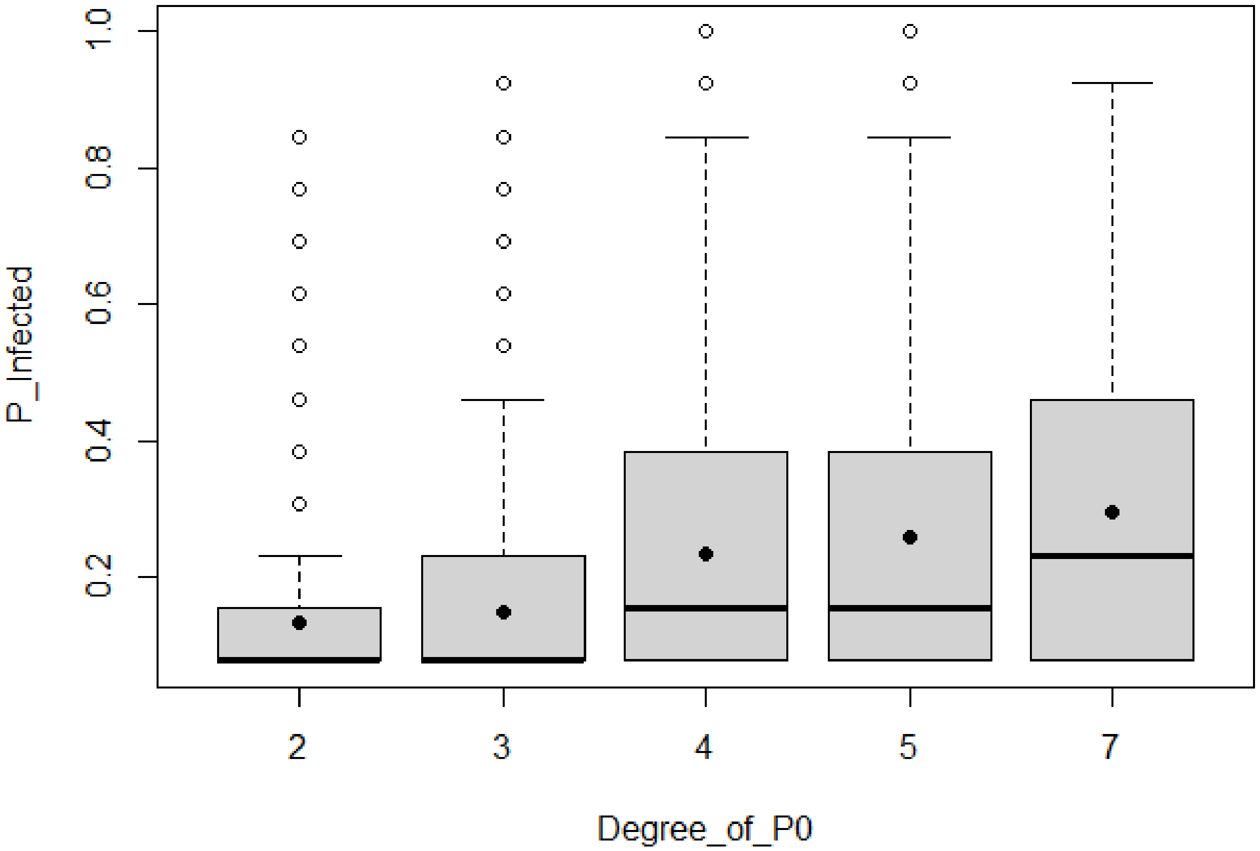
In the table below, we provide 13 student profiles that could be used for courses where students may not share common courses, extracurricular activities, or residential housing. You can replace Student 1-13 with names of students in your class.

	Course	**Extracurricular Activity**	**Residential Housing**
Student 1	General Chemistry	Theater	Macelwane Suites
Student 2	Calculus I	Chess Club	Hopkins Complex
Student 3	Italian 100	Flag Football	Comstock Lane
Student 4	Art History	Chess Club	Yulee Hall
Student 5	Introductory Biology	Flag Football	Macelwane Suites
Student 6	Introductory Biology	Newspaper	Magnolia Apartments
Student 7	Calculus I	Archery	Comstock Lane
Student 8	General Chemistry	Flag Football	Macelwane Suites
Student 9	Art History	Chess Club	Yulee Hall
Student 10	Introductory Biology	Theater	Magnolia Apartments
Student 11	General Chemistry	Theater	Magnolia Apartments
Student 12	General Chemistry	Flag Football	Macelwane Suites
Student 13	Introductory Biology	Theater	Magnolia Apartments

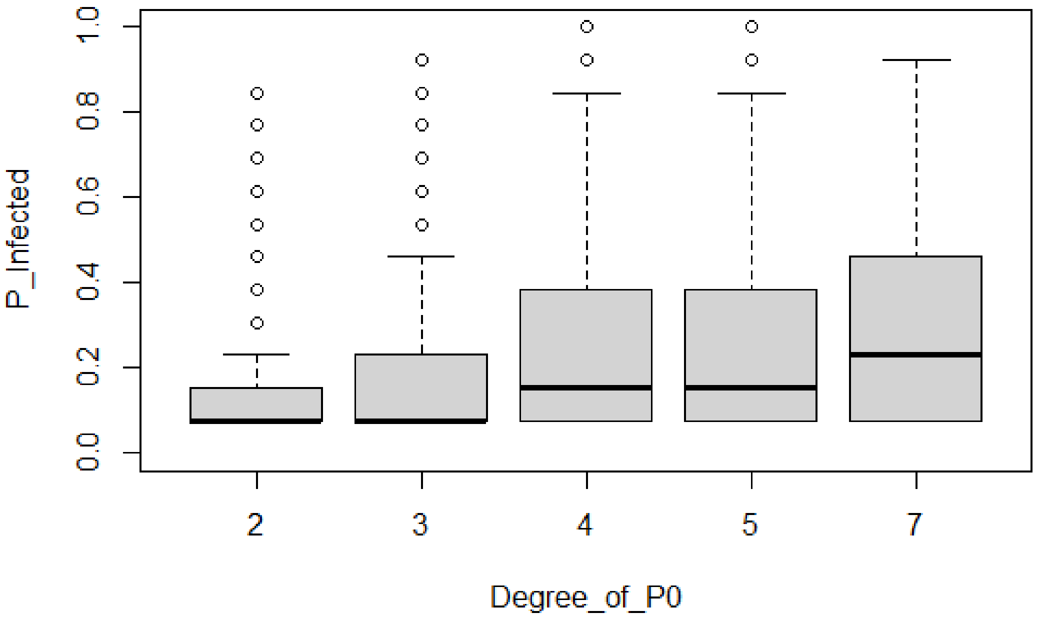
The adjacency matrix based on the mock student profiles in the table can be found in the Supplementary Materials as an .xlsx file. Below is the contact network that is produced by the mock student profiles.

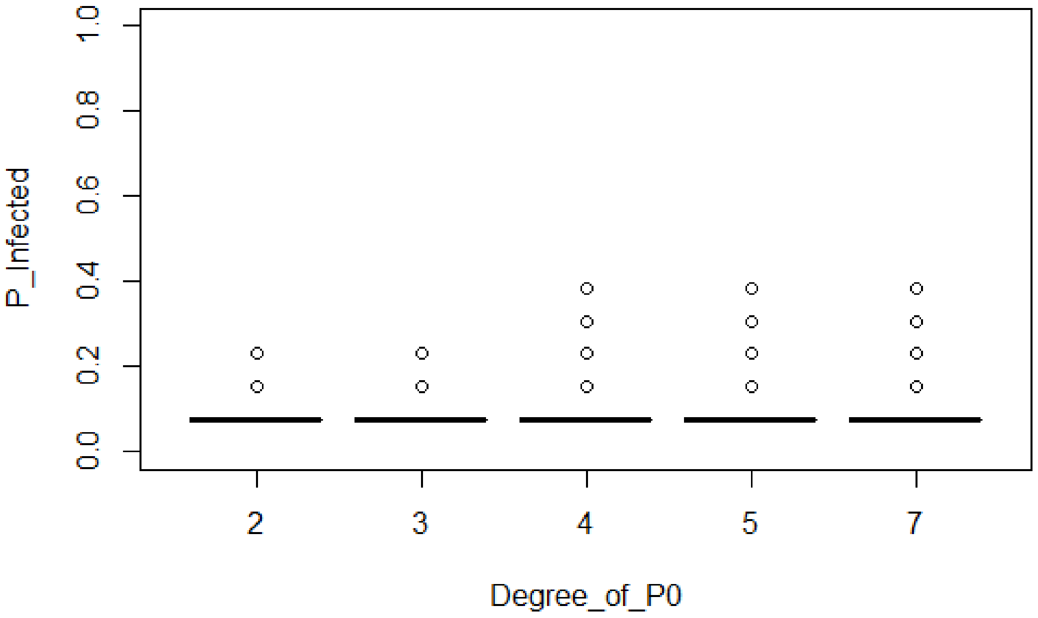


Below is a boxplot that shows, out of 10,000 simulations, how the degree of “Patient Zero” impacts the number of individuals that become infected. The bold lines represent the median, the black points represent the mean, and the black circles represent outliers. As expected, the higher the degree of “Patient Zero”, the more likely it is that individuals will become infected.

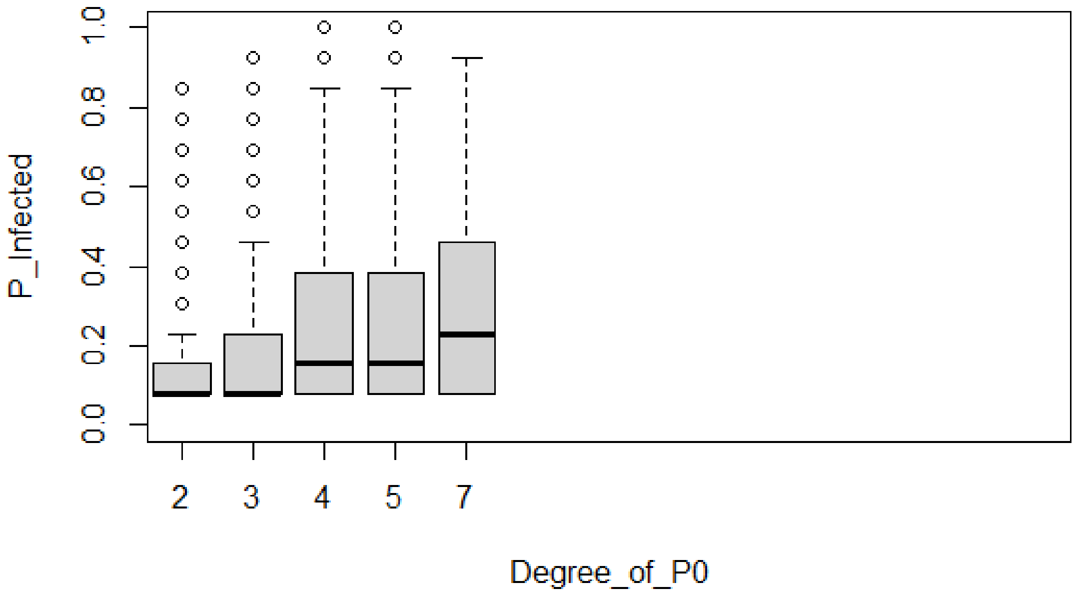


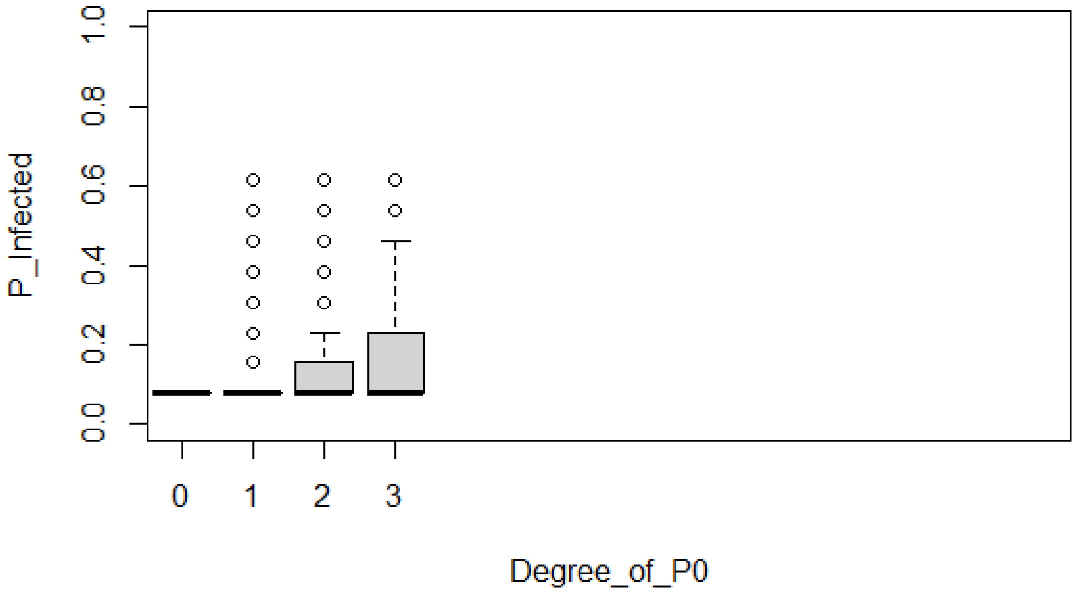
Below is a boxplot that shows, out of 10,000 simulations, how the degree of “Patient Zero” impacts the number of individuals that become infected (top panel) and the impact of masking on the number of individuals that become infected (bottom panel). The bold lines represent the median and the black circles represent outliers. As expected, masking dramatically reduced the number of infected individuals, regardless of degree.





Below is a boxplot that shows, out of 10,000 simulations, how the degree of “Patient Zero” impacts the number of individuals that become infected (top panel) and the impact of social distancing on the number of individuals that become infected (bottom panel). The bold lines represent the median and the black circles represent outliers. As expected, social distancing drastically reduces the degree to which individuals are connected, thereby decreasing the likelihood that outbreaks occur if individuals become infected.





Mission, Review Process & Disclaimer

The Association for Biology Laboratory Education (ABLE) was founded in 1979 to promote information exchange among university and college educators actively concerned with teaching biology in a laboratory setting. The focus of ABLE is to improve the undergraduate biology laboratory experience by promoting the development and dissemination of interesting, innovative, and reliable laboratory exercises. For more information about ABLE, please visit http://www.ableweb.org/.

Advances in Biology Laboratory Education is the peer-reviewed publication of the conference of the Association for Biology Laboratory Education. Published articles and extended abstracts are evaluated and selected by a committee prior to presentation at the conference, peer-reviewed by participants at the conference, and edited by members of the ABLE Editorial Board. Published abstracts are evaluated and selected by a committee prior to presentation at the conference.

Citing This Article

Karena H. Nguyen, Evelyn C. Rynkiewicz, and David J. Civitello. 2022. Simulating Parasite Transmissino and Epidemics Using Contact Networks. Article 13 In: Boone E and Thuecks S, eds. *Advances in biology laboratory education*. Volume 42. Publication of the 42nd Conference of the Association for Biology Laboratory Education (ABLE). https://doi.org/10.37590/able.v42.art13