Using a Hybrid Agent-Based and Equation Based Model to Test School Closure Policies

Elizabeth Hunter (✉ elizabeth.hunter@tudublin.ie)

Technological University Dublin <https://orcid.org/0000-0002-1767-4744>

John Kelleher

Technological University Dublin

Research article

Keywords: Agent-Based Model, Infectious Disease, School Closure, Measles, Simulation, Hybrid Model

Posted Date: April 13th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-20889/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published on March 12th, 2021. See the published version at <https://doi.org/10.1186/s12889-021-10513-5>.

RESEARCH

Using a Hybrid Agent-Based and Equation Based Model to Test School Closure Policies

Elizabeth Hunter* and John Kelleher

*Correspondence:
elizabeth.hunter@tudublin.ie
Technological University Dublin,
Grangegorman, Dublin, Ireland
Full list of author information is
available at the end of the article

Abstract

Background: In order to be prepared for an infectious disease outbreak it is important to know what interventions will or will not have an impact on reducing the outbreak. While some interventions might have a greater effect in mitigating an outbreak, others might only have a minor effect but all interventions will have a cost in implementation. Estimating the effectiveness of an intervention can be done using computational modelling. In particular, comparing the results of model runs with an intervention in place to control runs where no interventions were used can help to determine what interventions will have the greatest effect on an outbreak.

Methods: To test the effects of a school closure policy on the spread of an infectious disease (in this case measles) we run simulations closing schools based on either the proximity of the town to the initial outbreak or the centrality of the town within the network of towns in the simulation. To do this we use a hybrid model that combines an agent-based model with an equation-based model.

Results: Our results show that closing down the schools in the town where an outbreak begins and the town with the highest in degree centrality provides the largest reduction in percent of runs leading to an outbreak as well as a reduction in the overall size of the outbreak compared to only closing down the town where the outbreak begins. Although closing down schools in the town with the closest proximity to the town where the outbreak begins also provides a reduction in the chance of an outbreak, we do not find the reduction to be as large as when the schools in the high in degree centrality town are closed.

Conclusions: Thus we believe that focusing on high in degree centrality towns during an outbreak is important in reducing the overall size of an outbreak.

Keywords: Agent-Based Model; Infectious Disease; School Closure; Measles; Simulation; Hybrid Model

Background

Infectious disease outbreaks are a major threat to global health while emerging diseases such as COVID-19, SARS, Ebola, and MERS tend to grab the biggest headlines infectious diseases such as influenza, measles and mumps are becoming more prevalent due to a reduction in vaccination rates and global travel. In 2018 there were almost 10 million cases of measles with 142,000 deaths around the world and in 2019 there were three times as many cases reported as in 2018 [1]. As outbreaks get larger, many governments try to impose control measures by closing schools, starting vaccination campaigns, or limiting public events. In order to help mitigate outbreaks and choose the best interventions it is important to understand as much as we can about outbreak dynamics. An intervention is not useful if it

takes up resources and does nothing to alter the course of an outbreak or makes an outbreak worse. Additionally an intervention that only has a minor impact on an outbreak but uses a considerable amount of resources might not be the best strategy. However, it is difficult to test if an intervention works during a real outbreak. One main reason for this is there is no control scenario to compare what would have happened if the intervention was not implemented. So it is difficult to determine what was a result of the intervention and what would have happened if the intervention was not implemented.

One way to better understand how interventions influence an outbreak is by using models. Models are often used to study a system when it would be infeasible to run an experiment to study the real world system. Two of the main types of epidemiological models often used to study infectious disease outbreaks are agent-based and equation based models. We propose using a hybrid agent-based and equation based model to simulate the spread of measles through a county in Ireland. Our motivation for using a hybrid model is that it enables us to scale the model to include a larger population while at the same time controlling the computational cost of running the simulation. To test the model and show how it can be used to determine the best intervention policies in particular school closure policies, we run a number of different scenarios to determine which policy is the best in reducing outbreaks. The paper starts by reviewing the literature for agent-based models used to analyse interventions in particular school closure policies. Then we discuss the model used to look at interventions, and then outline our experiments. Finally, we present the results.

Modelling the Effect of School Closures on Infectious Disease Spread

An intervention strategy that is occasionally implemented but its effectiveness is still debated is school closure policies. The policy of closing schools during disease outbreaks is used in Japan, Bulgaria and Russia to lessen influenza outbreaks [2]. Its usefulness was debated in New York during the 1918 Spanish flu pandemic [3] and is being used today in an attempt to reduce the severity of measles outbreaks in Samoa [4] and to reduce the spread of COVID-19 in countries around the world. However, there is no clear evidence to show that closing schools helps to reduce the size of an outbreak. In fact, Lee et al. [5] find that shorter school closures of two weeks or less end up increasing the overall attack rate^[1] and school closures may only be effective if they last for the entire duration of the epidemic. Similarly, Grefenstette et al. [6] find that while the epidemic temporarily slows when schools close as soon as they reopen the epidemic peaks again. However, there are cases where studies have shown that school closure policies can play a significant role in reducing an outbreak. Litvinova et al. [2] use real contact data to simulate the effects of a Russian school closure policy and find that reactive strategies, closing down classes and schools when a given percent of students show symptoms reduces the severity of an outbreak. One of the reasons why the effects of school closure policies is hard to determine is that not all policies are the same. Many factors help to define a school closure policy such as the number of students needed to be infected before implementing the policy, the length of the closure and determining

---

^[1]Attack rate is defined as the risk of getting a disease during a specified period.

what schools close down. Is it enough to only close the effected schools or should all schools in the town or all schools within a certain radius be closed to reduce the spread of the infectious disease. We aim to use a model to test school closure policies that take into account a towns place in a network of other towns: focusing on the centrality of other towns in the network and the physical distance to other towns.

Methods

The model used for the study is at its core an agent-based model. The basic building blocks of an epidemiological agent-based model are an environment component, a transport component, a society component and a disease component [7]. The distinctive aspect of the model is that it has a hybrid agent-based and equation based disease component. The following sections give an overview of the hybrid agent-based model breaking the model down into the four main components of an agent-based model for infectious diseases (society, transportation, environment and disease).

Environment Component

In an agent-based model the environment can be created with a high level of detail. For example, in the Hunter et al. [8] model a town is made up of multiple small areas: from Irish Census data small areas are the smallest geographic region over which the census is aggregated. Each small area contains between 50 to 200 dwellings [9]. Within the small areas the model uses real zoning data to assign agents' homes to residential areas and their workplaces to industrial or commercial areas. In addition, each school is placed in the correct location. The model used in our experiments is based on the Hunter et al. [8] model; however, in order to reduce the computational cost of scaling the model over a larger environmental area (county as opposed to town level), the version of the model used in our experiments reduces environmental fidelity so that a small area is represented by one grid cell or Netlogo patch. When in a small area an agent has access to certain information such as the number of schools or workplaces in the small area and the real world distance to each other small area in the model. Agents can move between small areas and when in a small area all agents are physically in the same location, however, the agents keep track of where they are in that small area: home, work, school or community and restrict their interactions with other agents accordingly.

Society Component

The society is based off of real world census data from the Irish Central Statistics Office [9]. The population in each small area in the model reflects the population statistics of the real small area including age, sex, household size and economic status. Social networks are included in the model. Agents have a family social network that is made up of any agents living in their household. Agents also have a work or school social network that is made up of other agents in their workplace or their school and students are given an additional social network which is a class network that is made up of agents who are in their school and of the same age. Social networks help to determine the contacts an agent has in the model. Irish

vaccination data is used to determine the percentage of each age group that have received vaccinations for the infectious disease being modelled. For example, if 90% of 1 year olds in Ireland had been given the MMR vaccination in 2011 and we are running a model for 2012, we give each agent in the model with an age of 2 a 90% chance of having been vaccinated. If an agent is vaccinated they are given a 97% chance of being immune to the disease. This takes into account vaccination failure and is based on the vaccine effectiveness rate for measles [10]. Half of the agents with age less than 1 are given immunity to a disease to mimic passive immunity infants receive from their mothers [11]. For any agents that have an age corresponding to a vaccination year not in our data we give a 99% chance of being immune. Prior to vaccination campaigns the majority of the population would have either had or been exposed to childhood diseases, such as measles, leaving them immune in later life.

Transportation Component

Within the model agents move in one step from their location to their desired destination. Some of the agents movements are predetermined by the model while others are decided on by the agent during the runtime of the model. The predetermined movements involve the agents moving between home and school or home and work at given times. Commuting patterns within the model are determined using CSO Place of Work, School or College - Census of Anonymity Records (POWSCAR) data [12]. This data set provides information on the commuting patterns of people in Ireland and gives the number of people that commute from one electoral division to another. Electoral divisions are the census geographic area one step above the small areas.

When agents are not moving from home to school or work they are allowed to decide where they want to move in the model. To capture realistic movement patterns we use a gravity type model for transportation. A traditional gravity model gives the interactions between two location pairs and determines those interactions based on the characteristics of a location and the distance between locations [13]. The probability of an agent moving to another small area is proportional to the population density of the small area, an area that has a lot of other agents is more attractive, and inversely proportional to the distance to the small area from the agents current location, areas that are farther away are less attractive. We feel that this transportation model provides a more accurate simplification of movement within a larger area than that in the original [8] town model.

Disease Component

From a modeling perspective, the biggest difference between the Hunter et al. [8] model and the model we actually used in our experiments was that our experimental model uses a hybrid disease component (the disease component in the Hunter et al. [8] model was entirely agent-based). The disease component of the model is made up of two different types of models: an agent-based disease component and an equation based disease component. It is set up so that the model can be run with a completely agent-based disease model, a completely equation based disease model or switch between the two based on certain criteria. The idea behind using

a switch is that the agent-based models are especially important when a few agents are sick because at this stage the individual movements are what drive the spread of the disease so the heterogeneous movements of agents are more important. For example, if the one infected agent decides to stay home the outbreak might not take off versus if they decide to go to school or work every day. However, once the number of infected individuals reaches a certain number the individual movements should not matter as much because there are so many agents infected.

To capture this in the model the towns are allowed to switch between the agent-based model and the equation based model. The decision of which model is used in a geographic area in a given time step is determined by the number of agents infected in that town. Note, that if the number of infected or exposed agents in an area drops back below this threshold the model reverts back to an agent-based model. For the purpose of this model the towns will switch when 1% of agents are infected or exposed. The following sections discuss the agent-based model and the equation based model in more detail.

Disease Component: Agent-based Model

The agent-based element of the disease component is based off of a compartmental Susceptible, Exposed, Infected, Recovered (SEIR) model. Where agents start in one of four different compartments, they are either susceptible, exposed, infected or recovered and based on their interactions they move between the compartments. The agent-based disease component remains unchanged from the Hunter et al. [8] model. When an infectious agent comes into contact with a susceptible agent they have a given percent chance of passing the disease to the susceptible agent. If they do pass on the disease, the susceptible agent then moves to the exposed state for a given period of time before moving to the infectious state. They will remain in the infectious state for a set period of time before recovering. For most diseases, once an individual has recovered they cannot be reinfected.

The disease dynamics are set to mimic measles. An individual will stay in the exposed state for an average of 10 days [10]. The time an agent remains exposed is determined for each agent from a normal distribution with a mean of 10 and a standard deviation of 0.5. Once infectious an individual remains infectious for an average of 8 days [10]. The time an agent remains infectious in the model is determined for each agent from a normal distribution with a mean of 8 and a standard deviation of 0.5. The infection rate, the percentage chance that a susceptible agent will be infected after contact with an infectious agent, is determined using the basic reproductive number R_0 for measles (12-18) [10]. The basic reproductive number is the number of individuals infected by one infectious individual in a completely susceptible population and is a standard measure of transmission of a disease. The parameter is made up of three components, number of contacts per unit time (c), the transmission probability per contact (p), and the duration of the infectiousness (d) [14]. As we can determine the number of contacts per unit time from our model and we know the duration of the infectiousness, we can determine the transmission probability per contact.

Disease Component: SEIR Model

The equation based part of the disease component uses an SEIR difference equation model. Difference equations were chosen over the more common differential equation models because of the discrete time steps that are used in difference equations. This is more analogous to the agent-based model and will allow for a more seamless transition between the two models. In the simulation, each geographic area runs its own SEIR difference equation model. The model can be run at the small area level or the town level. The equations are as follows:

$$S_{i+1} = S_i - \frac{\beta I_i S_i}{N} \quad (1)$$

$$E_{i+1} = E_i + \frac{\beta I_i S_i}{N} - \sigma E_i \quad (2)$$

$$I_{i+1} = I_i + \sigma E_i - \gamma I_i \quad (3)$$

$$R_{i+1} = R_i + \gamma I_i \quad (4)$$

Where S_i is the number of susceptible agents in the geographic area in the previous time step and S_{i+1} is the number of susceptible agents in the geographic area in the current time step. E_i and E_{i+1} are the number of exposed agents in the geographic area in the previous and current time steps, I_i and I_{i+1} are the number of infected agents in the geographic area in the previous and current time steps, and R_i and R_{i+1} are the number of recovered agents in the geographic area in the previous and current time steps. β is the infection rate or the probability of infection per contact between agents, σ is the rate of moving from exposed to infected, γ is the recovery rate and N is the total population across all of the compartments.

In a fully equation based disease component, each geographic area starts its difference equation model when an infected or exposed individual enters the area. This could happen in two ways, either an agent from outside the area who is already exposed or infected moves into the area or an agent who is from the area becomes infected outside and returns home. Once the difference equation model has started it continues until there are no longer any more exposed or infected agents in the model. In a hybrid model the difference equation model will start when the number of infected or exposed individuals is over a certain threshold and will finish reverting back to an agent-based disease component when the number of infected or exposed individuals falls below the threshold.

At each time step, each area will calculate the values for the difference equations and adjust the number of agents in the area in each category. If the rounded difference between E_{i+1} and the number of exposed agents in the area is greater

than 0, that number of susceptible agents in the small area will randomly be selected to move from the susceptible category to the exposed category. Similarly if the rounded difference between I_{i+1} and the count of infected agents in the area is greater than 0 than that number of exposed agents will be randomly selected to move from exposed to infected, and if the rounded difference between R_{i+1} and the count of recovered agents in the area is greater than 0, than that number of infected agents in the area will recover.

Because movement is possible, there are times when the total number of agents in the area in one of the four categories is less or greater than the value predicted in the model. Adjustment are made to account for this. If the value for E_i , I_i , or R_i is less than one and the count of agents exposed, infected or recovered in the area is greater than one then the value for E_i , I_i , or R_i is changed to the count of agents in that area who are exposed, infected or recovered. If the values for the difference between E_i , I_i , or R_i and the number of agents exposed, infected or recovered respectively in the geographic area is greater than the number of agents who could potentially move into the compartment (if the difference between E_i and the count of agents exposed is greater than the number of susceptible agents) the value for E_i , I_i , or R_i are adjusted down to reflect the actual counts of agents in the geographic area.

Model Implementation

For the model implementation we use the county of Leitrim Ireland. Based on the 2016 census the county has a population of approximately 32,000 people over an area of 1,590km². The county is made up of 173 small areas and 69 different electoral divisions. About 45% of the population are students and using the vaccination rates discussed in Section: Society Component 12.3% of the population is not vaccinated or otherwise immune to measles. Although we use Leitrim as an example in this paper, any county in Ireland can be used for the model or any region another country as long as we have the appropriate data.

Experiments

Having a model that can recreate an outbreak allows us to learn interesting things about the dynamics of an infectious disease outbreak, for example Hunter et al. [15] examines how the centrality of a town within a network of towns influences the spread of an outbreak. In the study the degree centrality of towns in a network is considered and is determined based on the number of agents commuting into and out of a town. There are multiple types of degree centrality: *total degree centrality* weights agents commuting in and out of towns equally, *in degree centrality* is based on the count of agents commuting into a town, and *out degree centrality* is based on the count of agents commuting out of a town. They find that the agents commuting into a town are more important in spreading an infectious disease than the agents commuting out of a town and that the higher the in degree centrality of the town the outbreak starts in the less important the centrality of the other towns in the network is in determining if the outbreak will spread to those other towns. However, even though this is an interesting finding, the question remains how can this help us in stopping or slowing down an outbreak. We propose using the findings from the

Hunter et al. [15] model to test out different intervention strategies. For example, while it might seem to make sense to close the schools in a town when an outbreak of a childhood disease begins to take off, there is evidence to show that this does not always help and in some cases actually makes an outbreak worse. As it was determined that a town with higher in degree centrality will result in greater spreading of the outbreak across all towns in the network, we run experiments to look at the effects of closing down schools in the high in degree centrality towns as opposed to the town the outbreak starts in. The thought behind this is that it is the high in degree centrality towns that results in spreading to more towns throughout the network and that by stopping agents students from going into these high centrality towns we will stop them from bringing the disease into the high centrality town and then out to other towns.

In order to test the effects of closing schools in different towns we run two experiments using the hybrid model. The first experiment involves four different interventions. The first is with no interventions but vaccination rates based off of Irish vaccination rates and with the outbreak starting in the town of Drumkeeran, Ireland in County Leitrim. Drumkeeran was chosen as it is a smaller town in Leitrim County with relatively low in degree centrality and it has both a primary and a secondary school in the town. The second set of runs uses the hybrid model with vaccination rates and with schools in Drumkeeran closing down when more than two students are infected in the town. The third scenario again uses vaccinations and when more than two students are infected in Drumkeeran, the schools in Manorhamilton are shut down. Manorhamilton is the town in Leitrim with the highest in degree centrality and is approximately 20 km away from Drumkeeran. The final intervention scenario has the same vaccination rates and involves closing schools down in both Drumkeeran and Manorhamilton when more than two students are infected. The model is run for each set of initial conditions 300 times to account for stochasticity and we look at a number of different measures to compare the outbreaks.

The second experiment involves looking at different combinations of school closures based on the centrality of the towns versus the distance of the town to the initial outbreak. The outbreak begins in Drumkeeran for each set of runs. The model was run to test each of the following intervention scenarios based on in degree centrality: (i) only closing the schools in Drumkeeran; (ii) closing the schools in Drumkeeran and the schools in the town with the highest in degree centrality that has at least one primary or secondary school, this second town was Manorhamilton (see Table 1 for a listing of the five town in Leitrim containing at least one school ordered by in degree centrality); (iii) closing the schools in Drumkeeran and the schools in the town with the highest in degree centrality (i.e., Manorhamilton) and the town with the second highest in degree centrality (i.e., Mohill); and so on, with runs (iv), (v), and (vi) each progressively adding the towns included in the school closures in order of decreasing in degree centrality.

The second set of runs looks at closing schools based on distance. This is done to help determine if any results from closing down schools by centrality is simply because additional schools are closed or if it can be attributed to the centrality of the town that is closed. The first run closes school in only Drumkeeran, the town where the outbreak starts; the next set of runs closes schools in Drumkeeran and the

Table 1 Towns in Leitrim with at least one school ordered by centrality

Town	Centrality
Manorhamilton	197.8
Mohill	171.7
Drumshanbo	151.8
Carrigallen	93.7
Ballinamore	90.8

next closest town with at least one school in it; the third set of runs closes schools in Drumkeeran, the closest town with at least one school in it, and the second closest town with at least one school; and so on. Table 2 shows a list of the five closest towns to Drumkeeran.

Table 2 Towns in Leitrim with at least one school ordered by distance to Drumkeeran

Town	Distance (km)
Drumahaire	14
Killanummery	14
Drumshanbo	18
Manorhamillton	20
Leitrim	20

For each set of towns the model is run 300 times and the results are compared between closing schools based on distance and centrality.

Results

We first look at the results for the first experiment: vaccination, closing schools in Drumkeeran, closing schools in Manorhamilton and closing schools in Drumkeeran and Manorhamilton. The first measure we look at to compare the results of the different interventions is to look at the percent of runs that lead to more than three agents infected. In a model without interventions we typically look at the percent of runs that lead to an outbreak, using the World Health Organization's (WHO) definition of a measles outbreak which is two or more connected cases of measles, however, because we want to look at the effects of the interventions, and the interventions do not start until we have at least two agents infected, we look at the runs where there are more than three agents infected. Table 3 shows the percent of runs that have three or more infected agents for each of the versions of the model along with the 95% confidence intervals for the statistics and the p-value for a test of equal proportions to show if the percent of runs resulting in three or more agents being infected is significantly different from the percent of runs resulting in three or more agents being infected in the vaccinations only model.

Table 3 The percent of runs that result in three or more agents becoming infected based off of the intervention strategies used in the model. The 95% confidence interval and the p-value for a test of equal proportions. The first intervention is vaccinations with no school closures and the next three rows give results for the schools that are closed in the simulation to stop the spread of measles.

Intervention	Percent of Runs	Confidence Interval	P-value
Vaccinations Only	51.3	(45.7, 57.0)	-
Drumkeeran	51.3	(45.7, 57.0)	1.00
Manorhamilton	47.3	(41.7, 53.0)	0.37
Drumkeeran and Manorhamilton	43.3	(37.8, 48.9)	0.06

From Table 3 it can be seen that the percent of runs that lead to three or more infected agents is the same for the model with only vaccination as an intervention

and the model closing schools in Drumkeeran. This further emphasises the findings showing that closing schools in the town where the outbreak starts does not always reduce an outbreak. Looking at the effect of the other interventions on the model results we can see that even though the percent of runs with over three infected when schools are closed in Manorhamilton is slightly lower than the percent of runs when the schools are closed in Drumkeeran the results are not significantly different. However, when schools close in both Drumkeeran and Manorhamilton, the percent of runs with more than three infected agents is significantly different at a 7% significance level than when vaccinations are the only intervention or when only schools in Drumkeeran are closed down. These results show that there might be an advantage in closing down the schools in the high centrality towns nearby along with the initial town where the outbreak starts.

The results are further broken down to see if there are other effects of closing schools. We look at some summary statistics for the total number of infected agents at the end of the outbreak for the county wide outbreak for all versions of the model in Table 4.

Table 4 Summary statistics for the number of infected agents across model runs by intervention, including the confidence interval for the mean. The first intervention is vaccinations with no school closures and the next three rows give results for the schools that are closed in the simulation to stop the spread of measles.

Intervention	1st Quartile	Median	Mean	3rd Quartile	Max
Vaccinations Only	1	4	27.01 (23.95, 30.07)	44	492
Drumkeeran	1	4	43.47 (38.6, 48.4)	45	4253
Manorhamilton	1	3	25.57 (22.7, 28.5)	39	652
Drumkeeran and Manorhamilton	1	2	23.62 (21.0, 26.3)	37.25	210

From Table 4 we can see that there are some distinct differences between the results across the different interventions. In particular when looking at the mean value across the runs the mean number of infected agents is significantly lower when schools are closed in Manorhamilton and when schools are closed in both Drumkeeran and Manorhamilton. However, the maximum value for the total number of infected agents across the 300 runs is much higher for when the model runs with schools closing in Drumkeeran versus when schools close in Manorhamilton or both Drumkeeran and Manorhamilton. This extremely high value, 4,235, is an outlier in the Drumkeeran school closure runs and is what is driving the mean of those runs up. Looking at the medians, a statistic that is resistant to outliers, we can see that there is almost no difference between the runs.

Additionally we look at how the outbreak spreads beyond the initial town within the network. The first measure that we look at is the number of cases of the disease outside of the initial town. Table 5 shows the percent of runs that lead to an outbreak (two or more infected) anywhere in the model and the percent of runs that have at least one agent infected outside of Drumkeeran. The results show that for all intervention strategies the disease does not spread outside of Drumkeeran on every run. The p-values reported in Table 5 were calculated by comparing the outcomes of each of the school closure intervention strategies to the vaccination only strategy in terms of the number of runs when at least one agent outside of Drumkeeran is

infected. From the results we can see that there is no statistical difference between the runs when there is at least one agent infected outside of Drumkeeran for the model with only vaccination, the model where schools are closed in Drumkeeran or the model where schools are closed in Manorhamilton. However, we do see a smaller result for the model where schools are closed in both Drumkeeran and Manorhamilton that is significant at the 10% level.

Table 5 A comparison between the percent of runs that lead to an outbreak (2 or more infected agents) and the percent of runs where at least one agent is infected from outside of Drumkeeran the initial town and the 95% confidence intervals for those and the P-value comparing the results to the vaccination only model.

Intervention	Outbreak	Outside Drumkeeran	P-value
Vaccinations Only	63.3 (57.9, 68.8)	49.7 (44.0, 55.3)	-
Drumkeeran	62.0 (56.5, 67.5)	50.3 (44.7, 56.0)	0.95
Manorhamilton	61.7 (56.2, 67.2)	47.3 (41.7, 53.0)	0.61
Drumkeeran and Manorhamilton	57.0 (51.4, 62.6)	42.7 (37.1, 48.3)	0.10

In the second experiment we look at closing schools based on their in degree centrality. Schools initially close down in the town that the outbreak starts, Drumkeeran, then Drumkeeran and the town with the next highest centrality, Manorhamilton, then Drumkeeran, Manorhamilton and the town with the next highest centrality, Mohill. We look at the percent of runs where three or more agents are infected. The results for this experiment are found in Table 6. From the table we can see that when we close schools in two towns (Drumkeeran and Manorhamilton) there is a drop in the percent of runs that lead to an outbreak that is statistically significant at the 7% level but that when the school closure intervention is extended to more towns the percent rises and is no longer significantly different from the vaccination only model or the model where only schools in Drumkeeran are closed. This is an interesting finding but may be due to the fact that the students whose schools are closed do not change their behaviour in response to the outbreak. Instead of going to school they will treat the days off as if it were a weekend and thus will interact with each other potentially spreading the disease if an infected student decides to leave their home. For this network of towns and population distribution, closing schools in three towns may be the tipping point from which closing schools reduces the outbreak to a situation where closing schools does not have an effect.

Table 6 The percent of runs that result in three or more agents becoming infected when schools are closed by in degree centrality.

Additional Towns Closed	Percent of Runs	Confidence Interval	P-Value
Vaccinations Only	51.3	(45.7, 57.0)	-
Drumkeeran	51.3	(45.7, 57.0)	1.00
Manorhamilton	43.3	(37.8, 48.9)	0.06
Mohill	47.3	(41.7, 53.0)	0.37
Drumshanbo	51.3	(45.7, 57.0)	1.00
Carrigallen	49.0	(43.3, 54.7)	0.62
Ballinamore	54.0	(48.4, 59.6)	0.57

To determine if the results we found have to do with the centrality of the town and not just the number of towns that the schools are closed in we also look at closing schools progressively by distance to the town where the outbreak starts. Table 7

shows the results for the percent of runs that lead to three or more infections closing towns by distance. The results show that similar to when we close schools by centrality, there is an initial decrease in the percent of runs that have at least three agents infected but this is not as statistically significant as the drop when schools are closed based on centrality (the drop is not significant at the 10% level). This leads us to the conclusion that closing a second town might have a beneficial affect on reducing an outbreak but that closing schools in a town with high in degree centrality is more beneficial than closing schools in a town with the closest distance.

Table 7 The percent of runs that result in three or more agents becoming infected when schools are closed in towns by distance

Additional Towns Closed	Percent of Runs	Confidence Interval	P-Value
Vaccinations Only	51.3	(45.7, 57.0)	-
Drumkeeran	51.3	(45.7, 57.0)	1.00
Drumahaire	45.3	(39.7, 51.0)	0.16
Killanummery	46.3	(40.7, 52.0)	0.25
Drumshanbo	49.3	(43.7, 55.0)	0.68
Manorhamilton	49.0	(43.3, 54.7)	0.62
Leitrim	46.0	(40.4, 51.6)	0.22

Discussion

We have used a hybrid agent-based and equation based model to look at the effects of different interventions on the outbreaks created in our model. It is important to be able to show that our hybrid model can not just produce model results but can also show us how an outbreak will change when different intervention strategies are used. Because testing intervention strategies is where we will be able to learn the most about an outbreak and how to prevent and slow one that occurs.

Although the literature is undecided about the usefulness of school closure policies on lessening the severity of an outbreak, it is still a commonly used strategy. In this paper we tested different school closure strategies: closing schools in town with high in degree centrality and towns close in distance to the initial town. We found that there was a reduction in the outbreak when closing schools in the initial town and a second town but that selecting the second town using in degree centrality resulted in a larger reduction than selecting the second town using distance. This is likely because the high in degree centrality makes it more likely for an agent to commute to a town with higher centrality bringing the disease into the town. With the schools closed other agents will not become infected at school and bring the disease back to their own town thus reducing the spread. However, as we close schools in more towns, whether this decision is based on centrality or distance, we see that after a certain number of towns have their schools closed then closing schools in more towns actually reduces the overall reduction in outbreaks that the school closure policy achieves when only two towns have closed schools (i.e., resulting in a similar percent of runs leading to an outbreak as only closing schools in Drumkeeran or just vaccinating and closing no schools). This is likely because of an effect that is often cited as reason to not use a school closure policy, that the uninfected and asymptomatic students will still interact just outside of school and will still spread the disease. Introducing some level of restricted movement or self isolation for students who are not in school could lead to different results and greater reductions in outbreaks.

Our findings show that closing down the schools in the town where an outbreak begins might not have as much of an effect on reducing the outbreak unless schools in another town are also closed: in particular the schools in towns with the highest in degree centrality will result in the greatest decrease in the potential outbreaks. These findings are the first step in developing intervention strategies to reduce outbreaks based off of town centrality. Additional work can be done to look at the results for different counties. Showing that our results work for Leitrim is one thing but running the same tests for other counties in Ireland or regions in other countries will show that the findings are robust and could be applied anywhere. We could also look at different thresholds for closing down schools: instead of closing down when two agents are infected we could wait for a larger number of students to be infected before the schools close down to determine if this threshold has an impact on the results and which intervention strategies work the best. There is also the potential to look into changing the agents behaviours after schools close. Instead of moving as if it is the weekend agents could adjust their actions to prevent transmission knowing that there is an outbreak occurring. Similarly, in the current version of the model only students actions are changed when schools close down but the actions of adults whose children attend schools that closed down could also be adjusted. With such adjustment it might also be possible to calculate the economic impact of closing schools down versus letting an outbreak run its course without interventions. This could include the cost of paying teachers salaries while the school is closed and the cost for parents taking the days off from work compared to the cost of treating the number of agents who would be infected. This type of economic analysis is important because although closing down schools in two towns seems to have a beneficial affect on reducing an outbreak, in order to adopt such a policy it would need to be shown that the reduction in the outbreak was not outweighed by the cost of closing down the schools.

Further the work can be extended to show how school closure policies might differ based on the disease. For example, there might be other factors to consider with an emerging disease such as COVID-19 where there is not already a level of immunity in the population. While school closure polices for an infectious disease such as measles might be set into place in order to stop the spread of the outbreak completely, school closure policies for COVID-19 might be put into place to slow the spread of the outbreak and reduce the number of social contacts an individual has. Although the final number of infected in this case might not be reduced, if the outbreak happens over a longer period of time the maximum number infected at any one time is reduced leading to less of a strain on the health care system. In order to evaluate the ability of an intervention to flatten the outbreak curve, different measures would need to be looked at such as the maximum number of infected agents or the length of the outbreak.

Conclusion

We aimed to test interventions that look into stopping the outbreak from spreading out of the town of the initial case by looking at the schools in towns that have high levels of in degree centrality, and schools that are close in distance to the initial town and found that we were able to reduce the severity of the outbreak

spreading from Drumkeeran when we closed schools not only in the town where the outbreak begins but also in the town with the highest in degree centrality and the towns closest to Drumkeeran. From our finding we determined that closing schools in the town where the outbreak begins and then closing a second town is better at reducing the outbreak than just closing schools in the town where the outbreak initially occurs. In addition, we find that when selecting a second town in which to close the schools, it is more beneficial to choose the town in the region with the highest in degree centrality than it is to choose the nearest town.

Competing interests

The authors declare that they have no competing interests.

Author's contributions

Conceptualisation: EH JK; model development: EH; analysis: EH; writing and editing: EH JK; All authors have contributed to the manuscript and approved the final manuscript.

Acknowledgements

This work was partly supported by the Fiosraigh Scholarship Programme of the Technological University Dublin, and by the ADAPT Centre for Digital Content Technology which is funded under the SFI Research Centres Programme (Grant 13/RC/2106) and is co-funded under the European Regional Development Funds.

References

1. Boseley, S.: 142,000 died from measles last year, who estimates. *The Guardian* (2019)
2. Litvinova, M., Liu, Q.-H., Kulikov, E.S., Ajelli, M.: Reactive school closure weakens the network of social interactions and reduces the spread of influenza. *Proceedings of the National Academy of Sciences* **116**(27), 13174–13181 (2019). doi:10.1073/pnas.1821298116. <https://www.pnas.org/content/116/27/13174.full.pdf>
3. Spinney, L.: *Pale Rider: The Spanish Flu of 1918 and How it Changed the World*. PublicAffairs (2017)
4. Kelly, L.: Samoa measles outbreak kills 20, mostly children. *Reuters* (November 23, 2019)
5. Lee, B.Y., Brwon, S.T., Cooley, P., Porter, M.A., Wheaton, W.D., Voorhees, R.V., Stebbins, S., Grefenstette, J.J., Zimmer, S.M., Zimmerman, R., Assi, T.-M., Bailey, R.R., Wagener, D.K., Burke, D.S.: Simulating school closure strategies to mitigate an influenza epidemic. *Journal of Public Health Management and Practice* (2010). doi:10.1097/PHH.0b013e3181ce594e
6. Grefenstette, J.J., Brown, S.T., Rosenfeld, R., DePasse, J., Stone, N.T., Cooley, P.C., Wheaton, W.D., Fyshe, A., Galloway, D.D., Sriram, A., Guclu, H., Abraham, T., Burke, D.S.: Fred (a framework for reconstructing epidemic dynamics): an open-source software system for modeling infectious diseases and control strategies using census-based populations. *BMC Public Health* **13**(1), 940 (2013). doi:10.1186/1471-2458-13-940
7. Hunter, E., Mac Namee, B., Kelleher, J.D.: A taxonomy for agent-based models in human infectious disease epidemiology. *Journal of Artificial Societies and Social Simulation* **20**(3), 2 (2017). doi:10.18564/jasss.3414
8. Hunter, E., Mac Namee, B., Kelleher, J.: An open-data-driven agent-based model to simulate infectious disease outbreaks. *PLOS ONE* **13**(12), 1–35 (2018). doi:10.1371/journal.pone.0208775
9. CSO: Census 2011 Boundary Files. Date accessed 26.05.2016 (2014). <http://www.cso.ie/en/census/census2011boundaryfiles/>
10. Nelson, K.E., Williams, C.M.: *Infectious Disease Epidemiology Theory and Practice*. Jones and Bartlett Publishers (2007)
11. Nicoara, C., Zach, K., Trachsel, D., Germann, D., Matter, L.: Decay of passively acquired maternal antibodies against measles, mumps and rubella viruses. *Clinical and Vaccine Immunology* (1999)
12. CSO: Census 2016 place of work, school or college - census of anonymised records (powskar). Central Statistics Office (2017)
13. Rodrigue, J.-P., Comtois, C., Slack, B.: *The Geography of Transport Systems*. Routledge, Taylor and Francis Group, London (2006)
14. Thomas, J.C., Thomas, J.C., Weber, D.J.: *Epidemiologic Methods for the Study of Infectious Diseases*. Oxford University Press, USA (2001). <https://books.google.ie/books?id=m4IRDAAAQBAJ>
15. Hunter, E., Namee, B.M., Kelleher, J.D.: Degree Centrality and the Probability of an Infectious Disease Outbreak in Towns Within a Region (2019)