Article

Validating and Testing an Agent-Based Model for the Spread of COVID-19 in Ireland

Elizabeth Hunter *ID and John D. Kelleher *ID

ADAPT Centre, Technological University Dublin, Grangegorman, Dublin 7, D07 H6K8 Dublin, Ireland
* Correspondence: elizabeth.hunter@tudublin.ie (E.H.); john.d.kelleher@tudublin.ie (J.D.K.)

Abstract: Agent-based models can be used to better understand the impacts of lifting restrictions or implementing interventions during a pandemic. However, agent-based models are computationally expensive, and running a model of a large population can result in a simulation taking too long to run for the model to be a useful analysis tool during a public health crisis. To reduce computing time and power while running a detailed agent-based model for the spread of COVID-19 in the Republic of Ireland, we introduce a scaling factor that equates 1 agent to 100 people in the population. We present the results from model validation and show that the scaling factor increases the variability in the model output, but the average model results are similar in scaled and un-scaled models of the same population, and the scaled model is able to accurately simulate the number of cases per day in Ireland during the autumn of 2020. We then test the usability of the model by using the model to explore the likely impacts of increasing community mixing when schools reopen after summer holidays.

Keywords: agent-based model; epidemiology; infectious disease; simulation; COVID-19

1. Introduction

During an infectious disease outbreak, modeling can be an essential tool to help understand how a disease might spread and the possible impact of any interventions [1]. Modeling has been used to respond to the UK foot and mouth epidemic in 2001 [2], the H1N1 pandemic in 2009 [3] and more recently during the COVID-19 pandemic. Infectious disease modeling has shown to be an important part of many government responses. For example, models have been used as evidence for lockdowns in the UK and USA [4], models have also been used in the Irish response to the pandemic [5], and Australia has used models to understand the impacts of lifting restrictions [6].

Equation-based models, in particular, compartmental models, are the most common type of model used for infectious disease modeling. A compartmental model is made up of a set of differential equations. The simplest compartmental model is the SIR model that is made up of three compartments: susceptible (S), infected (I) and recovered (R). Variations of the model can include additional compartments, such as the SEIR model, which includes an exposed (E) compartment [7]. The population in a compartmental model is assumed to be homogeneous and well mixed [8].

While the homogeneous SEIR model is often used and is able to accurately predict infectious disease dynamics, in some scenarios, more detail is needed, and the heterogeneity of the population needs to be taken into account. One common method for this is to add additional compartments to the SEIR model that represent different cohorts of the population. This can be done for age groups [9] or vaccination status [10]. These models can play an important role in understanding how an infectious disease will spread when the heterogeneity of the population is important, such as when vaccinations are implemented by age group. Although a cohort SEIR model allows for heterogeneity in the mixing between compartments, there are some drawbacks. Each compartment is still homogeneous, and the model might not be able to capture the individual actions and variations in characteristics that drive a pandemic [8].

Agent-based models (ABMs) can be used when this greater level of detail is necessary. ABMs are a type of computer simulation made up of agents who interact with each other and their environment [11]. Those agents can be created so that they emulate a real population. This synthetic agent population can match the real population on different characteristics and geographic distribution, and agent networks can produce results due to the interactions between these factors [12].

There are a number of situations where there is an advantage to using ABMs compared to equation-based models. They allow us to create a baseline scenario where we can look at what would have happened if a certain intervention had not occurred, for example, if there was no lockdown, or if vaccinations were not introduced. Additionally, as ABMs are stochastic and are driven by agent decisions, each model run is slightly different, and this stochasticity gives a range of outcomes that could occur within the system. Because of these advantages, a number of ABMs have been developed to help respond to and better understand the COVID-19 pandemic [4,6,13–16]. The model we present in this paper was part of a suite of modeling tools developed by the Irish Epidemiological Modeling Advisory Group (IEMAG). IEMAG was a group formed to provide statistical and mathematical modeling support and advice to the chief medical officer and the national public health emergency team during the COVID-19 pandemic.

ABMs, however, can be computationally intensive, which can limit their usefulness when speed of results is necessary [17]. A number of studies have attempted to reduce the computing power while still retaining model fidelity. One method is using a hybrid agent-based and equation-based method designed to save computing power by switching between the two modeling types, depending on the number of infected agents in the simulation [18]. Alternatively, computing power can be saved by only simulating infected agents and modeling healthy agents as a property of the environment [19]. Another method of decreasing the amount of computing power needed to run a model while still increasing the population size is using scaling methods. Dynamic re-scaling is another method that scales agents when certain thresholds of infection are reached [13,16]. In this paper, we present an agent-based model for the Republic of Ireland with a population that has been scaled so that 1 agent represents 100 people. The model can be used to understand how changes in behaviors and movements impact the spread of COVID-19 in Ireland. In the next section, we discuss the different components of the model. We then discuss the validation of the model to show that the model results are trustworthy, and finally we test the scaled model, running an experiment looking at the impacts of different levels of community mixing when schools re-open after summer holidays, and present the result of the experiment to show the model's usability.

2. Materials and Methods

We use an ABM to model the spread of COVID-19 through Ireland. In this section, we provide a brief description of the model that is used in our work and then discuss the experiments run to validate and test the model.

2.1. Model Description

The model presented here is a version of a previous model [20] that has been scaled to simulate the entire population of the Republic of Ireland. It was created and implemented in the modeling environment, Netlogo [21]. The Republic of Ireland is a country in Northwestern Europe. In 2016, when the last Irish census was completed, there were 4,757,976 people in Ireland; in the scaled model, this would equate to approximately 47,580 agents. There are 26 counties that make up the country. Approximately 40% of the population of the country live in the region of the nation's capital city, Dublin. The model is an agent-based model that was created for the spread of measles in an Irish county [20] and has been adapted to simulate the spread of COVID-19 [22]. The model is made up of four main components: environment, transportation, disease, and society [23]. It uses census data from the Irish Central Statistics Office to create a population that matches the demographic

characteristics of the county at the small area level (small areas are the smallest geographic area over which census statistics are aggregated in Ireland, and contain between 50 and 200 dwellings) [24] and transportation patterns [25]. The scaled-up model presented here has the same four components and a similar structure. The largest difference between the two models is that the scaled-up model is designed to simulate the entire population and capture the interactions between counties. There are, however, other differences that are mainly meant to reduce the computing power necessary to run a county-level agent-based model. The more complex an agent-based model is and the more agents in the simulated population, the more computing power that is needed [17]. Thus, to simulate at a larger scale while still being able to produce timely results, we reduce the complexity of the environment matching population distributions at the county level instead of the small area level, and we scale the model so that each agent represents 100 similar people in Ireland. The following sections discuss the different model components used in the model in more detail as well as the model schedule, the initial conditions of the model and the interventions that can be implemented in the model. Further details including model parameters can be found in the model ODD [26].

2.1.1. Agents

At its heart, an agent-based model is made up of agents. The agents in this model represent people in Ireland and have a number of characteristics that define them. Agent characteristics fall into different categories. They have demographic information, such as age, sex and employment status; a set of social networks, including a family network and a work network; information on their disease status, such as if they are susceptible, exposed, infected or immune; and information on their location (home, school, work or the community). For more detail and for the full list of variables that define an agent, see the model ODD [26].

2.1.2. Environment Component

The model environment is made up of grid cells or patches in NetLogo. Thirty-one patches in the model are designated as counties (e.g., Leitrim County, and Cork County) or city and city/county areas (e.g., Dublin City, Cork City, Waterford City and County). These county and city labels are defined by the CSO and are referred to as county patches going forward. The number of primary and secondary schools in each actual county in Ireland can be found from the data from the Irish Department of Education. However, because we are scaling the model for 1 agent to represent 100 people, we reduce the number of schools in the counties. We do this by dividing the number of schools in the county by 50 (half the scale for agents to people) and rounding up. This gives us at least one primary and one secondary school in each county and also approximately reproduces the class and school network sizes that are found in the 1 to 1 scaling in the county model.

Agents will only occupy the county patches in the model and can move between them. Although all agents on a county patch will physically be coded in the same location, the agents keep track of their location within the county and will be at either home, work, school, or in the community. Agents will only come into contact with other agents in the same location as them in the same county. For example, an agent in the community will not be in contact with an agent at home. Additionally, all agents in the community on the same patch will not be in contact with every other agent in the community. Instead, a set of parameters derived using the POLYMOD contact matrix data [27] is used to determine the number of community contacts that an agent will have. Agents can move between county patches. Agent movement is discussed in more detail in Section 2.1.4. While on a patch, we can access certain information about the patch, including the number of agents on the patch and the physical distance to the other patches.

2.1.3. Disease

The disease component in the model is set to simulate the dynamics of COVID-19 and follows the dynamics of an SEIR model with the agents moving from susceptible to exposed then infected and finally recovered. In addition to the four stages of infection, we match the disease component of the model to the Irish population-based SEIR model [5]. In the model, individuals start in a susceptible component, then if infected, they move to an exposed component, and from the exposed component, individuals will move to one of two infectious components: asymptomatic or presymptomatic. If the individual is presymptomatic, they will then move to one of the following components: isolating, not isolated, waiting for a test, and tested. Once recovered, individuals will move to a recovered state. Thus in our model, an infectious agent can be in one of the following states: asymptomatic, presymptomatic, isolating, not isolated, waiting for a test and tested. When a susceptible agent comes into contact with an infectious agent, they have a chance of becoming exposed. Once exposed, the agent will stay exposed for a predetermined period of time before becoming infectious. This predetermined period of time is different for each agent. When an agent is exposed, the agent is assigned a length of time for their exposure period, using an exponential distribution with a mean of the exposure period taken from the literature. Agents will remain infectious for a predetermined period of time before recovering. Similar to the exposure period, this predetermined period of time is different for each agent. When an agent becomes infectious, the agent is assigned a length of time for their infectious period using an exponential distribution with a mean of the infectious period taken from the literature. Once recovered, the agents can no longer be infectious and cannot be re-infected. This assumption was made, as the risk of reinfection early in the pandemic was low [28], and the model was initially designed to look at short term outcomes of the pandemic (i.e., during a single wave). When an agent becomes infectious, the agent will either become asymptomatic or symptomatic. If they are symptomatic, they can either be isolating, not isolating, or waiting for a test. If waiting for a test, the agent will be waiting for a test for a predetermined time and then have tested positive. The infectious component that an agent is in (asymptomatic, isolating, etc.) determines how infectious an agent is. The base infectious rate is determined using the basic reproduction number, R_0, the average contacts an agent has, and the infection period [29]. The method for calculating the infectiousness of the agents and all the parameter values are discussed more in the model ODD [26]. Those who are presymptomatic, asymptomatic, and isolated have a reduction in their infectiousness. We implement these reductions to match with those in the SEIR model. The main disease parameters needed to initialize the model are the basic reproduction number, R_0, the length of the exposed period, the length of the pre-symptomatic period, and the length of the infectious period, and are taken from the literature.

Additionally, as there have been multiple variants of COVID-19 during the course of the pandemic, the model allows for agents to pass on specific variants. For example, if an agent is infected with the alpha variant, any agents they infect will also be infected with the alpha variant. As the variants that are widely circulated in the population tend to be more infectious, we also include a variant multiplier that adjusts the infectiousness of an agent for a specific variant.

2.1.4. Transport

The transportation component of the model is not altered compared to the transportation component in [20]. In the model, there are two drivers of agent movement within the model: scheduled movements determined by the time of day and the agent type and community movements determined from a gravity model. The schedule of movements and the parameters that define the community movements are defined further in the model ODD [26], but in this section, we provide a brief description of both, as well as an analysis of the contact patterns that are generated from the model and the process of updating the contact patterns to be more realistic using real contact data.

Scheduled movements define the movement of agents who are students or workers. On weekdays, these agents will move from home to school or home to work at certain times of the day and then will return home after work or school. Additionally, all agents will return home at a certain time of the day and will remain at home until the morning. Commuting patterns are determined using data from the CSO Place of Work, School or College—Census of Anonymity Records (POWSCAR) [25].

On weekdays, the movements of all agents who are not students or working and on the weekends, the movements of all agents are determined using a gravity model. A gravity model uses the characteristics of a location and the distances between locations to determine interactions between location pairs [30]. In the case of our model, the interactions between the location pairs is the probability that an agent at one location will move to the other location in the next time step. The idea behind using a gravity model for movements is that agents are pulled toward areas or counties that are closer to their current location and areas that have a high population density and pushed away from areas that are farther away and areas that have a low population density. Although not a perfect model of human transportation, it is a proxy for human mobility.

When an agent's movements are determined by a gravity model, they can be either at home or in the community. In reality, there are many locations within a community that agents might travel to, such as parks or shops; however, for simplicity, we only include a "community" location. While this makes the model slightly less realistic, the model is built so that not all agents in the community in the same county are in contact. Instead the contacts are determined by agent networks. If two agents in the community in the same county are in the same family network, they are more likely to be in contact than two agents in the community in the same school or work network who are more likely to come into contact than two random agents. The time that agents spend in the community and the likelihood of coming into contact with other agents in their networks was originally determined by parameterizing the model to match contact patterns and infection rates with the [12] model, where the agents moves in steps around a town and only comes into contact with another agent in the community if they are on the same patch. While this was done to preserve model fidelity when scaling up the model, to make the model more realistic, we adjust the parameters in the model so that the contact patterns simulated in the model match those from real-world studies of contact patterns. We use the contact patterns found in the POLYMOD study that examined the social contacts of people in eight European countries. The participants of the study recorded their contacts and locations of the contacts including home, work, school, leisure, transportation or other. As Ireland is not one of the eight countries, we use the data from the Netherlands to estimate the Irish contact rates [27]. The parameters determined from this analysis can be found in the model ODD [26].

2.1.5. Society

As we are simulating the spread of COVID-19 in the Republic of Ireland, we aim to match the characteristics of our agent population to the characteristics of the Irish population at the county level. Thus every county in the model has the correct portion of agents of the following characteristics: age, sex and economic status (student, working, retired, unemployed, etc). The counties also have the correct portion of households by type (single, couple, couple with children, etc.), and by number and ages of children (under 15, over 15 and both under and over 15). To make the contact networks in the model more realistic, we allow for agents to have an extended family network. This network connects two households one with agents over 65 and one with agents under 65. The reasons for creating these extended networks are two-fold. The first is to capture interactions and thus transmission between children and grandparents who act as carers. The second is to capture interactions between families during holiday periods, such as Christmas, when inter-generational mixing typically increases. Further discussion of the creation of these networks can be found in the model ODD [26].

From the 2016 census, the last Irish census collected at the time of model creation, there were 4,757,976 people in Ireland. An agent-based model with that many agents requires a large amount of computing time and power to run such that the results may take days or weeks. Thus we introduce a scaling factor into the model that equates 1 agent to 100 people. This greatly reduces the number of agents needed to simulate the population, from 4,757,976 to 47,580, which allows us to model the social and economic structure of Ireland without decreasing the level of detail in the model or taking an overly large amount of time to run the model.

Although we still keep the same population structure with the scaling factor, we are changing the total number of agents and need to make sure that the reduction in agents does not impact the model results. Section 2.2 discusses the scaling factor in more detail and the tests run to validate the changes introduced when scaling the model.

2.1.6. Schedule

The model is run on discrete time steps. Each time step represents two hours of the day, thus 12 time steps make up a day and 84 time steps make up a week. During "nighttime" hours, agents are not moving around the environment but are instead at home, and transmission can only occur between others in their household network. The model keeps track of the week of the year, as this determines when schools are open or closed for the summer and for winter holidays. From week 26 to week 34, schools are closed for the summer, and agents who are students do not attend school but instead move throughout the community as they would on weekends for the summer. Additionally, schools are closed down for weeks 51 and 52 to simulate the impacts of Christmas holidays. During two days in week 51, agents move to a household within their extended family network for the day and spend time to simulate family gatherings for the holidays.

2.1.7. Initial Conditions

To start the simulation the initial conditions for the disease component need to be set. As the society and environment are set based off of real census data, these do not change between model scenarios. For each scenario of the agent-based model that is run, we determine the number of agents who are vaccinated, exposed, infectious, and recovered. If vaccinations are included in the model, the number of initially vaccinated agents are determined first from the vaccination model discussed in Section 2.1.8. The start week of the model is used to select the number of agents in each age group that have been vaccinated. After vaccinations, the number of infectious, exposed and immune agents are determined. In order to run the model, there needs to be at least one agent who is either exposed or infectious. A given number of agents, determined by the user, who are not fully vaccinated are assigned to be sick. If variants are included, of those agents who are sick, a certain percentage of them are assigned to being infected with each variant. Then a given number, again assigned by the user, of sick agents are assigned to be asymptomatic, isolating, not isolating, or waiting for a test. Following the assignment of the initially infectious agents, we determine the exposed agents. Half of the predetermined number of exposed agents are chosen to be agents in the households of those agents who are sick and infectious. This is to make the distribution of exposed agents more realistic, as it is more likely that an infectious agent will have infected a member of their close contacts versus a random agent outside of their networks. Finally, a certain number of agents are set to be immune. We first ask a number of agents who are infectious to set at least one agent in their networks to be immune and then if the number of desired immune agents is greater than those selected in the infectious agents networks, the additional immune agents are selected from random agents in the model.

2.1.8. Interventions

The above sections describe the agent movements and patterns when the model is not adapting its behaviors to a pandemic situation. This would be a likely scenario for

the spread of an endemic disease such as influenza or measles where the population does not greatly adjust their behaviors during an outbreak. However, to better capture the response to the COVID-19 pandemic, the ability to simulate interventions and behavioral adaptations is important. Thus, we have included a number of interventions in the model.

Lockdowns and School Closures

Prior to the introduction of vaccinations, one of the main measures used to reduce the spread of the virus was change in behaviors. The fewer contacts individuals have, the less the virus will spread. In the model, the agents' behaviors are changed in two ways: schools are closed, or lockdowns are introduced.

Primary and/or secondary schools can be closed: this requires all agents in the school to stay at home during school hours instead of attending school. A lockdown can be introduced that restricts agents' movements. During a lockdown, schools can be either opened or closed and the user can determine the percentage of agents who are working from home or no longer working as well as the reduction in movement around the community, compared to "normal" movements.

Contact Tracing

Ireland implemented a comprehensive contact tracing program during the pandemic, where those who tested positive were contacted by contact tracers and their close contacts were identified and referred to a test [31]. The contact tracing program successfully completed contact tracing for 96% of the cases notified in Ireland between 17 March 2020 and 30 April 2021 [31]. However, due to the changing nature of the pandemic, the number of close contacts identified, the length of time to complete calls and the number of calls made varied throughout the pandemic. Additionally, not all contacts who were identified through contact tracing attended testing [32].

In the model, when an agent tests positive, a probability determines if they take part in contact tracing. This probability is estimated using the percent of cases in Ireland where contact tracing has been completed for a given month as a proxy. If the agent participates in contact tracing, then their contacts are identified, and a probability determines for each contact identified if that contact also participates in contact tracing. The probability is estimated using the percent of contacts that receive a test as a proxy. If an agent who is a contact of an infectious agent does participate in contact tracing, the agent isolates for a set number of days. In the first year of the pandemic, this isolation period was set to 14 days, as this was the suggested time to restrict movements when identified as a close contact.

Contact tracing in the model can be turned on and off, and the parameters defining contact tracing (probability of a case participating, probability of a contact isolating and the number of days a contact isolates) can be adjusted to match what occurred at different times during the pandemic or to investigate potential scenarios.

Vaccinations

Vaccinations can also be turned on or off to investigate their impact on the pandemic. We include vaccinations in the model, but allow the user to turn vaccinations on and off. If vaccinations are turned on, then a vaccination spreadsheet is used as input to the model that provides information on the number of individuals by age group who have been vaccinated, who have successfully been protected against severe disease, who have successfully been protected against symptomatic disease and who have successfully been protected against asymptomatic disease. There are also different levels of protection against the two main variant, alpha and delta, that were circulating in Ireland in 2021. Thus an individual could be protected against asymptomatic disease from the alpha variant but only against symptomatic disease from the delta variant. This was calculated based off the supply of COVID-19 vaccines to Ireland and the predicted efficacy of the different vaccines supplied.

2.2. Model Validation

The model presented in the previous sections is a scaled version of the Irish county agent-based model [20]. Although it is based off a previously validated model, we made fundamental changes to the society component of the model with the introduction of the scaling factor and altered the level of detail in the environmental component. Thus, it is necessary to validate the new scaled-up version of the model.

To validate the model, we follow the framework laid out in [33]. Agent-based model validation has three main steps: cross validation, where we compare the model output to a previously validated model, sensitivity analysis, and comparing model output to real data.

2.2.1. Cross Validation

The first step of validating an agent-based model is using cross validation, where the model output is compared to the output from a previously validated model. To show that the scaling factor does not have a major impact on the output of the model, we run tests comparing the output of the scaled model for a single county to the output of the original county model output presented and validated in [20]. The idea is that in the scaled-up model, the output from a single county should not change from the county model if we consider the county in isolation. In this cross-validation experiment, we are not aiming to compare the results to real data, which is done later in the validation, but to show that the assumptions we made to scale the model do not impact the output. Thus, the county model is a good benchmark for cross validation because it serves as the base model that was scaled up to obtain the scaled Irish country model. In the county model, 1 agent equates to 1 person. The model is run for Leitrim in the original county-level model, and also in the scaled version, where we equate 1 agent to 100 people. This means that in the original county model, there will be approximately 36,000 agents and in the scaled model there will only be 360, with each agent representing 100 agents. The initial conditions of the model will be adjusted accordingly. The model is run with no interventions in place. We study three different scenarios with different initial conditions: the first starts with 300 agents infected in the county model and 3 in the scaled model, the second has 1000 agents infected in the county model and 10 in the scaled model, and the third starts with 10,000 agents infected in the county model and 100 agents infected in the scaled model. We do not include any exposed or immune agents in the initial conditions for any of the three scenarios. For both the county model and the scaled model, we run the model 30 times. The reason why we run the model multiple times is that agent-based models are stochastic in nature, with each model run producing different results. The stochasticity is introduced in a number of ways: agents sample from distributions to determine the number of days they are exposed before becoming infectious and the number of days before they recover. Agents will also make decisions that determine their actions, which affect their contacts. Thus, there is a question of how many times should the agent-based model be run to accurately capture the true results of the model. Too few runs and the model results might not be accurate, and too many and it could waste computing time. To determine the number of runs necessary, we use the method outlined in [33]. The method looks at the size of a confidence interval around a statistic produced by the model to determine the number of runs necessary to account for the stochasticity in the model. We choose to look at the R_e statistic as the size of R_e can help to determine the speed at which the outbreak is spreading. We aim to pick the number of runs where the size of the confidence interval around R_e is determined to be stable and small enough. The methodology in [33] does not specify an exact metric to determine when the confidence interval is small enough but leaves it to the modeler to make the decision. Using this method, we determine that 30 runs are enough to capture the variability in the model.

2.2.2. Sensitivity Analysis

A sensitivity analysis is performed to determine if changes in the model inputs impact the model output as expected. In our sensitivity analysis, we look at changes in agent

behavior and determine how increased community mixing impacts the number of infectious agents. It is expected that as an agent mixes more in the community, there will be more cases. If we do not see an increase in cases as mixing increases, there is likely something fundamentally wrong with the model.

For the sensitivity analysis, we run four different scenarios, looking at how changes in agent movement impact the model results. In all scenarios, we start with 156 infectious agents, and because of the scaling factor, this equates to 15,600 people. Eighty-eight of those infectious agents are asymptomatic, ten are isolating, nine are not isolating and twenty-one are isolating and have tested positive. There are 92 agents exposed to the virus and 5784 agents who have recovered. These initial conditions are similar to the situation in Ireland prior to the start of the school year in 2020. Vaccinations are included in the model, schools are not open, and there is no contact tracing. Each scenario starts with movement at a 50% reduction from the pre-pandemic movement. In order to account for any model burn-in phenomenon, in this experiment, in each run, we allow two weeks of the simulation to pass before we include any changes in mixing. Four different mixing scenarios are considered: (1) agents stay at a 50% reduction from the pre-pandemic movement, (2) agents reduce their movements by 66% of the pre-pandemic movement, (3) agents increase their movements by 33% of the pre-pandemic movement, and (4) agents return to the pre-pandemic movement.

2.2.3. Comparison to Real Data

After cross validation, the next step to validate an agent-based model is to compare the output of the model to real data. To do this, we simulate the time between February and December 2020 in Ireland and compare the average model output across 30 runs to the real number of cases per day in Ireland. We start the model from 1 February 2020. There were no cases notified in Ireland until 29 February 2020; however, starting the model prior to the initial cases allows for a model burn-in period. At the start of the model runs, there are three infectious agents: one is asymptomatic, and one is presymptomatic. Two agents are exposed but not yet infectious, and all other agents are susceptible. We match all agent movements and intervention strategies to those that were in place in Ireland during that time period. Initially, schools are open and agents are mixing at regular pre-pandemic levels. In the third week of March, schools are closed. In the fourth week of March, a lockdown is introduced; agents reduce their mixing to approximately 50% of pre-pandemic levels and 70% of agents who would normally be working are either working from home or not working. In April„ agents reduce their mixing levels again to about 17% of pre-pandemic mixing. Lockdowns and other restrictions start to lift by June, and thus agents increase their mixing to 33% of pre-pandemic levels in June. In the model, schools reopen in September, and agents remain at the same community mixing level. In the beginning of October, agents increase their mixing back to approximately 60% of pre-pandemic levels; however, movements are reduced to 33% of pre-pandemic levels in the last week of October, as a new lockdown is introduced in Ireland. At the start of December, the lockdown is lifted, and agents increase their movements to near pre-pandemic levels. Table 1, gives a summary of the dates that interventions were implemented and when agent mixing changed.

Vaccinations are not included in the model, as vaccinations did not begin until late December 2020. However, we do include contact tracing, as this was an important part of the response to the pandemic and will impact the case numbers.

The number of new cases per day in Ireland during the summer of 2020 was low, under 100 cases. This can be a problem with the scaling factor of the model. Fewer than 100 cases equates to less than 1 agent. This resulted in a number of runs where COVID-19 would have gone extinct in Ireland. As this did not happen, we included a scenario, where if the number of COVID-19 cases in the model was 0, a new case would be imported into the model. This would represent a case that occurred due to travel.

Table 1. Interventions and changes in agent mixing implemented in the model.

Date	Lockdown	School Closure	Vaccinations	Reduction in Mixing
1 February 2020	No	No	No	Normal
12 March 2020	No	Yes	No	Reduced to 50%
27 March 2020	Yes	Yes	No	Reduced to 33%
20 April 2020	Yes	Yes	No	Reduced to 17%
1 June 2020	Yes	Yes	No	Reduced to 33%
29 June 2020	No	Yes	No	No Change
1 September 2020	No	No	No	No Change
1 October 2020	No	No	No	Increased to 60%
21 October 2020	Yes	No	No	Increased to 33%
1 December 2020	No	No	No	Normal Mixing

2.3. Model Testing

Once a model is validated, it is important to test the model and understand if it can help the user to learn something useful about an infectious disease outbreak or pandemic scenario. Thus, the final step in the validation and testing framework is to test the model and show model usability [33]. Unless the user is able to learn something about the system from the agent-based model, it is not a useful model.

To test our model, we run an experiment to better understand how changes in movements can interact with other interventions or the lifting of measures and impact the pandemic in the short term. We perform 30 runs of two scenarios of the model that focus on the period of time when schools reopen in the autumn: (1) schools re-open and community mixing remains the same at a 50% lower rate of mixing than pre-pandemic levels (2) schools re-open and when schools reopen community mixing returns to pre-pandemic levels. We run the model for the equivalent of 60 days, starting a month before schools open and running for a month after schools open. The model is only run for a month after schools open to determine the immediate and short term impacts of opening schools. These are the impacts that will likely occur before an intervention is put in place. For each set of runs, we compare the total infected agents at a given time step and the number of newly infectious agents per day. As we did in the sensitivity analysis, we expect that increased movement and opening schools will lead to an increase in cases. Understanding the magnitude of the increase is important, and we can learn this from the model; however, simpler infectious disease models, such as the SEIR model, will also be able to model this increase in cases. To emphasize the importance of using an agent-based model and the additional information we can learn from a model that simulates at the individual level, we also look at the location of the infection (home, school, work or the community) and the age groups of those infected, and discuss possible implications from the model output.

The initial conditions for our model testing experiment are the same as those used in the sensitivity analysis.
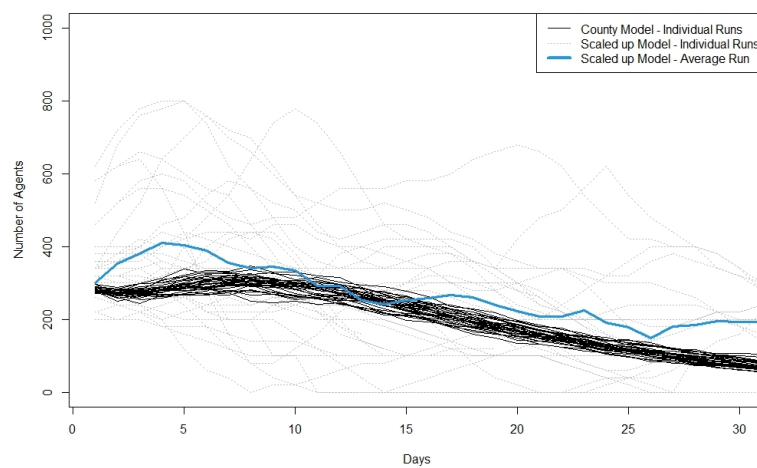
3. Results

The results we present in this section are from two distinct experiments. The first set of results presented are from a validation experiment designed to show the validity of the model. The second set of results are from an experiment designed to test the model to show its potential usefulness.
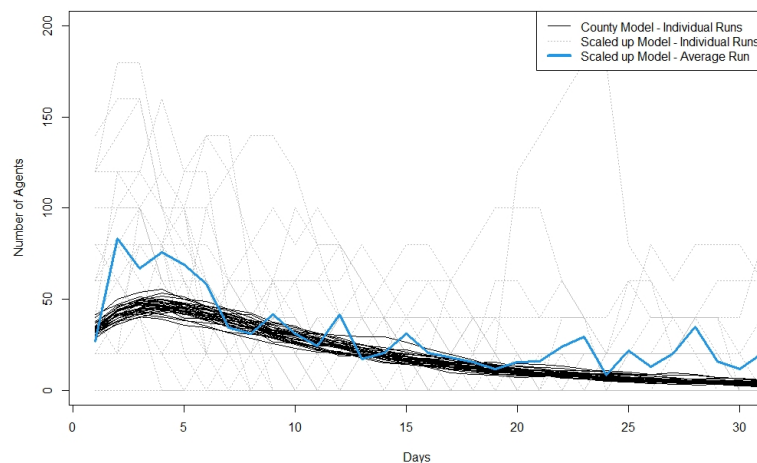
3.1. Validation

In the following sections, we discuss the results from the experiments run for each part of the agent-based model validation framework (cross validation, sensitivity analysis, and comparison to data).

3.1.1. Cross Validation

Looking at the cross validation results, we first start with 300 infected agents, and in the scaled version we start with 3 infected agents. Figure 1a,b show the total number of agents infected at a given time point and the number of newly infectious agents each day in both models for individual runs and the average across all runs for the scaled model. Looking at the figures, we can see that the scaled model produces a much greater variation in results than the county level model in both total infectious and newly infectious per day. This is likely due to the scaling factor in the model, which means that only three agents are infected at the start of the scaled up model. When few agents are infected, the individual actions of those agents play a larger part in the course the outbreak takes. It is important to note, however, that even though the scaled-up model produces more variation, the runs from the county level model are within the range of the runs from the scaled-up model and are close to the average across the runs from the scaled-up model.
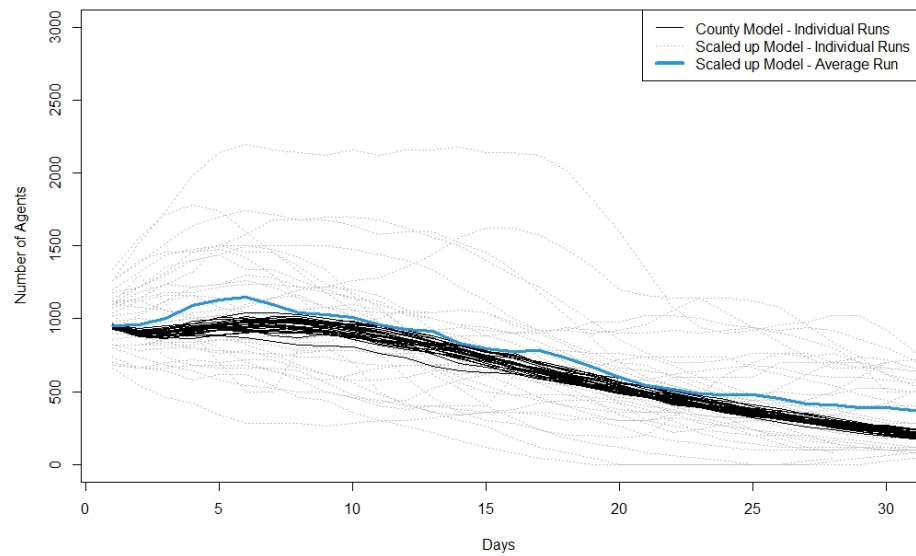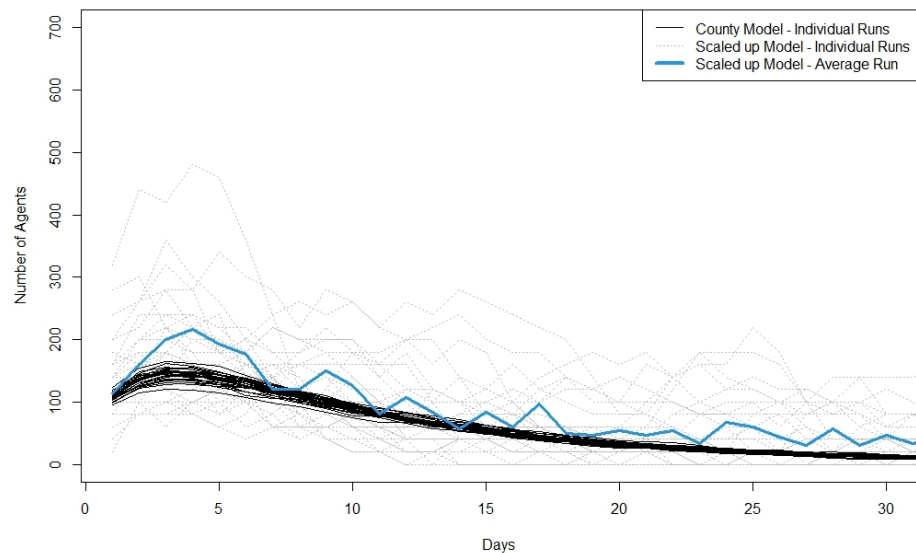

(**a**) Total Infected

(**b**) Newly Infectious per Day

Figure 1. Plots showing cross validation results when 300 agents are infected at the start of the run in the original model and 3 agents are infected at the start of the run in the scaled model. (**a**) shows the total number of infected agents per day and (**b**) shows the newly infectious agents per day. For the scaled model the *y*-axis is the number of infected ×100. We also plot the average value across the 30 runs of the scaled model. Best viewed in color.

We then compare runs for scenarios with 1000 agents infected in the county model and 10 agents infected in the scaled up model. Although there is still significant variation in the scaled model compared to the county model, we see less variation in model runs than when we started with only three agents infected. Figure 2a,b plots for this second scenario the total infectious and the new infectious agents per day, respectively. We see that in the first approximately 5 days, the scaled model has higher peaks for new infectious cases compared to the county model and a slightly higher average for total. However, by 10 days, the average curve for the scaled model appears to match well with the runs for the original county model.
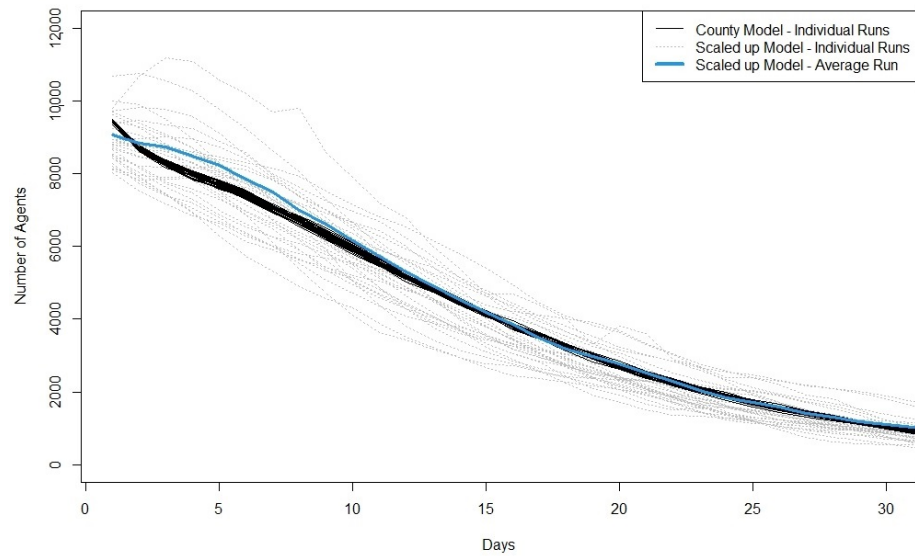

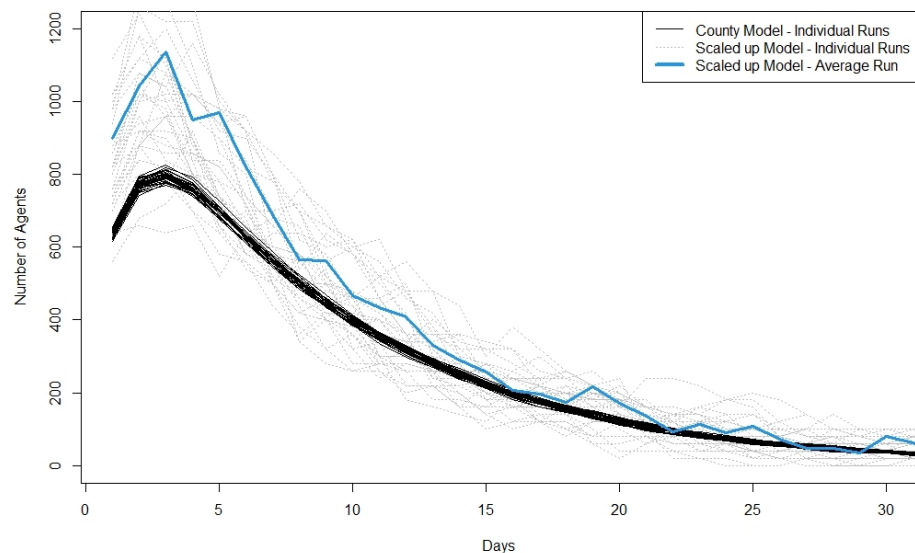(**a**) Total infected.

(**b**) Newly infectious per day.

Figure 2. Plots showing cross validation results when 1000 agents are infected at the start of the run in the original model and 10 agents are infected at the start of the run in the scaled model. (**a**) shows the total number of infected agents per day and (**b**) shows the newly infectious agents per day. For the scaled model the *y*-axis is the number of infected ×100. We also plot the average value across the 30 runs of the scaled model. Best viewed in color.

Finally, we run scenarios with 10,000 agents infected in the original county model and 100 agents infected in the scaled model. Figure 3a,b plots the total infectious and the new infectious per day for this scenario. Here, we see a much greater match between the two models. Although there is a greater variation in the scaled model, the trajectories for total infected match well across the simulation. Similar to scaling from 10 to 1000 agents, we still see a higher peak of new infectious cases in the scaled model compared to the county model, but by day 10 of the simulation, the average curve for the scaled model is much closer to the runs of the county model.

(**a**) Total infected.

(**b**) Newly infectious per day.

Figure 3. Plots showing cross validation results when 10,000 agents are infected at the start of the run in the original model and 100 agents are infected at the start of the run in the scaled model. (**a**) shows the total number of infected agents per day and (**b**) shows the newly infectious agents per day. For the scaled model the *y*-axis is the number of infected ×100. We also plot the average value across the 30 runs of the scaled model. Best viewed in color.

As the average runs for the scaled model appear to match well with the county model, this suggests that even though we see more variation in the scaled model any statistics

determined from the model output (total infected, new cases, exposed, recovered, R_f, etc.) will likely tend toward the statistics from the county model, meaning that we do not lose any information when we use the scaling factor. Across the different levels of scaling, we do consistently see that there tends to be a higher peak in newly infected cases in the first 5 days of the simulation. This might signify that there is a burn-in period with the scaled model that should be noted in any analysis of model output.

Our results show that for the first week to two weeks, the scaled model produces similar results for a single county to the original county-level model. However, the scaled model produces more varying output from each run when a smaller number of agents is infected. We think that when the number of infected is close to the number used in scaling (i.e., if the scale is 1 agent to 100 people and the initial conditions are that there are 100 people infected scaled down to 1) then the model will not perform as well.

3.1.2. Sensitivity Analysis

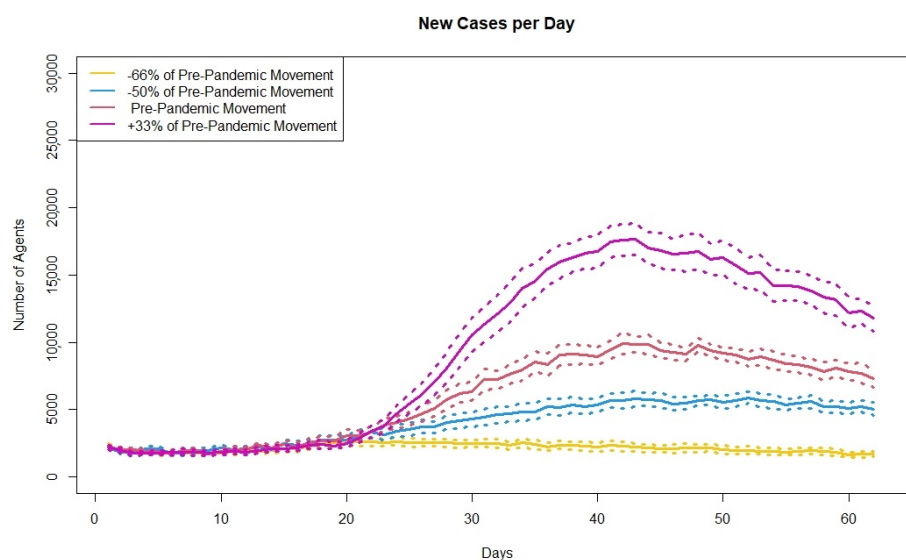Figure 4 shows the new cases per day for each of the four scenarios.

Figure 4. Plot showing the average simulated new cases per day across 30 runs for four different levels of movement. Dashed lines represent upper and lower bound of confidence intervals.

As the model output changes as expected, with the lowest number of new cases corresponding to the largest reduction in movements and then cases increasing as the agents level of movements increases, the model is considered validated through sensitivity analysis.

3.1.3. Comparison to Real Data

Figure 5 shows the average new cases per day across 30 simulation runs and the real Irish case counts during the same time period. Real cases are taken from the COVID-19 HPSC detailed statistics profile published by Ordnance Survey Ireland [34]. From the start of the model until April, the case counts are higher than the real cases and is likely due to a burn-in period of the model. During the summer months, from June to August, the case counts for the model are also higher than the real cases. This is likely due to the scaling of the model. From June to September, the number of new cases per day in Ireland is under 100, which is less than one agent with scaling. However, in order for the model to simulate the pandemic, there needs to be more than one agent infected and this likely leads to the difference in cases during the summer of 2020. However, we see a good match in the cases from October through December. Figure 5 also shows the confidence intervals around the average new cases per day. Although the average number of new cases per day matches closely with the real cases, the confidence intervals are relatively wide at the peaks. This

is likely due to the fact that the case counts are around 1000, and as we saw in the cross validation, the scaling factor leads to large variations between runs when there is a smaller number of cases. However, based on the comparison between the average simulated cases and the real case counts and that the real cases are within the confidence intervals produced by the agent-based model we can consider that the model is validated through comparison to real data when there are 100 or more cases per day.
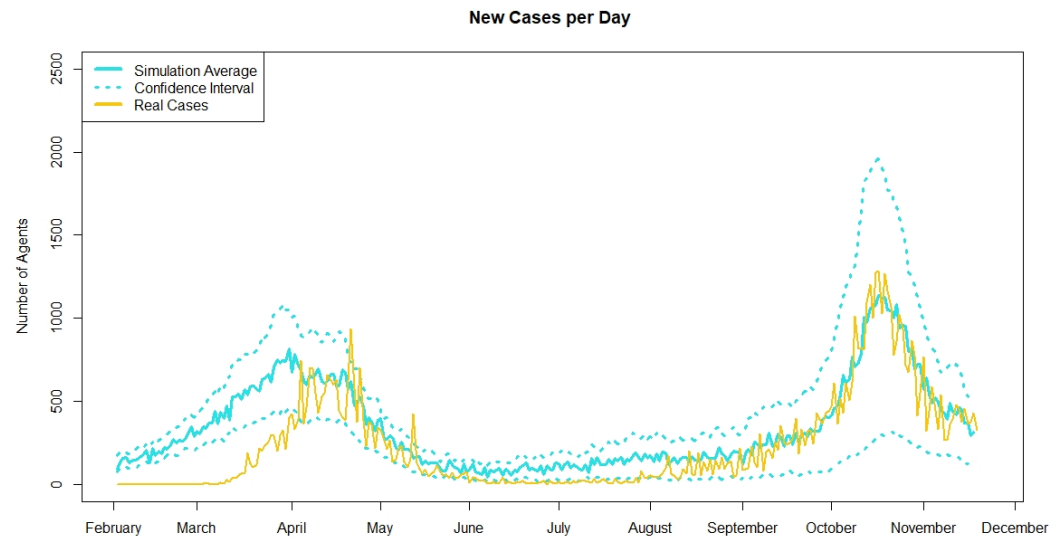

Figure 5. Plot showing the average simulated new cases per day across 30 runs in blue and the real Irish case counts in yellow. Upper and lower bounds of confidence intervals for the average simulated new cases are represented in blue dashed lines.

3.2. Testing

The previous sections discuss the scaled agent-based model for COVID-19 spread in Ireland and the validation of the model [33]. To show how the agent-based model is useful and can be used to better understand the pandemic we run an experiment that looks at opening schools after a period of closure, such as summer break, and the impact different levels of community mixing and return to work have on case numbers. During any infectious disease outbreak it is important to understand how changes to the current mixing and movement patterns will impact the number of cases. Re-opening schools post summer holidays is an important period of time with countries wanting to safely open the schools. This reopening will likely lead to an increase in cases but the size of the increase will be dependent on the movements of the other agents in the model. It is possible that as students go back to school, adults will see this as a return to normality and increase their movements as well.

To analyze the results of our model testing experiment, we first look at traditional output for infectious disease models, the number of new infectious cases per day, and then look at the output in more detail, analyzing the location where agents were infected and the age groups of the infected. Figure 6 shows the individual runs for the two scenarios and the average number of new infectious agents per day across the 30 runs in bold. In both plots the vertical lines indicate the day when schools reopen for the autumn term.

Looking at both plots, it can be seen that the increase in cases when schools reopen is much greater when the reopening coincides with an increase in agent movement in the community. The increased movement also appears to lead to a greater variation in individual runs as seen in Figure 6.
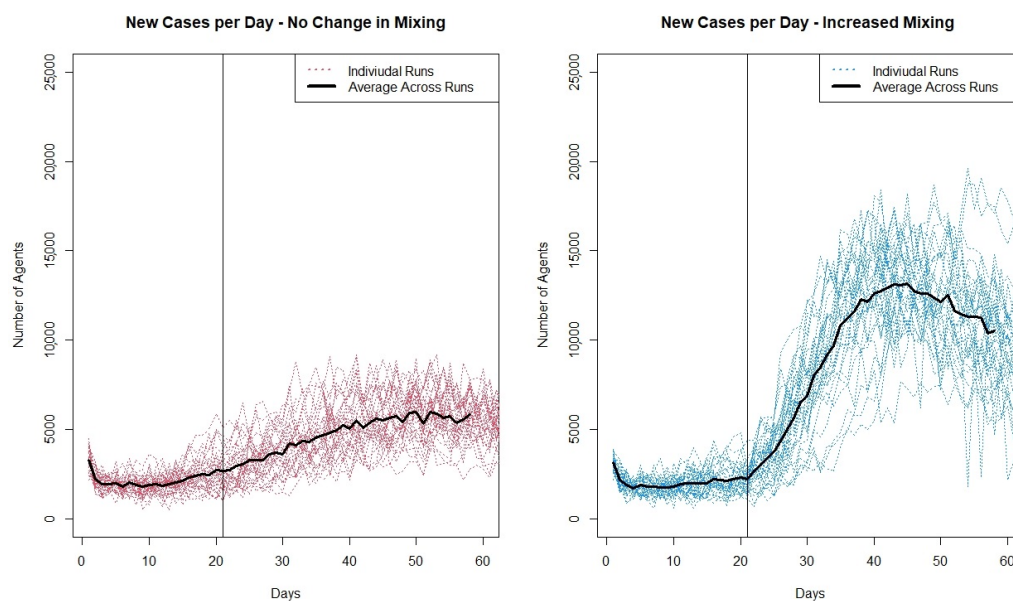
Figure 6. Plot showing the number of new infectious cases per day for the individual runs from the two scenarios.

This difference in increase in movement leads to a number of questions that the new cases per day metric cannot provide. While it is important to know that this increase in movement combined with the opening of schools leads to a greater number of cases, it might be important to understand where those cases are being infected. Are they all cases in the community? Or as community cases increase do cases that originate within the schools and home increase as well? Additionally, are certain age groups impacted more than others? These questions can help better understand the drivers of the spread of COVID-19 when schools reopen and can be better understood using the agent-based model. In the model agents keep track of if they were infected at home, school, work or in the community. Figure 7 shows the average number of infectious agents at a given time by their location of infection. When schools open but mixing does not increase, there is a slow, gradual increase in cases in schools, home and the community. There is little increase in the work setting, this is because in the model there are only a small percentage of agents who have returned to work in this scenario. An increase in agents returning to work may result in a larger increase in work cases. When school openings occur with an increase in community mixing, we still see little increase in cases at work, but the increased community mixing leads to a rapid increase in cases originating in the community and in homes. While there is a greater increase in cases originating in schools compared to when mixing does not increase, the greater levels of cases originating in schools only occur about 20 days after schools are open (the red lines for cases originating in schools only start to diverge at day 40). This suggests that there might be a delay in the impact of increased community mixing on cases in schools. There may be a threshold of cases that occur outside of schools before the higher cases spill over into the school setting.

The model output can also be used to look at the age groups of those infected. Figure 8 shows the average number of infectious agents by age group across the 30 runs for the two scenarios. The age groups are 0–9, 10–19, 20–29, 30–39, 40–49, 50–65, 66 plus. In the scenario where there is no change in community mixing when schools open while there is no considerable increase in the 20–65 year old age groups. However, when there is an increase in community mixing there is an immediate increase in cases in these age groups. In both scenarios, we see increases in the under 10 age group, the 10–19 age group and the over 65 age group. The increase in all three age groups is greater when community mixing increases. The increase in the over 65 group along with the student population is an interesting outcome from opening schools, especially as the cases in the other adult age

groups do not increase in the no-change-in-mixing scenario. One possible reason for the increase in the over 65 population as schools open is that the grandparent caring that is built into the model. This could lead to the direct increase in transmission between children and the elderly that is seen in the model output. When the increase in cases due to schools opening in the over 65 age groups is paired with an increase in community mixing, we see the highest cases in any age group in the over 65 group.
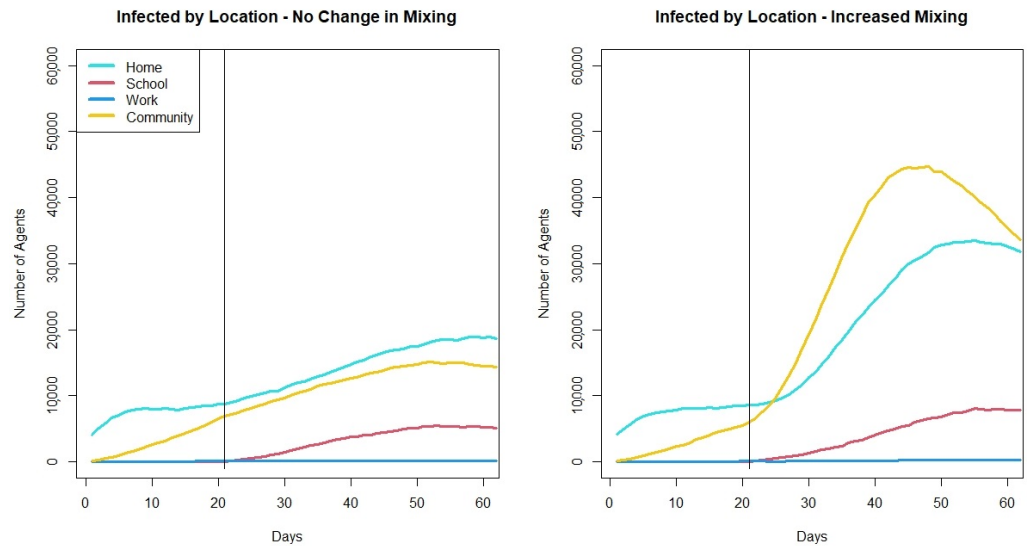

Figure 7. Plot showing average number of infectious agents at a given time based on their location of infection across 30 runs of the two mixing scenarios.
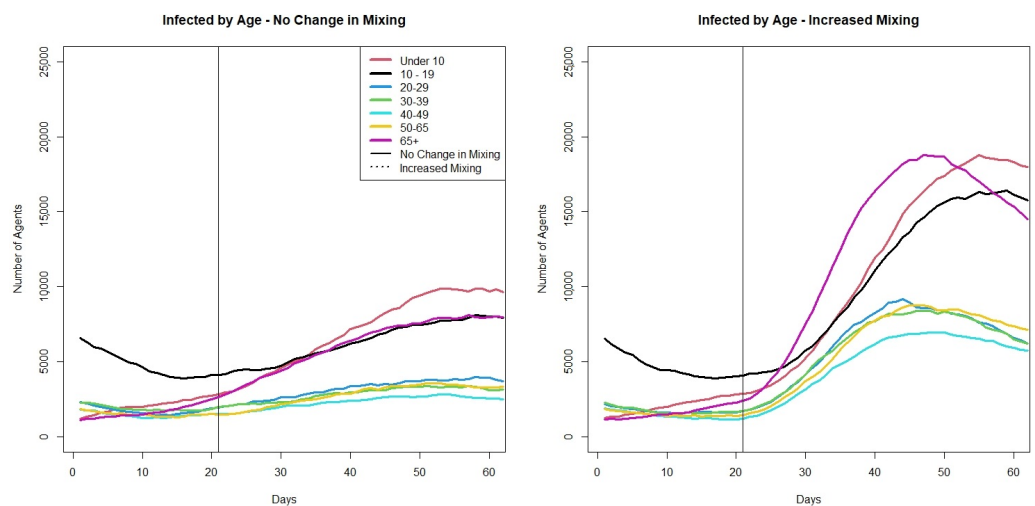

Figure 8. Plots showing average number of infectious agents at a given time based on their age across 30 runs of the two scenarios.

4. Discussion

Modeling the spread of COVID-19 has been a crucial piece in the pandemic response and agent-based models have played an important part of the modeling work done on COVID-19. The model presented in this paper is part of the literature surrounding models that were used during the COVID-19 pandemic. Models have been used to look at the effectiveness of contact tracing apps [15], or the impacts of interventions on ICU beds [14]. However, comparisons between model performance and model output are difficult to make, as there are no other agent-based models that simulate the spread of COVID-19

within Ireland, and it is not just the viral dynamics that allow for COVID-19 to spread as it has, but also the interactions of other factors, including the specific interventions a country put into place, the demographic makeup of a country, and the behaviors of the individuals. Additionally, as there is no set way to create an agent-based model for the spread of an infectious disease, parts of our model were designed to answer specific questions. For example, what was the impact of contact tracing or how does childcare by grandparents impact the number of cases in the over 65 population? Different models in the literature have included factors to address other questions. For example, Covasim looks at the health system capacity and thus has agents move from presymptomatic to mild, severe, and critical stages of infection [13]. We do not consider severity within our model but look at behaviors instead (not isolating, isolating, and waiting for a test). The Hoertel et al. model for France [14] also looks at health capacity and thus includes an agent's likelihood to have certain health conditions that might make COVID-19 more severe, such as obesity or diabetes. These factors are not included in our model, as we were not aiming to look at disease severity or health system capacity.

Agent-based models are computationally intensive and in some cases, work needs to be done to reduce the computing power needed to run the models. The scaling factor used in the model presented here reduces the number of agents needed to run the model which in turn decreases the memory needed to run the model. A scaling factor is used in other COVID-19 agent-based models, for example, the Covasim model [13] or the Thompson et al. model [16]; however, our scaling factor remains static while the factor used in the other models is dynamic. Although the scaling factor reduces the real-world fidelity of the model, we have shown through our model validation that the scaling of the model does not impact the average results. Thus the results produced by the model can help to better understand the COVID-19 pandemic and the impacts of different interventions.

The scaling factor does, however, lead to an increase in the variability in model output. The level of variability as well as how well the scaled model matches the output of an un-scaled model depends on the number of agents infected in the model. This suggests that there are thresholds above which the scaled model performs better. Future work on the scaling factor may involve investigating these thresholds in greater detail. For example, looking at different scaling factors, such as 1 agent to 50 people or 1 agent to 200 people, or alternative scaling methods, such as performing a clustering analysis on a set of individual characteristics, such as age, gender, family size, socioeconomic status, and geographic location, to find clusters of the population that can be represented by a single agent. This method would allow for variation in the 1 agent for X number of people scaling, such that 1 agent might represent 75 teenagers with one sibling in Dublin, but 1 agent might also represent 115 people who are 80 plus and living on their own in Leitrim. This clustered scaling would allow for a better representation of behavioral patterns among agent types and their importance to the course of the pandemic. For example, the 80 plus year old agents living in Leitrim might be greatly restricting their movements and thus not contributing as much to transmission, so their representation in the model would be down-weighted compared to the teenagers in Dublin who have higher levels of interactions and a greater contribution to transmission. A new method for scaling the results of the clustered model would need to be developed along with the model so as to not introduce artefactual results.

The results from our school opening experiment with the model shows how the agent-based model allows us to look at different scenarios and gain a deeper understanding of how changes in movements or interventions can impact an outbreak. With the agent-based model we can hold all other factors in the model constant and see how an increase in agent community mixing when schools reopen in the autumn impacts the number of cases. The agent-based model allows us to learn additional information about what might happen when schools open. When we do not have an increase in community mixing, we do not see a rapid increase of cases when schools open; as the model output when we do have increased community mixing does not show a large difference in cases originating within schools, this might suggest that opening schools might be relatively safe, and the resulting

increase in cases around school opening might be more impacted by the actions outside of schools. This is likely because in Ireland during 2020 and 2021, schools were a controlled environment where students were required to wear masks and interact with pods of other students. Outside of school, students do not have to restrict their interactions with those within their school pod, leading to more contacts and increased chances of infection and are not required to wear masks. However, in both scenarios, we see an increase in cases in the over 65 population, which is one of the most vulnerable populations. This might suggest that in such a scenario, with schools reopening after summer holidays, it might be worthwhile to include some preventative measures to reduce cases in the older age groups. We can also compare the output of the test on school openings and background mixing to what really occurred in Ireland in September 2020. Real cases went up very gradually in the first 30 days after schools re-opened but then there was a large increase in cases by the end of October 2020. Based on our modeling results, this might suggest that there was an increase in mixing but the increase was between the two levels tested. Another possible explanation could be that initially when schools re-opened, the rest of the population continued to restrict their movements out of concern that cases would increase but when cases did not immediately increase in September, community mixing increased, leading to a rise in cases later.

The results of our validation and test experiment show that the model is both validated and useful. This gives a level of confidence in the model output, showing that the model results can be used in a pandemic or outbreak situation to help understand the spread of the disease. The agent-based model is not only able to provide information on the numbers infected, but also on the location of infection and the age groups infected. This additional information may be helpful in shaping policies or implementing restrictions.

There are a number of limitations of our model. No model is an exact replica of real life, and all models make assumptions. Some of the assumptions we have made should be considered when analyzing the results of the model. One example is that in the schools we model, we do not assign adult agents to be teachers. This could impact transmission between children and adults. In the real world, there is potential for transmission between students and staff in a school setting; however, some studies of COVID-19 transmission in schools in 2020 showed limited transmission in schools [35,36]. Additionally, the model only simulates the spread of COVID-19 in the Republic of Ireland and not in Northern Ireland. Because we do not simulate the spread of COVID-19 across the entirety of Ireland, there might be some key transportation patterns between border counties that are not included in the simulation that played a part in driving the pandemic in those border areas. Another limitation is in our generated contact networks. There is no existing data set for Ireland that provides contact patterns between age groups. We use the Netherlands POLYMOD data [27] to determine the number of contacts agents have. While we think these data are an OK approximation for Irish contact patterns, we are introducing some uncertainty into the model with this assumption. Model uncertainty is also introduced in the choice of parameters determining the dynamics of COVID-19. The parameters we use come from the literature and, where possible, from literature on the spread of COVID-19 in Ireland. However, even in the literature, there is a wide range of values for the different parameters, for example, in a review from 2020, published values for R_0 were found to range from 1.5 to 6.49 [37]. The selections we made for the model will impact the model output. The model was also created with the idea of looking at the short-term impacts of behavior changes or other interventions, so some assumptions, including assumptions around reinfections, were chosen with this in mind. This will likely impact the outcomes of the model if run for longer terms and for periods of the pandemic beyond 2020; thus, these assumptions would need to be adjusted.

It is important to note that this model was created during the COVID-19 pandemic as a tool designed to simulate the spread of COVID-19 in Ireland to better understand what might happen in the short term. In this paper, the model is validated off 2020 data, and we test the model on an example where initial conditions surrounding the model are

based off of cases in 2020. For much of 2020 and 2021, Ireland was in strict lockdown, where non-essential workers were asked to work from home, and at times individuals were not allowed to go beyond 5 km from their homes. The lockdown interventions that can be implemented in the model reflect the strict nature of Ireland's lockdowns. At the time of publication, the situation is vastly different from that of 2020. While there are still many cases and deaths worldwide, vaccinations and previous infections have reduced the likelihood of contracting COVID-19 and the severity of the disease when contracted. Additionally, the strict lockdowns and travel restrictions have been lifted in most countries. While the model was validated on 2020 data, we included the ability to turn on a number of other other characteristics of the model, for example, multiple variants with different levels of infectiousness, or vaccinations that can take into account waning immunity. The model has and is evolving as the pandemic evolves. However, we feel that while having a model that is validated off 2020 data is a limitation, it is also important to have such a model and understand what we can learn from it to better prepare ourselves for the next potential pandemic.

5. Conclusions

Combined with other modeling techniques, such as a population level SEIR model, agent-based models can help to provide a greater picture of the future course of a pandemic and what the actual impact of interventions will be.

As it is possible that another pandemic will occur, it is necessary to take the lessons learned from the COVID-19 pandemic and create a better understanding of how modeling can help to respond to a pandemic situation. Additionally, looking at ways to make models that require high levels of resources, such as agent-based modeling, easier to run and use will improve their usability and uptake for future health crises.

Author Contributions: Conceptualization, E.H. and J.D.K.; methodology, E.H. and J.D.K.; software, E.H.; validation, E.H.; formal analysis, E.H.; writing—original draft preparation, E.H.; writing—review and editing, J.D.K. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data available on request.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

ABM	Agent-Based Model
IEMAG	Irish Epidemiological Modeling Advisory Group
SIR	Susceptible, Infected, Recovered
SEIR	Susceptible, Exposed, Infected, Recovered
CSO	Central Statistics Office
ODD	Overview, Design Concepts and Details

References

1. Department of Health. *Pandemic Flu—A National Framework for Responding to an Influenza Pandemic*; Technical Report; The Pandemic Influenza Preparedness Team Department of Health: London, UK, 2007.
2. Kao, R.R. The Role of Mathematical Modelling in the Control of the 2001 FMD Epidemic in the UK. *Trends Microbiol.* **2002**, *10*, 279–286. [CrossRef]
3. Simpson, C.R.; Beever, D.; Challen, K.; Angelis, D.D.; Fragaszy, E.; Goodacre, S.; Hayward, A.; Lim, W.S.; Rubin, G.J.; Semple, M.G.; et al. The UK's pandemic influenza research portfolio: A model for future research on emergin infections. *Lancet Infect. Dis.* **2019**, *19*, e295–e300. [CrossRef]
4. Ferguson, N.; Laydon, D.; Nedjati Gilani, G.; Imai, N.; Ainslie, K.; Baguelin, M.; Bhatia, S.; Boonyasiri, A.; Cucunuba Perez, Z.; Cuomo-Dannenburg, G.; et al. *Report 9: Impact of Non-Pharmaceutical Interventions (NPIs) to Reduce COVID19 Mortality and Healthcare Demand*; Technical Report; Imperial College: London, UK, 2020.
5. Gleeson, J.P.; Brendan Murphy, T.; O'Brien, J.D.; Friel, N.; Bargary, N.; O'Sullivan, D.J.P. Calibrating COVID-19 susceptible-exposed-infected-removed models with time-varying effective contact rates. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* **2022**, *380*, 20210120. [CrossRef]
6. Scott, N.; Palmer, A.; Delport, D.; Abeysuriya, R.; Stuart, R.M.; Kerr, C.C.; Mistry, D.; Klein, D.J.; Sacks-Davis, R.; Heath, K.; et al. Modelling the impact of relaxing COVID-19 control measures during a period of low viral transmission. *Med. J. Aust.* **2021**, *214*, 79–83. [CrossRef] [PubMed]
7. Hethcote, H.W. The Mathematics of Infectious Diseases. *Soc. Ind. Appl. Math. Rev.* **2000**, *42*, 599–653. [CrossRef]
8. Duan, W.; Fan, Z.; Gang, G., Zhang, P.; Qiu, X. Mathematical and Computational Approaches to Epidemic Modeling: A Comprehensive Review. *Front. Comput. Sci.* **2015**, *9*, 806–826. [CrossRef] [PubMed]
9. Andrade, J.; Duggan, J. An evaluation of Hamiltonian Monte Carlo performance to calibrate age-structured compartmental SEIR models to incidence data. *Epidemics* **2020**, *33*, 100415. [CrossRef]
10. Pang, L.; Ruan, S.; Liu, S.; Zhao, Z.; Zhang, X. Transmission Dynamics and Optimal Control of Measles Epidemics. *Appl. Math. Comput.* **2014**, *256*, 131–147. [CrossRef]
11. Adam, D. Special report: The simulations driving the world's response to COVID-19. *Nature* **2020**, *580*, 316–318. [CrossRef] [PubMed]
12. Hunter, E.; Mac Namee, B.; Kelleher, J. An open-data-driven agent-based model to simulate infectious disease outbreaks. *PLoS ONE* **2018**, *13*, 1–35. [CrossRef]
13. Kerr, C.C.; Stuart, R.M.; Mistry, D.; Abeysuriya, R.G.; Rosenfeld, K.; Hart, G.R.; Núñez, R.C.; Cohen, J.A.; Selvaraj, P.; Hagedorn, B.; et al. Covasim: An agent-based model of COVID-19 dynamics and interventions. *PLoS Comput. Biol.* **2021**, *17*, 1–32. [CrossRef]
14. Hoertel, N.; Blachier, M.; Blanco, C.; Olfson, M.; Massetti, M.; Rico, M.S.; Limosin, F.; Leleu, H. A stochastic agent-based model of the SARS-CoV-2 epidemic in France. *Nat. Med.* **2020**, *26*, 1417–1421. [CrossRef]
15. Almagor, J.; Picascia, S. Exploring the effectiveness of a COVID-19 contact tracing app using an agent-based model. *Sci. Rep.* **2020**, *10*, 22235. [CrossRef]
16. Thompson, J.; McClure, R.; Blakely, T.; Wilson, N.; Baker, M.G.; Wijnands, J.S.; De Sa, T.H.; Nice, K.; Cruz, C.; Stevenson, M. Modelling SARS-CoV-2 disease progression in Australia and New Zealand: An account of an agent-based approach to support public health decision-making. *Aust. N. Z. J. Public Health* **2022**, *46*, 292–303. [CrossRef]
17. Bobashev, G.V.; Goedecke, D.M.; Yu, F.; Epstein, J.M. A Hybrid Epidemic Model: Combining The Advantages of Agent-Based and Equation Based-Approaches. In Proceedings of the 2007 Winter Simulation Conference, Washington, DC, USA, 9–12 December 2007; pp. 1532–1537. [CrossRef]
18. Hunter, E.; Mac Namee, B.; Kelleher, J. A Hybrid Agent-Based and Equation Based Model for the Spread of Infectious Diseases. *J. Artif. Soc. Soc. Simul.* **2020**, *23*, 14. [CrossRef]
19. Montañola-Sales, C.; Gilabert-Navarro, J.F.; Casanovas-Garcia, J.; Prats, C.; López, D.; Valls, J.; Cardona, P.J.; Vilaplana, C. Modeling tuberculosis in Barcelona. A solution to speed-up agent-based simulations. In Proceedings of the 2015 Winter Simulation Conference (WSC), Huntington Beach, CA, USA, 6–9 December 2015; pp. 1295–1306. [CrossRef]
20. Hunter, E.; Mac Namee, B.; Kelleher, J.D. A Model for the Spread of Infectious Diseases in a Region. *Int. J. Environ. Res. Public Health* **2020**, *17*, 3119. [CrossRef]
21. Wilensky, U. *Netlogo: Center for Connected Learning and Computer-Based Modeling*; Northwestern University: Evanston, IL, USA, 1999.
22. Hunter, E.; Kelleher, J.D. Adapting an Agent-Based Model of Infectious Disease Spread in an Irish County to COVID-19. *Systems* **2021**, *9*, 41. [CrossRef]
23. Hunter, E.; Mac Namee, B.; Kelleher, J.D. A Taxonomy for Agent-Based Models in Human Infectious Disease Epidemiology. *J. Artif. Soc. Soc. Simul.* **2017**, *20*, 2. [CrossRef]
24. CSO. Census 2011 Boundary Files. 2014. Available online: http://www.cso.ie/en/census/census2011boundaryfiles/ (accessed on 26 May 2016).
25. Central Statistics Office (CSO). Census 2016 Place of Work, School or College-Census of Anonymised Records (POWSCAR). 2017. Available online: https://www.cso.ie/en/census/census2016reports/powscar/ (accessed on 5 September 2018).
26. Hunter, E.; Kelleher, J.D. An ODD-Protocol for Agent-Based Model for the Spread of COVID-19 in Ireland. *Reports* **2022**, *1*. Available online: https://arrow.tudublin.ie/sknowmanrep/1/ (accessed on 14 April 2022).

27. Mossong, J.; Hens, N.; Jit, M.; Beutels, P.; Auranen, K.; Mikolajczyk, R.; Massari, M.; Salmaso, S.; Tomba, G.S.; Wallinga, J.; et al. Social Contacts and Mixing Patterns Relevant to the Spread of Infectious Diseases. *PLoS Med.* **2008**, *5*, e74. [CrossRef]

28. Hansen, C.H.; Michlmayr, D.; Gubbels, S.M.; Mølbak, K.; Ethelberg, S. Assessment of protection against reinfection with SARS-CoV-2 among 4 million PCR-tested individuals in Denmark in 2020: A population-level observational study. *Lancet* **2021**, *397*, 1204–1212. [CrossRef]

29. Thomas, J.C.; Weber, D.J. Concepts of Transmission and Dynamics. In *Epidemiologic Methods for the Study of Infectious Diseases*; Oxford University Press: New York, NY, USA, 2001; pp. 61–62.

30. Rodrigue, J.P.; Comtois, C.; Slack, B. *The Geography of Transport Systems*; Routledge, Taylor and Francis Group: London, UK, 2006.

31. Martin, J.; Carroll, C.; Khurshid, Z.; Moore, G.; Cosgrove, G.; Conway, R.; Buckley, C.; Browne, M.; Flynn, M.; Doyle, S. An overview of the establishment of a national contact tracing programme: A quality improvement approach in a time of pandemic. *HRB Open Res.* **2022**, *5*, 12. [CrossRef]

32. Carroll, C.; Conway, R.; O'Donnell, D.; Norton, C.; Hogan, E.; Browne, M.; Buckley, C.M.; Kavanagh, P.; Martin, J.; Doyle, S. Routine testing of close contacts of confirmed COVID-19 cases—National COVID-19 Contact Management Programme, Ireland, May to August 2020. *Public Health* **2021**, *190*, 147–151. [CrossRef] [PubMed]

33. Hunter, E.; Kelleher, J.D. A Framework for Validating and Testing Agent-based Models: A Case Study from Infectious Diseases Modelling. In Proceedings of the 34th annual European Simulation and Modelling Conference, Toulouse, France, 21–23 October 2020. [CrossRef]

34. OSI. COVID-19 HPSC Detailed Statistics Profile. 2022. Available online: https://data.gov.ie/dataset/covid-19-hpsc-detailed-statistics-profile?package_type=dataset (accessed on 10 March 2022).

35. Larosa, E.; Djuric, O.; Cassinadri, M.; Cilloni, S.; Bisaccia, E.; Vicentini, M.; Venturelli, F.; Rossi, P.G.; Pezzotti, P.; Bedeschi, E.; et al. Secondary transmission of COVID-19 in preschool and school settings in northern Italy after their reopening in September 2020: A population-based study. *Eurosurveillance* **2020**, *25*, 2001911. [CrossRef]

36. Heavey, L.; Casey, G.; Kelly, C.; Kelly, D.; McDarby, G. No evidence of secondary transmission of COVID-19 from children attending school in Ireland, 2020. *Eurosurveillance* **2020**, *25*, 2000903. [CrossRef]

37. Liu, Y.; Gayle, A.A.; Wilder-SMith, A.; Rocklov, J. The reproductive number of COVID-19 is higher compared to SARS coronavirus. *J. Travel Med.* **2020**. [CrossRef]