The Anatomy of Simulated Covid-19 Epidemics in

Multiple Interacting Populations and Spaces

Ernie Chang MD PhD, Ken Moselle PhD

February 5, 2021

**INTRODUCTION**

CovidSIMVL is an agent-based simulation tool that models complex interactions of multiple spaces through populations that have different movement characteristics based on their roles in these Universes. The totality of the Universes makes up the Multiverse, whose complex behaviour can be studied in detail.

The simulation model is built on three abstract layers: the first is the action of the virus on agents, following the temporal dynamics of viral transmission in Xi,He [Nature Medicine Apr 2020, revised Aug 2020]. The second layer is the interaction between agents in states that progress from susceptible, to incubating, pre-symptomatic, symptomatic and then inert. The rules governing these interactions concern virtual movements in fixed arenas, and the size and mobility of agents are represented by Hazard Radius and Mingle Factor.

The third level of interaction is that of population cohorts that move between Universes based on a flexible schedule that encompasses different roles, mingle factors, age groups, risk factors, and family memberships.

The CovidSIMVL is a classic Monte Carlo Markov Chain (“MCMC”) microsimulation system, in which the dynamics of smaller populations are proxies for larger “In Real Life” (IRL) situations. As a simulation model, it facilitates the sandbox approach to scenario exploration, with a multiplicity of parameters and data collection tools, to capture epidemic processes and relationships difficult to obtain IRL.

In this paper, we present the definition of several cohorts who interact with one another through sharing common space and time according to defined schedules. The movement of agents, their viral load infection, transfer, growth In size are stochastic, relating to distributions for movements that are Pareto-like, and to standard statistical variations for random variables.

The fundamental concepts in CovidSIMVL, and related metrics of theta (time interval between infections), and Q-value for transmission trees, have been described in [medrxiv.org], and the relationships between intensity of an epidemic and the parameters of Hazard Ratio and Mingling Factor has been discussed in [medrxiv.org].

Appendix I is a brief summary of these parameters and metrics.

**GOAL OF THE MULTIVERSE MODEL**

The intent is to create a tool with flexible representations of populations that interact through shared spaces that may themselves have different characteristics. By doing so, and capturing the flow and temporal dynamics of infections in the entire Multiverse, we aim to gain insights into what populations are most vulnerable, and the timing and intensity of infections in the separate universes, and their relationships and influences on one another.

We do this by defining precisely an exemplar population of 100, described below, and their movements on a repeated 24-hour basis on schedules from one space to another, with varying durations for individuals, who may have different roles and mingling Factors in a different place.

Then we use the console.log which is a trace of every infection, describing who in what family group infected whom in what Universe at which time (referred to as “generation”, 24 making a day). The analysis will be described in the section below.

**METHODOLOGY**

We define nine Universes (below), and the demographics of a population of 100 persons, followed by the schedule in which these this population moves, and their activity In their new places. We then take the captured console.log, and find the time pattern of infections across the universes, and examine using the theta measure, the relative speed at which cohort members were infected (which is a proxy for their risk of infection within a period of time). Through the console.log, we derive the transmission trees, their Q values.

***UNIVERSE DEFINITION***

There are 9 Universes in this Trial, called U0 to U8. The names assigned to them are:

U0 Classroom 1

U1 Project/Lab Room

U2 Lunchroom

U3 Playground

U4 Classroom 2

U5 Teacher’s Lounge

U6 LTC (Long Term Care)

U7 High Mingle site (Bar, reception, party)

U8 HOME

HOME is a unique Universe, in that transmissions among agents who are in the HOME universe may only take place if they are members of the same family.

There is a resident population within Long Term Care who do not leave U6, and there are persons in U6 who are not members of any of the families pertaining to the rest of the population.

**POPULATION DEFINITION**

The population structure is defined with Person ID, roles and Family Identifications Fnn:



***The Family Structure***

The family structure is reflected above, and follows a representational mix of various parental, multigenerational mixes. The details are described in Appendix 2.

***The Schedule***

The daily schedule, which repeats is captured in the table below.



The specific details of how the Population.csv file represents the schedule is found in Appendix 3.

***Running a Simulation Trial in CovidSIMVL***



This is a screen capture of the Multiverse interface. The top left button “MV Toggle” would bring us to a Fixed Universe screen, in which Universes can be selected and parameters for Hazard Radius, Mingle Factor and Days of Symptomatic transmission set, as well as for vaccination schedules in other scenarios.

Appendix 4 is a brief description of the Fixed Universe interface, and the initial case file. VLfive.csv, which has 5 infected students to start the simulation.

During the execution of a simulation, the trial can be stopped with the yellow “AUTO” button which otherwise permits the system to run until TERMINATION when there are no further infective or potentially infective agents left, which may or may not be with survivors (susceptibles) that have not been infected.

Each panel shows the statistics of the occupants of a Universe in counts of Green (susceptible), Yellow (incubating), Blue (pre-symptomatic), Red (symptomatic transmitters), and Orange (inert) agents, and the charts in each panel show the counts of these agents each hour with Days shown on the X-axes.

The charts on the right show the counts for the total population of 100.

***Console.log***

The trace of each infection is produced in the the console.log and retrievable at any time, but we use the console.log at Termination for analysis. The log appears like this:



These are captured as single lines of text, and in Excel one can use the Data to Column function to break out the fields and end up with a data set that is easily sortable, by time (generation), by ID of Infected persons, and so on. In addition, transmission trees can be derived, as seen below.

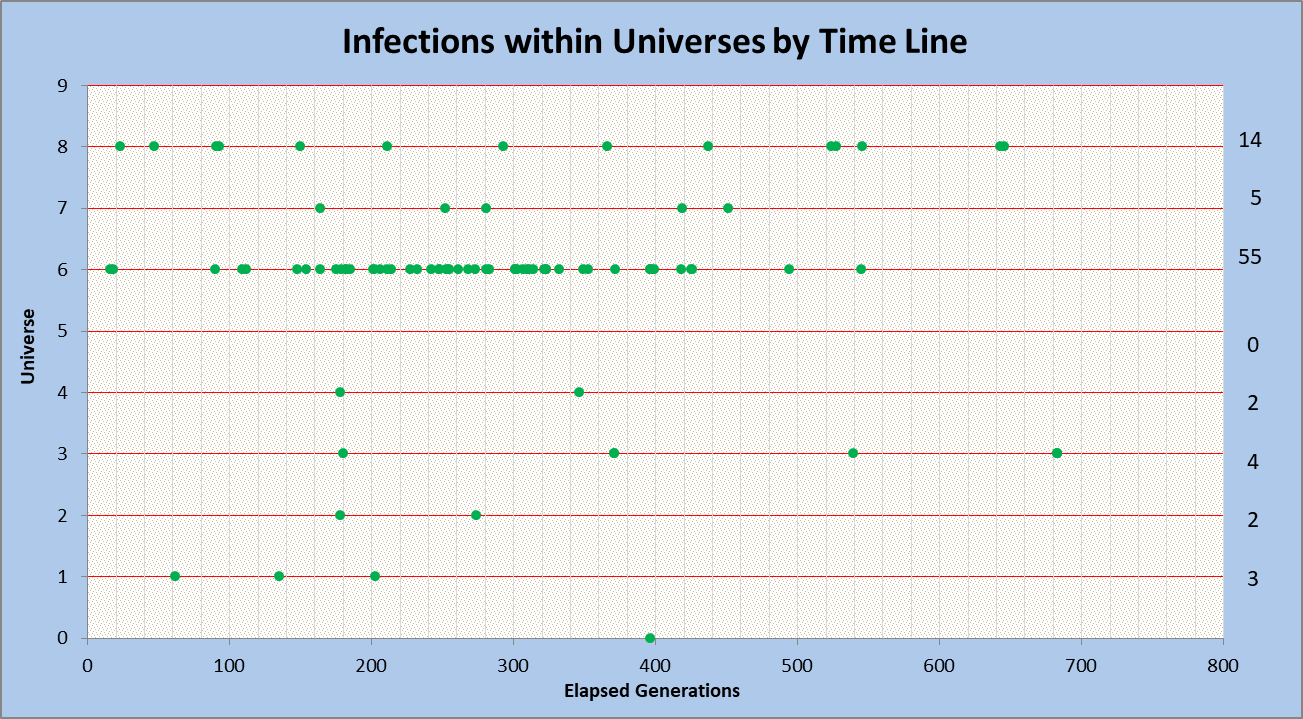
***ANALYSES***

The analytics below relate to three aspects, and their relationship to the dynamics of the epidemic in the Multiverse will be subsequently discussed.

1. The simultaneous infection time-line for the Multiverse
2. The theta values for various time frames, reflecting the intensity within cohorts
3. The transmission trees, showing the overall progression of the epidemic.

***Infection Time-Line***

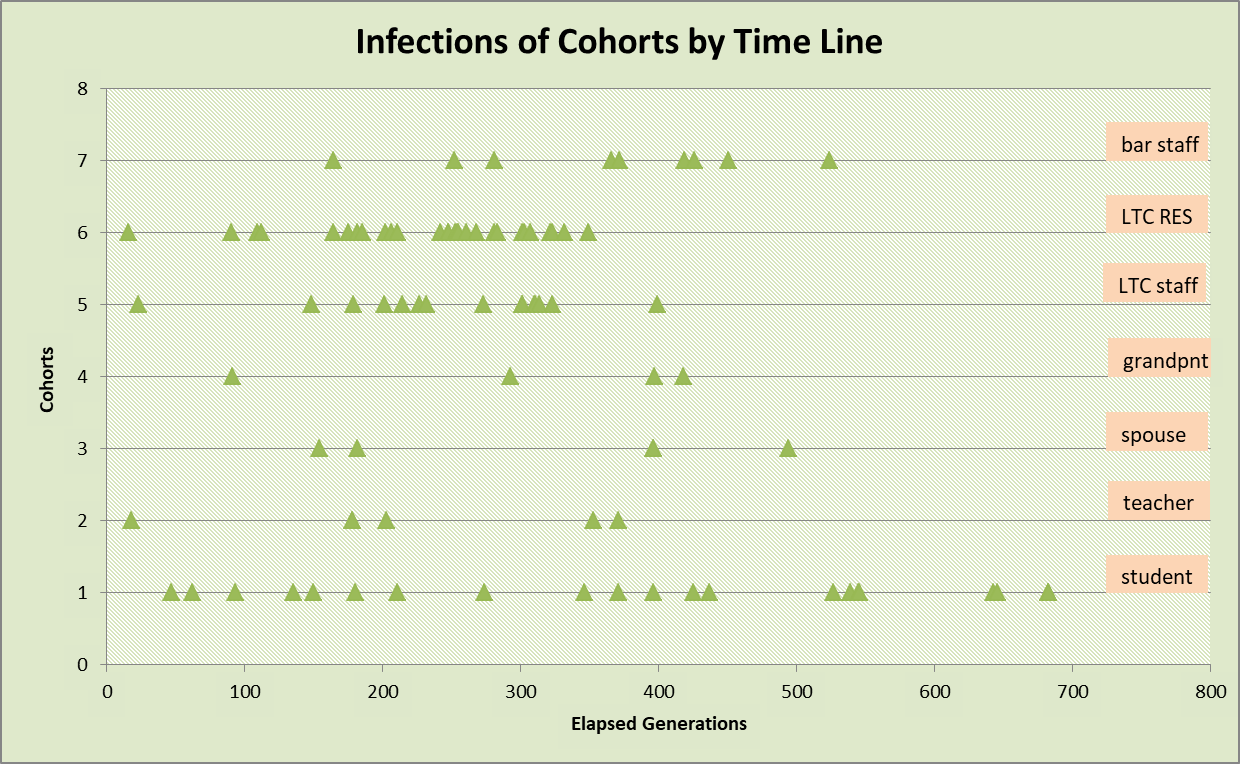
By sorting the console.log by Universe, and counting infections for each, we derive this chart, which shows the progression of time on the X-axis, and the Universes on the Y-axis, each represented by a horizontal line. Green dots on that line represents one or more infections at the time (generation) corresponding to its X-position. The actual count of infections for the Universe is shown on the right, for each horizontal line.



Going from left to right, we see the first infections in U6 and U8; that U6 has most of the infections (55), and that U0-4 are quite sparse, despite 30 students being there all day. Furthermore, the high mingle factor bar only saw five infections within it for the entire duration.

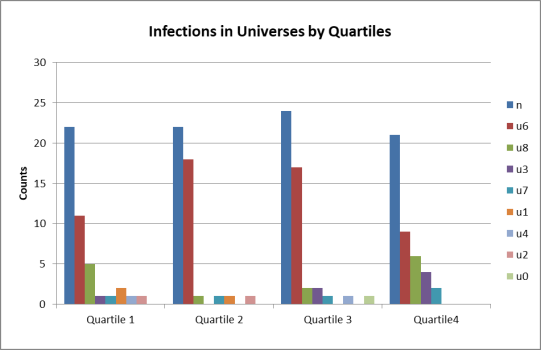
***Distribution of Infected and Survivors***

This simulation terminated with nine survivors. The following is equivalent TimeLine chart of infections by cohorts.



We observe here that LTC infections cluster between 150 to 350 generations, while student infections are spread out fairly evenly till saturation (fewer susceptible students) at 550. This is likely related to LTC residents being there 24hrs per day while students are at school from 9am to 4pm.

Another way of looking at when infections occur, and where is by examining them by quartiles (of generations), as in the table below.



These relationships are quite apparent when charted. The blue bar, ***n*** is the total number of infections in that quartile. It is interesting that the diversity of infections among universes is largest in Q1 and Q4, but really focused in Q2 and Q3 on U6.

***Theta Values***

Recall that theta values represent the average time in generations between infections in a particular time frame.

1. Thetas of the Universes  
   Using console.log which are lines of text, Excel’s “Text to Columns” function can separate the fields using “space” and “:” as delimiters. Then we can sort the resulting table as we wish. In this case, we use generations as the primary field, and select the specific Universes and their infections. With the time stamps, we can calculate theta for each Universe, as shown in the table on the left.   
     
   The theta values for the Universes seen together are seen in the table on the right, and it is obvious that the most active Universe is U6, Long Term Care with theta of 9.6, at least 4x more than any other.  
     
   Returning to Universe 7 (bar/reception) in the first table, the data show that there were five infections from gen=164 to gen=451 for a delta of 287, which yields a theta value of 57.4.  
     
   Those infected in U7 include one student (ID 19), a grandparent (ID 41), a LTC staff (ID 49), and two LTC residents who came out (ID 76 and 81). We can also see from the table who they were infected by. Three infections were by IDs in red states (symptomatic cases), which implies that they came into U7 already infected.
2. Thetas of Infected Cohorts (table shows ID 90-98 – bar staff)

The last infection was at gen 524, and the first at 164. Thus, the delta is 360. Nine agents were infected, so theta is 360/9 = 40. This is the average time between infections for this cohort.

Also relevant are the Universes in which the infections took place (column “U”). Note that there were two family-member infections (U8 HOME), 2 infections that took place in U6 (LTC), and 5 infections in U7 (the bar/dance high-mingle place). Of the LTC infections, one was from ID 57, and the other from ID 69 the first being LTC staff, and the other a LTC resident.

The values for all role-based cohorts in this simulation run are shown below. The table on the left shows theta values for all groups. Note that LTC residents have the lowest theta values – their infections were more rapid (less time between) than for grandparents, at 81.75, or for teachers, at 58.83.

Students overall had theta of 28.91 (time from earliest to latest divided by number), but when we look at each group of students (table on the right), we see that individually, the groups had much higher thetas. This is the difference between any of 30 being infected in a time interval theta, versus any of student group A being infected in the same time frame (the simulation duration).

1. Theta by Infective Agents  
   In this analysis, we set the reference frame for theta to be the first and last infections caused by an agent who was infective. This was simply done by sorting using “transmitter” as the primary field.  
     
   The tabulation showed one agent (ID 75) with 10 infections, one with 7 (ID 73), three with 4 infections (IDs 4,10,50,84), five with three, and the rest two or one infection each. Here are the data for the most active (“super-spreader”?) agent. Both ID 73 and 75 are LTC residents, and do not leave or go to HOME (U8).

Note that each agent only has a limited duration for transmission, and so the delta for generations can be small, but not exceed a maximum of about 11 days, counting pre-symptomatic and symptomatic days.

ID 75, a LTC resident, infects 5 LTC staff (51,54,55,56,60), 4 LTC residents (62,65,66,67) and one spouse (ID 40). They all occur in Universe 6 (LTC).  
  
It can be seen that some multiple infections occur close together, to give a smaller theta. In the table on the right, the two yellow entries on the bottom each have only 3 infections, but they are space 33 and 27 generations apart, and therefore have a theta of 11..00 and 9.00 respectively. Finding the IDs with the highest infections gives us an insight into the distribution of infections, which can be more closely looked at with the following analysis, on transmission trees.

***TRANSMISSION TREES***

From the console.log, we can trace definitively the sequence of infections, and thus the spread of the infections as time progresses, attributed to infecting agents all way to the index case(s).

In this simulation trial, we started with 5 initial infections (ID 10-14). Each of them is therefore the ancestor of a sequence of infections which can be represented (in raw and time-colored formats) below.





The most important row entries are the infection instances, which have the format gen.ID.U. Thus, 112.80.6 means “at gen 112, ID 80 was infected in U6”, with the infector being in the column to the left of the entry. It is obvious in this tree that U6 dominates, then U8, then U3.

The colored table on the right uses the generation to color coding as shown by the color strip. Therefore, looking at the right hand table, the first column shows infection sequence numbers, the second the index case (root node) for this transmission tree.

The third column contains repeated entries for each node that has descendants, for each leaf. Thus, the pale green entries 16.75.6 row 1 shows that ID 75 was infected by ID12 at gen 16, and the next column that at gen 148, ID 75 infected ID 60. The rows 1 and 2, in the next columns, show that ID60 infected ID86 and ID59 at gen 268 and 311 respectively.

The rightmost column in the colored table are indices of this transmission tree, and with them, we create the metric

Q = ((Leafs/AvDepth)/Gen) x New Infections x 100 = 28.12

This is a tree with more breadth than depth (which means more simultaneous infectors), and this Q value is consistent with an aggressive epidemic.

The other four transmission trees (recall there were five initial infected students) have Q values of 18.9, 16.6, 2.7 and 0.7. Appendix 5 shows the details of the other transmission trees, and the process of linking them to a common ancestor node gives a Q value to the entire Multiverse epidemic.

**DISCUSSION**

In this paper, we have presented the interface to an agent-based simulation model that supports multiple interacting spaces through the movements of population cohorts between those Universes. The CovidSIMVL modeling system is intended to be a sandbox with many flexible parameters that permit the exploration of different scenarios, with the ability to record and analyze in detail the dynamic processes in this complex stochastic system that is not possible In Real Life.

The purpose of scenario simulations is not just to perform trials that would be dangerous, or unethical, or impossible to perform IRL, but to attempt to understand in more depth the relationships between process parameters and outcomes so that policies, strategies and tactics can be informed (not determined) by broader and deeper knowledge.

In the scenario that we have described above, we use the console.log to provide exact histories of infections in specific universes, the times and the agents involved, and their family membership, since HOME is a Universe in which only family members are capable of infection one another.

We presented TimeLine charts showing the progress of infections for Universes, and for population cohorts. These charts show not just in which Universe, and in what cohort, the epidemic was most active, but also the time-based distribution of infections in those dimensions. The quartile chart on infections also showed where the majority of infections took place over time.

Theta values which show the quantitative nature of the intervals between infections for different groupings and time frames were explored. Again, these showed that the theta for Universe 6, and for LTC residents, were highest. The theta for infecting persons revealed the stochastic creation of a super spreader.

The transmission trees provided another way of assessing the distribution of infections and in this case, we can see that stochasticity produces heterogeneous progression. However, the transmission tree analysis is a precise way of characterizing the dynamics of the simulation of scenarios, and when vaccinations are introduced into the model, we expect to see changes in the resulting transmission trees.

The environment that we are modeling is complex, with persons of different roles, ages, risk profiles, family structures, moving between universes with varying inherent mingle characteristics, staying for varying periods of time. Understanding the nature of the processes in this environment and their temporal dynamics is the first step, the second being the introduction of various mitigating factors, including vaccinations, in order to explore the consequent changes in infection patterns.

We believe that this flexible parameter driven agent-based stochastic MCMC (Monte Carlo Markov Chain) simulation sandbox is precisely such a tool. It is in the public domain, in <https://github.com/ecsendmail/MultiverseContagion>, and is freely available to use as is, or to adapt.

**Appendix 1. CovidSIMVL Parameters and Metrics**

***Hazard Radius***

Each agent has a physical representation within simulation space, in terms of pixel size. The simulation arena is fixed at 800 x 600 pixels, and the default for the Hazard Ratio is set to 5 pixels as the radius.

Infection occurs when two agents overlap in space following a synchronous move (of all agents), if one of them is susceptible, and the other is in an infectious state.

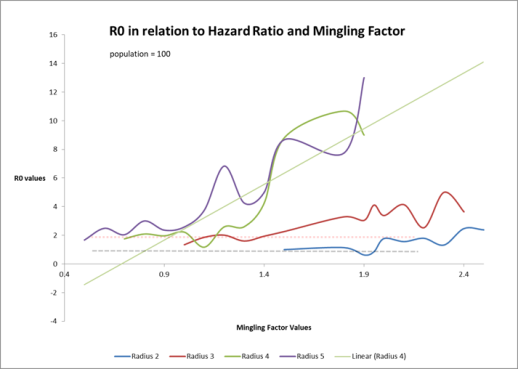
For the susceptible, a larger Hazard Radius makes them more likely to find contact with another agent after a move. Therefore, the Hazard Radius would increase with risks, such as age, premorbid conditions, while it could decrease with youth, masking, socio-economic factors. The mechanism therefore exists to represent increase or decreases in risk through the Hazard Radius. The difficulty is in calibrating a change in Hazard Radius with an intervention or a specific risk factor.

However, the viral temporal dynamics model of Xi, He specifies a viral growth curve for PCR lab values, peaking at onset of symptoms. We therefore in the simulation software code the viral load growth for agents, and directly change the Hazard Radius (stochastic factor added) with change in viral loads.

In general, the approach is to calibrate a scenario simulation by testing a combination of parameters till we obtain a base in which the agents reach some endpoint in the trials that approximate a related In Real Life situation. From this base line we can then introduce mitigations and observe their impact.

***Mingle Factor(“mF”)***

The extent to which an agent moves from its present location governs whether its infective influence is localized or dispersed. Each agent can have its own mF, which may be specific to a role such as a waiter in a restaurant, or a student sitting at a desk, but a Universe can also have a Mingle Factor.

The intensity of an epidemic is related to the probability of contact between agents, which is influenced by the Hazard Radius, and by the Mingle Factor, which governs the degree of mixing of agents within the Universe. The following graph shows the relationship between HzR, mF and the computed R0 for simulated trials.

CovidSIMVL can track everyone infection of every agent through its lifetime, so at the end of a simulation, it can average the number of infections of those agents that have been transmitters to find the true R0 for that epidemic.

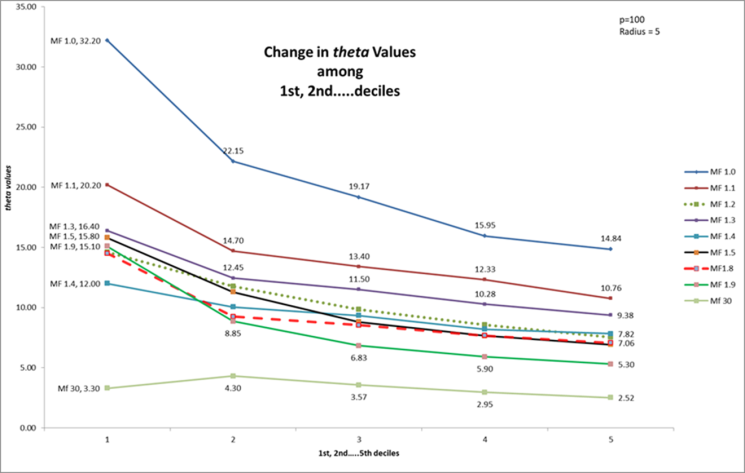
The graph shows that when HzR is large (4 and 5) R0 will increase with increasing mF, but when HzR is small, R0 remains small.

***Theta as a metric of rate of Infections***

Since the console.log gives precisely when and where an infection occurred, by which agent, to which agent, in what Universe, there is considerable flexibility in comparing the time interval between successive infections in any desired context.

If the time interval theta is large, the corresponding risk of being infected is smaller. Contexts that may be relevant range from: what Universes are most active, what cohorts are most active, are infections more frequent during daytime or in the evenings, how do infection rates change from the start to the middle of the epidemic period, are age groups infected at rates different from their population distributions and so on.

The following graph shows the theta for different mingle Factors, with HzR held constant at 5, looking at the values of theta from time T=0 to 10% of infections, 20% of infections, etc to 50%.



The top line (blue) is for mF=1.0, and to 10% of infections, theta.10% = 32.2. As the simulation proceeds, the theta values drop to 14 at theta.50%. This shows that the interval between infections decrease to the 50% mark.

As we increase mF (the other lines), we see that the theta values have the same characteristic – a slower start (higher values of theta) but faster with time.

However, increasing mF produces smaller theta values, till the mF=30 (bottom green line) gives initial theta.10% of 3.3, and this value stays relatively the same to theta.50%.

Decreasing values of theta imply accelerating rates of infection, and increasing likelihood of being infected. When an epidemic slows down, or is going to be self-extinguished because of inability to find susceptible persons, theta values will be increasing.

In the normal course of epidemics that follow exponential growth, we would expect that the counts per day reflect a logistic frequency distribution, resulting in V-shaped theta graphs – slow at the beginning, faster mid-epidemic, and slower at the end.

Although we do not necessarily expect that In Real Life (“IRL”) data collection could produce accurate theta’s, we have not put this metric to the test. However, it would not be difficult, since 7-day averages of new cases are the standard metric, which is easily converted to time between cases…theta.

The Q-value for characterizing transmission trees is another metric that we have derived in CovidSIMVL. This will be elaborated upon in Appendix 5.**Appendix 2. Simulation Scenario - Family Structure**

In the scenario used to introduce the Multiverse above, we have a population of 100, most of whose members fall into families, so that we could use the HOME Universe 8 as a place where people went after work and recreation, but in which only family members would be able to infect one another.

The family structures were created with a mix of multi-generational, with or without children, to reflect some of the many social groupings found today as families. These were 19 in number, and accounted for all but 18 permanent residents (of the 28 in total) of Long Term Care.

The social roles for the population were:

1. Students in 3 groups of 10 each (A,B,C)
2. Teachers – 6
3. Spouses - 5
4. Grandparents – 7
5. Long Term Care Staff – 14 (in 3 shifts)
6. Long Term Care Residents - 28
7. Bar/Reception staff – 10

The general structure of a family is:

1. Heads of family (one or two) with social roles
   1. May be two-person eg Teacher + LTC staff
   2. May be single-person
   3. May be grandparents only (two-person with social role)
2. May have children (0 to 4)
3. May have multi-generational (grandparents not heads of family)
4. May have LTC resident as part of family (grandparents, aunts)

The 19 families have been carefully defined and structured with the use of analog tools.

F0 LTC(S), BAR, A, C LTC(S) – staff, BAR – bar staff, A,C – student groups A,C

F1 T, BAR, LTC(R), C T – teacher, LTC(R) – LTC resident

F2 T, LTC(S), LTC(R),G,B,C G – grandparent

F3 LTC(S),BAR,B,C,LTC(R),LTC(R)

F4 G,G,B Grandparents as head of family with one grandchild

F5 T,BAR,A,C

F6 T,LTC(S),A,A,B,C,G,G,LTC(R) biggest family unit – four kids, 2 grandparents, one LTC res

F7 BAR, LTC(S), A,B,B

F8 BAR,LTC(S),A,C

F9 BAR,LTC(S),A,C

F10 BAR,S,G,A,C Bar worker, spouse, grandparent, two children

F11 LTC(S), S Long Term Care staff, spouse – no kids

F12 LTC(S),S

F13 LTC(S),S, LTC(R) this will be among the visitors to LTC

F14 LTC(S), S, LTC(R), LTC(R)

F15 T,LTC(S), G, LTC(R)

F16 BAR,LTC(S),A,B,B three kids, two in same group

F17 BAR, LTC(S), A,B,C

F18 T,B, LTC(R) single parent (teacher), one child, one LTC resident

The assignment of individual IDs to the Fnn units are done at the time of creation of the population file, and the result is in the Population Definition section in the main body of this paper.

This microcosm is tightly coupled between the workers who are teachers, LTC staff, and BAR workers (BAR being a place-holder for high-mingle spaces). With so many being in the same family units, this scenario may have a higher rate of crossover transmission between the school and LTC unit not just because of frequenting the high-mingle bar, but also because of the family units, which have most LTC workers with children, and/or spouses who are teachers or work in the HIGH-MINGLE bar.

At the time that these structures were defined, we could not anticipate exactly how the process dynamics would reveal themselves. Part of the reason is that the interaction between these structures is highly dependent on who went where for how long, and the scheduling therefore of the population is a critical part of the scenario generation.

In summary, scenario generation for the Multiverse includes:

1. The purpose of the scenario simulation, and the goals for observation
2. The definition of the Universes
3. The definition of the population structures
4. The definition of the movements of the populations

Running through these definitions is the question of the mingling factors of Universes, and of individuals in their roles in the different Universes they visit.

For example, in classrooms, we expect there to be low movement, except for the teacher, and even then perhaps not, if they have a face the class approach. On the other hand, the playground would have a higher mingling factor, and the teacher in the playground would have a higher mF than the children, if they were proactive.

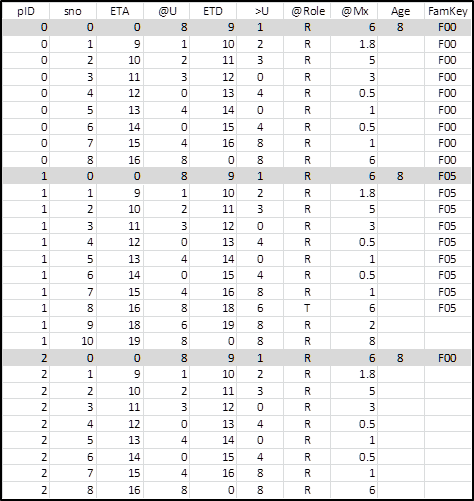
The LTC staff would have a higher mF in their roles, while LTC residents would have low mF.

The HOME Universe, restricted to interaction between family members only for infection, would be given a high mF, so that this might increase the opportunity for family members in the physical arena of 800 x 600 pixels, to contact one another during their 9 to 12 hours together daily.

The schedules for movements, and how they are described, follows in Appendix 3.

**Appendix 3. Population.csv Structure**

The population.csv file is the technical specification for the movements for each person, and has a structure like this.

****The first set of rows is for person with pID=0, and the second column describes the sequence number of movements (can think of as tickets).

In the first row, then, the ETA is at 00hrs (midnight) and they are in @U8 (they are “at U8), with an ETD departure of 0900 hours bound for (“>U”) Universe 1 (Project Room) where the role is that of “R” – with an mF of 6 – expected to move around in the project studio, an age of 8, and a family id of F00.

The second row shows them arriving at 0900 at U1 (@U), staying till 10 ETD to (“>U”) Universe 2 (Playground) where they will have a mF of 1.8.

The rows continue till they complete the schedule for a day, at which point the system repeats, unless the schedule defines specific days in the form of DDHH rather than straight HH for times.

Constraints include the number and distribution of teachers among classrooms and other facilities, and considerations like the following.

1. The HOME is a unique Universe in which transmission only happens between family members. The simulation space is 800 x 600 but might contain 70 persons at times. Depending on the mingle factors, the odds of any 3 persons in the same family affecting one another can be very small (as experienced in simulation trials of population vs size vs mingle factor) unless the sizes are quite large;
2. It is expected that if a family member is a LTC RESIDENT, that family head (at least one) and child (zero to one) will visit during the day. Of course, this is one of the parameters that can be changed for different trials of this scenario.
3. The scheduling for LTC staff of 14 is 7+4+3 for 7am-3pm, 3pm-11pm, 11pm-7am.
4. The scheduling for BAR staff of 10 is 10am-6pm 4 staff, 6pm-12 midnight 6 staff
5. Teachers, spouses, LTC staff are scheduled to be visitors to the HIGH-MINGLE universe U7 during the hours 12 to 12, including at least one of the grandparents, and LTC staff as they fit.
6. The students go home, and there is no venue for their physical interaction. This keeps the simulation “pure” in the sense that school is the only place for transmission between students, and from the student-teacher interactions.



**Appendix 4. CovidSIMVL Fixed Universe Interface**

The screen capture below is from the October 2020 version of CovidSIMVL used to run the simulation scenario described in this paper.



This screen can toggle to the Multiverse view shown earlier. However, in this interface, some parameters for this Universe (shown in the left top of the arena as U6) can be set, with the buttons shown in the red rectangle on the right. In particular, the PP radius, which is 5, can be set globally for the agents in the Multiverse.

The left red rectangle is a status box showing the population of Universe 6 – greens for susceptible agents, yellow for incubating, etc. The HR++ button above lets the operator advance one hour at a time (one generation) while the yellow AUTO button lets the system advance time automatically. Clicking on AUTO in automatic mode puts the system into manual mode.

The green ShowU button advances to the next Universe, cycling after U8 to U0, and with each new Universe selected, the display in the arena, and the counts and charts are appropriately updated.

In the arena, the circles represent roles of “Resident”, the plus shapes the role of “Attached” such as staff who are temporarily in a Universe, and the squares represent “Transients”, who have different mobility than staff.

The version that is shown here does not adjust mingle Factors from the interface, so all the variation in mingle Factors for individual agents are modified through the schedules in the **Population.csv** file, and changed dynamically as they move to their new destinations.

The scenario for the Multiverse was run with initial infective agents in the Case file, another .csv file. The system permits the user to invoke additional case files during a run if desired. The format is simple, but has to include the viral load, and the number of days since the agent was infected, in order to comply with viral dynamics.



This is the Case File that was used in the scenario simulation in the main body of this paper.

**The scenario simulation was run with the population file “Population.csv” and the Case file “VLfive.csv” on Oct.14, 2020 using the version of CovidSIMVL that was current at that date.**

**The parameters used were: HzR = 5, mF=10, and 9 survivors were present at TERMINATION (no further infectious agents).**

*Note: since then the mF values have been scaled down by a factor of 10 in current versions, to be 1 as default.*

**Appendix 5.Transmission Trees for this Simulation Trial**

We developed the notion of a Transmission Tree, since the simulation can capture every transmission between agents to show the continuity of infections. The metric Q was derived to capture the shape of Transmission Trees in terms of their average depth (over all path lengths), their breadth expressed as the number of leaves (which is the same as the number of distinct paths), and the time elapsed for the generation of the entire Transmission Tree.

Q is a measure of the aggressiveness of the epidemic as reflected in the particular Transmission Tree. The broader the tree, the more simultaneous transmitters there are, and the shallower the whole tree. Thus, we put the breadth of the tree in the numerator, and the average depth in the denominator (the smaller it is, the larger is Q). Similarly, the elapsed generations is a denominator term (the smaller, the larger Q is).

Q = (Leafs/AvDepth)/Gen \* New Infections \*100

Here are the five transmission trees that were generated by the initial five infective agents, which are 10,11,12,13,14. The syntax for each entry is gen.pID.U., and the color coding is for chronology.









Notice that the five trees are not similar in size – three of them have more than 10 leafs (distinct paths), and are broad, while two of them have less than 5 leafs. The Q values are as expected – 28.2, 19 and 16.6 for the three trees which are aggressive, while the other two have Q values of 2.7 and 0.67.

The Transmission Tree with root node 12 is shown below in tree form – broad and shallow. Some subtrees are color-coded to avoid confusion with parent-child relationships.

|  |  |
| --- | --- |
|  |  |



It is certainly possible to create a virtual ancestor for the five trees, to consider them as one single Transmission Tree for the entire Multiverse population. We only need to modify the AVERAGE DEPTH and the count of LEAFS.

Doing this by pen and paper (the reader is invited to do this at their own leisure), we end up with

48 leafs

184 total path length

87 new infections

683 generations

Average depth = 184/48 = 3.833

Q = (48/3.833)/683\*87 \*100 = 159.5

This virtual Transmission Tree is broad (five immediate descendants), and shallow with average depth being less than 4 including the new link to the virtual root node.