

PANDEMIC: To Simulate Infection Spread in a Complex Environment

JAY BORIS

*Chief Scientist
Material Science and Component Technology Directorate*

KEITH OBENSCHAIN

*Head, Laboratory for Advanced Computational Physics
Laboratories for Computational Physics and Fluid Dynamics*

DAVID BORIS

*Technical Collaborator
Webster New York*

September 17, 2021

REPORT DOCUMENTATION PAGE				Form Approved OMB No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.					
1. REPORT DATE (DD-MM-YYYY) 17-09-2021		2. REPORT TYPE NRL Memorandum Report		3. DATES COVERED (From - To) 01-03-2020 – 05-21-2021	
4. TITLE AND SUBTITLE PANDEMIC: To Simulate Infection Spread in a Complex Environment				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S) Jay Boris, Keith Obenschain, and David Boris*				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER 1G56	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Research Laboratory 4555 Overlook Avenue, SW Washington, DC 20375-5320				8. PERFORMING ORGANIZATION REPORT NUMBER NRL/6003/MR--2021/2	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) Naval Research Laboratory 4555 Overlook Avenue, SW Washington, DC 20375-5320				10. SPONSOR / MONITOR'S ACRONYM(S)	
				11. SPONSOR / MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION / AVAILABILITY STATEMENT DISTRIBUTION STATEMENT A: Approved for public release; distribution is unlimited.					
13. SUPPLEMENTARY NOTES *Technical Collaborator, Webster New York This is the first NRL Memorandum Report in a series about the PANDEMIC model.					
14. ABSTRACT PANDEMIC is a new, time-dependent simulation model treating the infectious spread of a disease in a complex, geographically varied region containing millions of people. In PANDEMIC, each person in the region moves from place to place and interacts with others they meet while commuting to work, going to meetings (i.e. church and other events), to shop or school, and returning to or remaining at home. This model takes advantage of parallel computing using an NRL-developed algorithm called the Monotonic Lagrangian Grid (MLG) to keep track of who is close to whom, as the people follow their usual routines. This model allows “what if” tests of different personal and societal public health behaviors and various mitigation strategies to compare their impact and effectiveness. The goal of this research is to predict in greater detail the dynamics of mitigating the current covid-19 pandemic while developing more efficient algorithms and computational techniques using new computer hardware designs. PANDEMIC simulations are presented predicting the three covid-19 infection surges observed through January 2021. Different levels of testing to mitigate the catastrophic ‘third surge’ are compared. Another prototype study, illustrating how the model can be used for hypothesis testing, is also reported. It evaluates the impact of largely unmasked “super-spreader” events showing how each such gathering ultimately increases the number of covid-19 related deaths by hundreds.					
15. SUBJECT TERMS Covid-19 Transmission of viral load Epidemic spread Pandemic mitigation techniques Detailed complex-environment simulation Monotonic lagrangian grid					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT U	18. NUMBER OF PAGES 27	19a. NAME OF RESPONSIBLE PERSON Jay Boris
a. REPORT U	b. ABSTRACT U	c. THIS PAGE U			19b. TELEPHONE NUMBER (include area code) (202) 767-3055

This page intentionally left blank.

1. Brief Introduction to Modelling Epidemics

This paper reports results comparing several PANDEMIC simulations that each span the first year of the covid-19 ‘pandemic’. The model treats a population of more than 6 million people in a 100 km by 100 km area around the Washington D.C. Metropolitan Region. Approximations to the population’s daily behaviors are simulated in the model including the people’s imperfect attempts to implement the societally recommended responses to mitigate the pandemic. The effects of masking, social distancing, telework, virtual meetings, isolation/quarantine, and testing are treated in the model. The results of PANDEMIC simulations are reported that predict the three covid-19 infection surges observed through the end of 2020. Different levels of testing to mitigate the current catastrophic ‘third surge’ are compared. Another prototype study, illustrating how the model can be used for hypothesis testing, is also reported. It evaluates the impact of largely unmasked “super-spreader” events showing how each such gathering can ultimately increase the number of covid-19 related deaths by hundreds.

Some epidemic evolution models are statistical in nature, trying to extrapolate curves adapted from previous epidemics and/or based on detailed data collected from the current situation. Because epidemics seem to be all different, modelers using previous epidemic data to model new epidemics are put at a disadvantage. Further, “The sparsity of such events (i.e. pandemics) and changes in outbreak recording, may make identifying suitable validation data difficult.” (Walters, et al., 2018) Other modeling approaches try to approximate individual interactions according to representations that break populations into broad classes or categories, e.g. SEIR models (IDM docs>EMOD, 2021). Yet other approaches try to treat the problem more atomically as a modified form of molecular dynamics. None of these models are very predictive because they do not model the necessary interplay between the spreading of the virus and the behaviors of individuals. A “simple model makes a gallon of faulty assumptions – including that human beings move around and contact one another randomly, like molecules in a gas, instead of making very nonrandom contacts within certain networks, as we actually do. You are much more likely to contact a specific group of folks in your local community, whether at your workplace, wherever you worship or in your own household.” (Hanage and Jenkins, 2020).

“The inevitable, but unpredictable, appearance of new infectious diseases has been recognized for millennia, well before the discovery of causative infectious agents. Today, however, despite extraordinary advances in development of countermeasures (diagnostics, therapeutics, and vaccines), the ease of world travel and increased global interdependence have added layers of complexity to containing these infectious diseases that affect not only the health but the economic stability of societies. HIV/AIDS, severe acute respiratory syndrome (SARS), and the most recent 2009 pandemic H1N1 influenza are only a few of many examples of emerging infectious diseases in the modern world; each of these diseases has caused global societal and economic impact related to unexpected illnesses and deaths, as well as interference with travel, business, and many normal life activities. Other emerging infections are less catastrophic than these examples; however, they nonetheless may take a significant human toll as well as cause public fear, economic loss, and other adverse outcomes” (Morens and Fauci, 2013).

The new PANDEMIC model described here is introduced through several examples related to the current covid-19 epidemic in the United States. The sections of this report are:

- 1. Brief Introduction to Modelling Epidemics

- 2. The PANDEMIC Simulation Model – A Brief Description
- 3. Testing Protocols to Mitigate the Third Covid-19 Infection Surge
- 4. Understanding the Variability of PANDEMIC Simulation Realizations
- 5. Hypothesis Testing: The Cost of Crowds, Rallies and Super Spreader Events
- 6. Discussion and Conclusions

Section 2 is a brief description of PANDEMIC. This is not intended as a comprehensive algorithm definition nor a user manual. Rather, the underlying components and assumptions of the current approach are introduced.

2. The PANDEMIC Simulation Model – A Brief Description

Figure 1, based on a Google Map, illustrates the geographic region of the PANDEMIC model used in this paper. PANDEMIC is not a lumped parameter or regionally averaged representation and it is not a “molecular dynamics” algorithm where the people move about randomly and bump into each other to interact. PANDEMIC, more closely, resembles an earlier Discrete Simulation Monte Carlo model (Cybyk, et al., 1994 & 1995). The simulated people in PANDEMIC all have different parameters and live in different places. These people each follow realistic movement and behavior patterns that vary throughout the day and week. The current version of the model treats an idealized 100 km by 100 km region including Washington DC, Baltimore, Annapolis and Frederick in Maryland, and Alexandria, Fairfax City, and Reston in Virginia, as shown below. These cities appear as local population concentrations, and are shown in Fig. 2 below, where the pixels representing people are clustered more closely in the urban and suburban areas. This approximation to geographic fidelity could certainly be improved given more time and resources. The current treatment, however, provides an improvement over more homogeneous historical or data-based models and affords opportunities to study the effects that differences in behavior patterns, both personal and mandated, will have on the evolution of the pandemic.

Each person is assigned a random age from a demographic probability distribution, available from gathered data. Each simulated person’s susceptibility, also randomly assigned, drives their mortality under covid-19 to closely approximate the population average for the person’s age. These are not intended to be specific people but rather to approximate the general population distributions reasonably. Separate economic, gender, race, ethnicity, and co-morbidity factors are not considered, though these complicating factors could be included with significant additional work. Each person has a viral load V_L with a value from 0.0 to 4.0 that is updated by an ‘exposure’ subroutine in several-minute timesteps. Different ranges of the viral load are associated with different stages of Covid-19 infection. These are plotted in the graphical displays from Fig. 7 onward, which are produced by PANDEMIC and described below.

The initial model did not include algorithms for following the effects of vaccination or include treatments of different behaviors for different groups of people considering the now recognized dependence of transmissibility on vaccination status, co-morbidities, and differences between indoor and outdoor venues. The model also does not yet treat in detail different regions within the overall domain implementing different requirements. None of these conditions are show stoppers but require implementation to update the model in the months since January 2021. All the capabilities needed for the covid-19 pandemic need to be added and verified for the current situation in order that the PANDEMIC model be ready for the next epidemic.

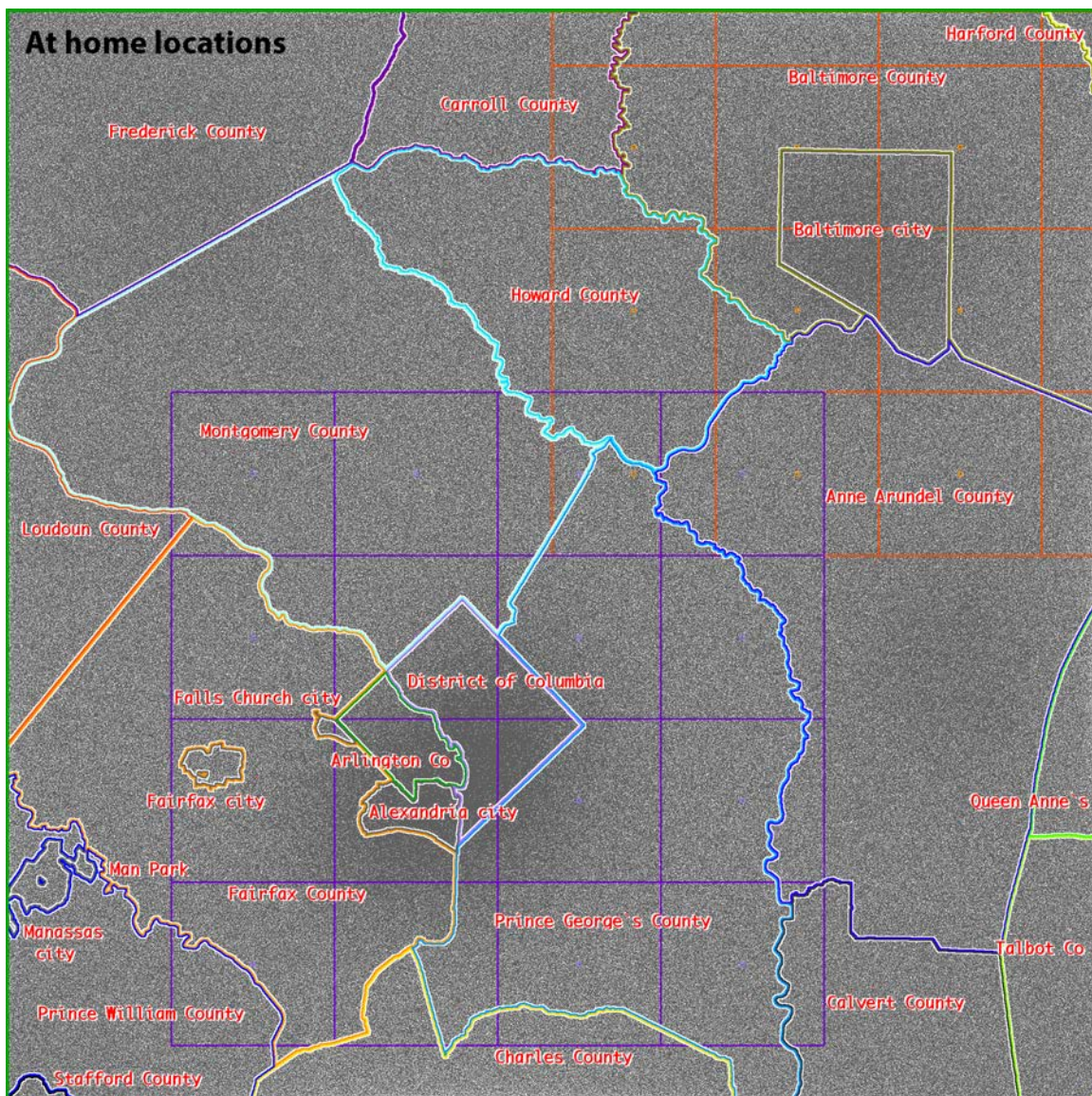


Figure 2. Geographic home locations of 6.25 million people in NRL's PANDEMIC simulation model approximating the Washington DC Metro region. These homes are arranged in a 2500 x 2500 Monotonic Lagrangian Grid to efficiently compute the near-neighbor exposures to the virus as the people move about. The grey pixel density shows the population density when everyone is at home.

Each person has a critical V_L value, based on their age and susceptibility. At this critical value, V_L begins to decrease due to assumed natural or treatment-induced immunity. The person may pass back down through the contagious range, again infecting others, until the person becomes non-contagious. There is still a serious shortage of information about how this all works for covid-19, so PANDEMIC, like other detailed models that aim to be predictive, makes a host of assumptions, many of which may be faulty, while actual data is being collected about the virus. If a person's critical V_L value, where cure/immunity would begin, is 3.8 or larger, the person dies and no longer interacts within the simulation.

Figure 2 above shows the at home location of every simulated person (actor) in the D.C. Metro region, plotted as grey pixels. This figure visualizes the initial condition of a set of simulations aimed at studying the covid-19 pandemic as it has evolved in this region from March

2020 to the present. Each pixel covers a 50 x 50 meter square so several people are often plotted in the same pixel. In the initial setup, here 1 March 2020, 81 infected people were randomly distributed around the region. These original carriers, assumed still to be asymptomatic but contagious, are represented by isolated red pixels scattered about the region. More red pixels are visible later in time in the insets of Figs. 5 and 6. The Maryland and Virginia county boundaries are shown as are the several separately incorporated cities.

The darker grey areas of the figure show the population concentrations around Washington D.C. including Alexandria, Falls Church and Reston in Virginia and, Baltimore Annapolis, and Frederick in Maryland. The 15 km by 15 km orange and purple squares mark local areas surrounding 16 Washington area (purple) and 16 Baltimore area (orange) notional meeting sites/churches. These sites are marked by small lavender and orange facilities drawn at the center of each square. On Sunday morning people attend the nearest site (church) to their home.

The goal is a fast but accurate simulation model based on computing person-to-person transmissions under a realistic range of circumstances. This allows an analysis of some of the complexities and possible mitigation strategies for the current corona virus epidemic. The current epidemic is also a timely vehicle for developing this model while researching efficient parallel algorithms to exploit new, many-core computer architectures. Our current project addresses the current crisis using a software algorithm originally developed at NRL for other many-body applications. The programming model, currently based on OpenMP parallelization, uses a 2500 x 2500 Monotonic Lagrangian Grid (MLG) dynamic data structure (Boris and Lambrakos, 1985; Boris, et al., 1986; Lambrakos and Boris, 1986; Kaplan et al., 2010, 2012). A uniform spatial grid of 40 by 40-meter cells would uniformly span this area, which is home to between 6 and 7 million people, but the separation between nearby people varies wildly from minute to minute. Representing each person by a single Lagrangian node in the MLG data structure provides the variable spacing necessary so the interactions between people can adjust in time and space with separations that vary locally from kilometers to fractions of a meter.

The MLG node locations are closer together near population centers without disturbing the regularly indexed 2D lattice of the MLG. Using the MLG as a way to store and arrange the data describing the state and condition of each person converts the time-consuming close-interaction computations to a form particularly well suited to parallel computing. These algorithms make the complexity and cost of tracking physically realistic, separation distances possible for quite large groups of people, even on a laptop. A potentially disorganized set of geographical data is now in a form that facilitates determining which people happen to be close to a particular person at any time and place.

Figure 3 shows an example of a 5 x 5 local template MLG embedded within the full 2500x2500 MLG array. This local template shows the 24 near neighbors of a contagious person (white node) used to evaluate whether one of the near neighbors becomes infected during each timestep. The number of near neighbors checked is a run-time variable, generally between 12 and 24, for

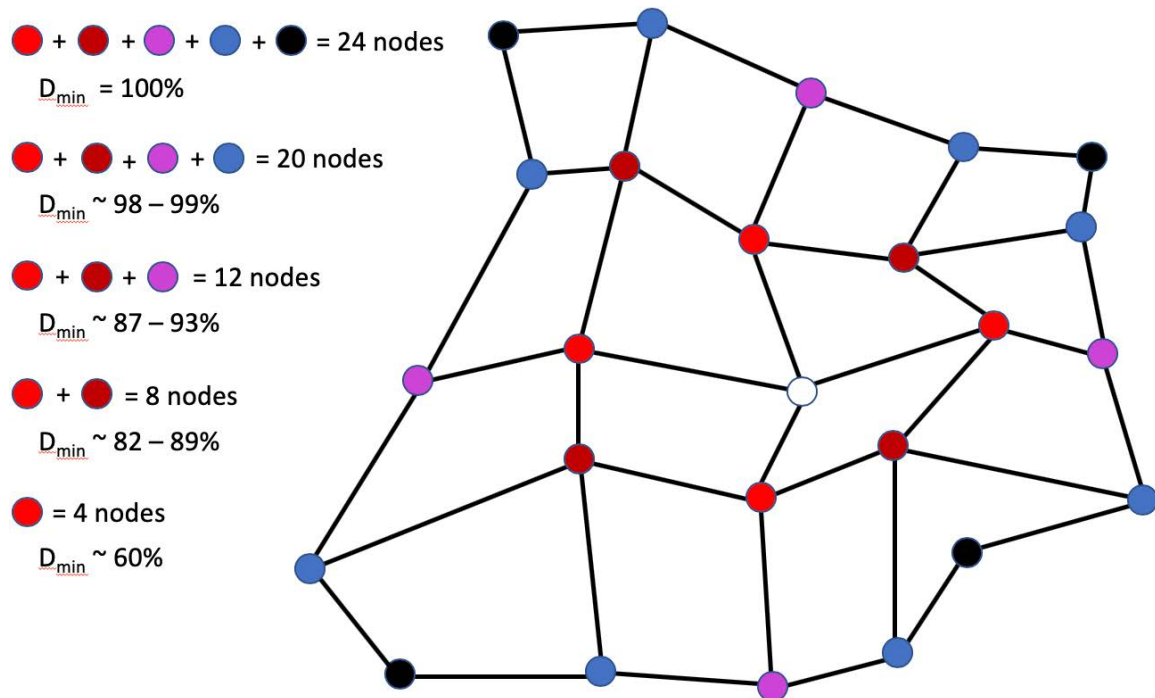


Figure 3. A Monotonic Lagrangian Grid (MLG) in the vicinity of a contagious person indicated by the white circle. This 5 by 5 template repeats throughout the domain for each node (person) because the MLG is logically rectangular across the entire domain regardless of people's actual locations.

each contagious node. Using 16 or more near neighbors generally ensures that 95% of the time the nearest node to the contagious person is being checked. A few can be missed but the probability is that most transmissions do not occur in one timestep. Further, usually the nearby people in the MLG are too far away to be risking infection.

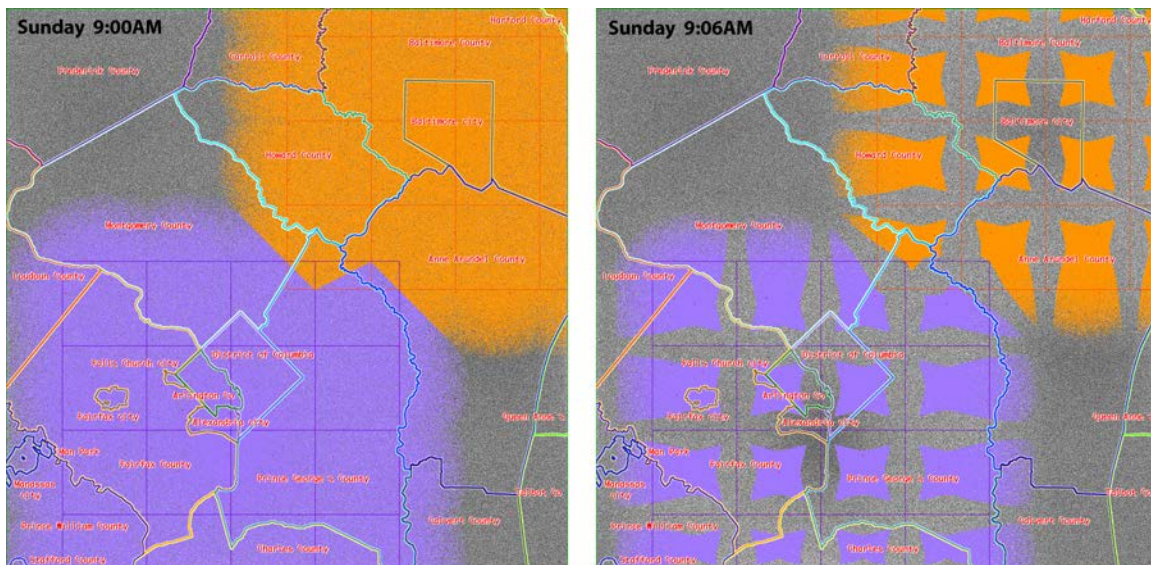


Figure 4. The MLG transforms as church members travel to church. Sunday 9:00AM: Home locations bound for D.C. (lavender) and Baltimore (orange) churches. (b) 9:06AM: People still in transit before reaching church. Many have already reached the church site while the remaining grey pixels represent people who stayed at home. Cars travel at 10 m/s with 6 min PANDEMIC timesteps.

Figure 4, immediately above, displays this overall MLG by plotting each person at their geographic coordinates in the 100 km by 100 km region. The figure shows two snapshots of the people's locations at 9:00AM Sunday (left), just before they leave for church, and at 9:06AM (right) after they have travelled 3.6 km toward the church of their choice. The people who go to church travel to the closest church to their home at 10 m/s. The lavender people attend one of the D.C. churches at the center of a purple square. The orange people attend one of the Baltimore churches at the center of an orange square. In the left snapshot, all the people within 3.6 km of their chosen church have reached their destination. The grey people who are not attending church are seen between the lavender and orange people still travelling. Fig. 4 also shows variations in the density of people staying home (grey) including the depletions in density between the orange and lavender folks still travelling at 9:06AM in the right panel.

Figure 5 below shows two snapshots of the 6.25 million people on a Sunday morning after church services have let out at 11:00AM. People going home (green) get in their cars at 11:00 and begin driving radially away from the church site, resulting in expanding green circles around each of the 32 church locations. In the left hand image at 11:06AM the cars have moved somewhat less than halfway to the edge of the 15 km by 15 km squares marking the local area around each church. The people whose homes are within the circles have already reached home and been marked in grey, accounting for the higher density (darker grey) regions inside the circles. The inset on the left shows a blow-up of the image for one of the 32 churches. The two red dots on the green circle are contagious people still travelling home. The two red dots inside the circle stayed at home or have just reached home. The red dot outside the circle was a previously infected person who stayed at home.

After 18 minutes of travel, on the right, the circles are beginning to merge beyond the 32 outlined church local areas and all of the people living in some of the local areas have reached their homes where their node changes from green to grey. The density of grey in the inner areas has returned to the pattern shown in Fig. 2 above. PANDEMIC moves all people along their daily

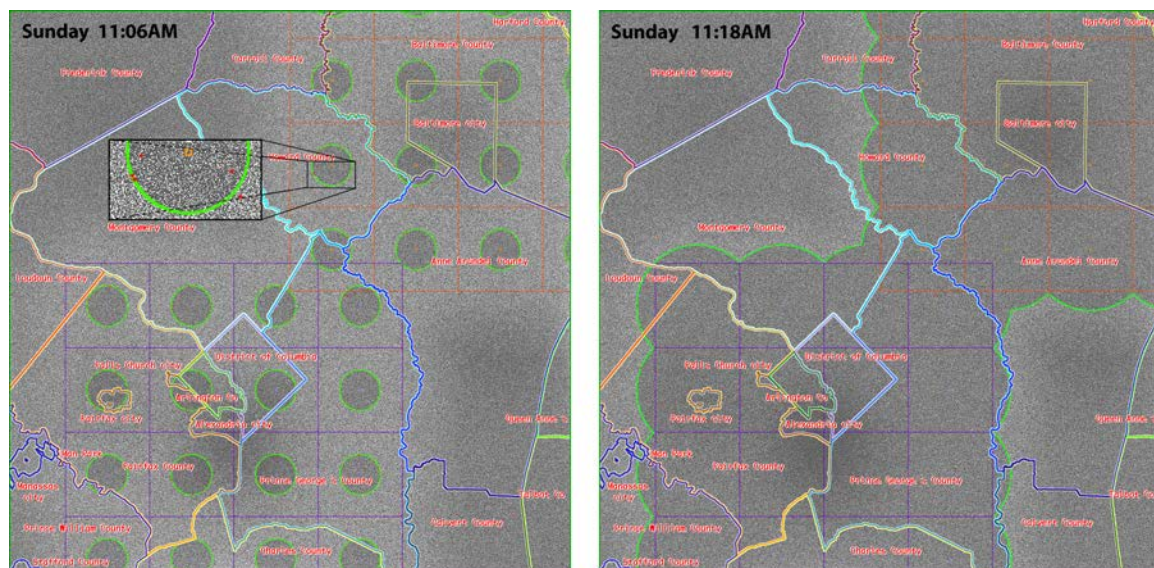


Figure 5. Church (green) members return home. Sunday 11:06AM: Cars have travelled 3.6 km radially away from the D.C. and Baltimore churches. (b) 11:18AM: Cars have travelled 10.8 km and most people have reached home.

routines whether their viral load exceeds the isolation level $V_L = 1.5$ or not. However, they are not allowed to interact with even their closest neighbors because they are marked as inaccessible to enforce quarantine or isolation. This is also done with people who are teleworking to simplify the movement and MLG maintenance algorithms.

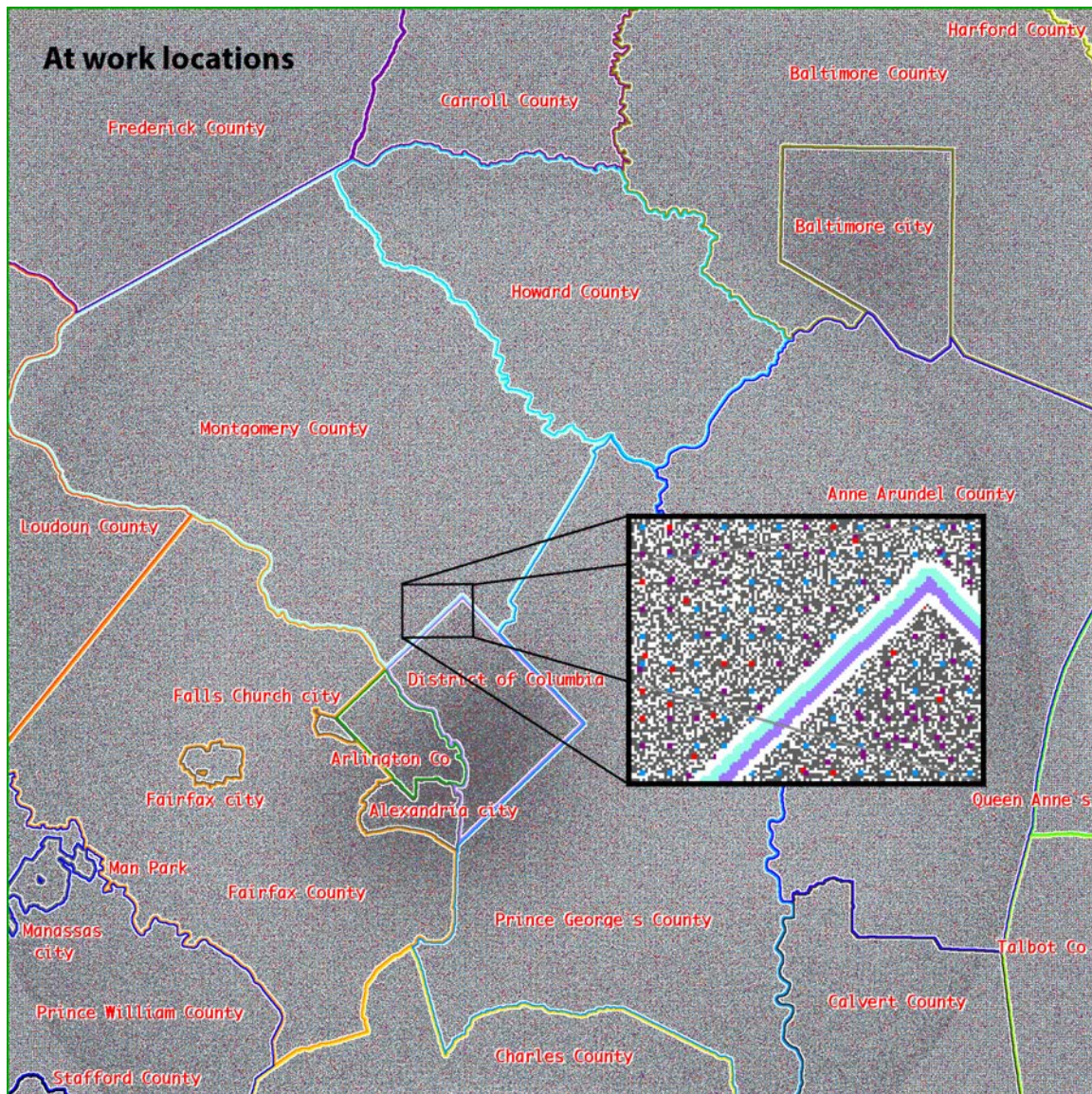


Figure 6. The weekday workplace locations from ~7:30AM to 4:00PM of commuters. People who remain at home are also shown. The shaded dark ring arises because people living more than 50 km from the district commute closer to the district. This depletes the distant areas, causing the darker ring. Blue squares in the inset mark the locations of some of the 40,000 workplaces.

Figure 6 above shows the locations of the entire population during weekdays once commuters have arrived at work. The pixels representing the people are colored grey (uninfected), red (contagious carriers with a viral load greater than 1.0 and less than 1.5). The magenta pixels mark contagious carriers whose viral load is 1.5 or greater and who are, therefore, effectively in isolation, quarantine or even dead. The inset in the figure above shows many 100 m by 100 m blue squares that depict the locations of occupied workplaces with one or more people. These workplaces are spaced on a regular lattice separated by 500 meters as seen. If

one of these squares is red or magenta, the last person plotted at that workplace has a viral load exceed 1.0 or 1.5 respectively. The snapshot in Fig. 6 was taken at two weeks after many people had been infected, as seen by the number of red and magenta nodes in the figure. Several hundred workplaces, mostly near the periphery of the metro area, are unattended and thus are not plotted in blue. They are unattended because the local population density is low and because the people tend to commute toward population concentrations.

About 46% of the people commute to work every weekday while the rest stay home (or go to school). When the commuters arrive at their destinations, randomly dispersed in the limited area designated as their workplace, the population data set is restructured into a Monotonic Lagrangian Grid (MLG) order adapted to these people at work sites. Thus each timestep of the exposure algorithm has immediate access to the near “targets” of each contagious carrier at the workplace because they are close index neighbors of the source in the MLG. If the viral load of either the source or the target exceeds 1.5, the potential transmission of the virus is prevented because any person with a $V_L > 1.5$ is defined to be in isolation or quarantine. These people with $V_L > 1.5$ move through their personal daily routines but are unable to transmit the virus because of their isolated/quarantined status. This condition also renders them non-contagious at home.

Figure 7 below shows the integrated viral load data from a PANDEMIC scenario for the Washington DC metro area. This scenario lasts 39 weeks (~ 9 months) into the third surge. Each vertical lavender line indicates a week and the nearly invisible grey dashed lines mark days. The vertical magenta lines separate months. The durations and the characteristics of the 5 stages making up this scenario are shown in black across the top of Fig. 7 and of Fig. 8 following. Because of the wide range of the population’s viral-load numbers integrated over the specific, color-coded ranges indicated at any time, the axis scale at the left is logarithmic.

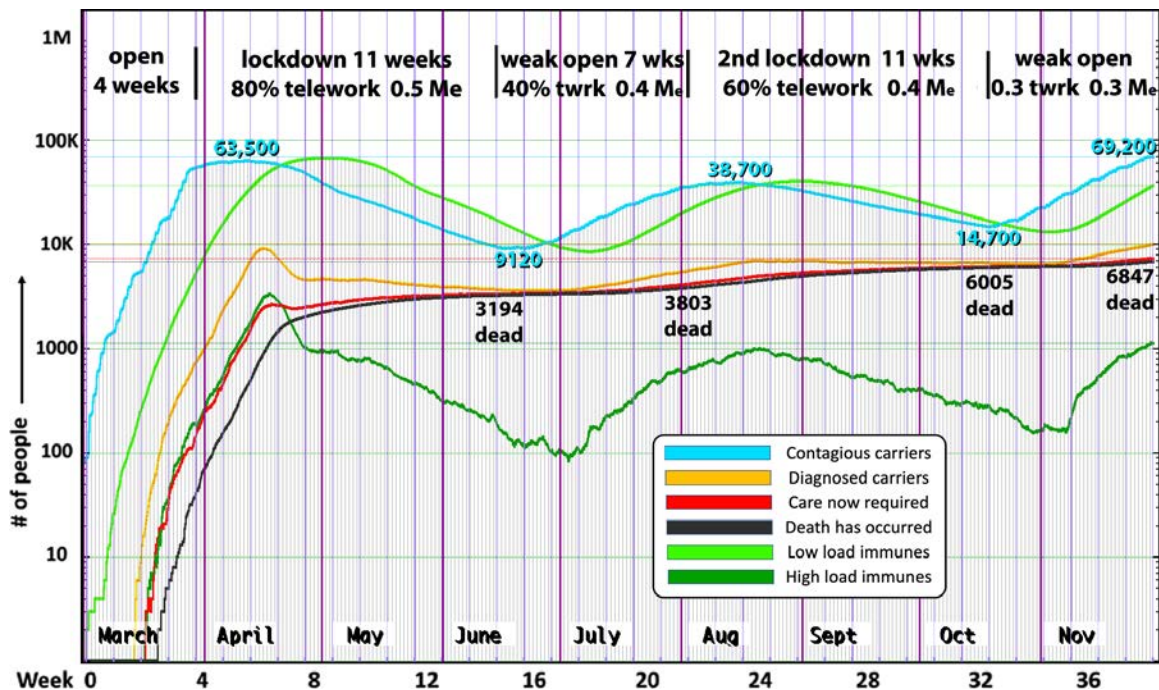


Figure 7. 39-Week simulation of the Covid-19 epidemic projected through November 2020. After an 11-week lockdown from April into June, a weak 7-week re-opening requires a 2nd lockdown through mid October. 6847 deaths for the Metro region scales to about 340,000 deaths for the nation.

To repeat, each person has a viral load V_L ranging from 0 to 4 with four ranges. 0 to 1 is infected but not yet contagious. People with $V_L \geq 1.0$ to 1.5 are contagious carriers (light blue). People with $V_L \geq 1.5$ and < 3.0 (yellow) are diagnosed or symptomatic and assumed to be isolated or in quarantine. People with $V_L \geq 3.0$ and < 3.8 (red) are quite ill and are assumed to be under care in isolation, and thus cannot transmit the virus. People with $V_L \geq 3.8$ are dead (black). PANDEMIC does not deal with the medical aspects of treating these people in hospitals or care centers. The rates of progression and mortality are intended as averages pending future work on the model when more data becomes available. This viral load scale is quite arbitrary; only the rate of progression along this scale matters. The rate used in PANDEMIC is based measured demographics and susceptibility.

Fig. 7 shows the number of people at each time with viral loads above 1.0 (light blue), above 2.0 (yellow), above 3.0 (red) and above 3.8 (black denoting death). The number of ‘contagious carriers’ is always at least as large as the number of ‘diagnosed carriers’ which is always at least as large as the number with ‘care now required,’ and that number is always at least as large as the number ‘dead.’ This scenario, beginning with the 81 initial carriers (light blue on the left hand Y axis), progresses in 12 minute timesteps. During each timestep, the people are moved to their individual destinations, as scheduled, and allowed to interact with their close neighbors, as constrained by their two viral load values, using the ‘EXPOSURE’ subroutine, based on small index offsets in the current MLG data structure.

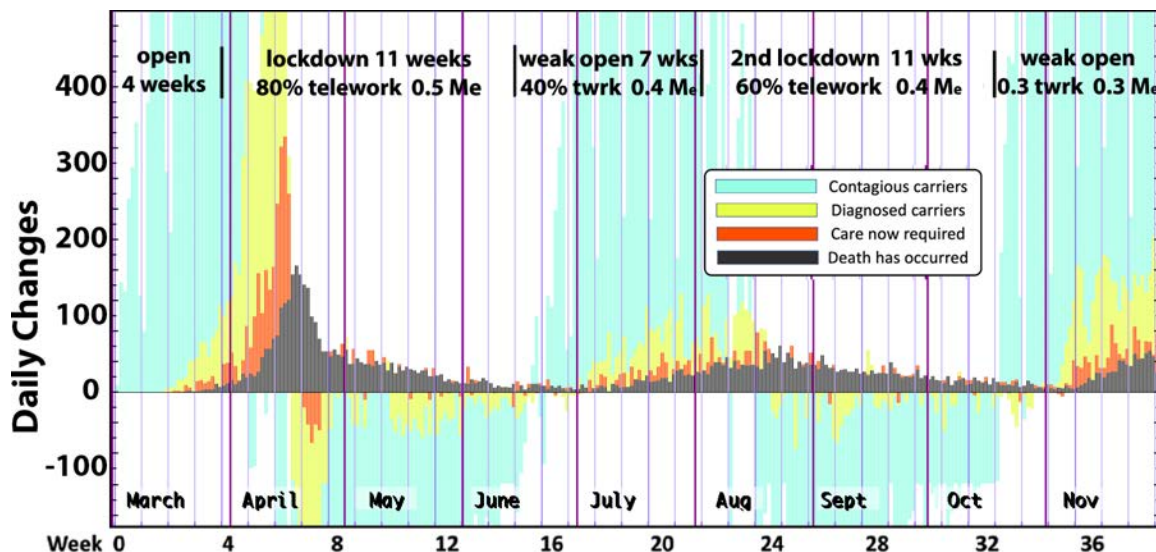


Figure 8. 39-Week simulation of the Covid-19 epidemic projected through November 2020. After an 11-week lockdown from April into June, a weak 7-week re-opening requires a 2nd lockdown through mid-October. The third covid-19 surge is beginning in November.

The numbers of contagious carriers (i.e. infected contagious people), diagnosed carriers, people under care, and fatalities are tallied every day and regularly plotted (e.g. Fig. 7). The day to day differences are saved as tables of ‘daily changes.’ Figure 8, above, shows that the daily change counts capture successive peaks as the virus spreads through the community and worsens. First yellow peaks (diagnosed carriers), then red peaks (care now required), then black peaks (deaths) in a ‘surge’ about 6 weeks after the introduction of the initial 81 contagious people. In this scenario the first surge decays after about 2-months. This ‘daily changes’ figure is PANDEMIC’s composite version of the display we have grown to recognize from reports in the media. With our choices for the timing of lockdowns enforcing higher levels of telework and

mask use, which are intended to approximate the societal measures implemented in response to the pandemic, the simulation reproduces the three surges of covid-19 cases observed in 2020. Since the infection can be transmitted several times during the 1-3 week delay between initial infection and the onset of symptoms, a reliable, cheap, and universally available viral ‘breathalyzer’ test is needed to get ahead of epidemics like this. For new viruses it is problematic to develop the test fast enough to stop the initial surge. In lieu of this, a combination of rapid testing, quarantine and contact tracing can help stop the spread.

PANDEMIC also has a simple representation of immunity, shown by the light green and dark green curves in Fig. 7. Depending on the person’s susceptibility, an increasing viral load reverses at a different non-fatal maximum V_L value for each person and then decreases through the contagious range again as V_L drops below the isolation level $V_L = 1.5$. The time rate of change of the several V_L ranges is evident in Fig. 8, including the negative daily changes that occur in all V_L conditions except death. Notice that the cumulative, integrated consequences highlighted in Fig. 7 are de-emphasized in the daily changes. The logarithmic scale, needed in Fig. 7 to display the integrated viral load data, is unnecessary for the ‘daily changes’ summary display.

Although the PANDEMIC scenario of Figs. 7 and 8 was relatively open for the first four weeks of March, it takes about two weeks for significant numbers of diagnosed carriers to appear and another week for significant numbers of people to require care. The peak of the daily fatalities is delayed by about six weeks, about three weeks after a strong ‘lockdown’ was instituted represented by significant mask use and 80% telework (8 of 10 people working from home). In the middle of June the decaying daily infections, hospitalizations, and deaths encouraged a weak opening that turned out to be premature. By the middle of August, a de-facto lockdown occurred as society responded to the worsening infections, though this was being opposed or ignored in some places. Nevertheless, the daily numbers began to decrease slowly in September in the simulation of the Washington D.C. Metro region. A third weak ‘opening’ in October sparked the beginning of the third covid-19 surge.

3. Testing Protocols to Mitigate the Third Covid-19 Infection Surge

In February 2021 we were deep into the third covid-19 surge in terms of cases, hospitalizations, and deaths. Masks and social distancing are beginning to blunt the surge after the holidays but new virus variants may make matters worse. This section presents four year-long simulations through February 2021 that display this third covid-19 surge under different simple testing protocols. Testing is a mitigation opportunity being now being added to help control the pandemic while waiting for vaccination to take effect. These runs show examples of the PANDEMIC scenario runs to concerning testing. These runs also spur the discussion in Section IV of the statistical “aleotonic” uncertainty of multiple Monte Carlo simulations of a given scenario. Every simulation uses a slew of random numbers to determine exactly where a person goes at work, to determine whether a particular close interaction transmits the virus, and to determine each person’s age and susceptibility. Running exactly the same background conditions will give a slightly different simulation depending on the particular set of random numbers chosen.

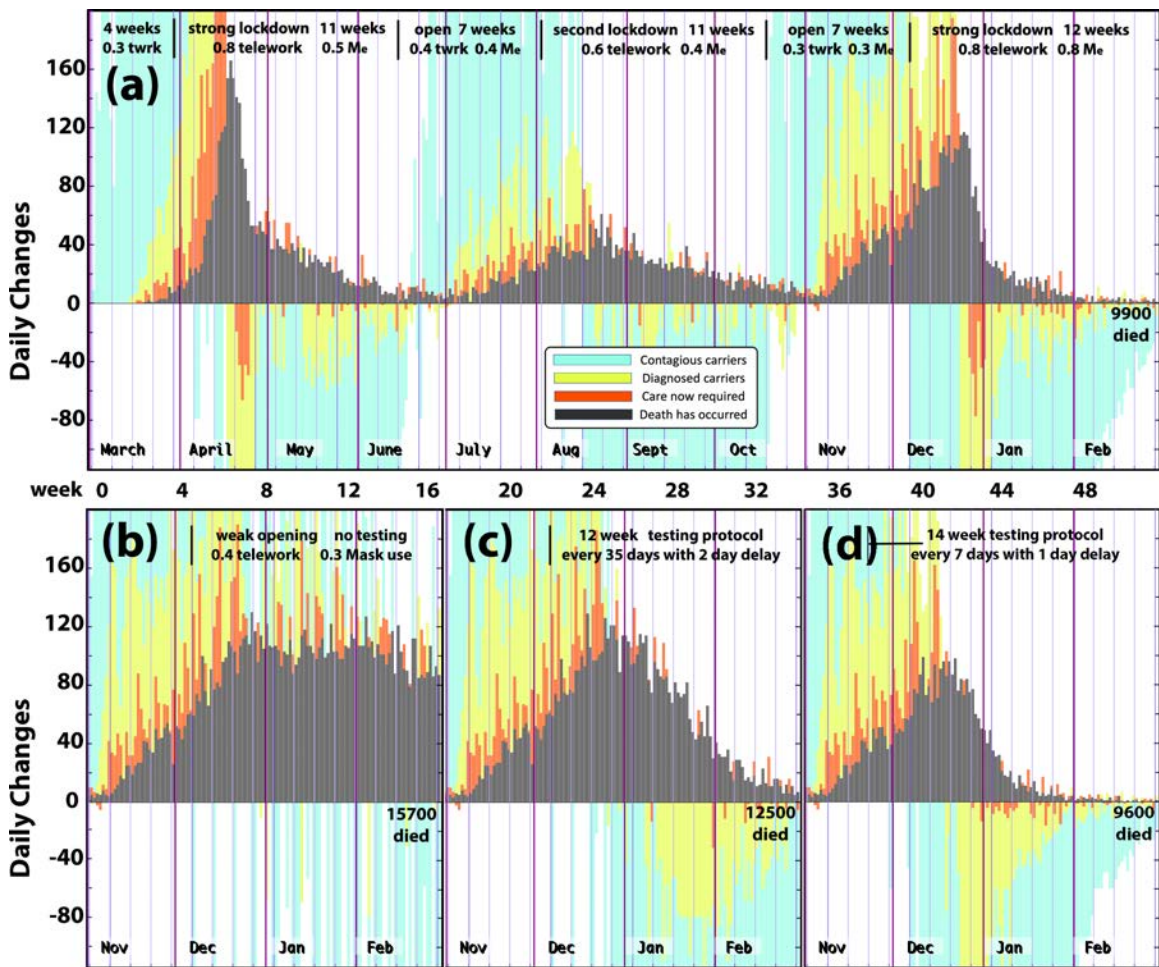


Figure 9. Four scenarios to mitigate the third covid-19 surge from November 2020 through February 2021. Scenario (a): no testing but strong (0.8) personal mask use and telework. (b) No testing with 0.4 telework and 0.3 mask use. (c) Testing every 35 days with a 2-day report delay. (d) Testing every 21 days with a 1-day report delay. Mask use and telework for runs (c) and (d) as in run (b).

The four 52-week testing scenarios presented here begin with the first 39 weeks illustrated in Figs. 7 and 8 above and are extended another 13 weeks until the end of February 2021. This extension allows enough time to see different evolutions of the third pandemic surge. Four different mitigation combinations were tried. The simulation in scenario (a) was run with no testing at all but with the mask efficiency ramped up to 80% and enforcing 80% teleworking as in the original strong lockdown beginning in April. Such diligent mask use and extensive telework almost certainly did not happen but the scenario suggests what could have been accomplished. The daily changes for the four principal viral load range indicators of scenario (a) are plotted for the full year in the top half of Fig. 9. The black bars, recording the fatalities each day, clearly show the three surges.

The rather precipitous drop in fatalities in scenario (a) from December into January results from increasing the mask use and confinement/telework by almost a factor of three from the relatively 'open' period 7 weeks before. Testing and contact tracing are not factors; the people are able to do this themselves. 9900 deaths have occurred in the DC Metro region by the scenario's end but the pandemic surge has been largely snuffed out. This 6.25 million Metro population contains almost exactly 2% of the US population so the 9900 fatalities extrapolate to about like 495,000 fatalities by the end of February. The region implemented here is

geographically and demographically diverse but it is not a perfect microcosm of the entire United States. The various implementation and mandate details and societal responses differ from state to state in the U.S. and these differences affect the local timing and severity of the pandemic.

In all three remaining scenarios (b), (c), and (d) the mask efficiency fraction was $M_{\text{eff}} = 0.3$ and the school, work and church average telework fractions were set at 0.4 (40%). Only the last 17 weeks of these three scenarios are plotted in the lower half of Fig. 9 since the first 39 weeks are the same as in the top half of the figure. The only differences between (b), (c), and (d) are in the virus testing protocol used. There are many ways to implement virus testing and test parameters such as time delay before results, cost, accessibility and test reliability/accuracy are major factors in the effectiveness of a testing protocol. In PANDEMIC we have chosen a single testing interval T_{freq} that is established for the entire population. Each person is tested every T_{freq} days regardless of where they have been or what they have done. Each person is assigned a testing day randomly so $1/T_{\text{freq}}$ of the population is tested every day. Each test is reported after a processing delay T_{delay} in days. Once a positive test is reported, the person is immediately marked as isolated/quarantined regardless of their viral load, This subsequently prevents tested contagious people from spreading the virus. The importance of rapid testing was investigated by varying, T_{freq} between 35 or 14 days and T_{delay} between 1 or 2 days.

In scenario (b) the testing interval is infinite (no testing). With the relatively low mask use and telework fractions employed, the pandemic is largely unchecked and results in 15,700 dead by the scenario's end in the DC Metro region. This scenario fatality count extrapolates to about 785,000 dead across the U.S. by the end of February, a kind of worst case. The IHME projection for this period is about 510 thousand deaths. Scenario (c) used $T_{\text{freq}} = 35$ days and $T_{\text{delay}} = 2$ days. 12,500 covid-19 deaths occurred, which extrapolates to 625 thousand deaths across the U.S. Testing at this level does not have a large enough impact to blunt the surge significantly. Scenario (d) use a more aggressive testing protocol: each person is tested every two weeks, unless they are already known to be infected, and the results are reported within a day. Testing was started two weeks earlier in the third surge for scenario (d). 9600 died by the end of February, the best result of the four. This extrapolates to 480,000 fatalities by the end of February, somewhat less than the two-week IHME projections. It is worth noting that these PANDEMIC simulations project forward several months making predictions for a region which, though not perfect, may be quite representative of the entire nation.

Figure 10. overlays the plots of the integrated viral load statistics for these four testing scenarios (a) through (d) for week 26 through week 51 (which is the 52nd week). There are four curves for each color representing the four scenarios. The lowest curves of each color are from scenario (a) and the top curves in each color are for scenario (b) with no testing coupled to low telework and mask use fractions. The lower black curves of covid-19 deaths for scenarios (a) and (d) are nearly fully overlapped, displaying only an apparent broadening. The low load immunes and high load immunes are shown for completeness but not discussed in the interests of brevity. The numbers of fatalities at the end of the simulations on February 28, 2021 are written in black at the right edge of Fig. 10 since the logarithmic scale makes reading the numbers difficult. The four scenarios started with 4880 fatalities in September, shown at the left edge of the plot. The third surge, November 2020 through February 2021, accounts for roughly half of all the deaths.

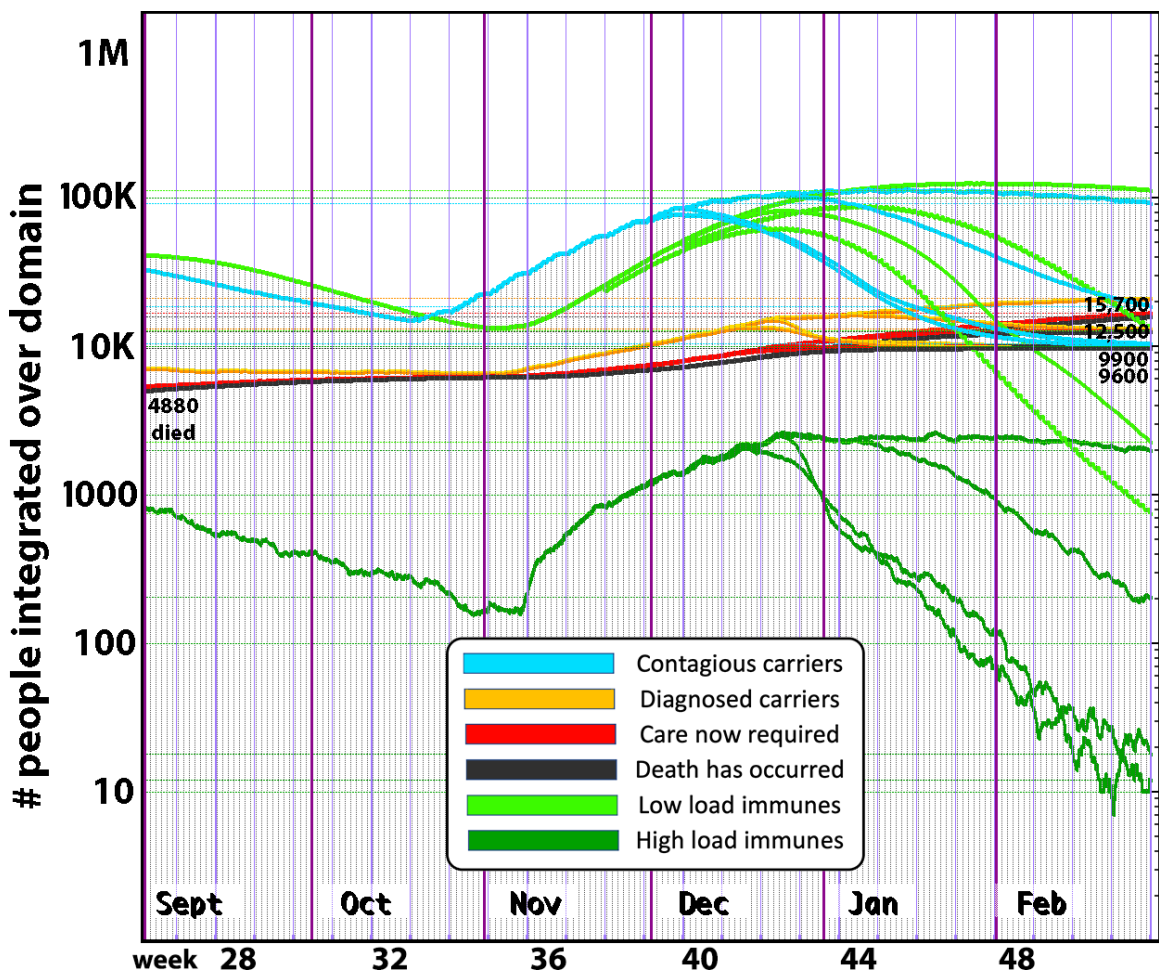


Figure 10. Testing reduces the impact of covid-19 on public health. The daily covid-19 domain- integrated statistics of the 4 scenarios from Figure 9 are superimposed for six different metrics of the time-varying viral load. Upper curves: scenario (b) no testing. 2nd curves: scenario (c) 35-day testing cycle. Lower two sets of curves: scenarios (a) strong masking/telework and scenario (d) 14-day testing cycle.

The testing protocol implemented was chosen for simplicity. Much greater care could be taken to reduce the number of tests or implement forms of personal testing. These are possible future extensions of this model and could also include reliability metrics for the various tests and a careful treatment of testing reliability as a function of the person's stage of infection. PANDEMIC currently does not have contact tracing capability implemented but all of the data necessary is available to do that, a major strong point for this type of detailed computer simulation. Different algorithms could be tried and optimized to seek optimal approaches. Once it is known how long vaccinated people and recovered people are immune to reinfection, these factors could be added to the simulation as well.

4. Understanding the Variability of PANDEMIC Simulation Realizations

In an epidemic the behavior of every individual potentially can affect the resulting minute-to-minute spread of the virus, which was noted previously. The details of an evolving scenario can only be specified using many random variables. It is impossible to know all of the ever changing input data in the real system, but we still can get a reasonable representation of the underlying average scenario variables when the system is large enough. We may understand the probabilities of the various possible conditions for the people even though we cannot predict the exact condition for each separate person. Random numbers are used to populate the many uncertain variables, making PANDEMIC an adaptation of a Discrete Simulation Monte Carlo (DSMC) method. The uncertainty due to the apparently random variability of model quantities, whose probability distributions may be known, is called aleatoric uncertainty. Aleatoric uncertainty can't be reduced; it always appears in the model solutions.

On the other hand, each PANDEMIC simulation of a specified scenario must be computationally reproducible. A later simulation with identical initial conditions must give exactly the same answers previously determined if the computer is working reliably. This is sometimes described as being "deterministic," a necessary condition for debugging and testing programs. In the presence of uncertain inputs, which are chosen from known or postulated probability distributions using random numbers, we must ensure that every case of the same run uses the same set of random numbers to be deterministic. Statistically independent "realizations" of a specified scenario can be found using different sequences of random numbers, often generated by changing a displacement in the random number sequence or by changing the 'seed' of the random number generator.

In setting up each simulation, random numbers are used to set locations, to determine the spatial offsets each day of people at work and at meetings, to determine who gets tested on which days, and to determine the age and susceptibility of each person. Care has been taken to ensure that, when effects are turned on and off or varied in strength in PANDEMIC, the simulation is only changed by the changed parameters and not by inadvertent changes in the random numbers used for other effects and processes. PANDEMIC does this by storing a large random number table initially and then by accessing it uniquely for each effect and process while never changing the access algorithm for that effect, despite how or when other parameters might be changed or other effects added to the simulation.

All the randomization algorithms have a different offset length controlled by a single integer N_{rlz} identifying the statistical "realization" of each scenario. The realization number N_{rlz} starts at 0 and can be increased to at least 100 without exceeding the overall random number table. Each different realization is a statistically independent simulation of the scenario and can be compared with other realizations to estimate the uncertainty in the results being computed. Figure 11, below, shows six realizations of an idealized 24-week simulation (nearly 5 months) computed with parameters approximating the third covid-19 surge seen in panel (c) of Fig. 9, the intermediate level testing case. An initial transient of about 3 weeks in these shortened PANDEMIC simulations approximates the 32-33 week initial 'transient' provided by the first two real-world surges in April and August.

The daily change plots for these six realizations, Realizations 0 through 5, are shown in Fig. 11 below. The six panels show day-to-day differences but these six realizations surely appear to represent the same scenario. However, the small parameter changes between similar underlying

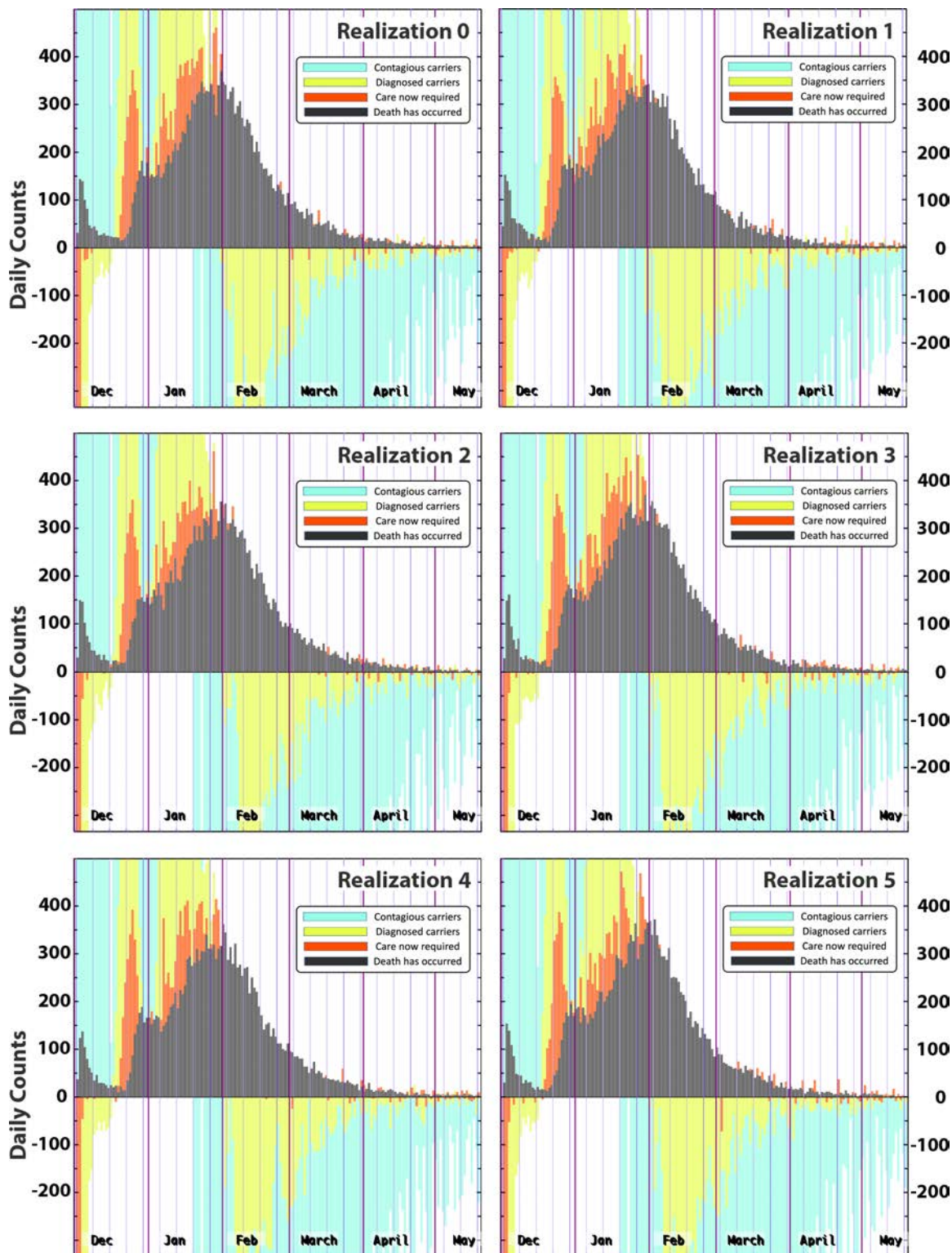


Figure 11. Six independent realizations of a single scenario. These 24-week runs used the baseline scenario with the 422 testing protocol. The “Daily Counts” displays for Realizations 0 through 5 show marked similarities but daily deaths, numbering in the low hundreds show 10% fluctuations and more.

scenarios, for example small changes in the telework fractions or the mask effectiveness fraction M_{eff} , may yield differences that are difficult to distinguish from the underlying aleotoric variability due to different random realizations. Building metrics to distinguish small differences between scenarios without a large number of realizations is a difficult task. Comparing the viral-load variables integrated over the domain is a good starting approach.

Figure 12 below shows the domain-wide integrated statistics from these six realizations overlaid in a single plot. Although the daily changes plots for the six realizations, such as seen in Fig. 11 above, show significant variations from day to day within a realization and between the corresponding daily values of the separate realizations, the overlaid integrated statistics are virtually identical. This is primarily because there are over 6 million people in the population. Noticeable differences only appear when the integrated numbers drop to 100 counts or less where 10% differences can be seen. An example in Fig. 12 is the relatively few “high load immunes” (dark green curves) encountered late in time and seen in the lower right of the figure. Two or three of the other integrated viral load range counts at higher levels are seen to be broadened somewhat by superimposing the six separate curves.

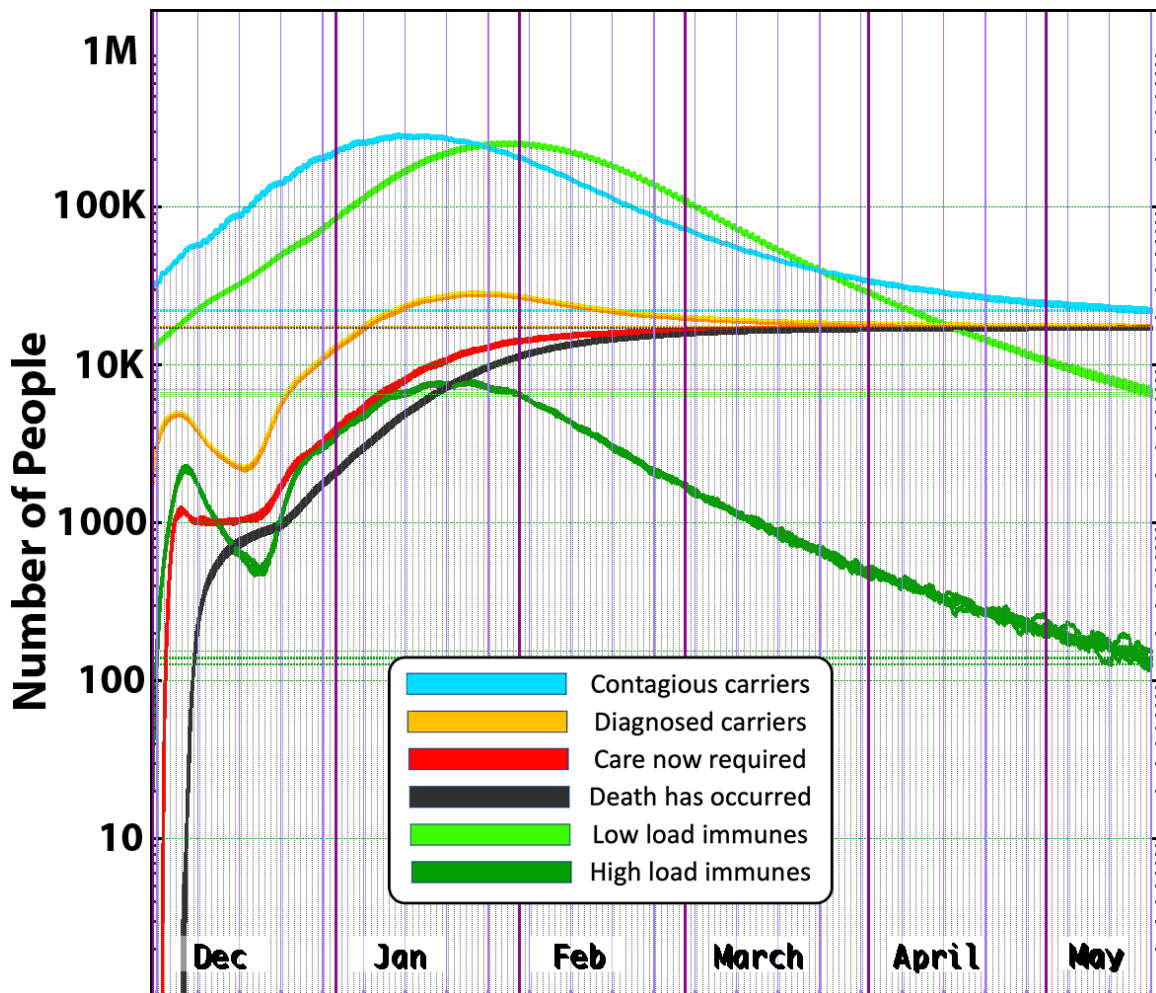


Figure 12. Six PANDEMIC realizations of the number of contagious and diagnosed carriers of Covid-19, people under care, and the daily death toll. These are the number of people in each group integrated over the entire 100 km by 100 km domain as a function of time. The width of the lines reflect the deviations between the six statistically independent realizations.

Being able to rely on the domain-integrated statistics to de-emphasize the statistical differences between individual realizations is reassuring but this assurance cannot be extended to problems that involve studying the differences between scenarios. This uncertainty is illustrated in Section V for an important application, evaluating the cost of super-spreader events.

5. Hypothesis Testing: The Cost of Crowds, Rallies and Super Spreader Events

The real public health cost of large unmasked meetings is an important issue in this time of pandemic. Crowds go to weddings, church services, political rallies, protests, sports events when permitted, and other gatherings. Adherence to established public health measures including the use of masks and adequate social distancing is often not honored at these events and thus public transmission of the virus is facilitated and as the rules are in flux due to vaccination. Some have hypothesized that rampant transmission is an advantageous strategy to reach “herd immunity.” This is an example of the sort of hypothesis that we can test using PANDEMIC. Based upon our simulations, we can now estimate the true cost in lives of a super-spreader event.

To this end, PANDEMIC was applied to two pairs of scenarios designed to approximate the large gathering held in Georgia on 5 December 2020. The attendees at that event came from all over Georgia with a number from out of state. In Georgia, about 1.7% of the people were infected at the time. Many of these were symptomatic, and would have been in isolation, quarantine or under care. Therefore, as an initial condition, we assumed that 0.5% of the entire population was out and about but contagious. This means there were over 31,000 active covid-19 carriers in the initial PANDEMIC simulation population. An initial 2 week period was allowed for the simulated pandemic to establish itself in the calculations before the possibly “super-spreader” event was launched computationally and followed for an additional 34 weeks. The crowd was approximated as a 150 x 150 person Monotonic Lagrangian Grid (MLG) array with a randomly varying separation between people that averaged 1.2 meters, about 4 feet. This simulated crowd is about the size of about 8 football fields. This also approximates the scale of the rally at Washington DC in January 2021.

At the end of simulation week 2, once the crowd was assembled, the 22,500 people were allowed to interact through the PANDEMIC exposure algorithm, which accounts for the separation between nearby pairs of people, their mask-use effectiveness, their viral load, and their duration in proximity. The average mask use effectiveness was assumed to be only 10% of maximum protection as most people in these crowds seemed to ignore mask mandates. After one hour, the people’s locations were shuffled randomly so everyone would get to sample other close interactions during the assumed 2-hour gathering. In reality, these events seem to last all more than 2 hours so the following may be under estimates of the transmission of the virus.

Three pairs of 36-week PANDEMIC runs were performed to estimate the number of additional deaths that result from covid-19 transmission of the virus between the people at such super-spreader events. The first run of each pair in a scenario is a control run where the crowd never assembled, i.e., the event never happened. Each of the 6.25 million people in the PANDEMIC domain went about their usual routines for 34 more weeks of home-work/church-school without attending the super-spreader event. The progress of the infections in the people selected to be in the crowd when the event was not actually “held” constitute a control to determine how many would have become infected, get sick, and possibly died anyway. The second run of each scenario pair uses exactly the same run parameters as the control but now the crowd gathers in

a small area for two hours at the beginning of simulation week 2, indicated by the vertical black line in Fig. 13. The model treated these as indoor meetings.

At week 4, an organized regimen of testing is also implemented in the each scenario pair. In the first scenario pair, comprising the two runs labelled ‘Skip’ in Fig. 13, testing was not used. This is rather like scenario (b) in Figs. 9 and 10. The second scenario pair of runs used a testing protocol, labelled ‘T422’ in the figure, with $T_{\text{freq}} = 42$ days and $T_{\text{delay}} = 2$ days. These runs have even less testing than scenario (c) in Figs. 9 and 10. The third scenario pair in Fig. 13 used a testing protocol, labelled ‘T211’ in the figure, with $T_{\text{freq}} = 21$ days and $T_{\text{delay}} = 1$ day. These runs used less frequent testing than scenario (d) in Figs. 9 and 10. As described earlier, any people testing positive go into quarantine and are no longer able to transmit the virus to anyone. The crowd deaths (colored squares at the bottom of the figure) and the other deaths in the non-attending population (colored circles) all level off roughly five months after the event was held. The data in Fig. 13 were determined by subtracting the fatalities in the control run of each pair from those in the corresponding super-spreader run.

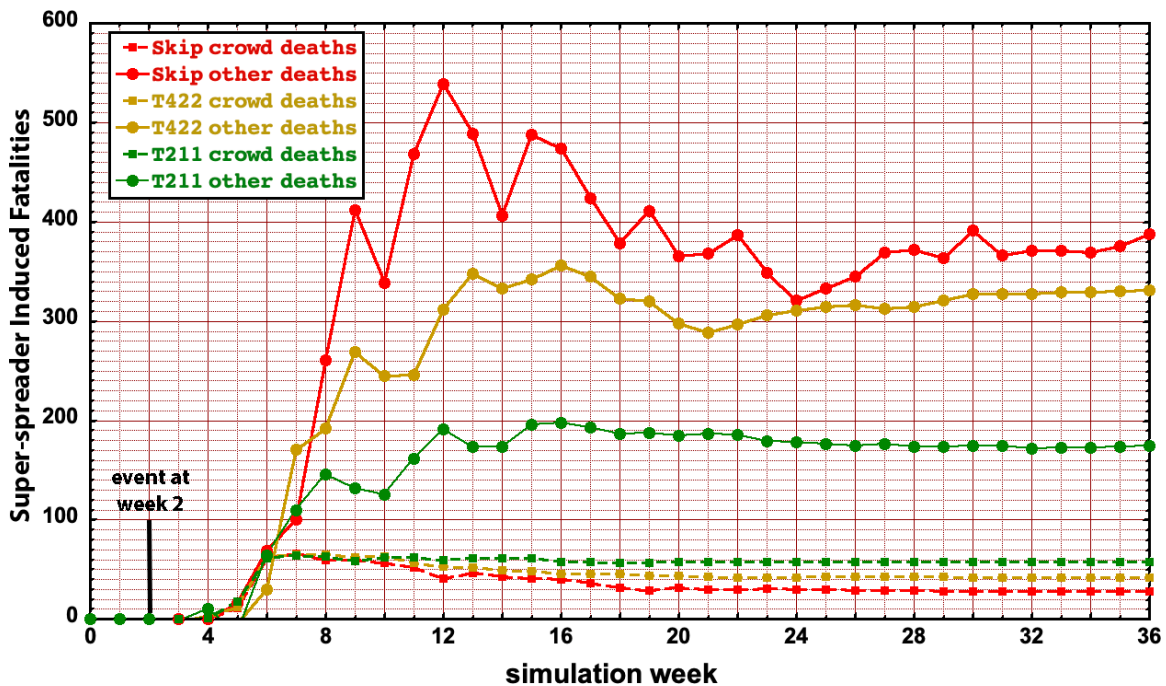


Figure 13. Super-spreader induced fatalities occurring among the 22,500-person crowd (150x150) and among the other background population (6.25 million) into which the crowd returned after the completion of the event. The overall death toll is 2 to 3 times the number of crowd members who died.

The first pair of runs for the ‘Skip’ testing protocol, which is probably closest to reality, are shown in the red data. Fluctuations in time are particularly severe in this scenario. About 400 people died, adding the two red curves together, who would not have died if the event were not held. Only 10% of these were people actually participating in the super-spreader event. The pair of runs denoted ‘T422,’ the intermediate testing case, shown in the gold data, had about 380 total fatalities as a result of transmission chains interrupted by even relatively sparse testing. Again about 10% if these were participants of the super-spreader event. The third pair of runs, with more frequent ‘T211’ testing to reduce the number of infections, whose fatality differences are plotted in green, showed a total of about 230 deaths due to the event, of which about 25% actually participated in the super spreader event. Testing had a major effect in this case.

At the end of the two-hour super-spreader event in week 2, the participants go home where they generally will not infect other participants, being well separated from the those in the crowd. In the most likely 'Skip' scenario, about 475 people were contagious when they arrived and about 2600 people left the rally infected. Of course, these people infected innocent non-participants in their immediate vicinities during their subsequent daily routines. Thus the pandemic expands. In a super-spread event of this scale, 22,500 people in a crowd that lasts a few hours, an estimate of the overall average number of deaths from the three testing scenarios in Fig. 13, is approximately 300. The current pandemic across the U.S. was claiming more than 2000 lives each day so this one hypothetical event could account for 10% to 20% of a one-day total for the whole nation. In the recent Washington DC indoor event, 6 deaths and a number of injuries, which occurred essentially immediately, have been documented. The actual death toll resulting from this gathering will be about 50 times larger with most of these hundreds of lost lives being innocent people who weren't even in Washington. After the event in Washington, 38 participating police officers were reported to contracted covid-19.

The number of additional covid-19 deaths, attributed to infections beginning in a super-spreader event, appears to vary greatly. This result is clearly evident in Fig. 13 and in Fig. 14 below. This variability arises in part because the number of super-spreader participants is much smaller than the total population and thus the number of infections is much smaller. Also, the cost of the event can only be determined by taking the difference between two computations

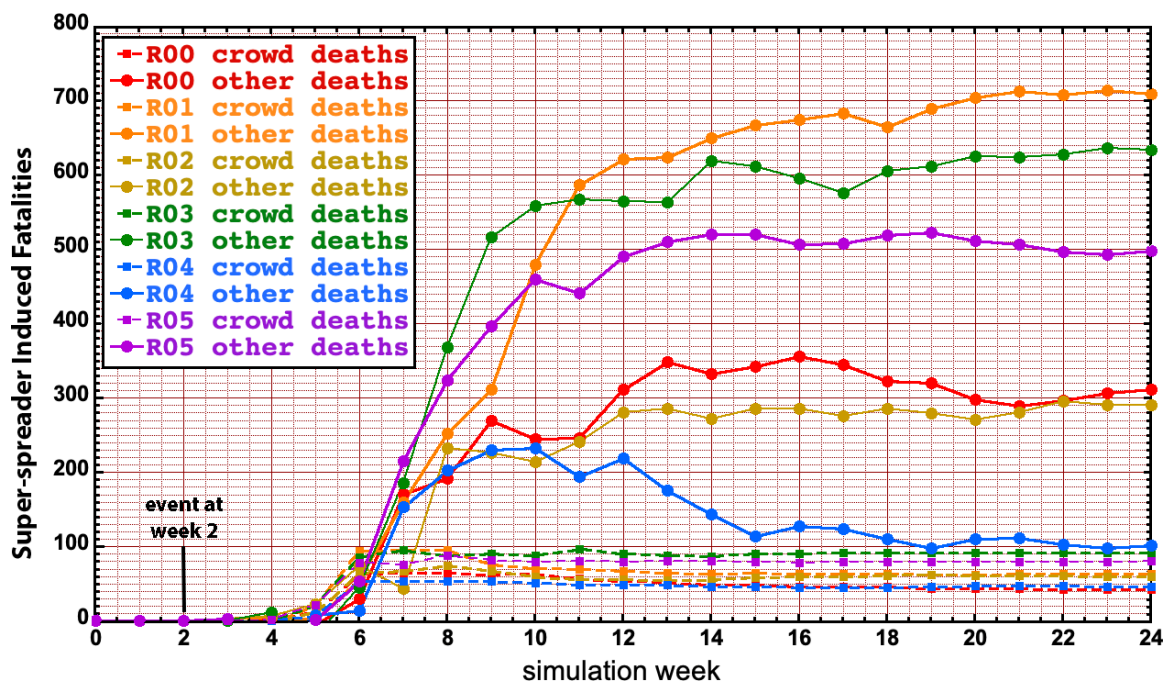


Figure 14. Covid-19 induced deaths in a large unmasked crowd and other people not in attendance as a function of time. 22,500 people in a population of 6.25 million congregate for two hours in a small area and then return to their usual daily routines at work, school, church and family. Six separate realizations of this event were compared with a control simulation where the crowd stays home.

that are separate Monte Carlo simulations. The number of deaths in the crowd participants, the squares in Figs. 13 and 14, seems to peak at around 6 weeks and then actually decrease. This decrease occurs because later in time, people in the control case, who would have been infected in the crowd or by people who participated in the crowd, are later infected by other transmission paths instead. Thus some of the fatalities, which could be attributed to the super-spreader later,

happen anyway in its absence. Those later fatalities are thus subtracted off the ‘cost’ of the event when they occur, reducing the ‘cost.’ Logically, if the pandemic spreads unchecked to its maximum lethality with the entire population contracting covid-19, everyone who could die will die, and the ‘cost’ of the super-spreader event is simply earlier deaths. Clearly the choice of the background control scenario is important to assessing the ‘cost’ of such an event.

To investigate these fluctuations further, six separate realizations were simulated of the T422 control and super-spreader scenario runs. The increase in the number of deaths due to the super-spreader event for these six PANDEMIC simulation pairs, each recorded in different colors, are displayed in Fig. 14 for 24 weeks rather than 36 weeks as in Fig. 13. These six realizations all use the same run parameters but the crowd deaths and the other deaths vary significantly from realization to realization because different sequences of random numbers are used. The much larger fluctuations seen in Fig. 14 arise because differences between large numbers are being taken to analyze each pair of runs constituting a control run/rally crowd pair scenario. The most extreme case is realization 4 (R04) in blue. Over 200 innocents in the ‘other’ population have died by week 10 but this count drops to just over 100 fatalities by week 24 as half of the other people infected by the participants are later infected in other ways in the background population control run. Clearly this sort of hypothesis testing is possible and instructive, but it must be done with care realizing that the intrinsic variability of PANDEMIC realizations needs to be appropriately accounted for in assessing the test results.

6. Discussion and Conclusions

PANDEMIC was developed at the U.S. Naval Research Laboratory following the onset of the covid-19 pandemic in the United States in February/March of 2020. It is a detailed Monte Carlo simulation model of the person-to-person viral transmission and progress of covid-19 based on the individual movements and interactions of millions of people in a finite geographic domain. The model implements the population’s daily routines to perform ‘what if’ hypothesis scenarios testing the range of public health personal advisories, public mandates, and epidemic mitigation strategies. PANDEMIC is not a population-based modelling, but rather a ‘person-based’ model, also described as ‘agent-based.’ “Population-based modelling is a top-down approach depicting disease dynamics on a system level. Agent (i.e. person)-based modelling is a bottom-up approach regarding each individual in the environment as an agent with their own movements and infection states. Population-based models are typically used for analyzing public-health questions from macroscopic perspectives, whereas agent-based models are used for analyzing microscopic behaviors.” (Kwok, et al., 2019).

There are many different population-based models. The most widely used population-based modelling approach is called SIER or SIERS. The SIERS acronym stands for ‘Susceptible’ – β → ‘Exposed’ – σ → ‘Infectious’ – γ → ‘Recovered’ – ξ → ‘Susceptible’. “The infectious rate, β , controls the rate of spread which represents the probability of transmitting disease between a susceptible and an infectious individual. The incubation rate, σ , is the rate of latent individuals becoming infectious (average duration of incubation is $1/\sigma$). Recovery rate, $\gamma = 1/D$, is determined by the average duration, D , of infection. For the SEIRS model, ξ is the rate which recovered individuals return to the susceptible state due to loss of immunity.” (IDM docs>EMOD, etc, 2021) This reference gives a solid description of the generic SIER model, implemented as a set of four ordinary differential equations for the four variables S , E , I , and R , the number of

people in each of the four SEIR categories. SEIR is the basis for the IMHE modelling effort (IHME^a,Murray,2020; IHME^b, et al.,2021).

Relating population-based models on the SEIR framework to the PANDEMIC approach, each of the four SEIR categories is, in effect, a range in the continuous ‘viral load’ value $V_L(t)$ assigned to each individual actor in PANDEMIC. Clearly each person in PANDEMIC is undergoing their own, asynchronous progression through something like the SIER stages using rate and demographic parameters appropriate for each person as determined by the scenario and the geographic domain set-up. The overall state of the system at any time must be determined by accumulating the status of the millions of individual ‘microscopic’ infections at a given time after the dynamics for each person has been computed. The ‘statistics’ of the scenario is only determined after the ‘physics’ of the individual actors have been calculated. This strategy always promises a more robust ability to predict when nonlinearity is such a large part of the competing effects and interactions. In science and engineering that approach is often characterized as “modelling from first principles.”

The leading covid-19 modelling effort based on the SIER approach appears to be the team effort at the Institute for Health Metrics and Evaluation (IHME) centered at the University of Washington. The IHME model results, coupled to an up-to-date database can be accessed under ‘projections’ at <https://www.worldometers.info/coronavirus/country/us/>. A modification that extends the SEIR framework to include some geographic variation via “spatial diffusion pathways”, rather than just at a national level, was developed by Klüsener, et al. (2020). The authors called this “microsimulation” but it is not yet a detailed actor-based model such as PANDEMIC. The coupled SEIR-like solutions were represented on three regional scales of aggregation from a few tens of regions to several hundred.

Actor-based models are better when studying highly inhomogeneous environments and scenarios that involve changes in conditions or behavior. Population-based models are usually much faster than actor-based models, and allow the import of aggregated public health data more readily. Actor-based models, on the other hand, do not require an extensive data base of actual cases, status and outcomes that needs to be kept accurately up to date and guaranteed to be interpreted uniformly over the domain. We chose to use the word ‘actor’ rather than ‘agent’ since agents can be human actors but also dangerous chemical or biological substances. In PANDEMIC the properties of the people (actors) are assigned from known or assumed distribution functions. The actors are drawn from demographic and geographic data but do not represent specific people with known health histories, jobs, routines, and families. This suggests the designation, used earlier, of PANDEMIC as a Discrete Simulation Monte Carlo model.

The objective of PANDEMIC is to calculate the spread and/or mitigation of an epidemic as different personal behaviors and public mandates are applied at different times in response to the -perceived overall spread of covid-19. The model encompasses evolving social considerations such as school, suggested or mandated lockdowns, limitations on public gatherings, telework restrictions, mask use, social distancing, and testing. Particular test cases discussed briefly in this paper are simulating the evolution of the main covid-19 surges in the United States and the actual cost of large, mostly unmasked crowds or ‘rallies’ in terms of additional fatalities. PANDEMIC reproduces the three surges of the U.S. covid-19 epidemic measured in cases, serious cases, and deaths when making reasonable assumptions about changing behaviors and ‘mandates’ before during and after the first two surges that peaked in April and August of 2020.

PANDEMIC scenario pairs of one ‘control’ simulation having no large unmasked crowd and a matching ‘rally’ simulation, in which the rally (crowd) is allowed to occur, show that the cost of such open gatherings is a few hundred additional lives lost. This is due to repeated retransmission of the virus, first at the gathering and then later after the infected attendees have gone back home and then to work, to school, and to church. Cast in terms of the January political rally in Washington DC, the actual death toll appears to be 50 to 100 times worse than the few lives lost in the event itself. Similar tolls can be expected of any large, unmasked crowds involving 10 to 20 thousand people crowded into a small area for several hours. Considering much larger mostly outdoors motorcycle rallies over much longer periods, the people spend many hours indoors in the evening hours, probably when much of the transmission take place.

The PANDEMIC project has been hurried along by the critical nature of the subject and was aided by the telework mandate we have been working under. Certainly the pull of finding more efficient parallel algorithms to be implemented under OpenMP for new computer architectures, given the computational constraints of working principally off a laptop connected to a state-of-the-art monitor at home, has also provided further incentive to get the model working reliably. Documentation has lagged, however, and many design decisions made early in development must be revisited before versions of the model would be ready to transfer to others. Considerable generalization is possible using the library of software that now exists but a streamlined framework is needed to keep further extensions from being too complicated to implement. PANDEMIC also needs a better interface, and perhaps even an interactive interface, to define the sequential stages of scenarios as more possibilities should be considered.

A number of extensions and refinements to PANDEMIC suggest themselves. The model implemented for virus testing is very simple. A more sophisticated model, with a number of variables and sensible conditions for testing and response based on the results, can and maybe should be written. Sophisticated testing and ‘Contact Tracing’ algorithms could be added and different strategies could be made the basis for computational campaigns. PANDEMIC currently does not change the fundamental parameters of the population depending on whether or not they have been vaccinated, which vaccine they have received, and when. PANDEMIC could also be extended to include different properties for the several virus variations that have been identified and the efficacy of the vaccines. Care and treatment, both in hospitals and at home, could be added as well.

In all these possible extensions, the detailed modeling approach permits capturing all the intermediate data generated by the model, such as who infected whom, where, and when, both for concurrent and subsequent analysis. In addition, the model stages and controllable parameters can be slaved to what is being measured and adjustments be made to the evolving scenarios automatically to test different strategies to mitigation. One important metric is the cost of detailed concurrent analysis and the costs of obtaining the data needed to perform the analyses. Further, case numbers and fatalities, are perhaps the primary metrics to be used. There are a number of directions that the development of PANDEMIC can go and the research that will use PANDEMIC can take. Such a model is never likely to ever be completely finished. As the covid-19 epoch winds down, there will certainly be new pandemics and flare-ups of old ones with new twists so refining and stabilizing this modelling technology should have a lasting value.

As in the present case, a new threat will have different properties so different ‘viral-load’ models will be needed, models that will be very uncertain in the early stages without firm

knowledge of the new virus properties. Using models based on pushing extrapolations forward based on extensive evolving data bases, which underpin much of the current projection technology, make conducting what-if scenarios difficult. We all need to address the “800 pound gorilla” in the room: the uncertain quantitative descriptions of the underlying processes and interactions, which have to be modelled to have a truly predictive basis. The uncertainty in these models is always going to be high at the beginning of a new threat. Detailed modelling approaches have had their origins in hard science and engineering and can be difficult to morph to the behavioral and social sciences. PANDEMIC is an attempt at bridging this transition.

Acknowledgements

Special thanks are due our colleagues Adam Moses and Gopal Patnaik for their numerous contributions to the technology we have used over the last ten years in herding this effort forward. Extensive discussions with Susan Bassin are also greatly appreciated. The research leading to this capability and preliminary demonstration was supported by the projects on “Highly Complex Fluid Dynamics” and “Advanced Computational Models that Exploit Many-Core Computer Architectures” within the ONR/NRL 6.1 basic research core program.

References

- J.P. Boris and S.G. Lambrakos (1985), “A Vectorized Near Neighbors Algorithm of Order N Using a Monotonic Logical Grid”, **Lecture Notes in Physics No. 238, The Free-Lagrange Methods**, Fritts, Crowley, and Trease (eds): 158–181, (Springer-Verlag, New York, 1985).
- J.P. Boris, J.M. Picone, and S.G. Lambrakos (1986), “BEAST: A High-Performance Engagement Area Simulator/Tracker”, Naval Research Laboratory Memorandum Report 5908, 31 December 1986.
- B.Z. Cybyk, E.S. Oran, J.P. Boris, and J.D. Anderson, Jr. (1994), “Application of the Monotonic Lagrangian Grid in Direct Simulation Monte Carlo Methods,” AIAA Paper 94–0354, 32nd Aerospace Sciences Meeting, Reno NV, 10 – 13 January 1994, AIAA, Reston VA, 1994.
- B.Z. Cybyk, E.S. Oran, J.P. Boris, and J.D. Anderson, Jr. (1995), “Combining the Monotonic Lagrangian Grid with a Direct Simulation Monte Carlo Model,” *Journal of Computational Physics* **122**(2): 323–334, December 1995.
- Willian Hanage and Helen Jenkins (2020), The Washington Post, April 19, 2020.
- IDM docs>EMOD (2021), “SEIR and SEIRS models,” <https://docs.idmod.org/projects/emod-hiv/en/latest/model-seir.html>, 2021.
- IHME^a COVID-19 Forecasting Team, C.J.L Murray (2020), “Forecasting COVID-19 impact on hospital bed-days, ICU-days, ventilator-days and deaths by US state in the next 4 months,” medRxiv preprint 27 March 2020. Doi: <https://doi.org/10.1101/2020.03.27.20043752>, 2020.
- IHME^b COVID-19 Forecasting Team (2021), “Modeling COVID-19 scenarios for the United States,” *Nature Medicine* **27**, pages 94–105, 2021, <https://doi.org/10.1038/s41591-020-1132-9>.
- C.R. Kaplan, E.S. Oran, N. Alexandrov and J.P. Boris (2010), The Monotonic Lagrangian Grid for Fast Air Traffic Evaluation, 48th AIAA Aerospace Sciences Meeting including the New Horizons Forum and Aerospace Exposition, AIAA Paper 2010–597, Orlando FL, 4–7 January 2010, (AIAA, Reston VA, 2010).

C.R. Kaplan, E.S. Oran, N. Alexandrov and J.P. Boris (2012), “Fast Tracking Methods and Systems for Air Traffic Modeling Using a Monotonic Lagrangian Grid”, US Patent 12/981432, issued July 2012.

S. Klüsener, R. Schneider, M. Rosenbaum-Feldbrügge, C. Dudel, E. Loichinger, N. Sander, A. Backhaus, E. Del Fava, J. Esins, M. Fischer, L. Grabenhenrich, P. Grigoriev, A. Grow, J. Hilton, B. Koller, M. Myrskylä, F. Scalone, M. Wolkewitz, E. Zagheni, M. Resch (2020), “Forecasting intensive care unit demand during the COVID-19 pandemic: A spatial age-structured microsimulation model,” medRxiv preprint 24 December 2020.

doi: <https://doi.org/10.1101/2020.12.23.20248761>.

K.O Kwok, A. Tang, V.W.I. Wei, et al. (2019), “Epidemic models of contact tracing: systematic review of transmission studies of severe acute respiratory syndrome and Middle East respiratory syndrome,” *Comput Struct Biotechnol J* 2019;17:196–194.

S.G. Lambrakos and J. P.Boris (1986), “Geometric Properties of the Monotonic Lagrangian Grid Algorithm for Near-Neighbor Calculations”, *Journal of Computational Physics* **73**(1): pp. 183–202, November 1987, also Naval Research Laboratory Memorandum Report 5761, 24 April 1986.

D.M. Morens and A.S. Fauci (2013), Emerging Infectious Diseases: Threats to Human Health and Global Stability. *PLoS Pathogens* **9**(7): e1003467. <https://doi.org/10.1371/journal.ppat.1003467>.

C.E. Walters, M.M.I. Meslé, and I.M. Hall, (2018), “Modelling the global spread of diseases: A review of current practice and capability “Modelling the global spread of diseases: A review of current practice and capability”, *Epidemic* **25**, pp 1-8, 2018.