Project Report
ACTA-5

The Simulation of Automated Exposure Notification (SimAEN) Model

D.W. Schuldt
E.H. Londer
J. Saunders
A.S. Norige
M.C. Schiefelbein
R. Yahalom
W.W. Streilein

15 April 2022

Lincoln Laboratory
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
Lexington, Massachusetts

This material is based upon work supported by the Under Secretary of Defense for Research and Engineering under Air Force Contract No. FA8702-15-D-0001.

This report is the result of studies performed at Lincoln Laboratory, a federally funded research and development center operated by Massachusetts Institute of Technology. This material is based upon work supported by the Under Secretary of Defense for Research and Engineering under Air Force Contract No. FA8702-15-D-0001. Any opinions, findings, conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Under Secretary of Defense for Research and Engineering.

Massachusetts Institute of Technology
Lincoln Laboratory

The Simulation of Automated Exposure Notification (SimAEN) Model

D.W. Schuldt
E.H. Londer
J. Saunders
A.S. Norige
Group 21

M.C. Schiefelbein
Group 46

R. Yahalom
MIT Sloan School of Management

W.W. Streilein
Biotechnology & Human Systems

Project Report ACTA-5

15 April 2022

This material is based upon work supported by the Under Secretary of Defense for Research and Engineering under Air Force Contract No. FA8702-15-D-0001.

Lexington Massachusetts

This page intentionally left blank.

ABSTRACT

Automated Exposure Notification (AEN) was implemented in 2020 to supplement traditional contact tracing for COVID-19 by estimating "too close for too long" proximities of people using the service. AEN uses Bluetooth messages to privately label and recall proximity events, so that persons who were likely exposed to SARS-CoV-2 can take the appropriate steps recommended by their health care authority. This paper describes an agent-based model that estimates the effects of AEN deployment on COVID-19 caseloads and public health workloads in the context of other critical public health measures available during the COVID-19 pandemic. We selected simulation variables pertinent to AEN deployment options, varied them in accord with the system dynamics available in 2020-2021, and calculated the outcomes of key metrics across repeated runs of the stochastic multi-week simulation. SimAEN's parameters were set to ranges of observed values in consultation with public health professionals and the rapidly accumulating literature on COVID-19 transmission; the model was validated against available population-level disease metrics. Estimates from SimAEN can help public health officials determine what AEN deployment decisions (e.g., configuration, workflow integration, and targeted adoption levels) can be most effective in their jurisdiction, in combination with other COVID-19 interventions (e.g., mask use, vaccination, quarantine and isolation periods).

This page intentionally left blank.

ACKNOWLEDGMENTS

The authors are grateful for the dedicated efforts and expertise of all who contributed to SimAEN's design, implementation, analysis, and technology transfer: Lauren Finklea, Jody McLean, Hammad Ali, Charlie Ishikawa, Jesslyn Alekseyev, Richard Gervin, Madeline Chiemelinski Malan, Sarah Miller, Greg Gianforcaro, Bob Hallowell, Nathaniel Hanson, Gwendolyn Gettliffe, Ben Davies, Joaquin Avellan, Chet Beals, Travis Riley, Maegan Jong, Elizabeth Bernardo, Brent Cassella, Tom Savel, Peter White, and Addam Driver.

TABLE OF CONTENTS

TABLE OF CONTENTS
(Continued)

LIST OF FIGURES

LIST OF TABLES

This page intentionally left blank.

1. INTRODUCTION

Coronavirus disease 2019 (COVID-19) is a viral respiratory disease, spread primarily through close contact with infected individuals who may not be aware of their infection status (pre-symptomatic or asymptomatic transmission).[1] Initial methods of disease transmission control relied heavily on traditional contact tracing methods, as well as profoundly disruptive stay-at-home orders and social distancing guidelines.[2] The success of contact tracing hinges on several factors, including (1) the ability of public health to identify and communicate with infected individuals, (2) the ability and willingness of infected individuals to share their contacts, (3) the ability of public health to notify these contacts, and (4) the willingness of these contacts to comply with public health instructions. The COVID-19 pandemic has particularly challenged contact tracing efforts:

1. In many jurisdictions, the size of the outbreak has exceeded local contact tracing capacity
2. Infected individuals may have more contacts than they can identify
3. Infected individuals may not know that they are infected
4. Distrust of public health institutions prevents some individuals from cooperating with contact tracing efforts

Automated exposure notification (AEN) was devised in March of 2020 [3] to complement traditional or "manual" contact tracing (MCT) activities for COVID-19. The overarching goal of AEN is the same as MCT: to notify individuals of their potential exposure to a disease to mitigate the disease's spread. AEN seeks to complement MCT by notifying contacts that MCT may have missed, and by notifying contacts faster, without compromising privacy. The Google-Apple Exposure Notification (GAEN) service is one example of an AEN system, and has been adopted by many countries and U.S. states[4], although it is not the only AEN-type system deployed during the COVID-19 pandemic. For the purposes of this paper, the term *AEN* refers to any automated exposure notification service that implements a privacy-preserving protocol such as PACT[5], and that relies on a proximity sensor on smartphones and the voluntary participation of individual citizens and public health jurisdictions in its operation. Prior work has found that AEN has the potential to control the spread of COVID-19.[6]

An AEN service uses Bluetooth messages to estimate when individuals were *too close for too long* (TC4TL) to an infected person and are therefore at risk for infection themselves. These individuals are notified of their exposure through alerts on their smartphones, which include instructions about testing or recommending quarantine. The process is anonymous: notified contacts are not made aware of which infected individual was nearby, and neither public health nor the infected person is aware of who receives notifications.

Public health institutions may determine how to set the TC4TL threshold for AEN—i.e., the sensitivity and specificity of the Bluetooth-based risk detector. The detector attempts to act as a dosimeter, recording exposure in both duration and distance. Lowering the detector threshold—i.e., increasing its sensitivity—corresponds to a nearer *too close* exposure and/or a briefer *too* *long* exposure. The intended effect of increasing sensitivity is to alert more infected individuals, with the goal of slowing disease spread. The side effect is that more uninfected individuals will be unnecessarily alerted (lower specificity), because the distance and time are not decoupled in the dose estimate comparison to the threshold. Many of these uninfected individuals will seek a COVID-19 test, potentially burdening their local test infrastructure. They may also quarantine unnecessarily, at social and economic cost to themselves and their communities. This trade-off between slowing disease spread and increasing public burden leads to a critical question: *How should public health set the sensitivity and specificity configuration of AEN?*

The answer depends on a variety of factors, including:

1. The properties and prevalence of the disease
2. The performance of the Bluetooth-based sensor
3. The behavior of individuals within the jurisdiction
4. The workflows and capacities of public health

These variables interact in complex ways, preventing obvious answers to the question above.

This paper presents SimAEN,[1] an agent-based simulation whose purpose is to assist public health in understanding and configuring an AEN system. SimAEN models a population of interacting individuals ("agents") in which COVID-19 is spreading. These individuals interact with a public health system that includes AEN and MCT. These interactions influence when individuals enter and leave quarantine and isolation, affecting the spread of the simulated disease. Over 60 user-configurable parameters influence the outcome of SimAEN's simulations. These parameters allow the user to tailor SimAEN to a specific public health context and to explore the predicted effects of various interventions, including different sensitivity settings of AEN.

[1] SimAEN, pronounced "SIM-ee-uhn," is short for "Simulation of AEN."

2. BACKGROUND

2.1 HEALTH CARE

Identifying exposed individuals, and advising them to test and/or self-quarantine, is only part of an AEN system's effect on a jurisdiction. This capability has to be integrated into the public health workflow around testing, test results, case investigation, and communication with the public, as well as with up-to-date epidemiological guidance on the transmission behavior of the virus, the effectiveness of vaccines, and the effectiveness of other non-pharmaceutical interventions (e.g., masking and reducing social interactions), and pharmaceutical interventions (e.g., vaccination), *in conjunction with each other*. Figure 1 shows a typical U.S. jurisdiction's workflow from 2020 through mid-2021, with the Google-Apple Exposure Notification service integrated into lab testing and case investigation/contact tracing activities. Note that although the public health authorities are mandated to carry out their work, individual citizens may be more or less likely to get tested, answer the phone, share their contacts' information with a contact tracer, or anonymously alert their close contacts through AEN.

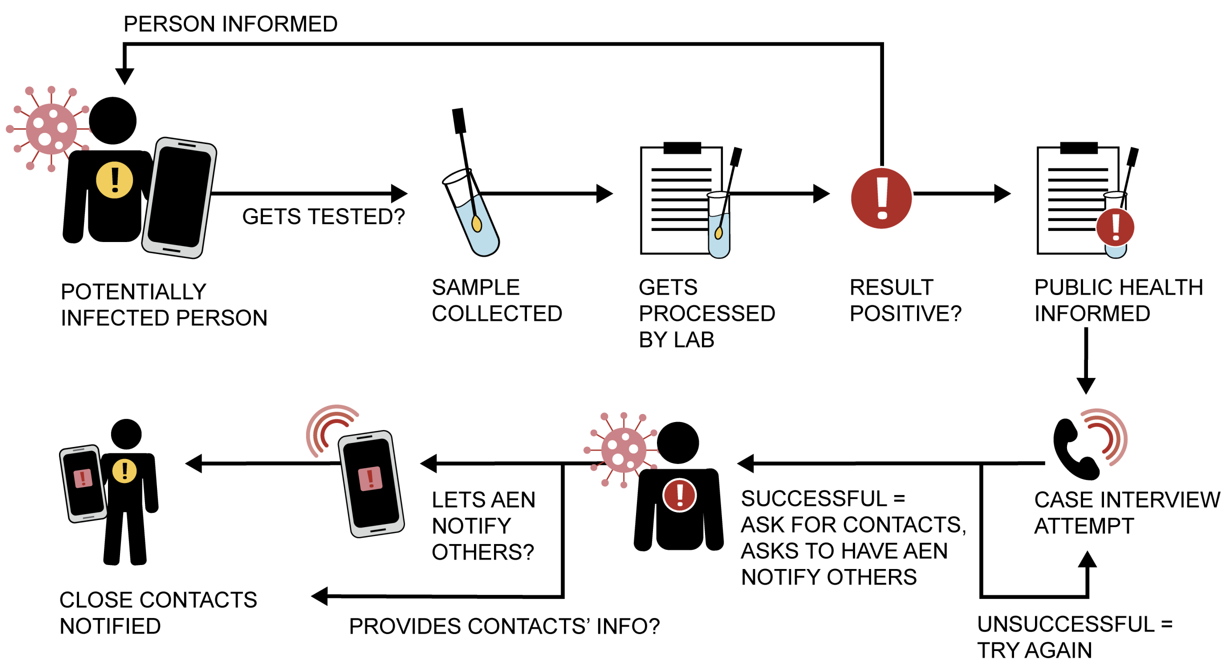

Figure 1. Original public health workflow incorporating Exposure Notification (EN) into testing, case investigation, and contact tracing activities.

In addition to integrating AEN with the testing and close contact workflow, a public health authority must decide how to configure the sensitivity and specificity of the Bluetooth sensor. For example, in GAEN, the configuration is a set of weights and thresholds that enable a public health authority to "bin" exposures into more or less risky categories, and alert and advise users according to their relative risk.[7] This will determine how many people are sent AEN alerts and subsequently how many people may test and/or quarantine. Choosing the “right” AEN configuration is not an exact science, due to variations in phone hardware as well as the effects of the phone’s use and environment on signal strength.[8] [9]

The next touch point between public health and AEN is through the testing procedure. Initially, states opted to authorize positive cases to upload their AEN keys via a call from a public health worker, but this was burdensome and states could not keep up with rising case loads, so many shifted to providing authorization via an automatic text message.

The public health jurisdiction also chooses the alert message to display to AEN close contacts. During different phases of the pandemic, states have chosen messages they found congruent with both available resource levels (e.g., for polymerase chain reaction (PCR) testing or home testing) and with the availability and palatability of other interventions (e.g., vaccines, mask mandates, and levels of lockdown). Therefore, the behaviors requested of close contacts have included calling public health, self-isolation, testing, vaccination, and simply monitoring for symptoms or wearing a mask if leaving home. The probability of compliance with any of these steps, and the concomitant effect on public health resources, is part of a complex feedback system that affects the spread of disease.

The choices that a jurisdiction makes in regards to these questions will determine the workload experienced by their contact tracers and other public health professionals. SimAEN aims to provide guidance to public health teams to help them select the AEN settings and structure their program in the way that best serves their constituents, in a variety of possible pandemic contexts.

2.2 HOW AEN WORKS

When two individuals, each running the AEN smartphone service, are close enough to each other that the Bluetooth signal from one phone is detected by the other, encrypted messages are exchanged between the phones. The message is stored by the AEN service on the phone for future comparison, along with the signal strength associated with the message. The encrypted message is rotated periodically as part of the privacy scheme.

After receiving a positive test, an individual may upload to a remote server all of the keys that they have used over some prior period (originally 14 days, but configurable by public health). This establishes a database of all recent keys for all positive individuals (stale keys are automatically removed). Several times a day, the AEN service on the smartphone downloads fresh keys.

When “unlocked” by the shared cryptographic key of the infected user, and combined with a weighting scheme set by the public health jurisdiction, the message history data generates a score for each interaction on the “close contact” phone. A cumulative daily score that exceeds the public health authority’s risk threshold will cause an exposure alert to appear on the phone screen, along with messaging selected by the public health authority regarding next steps for the individual.

2.3 MODELING APPROACHES

The prediction and estimation of disease spread have been performed using many different methods. In general, these methods land in three broad categories: compartmental models, data-driven models, and agent-based models.

2.3.1 Compartmental models

Compartmental models use a series of interdependent ordinary differential equations (ODEs) to describe the mean number of individuals in each of several categories. The total population is broken down into several subpopulations (susceptible, exposed, infectious, and recovered, among others) and the fraction of the population in each of the subpopulations changes over time according to the prescribed ODEs [10]. These models are typically described using a convenient shorthand (SIR, SEIR, etc.) indicating the subpopulations being considered. Owing to their relatively easy evaluation and understandable results, these models have gained wide adoption in the epidemiological community. To gain this tractability, the model makes many assumptions, such as that the entire population is homogeneously mixed. However, this assumption and other aspects of their formulation limits their ability to predict the real-world operational changes that occur when resources are limited and interactions are conditionally probabilistic. Small world models, such as those by Strogatz and Watts [11], show that mixing heterogeneity is omnipresent in human interaction and greatly affects the propagation of disease.

2.3.2 Data-driven models

The idea behind data-driven modeling is that given sufficient data about a population, you can discover a relationship among the inputs that minimizes the error signal between the model and some set of real-world outputs. This can take the form of neural networks [12], fuzzy logic systems [13], or regressions [14], which all use prior data to train the weights in the systems of equations that constitute the model. One problem with these models is that they can be opaque, meaning their internal workings are difficult to interpret and their outputs hard to justify.

This sort of model does not consider the population dynamics explicitly, instead focusing on measurable data and assuming that there is an underlying sensible reality generating them. By creating a system with enough freedom, the model is able to work as a surrogate for this reality. However, since they are only a product of the data, expert knowledge might not be considered.

Further, since they require initial data and only represent reality as it is currently operating, these models are not suited for investigating potential changes in the system or in the early stages of a disease before sufficient data has been collected.

2.3.3 Agent-based models

Agent-based models attempt to limit abstraction, instead depicting every member of the susceptible population and adjusting their condition based on the progression of the disease according to a set of rules. On the most abstract side of agent-based modeling are grid-based simulations known as cellular automata. This framework is most famously implemented in Conway's Game of Life [15] but has also been applied to disease dynamics [16]. These models treat the agents as nodes of a regular grid or irregular network [17] and update the state of these agents based on the state of their neighbors.

Less abstract agent-based models account for spatial variation in interactions. These models create agents that conform to schedules as they move through a virtual landscape. The population of agents can mirror reality in terms of demographics, behavior, and spatial distribution. Underlying these models is the assumption that if reality is intricately modeled, then the model will behave in the same way as reality.

The benefits of more realistic implementations do come at a cost. The amount of computation required to evaluate the model goes up as a function of the number of agents, and memory requirements expand as the list of agent parameters (e.g., features) grows. For these reasons, these models are reasonably recent developments, only just now being fully realizable thanks to modern, high performance computing architectures and tools.

3. THE SIMAEN MODEL

We aimed to augment the early literature on AEN's potential effects by creating an agent-based model that could be rapidly adapted to additional contact tracing protocols as they are developed, and to include new or refined parameters for disease transmission, human behavior, and public health resources. Our model also has a more realistic testing and quarantine model that takes delays of testing into consideration and allows for transition out of quarantine after negative tests. Further, because this effort focused on the public health response, the model provides probabilities associated with the subtleties of AEN implementation and configuration by the public health jurisdiction.

3.1 HIGH-LEVEL OVERVIEW

The workflow SimAEN uses is a simplified version of real-world actions surrounding COVID-19 exposure and interactions with public health. SimAEN has over 60 input parameters (described in Appendix A); all of these are variable in the model's configuration. We selected the model outputs to provide insight into the potential effects of combined interventions on public health outcomes (disease prevalence) as well as workloads. The specific model outputs are listed in detail in Appendix B. Our model validation focused on test counts, "cases identified" counts, and effective reproduction number (R_E), as those are most directly observable in the real world.

A high-level overview of the relationships between key model inputs (parameters) and outputs (metrics) is shown in Figure 2. Each multi-day model run generates and captures the complex interactions between probabilities, agent interactions, and the effects of interactions. The modules and connections in the center show a simplified view of SimAEN's internal state. The model's internal components include aspects of disease spread and intervention strategies, represented by white modules. Those modules that directly translate into model outputs are highlighted in blue.

The model includes several feedback loops. Positive (amplifying) effects are shown with green solid arrows, and negative (reducing) effects are shown with red dashed arrows. As each model run plays out over the specified number of days, the numbers of tests, calls, infections, isolations, and quarantines are tracked and reported out. As each model run is not identical, we run the module multiple times and average the outputs to produce a more accurate prediction for a given combination of input parameters.

The model was implemented in Python and is available on GitHub so that others may update parameter values and adapt it for their needs.[18] Some or all of the variables that define model behavior are accessible as input parameters to the model execution. This simplified the organization and execution of parameter sweeps over many runs of the model, for better fidelity across ranges of real-world conditions.

The model component was designed for low computational cost, but users with access to a high performance computing cluster will see significant benefit from parallelizing model runs. We constructed a small utility to generate the Cartesian product of the free variables, and distribute the model runs to multiple computational nodes. The utility is also responsible for maintaining the

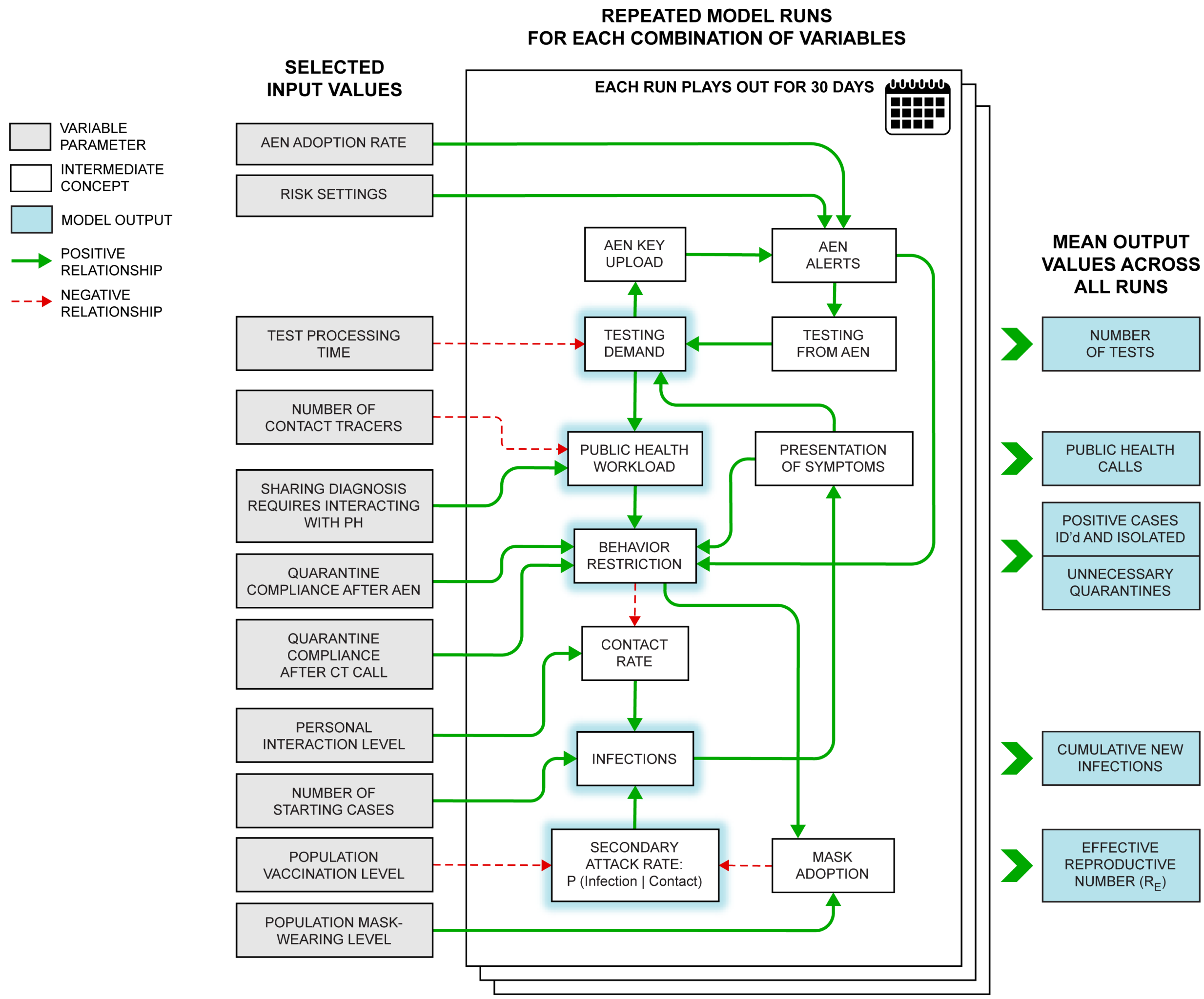

Figure 2. Simplified diagram of key feedback loops modeled within SimAEN, showing a subset of the input parameters and output values.

parallel outputs without collision from other processors, and finally, organizing them into a SQLite database for exploration.

3.2 MODEL DESCRIPTION

The SimAEN model definition is based on the Overview, Design concepts, Details (ODD) protocol as specified in [19]. This formulation establishes a standard way of communicating an agent-based model in a manner that allows for re-creation by other interested parties. The agents in this description represent people and we will refer to them as agents, people, or individuals interchangeably.

3.2.1 Overview

Purpose AEN is a new technology and there is need to understand its effectiveness. The SimAEN agent-based simulation was developed to explore and understand how AEN and its configuration may affect COVID-19 case rates and public health workloads, in the context of other available interventions. In addition to the independent effects of AEN, we also examined the effects of manual contact tracing, widespread testing, and mask use, as well as the results of deploying these strategies in combination with each other.

Entities, state variables, and scales Agents in the SimAEN model represent individuals whose interactions are guided by a collection of probabilities. The "world" in which individual agents operate is also implemented as a collection of probabilities and parameters (e.g, the duration of each phase of the disease, the number of contact tracers within the jurisdiction, etc.). Agent-based models are naturally suited to object-oriented programming methods, so both individuals and worlds can be thought of as objects—though for this simulation only a single world exists at any one time. Further information on the parameters of the individual and world objects can be found in Appendix A. These parameters determine the states that individuals traverse during the simulation.

The world in which individuals exist advances on a discrete schedule, where state changes occur once per day. We chose the 24-hour time frame because it captures the phenomena relevant to disease transmission, while also being long enough that it does not take an unacceptable amount of computation to determine results. In addition, longer time frames (e.g., once per week) would not permit the granularity associated with testing events or other characteristics of interest.

The individual is the smallest agent level considered. This model does not directly consider family units or workplace structures. The type of transmission events that take place in these settings are accounted for by modeling the events using a log-normal distribution. This distribution features a long tail, meaning that there is a relatively high probability of an event occurring where a large number of individuals contract the disease (e.g., in the tail of the distribution).

One of the key features of the SimAEN model is the low computational cost, achieved chiefly by not maintaining a fixed population. A typical agent-based model would track all members of the population, even if they never become infected or come in contact with an infected person.

However, these individuals have no impact on the results of the model, meaning that all of the computations required to process them are unnecessary effort. SimAEN avoids this by using the following approach.

At the start of the simulation, a chosen number of infected agents are created. For each of the agents, the model generates new people with whom they have the possibility of interacting. This group is called the agent's "neighborhood" and is distributed as

$$C_i \sim Lognormal(\mu_{neighborhood}, \sigma^2_{neighborhood})$$

where $\mu_{neighborhood}$ and $\sigma_{neighborhood}$ are the mean and standard deviation of the underlying Gaussian distribution. On each simulated day, the model draws from this pool to select the agents that the infected agent interacts with that day. The number of agents that are interacted with on a day is drawn from the distribution $C_i(t) \sim Lognormal(\mu_{int}, \sigma^2_{int})$. If any of these agents are infected, then the simulation will generate neighborhoods of agents for them to interact with.

Consider an agent A_i and their set of potential interactions $C_i = \{A_{i1}, ..., A_{in}\}$. Each day, A_i interacts with some subset of C_i. The subset it interacts with on day t is $C_i(t)$. For this model, $C_i(t_u) \subset C_i$ and $C_i(t_v) \subset C_i$, but there is no guarantee that $C_i(t_u) \subset C_i(t_v)$. Additionally, there is no enforcement that $\bigcup_{u=0}^{\infty} C_i(t_u) = C_i$; some individuals may never interact with others. When A_i recovers, all of the agents in C_i that were not infected are removed from the simulation.

Spatial units are not represented in SimAEN, although the chances of disease transmission are inherently spatially influenced. The duration and distance of the exposure are encoded in the BLEMUR model's prediction of the probability of detection and false discovery rate of the AEN detector. This simplification was made both because the disease transmission characteristics of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) were still under active research during the model development in 2020, and because more research is needed on contact characterization and distribution in various social contexts in order to inform a representative model of spatial behaviors.

3.2.2 Process overview and scheduling

The simulation advances one day at a time, during which each of the individuals in the simulation updates its status, potentially gets tested, is processed by public health, and changes its behavior.

Each day is broken down into a series of events where various aspects of the simulation are performed. These events always occur in the specified order, though this order was chosen arbitrarily.

The first event to be processed is transmission. During this event, each infected individual is evaluated to see whether they produced additional infections. Some number of uninfected individuals are also produced, based on the false discovery rate and the probability of transmission associated with the individual.

The number of agents interacting with individual A_i on day t is distributed as $Int(A_i, t) \sim Lognormal(\mu_{int}, \sigma^2_{int})$. Some percentage of the agents being interacted with will be infected based on the infected agent's stage of the disease (see: `Disease Properties` in Appendix A).

Following the transmission event is the testing event, during which all individuals are checked to see whether they test for COVID-19, based on a probability conditioned on the agent's traits (i.e., have they been notified of close contact status by AEN or MCT, as well as the baseline testing rate and symptomatic testing rate). If a test is performed, then a countdown is started, simulating the delay that occurs between test and results. The test is also entered into a queue to account for potential limits on the testing capacity of public health.

The next event is automatic tracing, in which close contact notifications are triggered by individuals who have tested positive on a prior day and have probabilistically decided to upload their keys to the AEN server. The delay between test results and automatic notification (in days) depends directly on the parameters for the number of contact tracers, the time each call takes, and whether sharing one's diagnosis with AEN requires interacting with public health to obtain authorization.[2]

In the final event of each simulation day, public health performs manual contact tracing. This step involves contacting a person who has tested positive to identify individuals they may have come in contact with and have potentially infected. Whether contact tracing succeeds in reaching the next generation of infected individuals depends directly on the parameters for call success rates (i.e., did the positive individual answer), the contact tracing identification success rate, and the maximum number of close contacts recalled. The timing of a successful MCT activity (i.e., length of delay after test result is available) depends directly on the parameters for numbers of contact tracers, call durations, number of call attempts to one individual, and length of the work day.

3.2.3 Design concepts

Basic principles This SimAEN model is based on the transmission and public health response associated with COVID-19. This means that transmission occurs through close contact between individuals, not through contact with a previously exposed surface, consumption of contaminated food or drink, or other methods of disease transfer. This affects the number of people that can be expected to become infected by an individual in any particular transmission event and the likelihood that a person would be able to identify the person who infected them or whom they may have infected.

One important assumption of our model is that the disease only ever infects a small portion of the population. In the standard SEIR model [20] the rate of change of the susceptible population is a function of the infected population as a proportion of the overall population. However, as long as this fraction is small, we can treat the susceptible population as a constant—a pool of individuals

[2] Initial GAEN deployments in the U.S. required public health teams to authorize AEN users to share their keys through issuing a one-time code, which the user would manually enter into the AEN service on their phone, before keys could be uploaded to the national server. The Delta variant caused case rates to rise more sharply and motivated public health jurisdictions to automate the delivery of authorization codes via SMS, instead of requiring a phone call between the positive individual and a case worker. It was a simple change to SimAEN to add a boolean parameter to support both workflows.

that we can always draw from. As such, we do not model the greater population of people who have not been directly affected by the disease. Instead, individuals are created at the time that they are needed and then disposed of once they are no longer integral to the simulation. The downside of this method is that it allows the number of people affected to grow without bound over long time scales. However, it saves significant computation by not simulating all of the people who are not impacted (which, per our assumption, is most of them). This also allows us to ignore the spatial aspects of individuals since transmission events do not have to occur at a given intersection of modeled individuals journeys.

The nature of COVID-19 also determines ranges of potential parameters for stages and duration of the disease. Infections are assumed to last 17 days [21] from initial exposure to recovery. In SimAEN individuals who are infected progress through the following stages of the disease: `EXPOSED`, `PRESYMPTOMATIC`, `SYMPTOMATIC` (or `ASYMPTOMATIC`), and `RECOVERED`. The SimAEN model does not account for the difference between recovery and death as both of these outcomes remove the individual from the system and do not have an impact on the methods of mitigation employed by public health.

Testing capability is also modeled based on what has been seen in the COVID-19 pandemic. Evidence from the current testing regimen suggests that there are very few false positive results. There is, however, a relatively high rate of false negatives.[22] These rates are dependent on what stage of disease the individual is in at the time that the test is performed.

Emergence We are interested in the proportion of individuals who need to be running the AEN service for there to be a notable effect on disease propagation. This fraction will be dependent on the expected time between when the individual contracting the disease and any of the people they infected receiving an alert from AEN service. It will also be affected by the other mitigation efforts taken, as they will have potential overlap with the population using AEN.

A quick analysis shows that there is a quadratic effect between the fraction of people who are running the AEN service and the fraction of people who will receive an alert. Assume that the probability of an individual running the AEN service is x. Since it requires that both members of an interaction be running the AEN service for either of them to receive an alert, this means that an interaction has a probability of $x * x = x^2$ of meeting this criteria.

Another emergent aspect of this model is the feedback between MCT and AEN. Individuals being contacted by MCT are more likely to get tested, and if they are running the AEN service and test positive, then more potentially infected individuals will be notified of this fact. This works in the other direction as well. However, there is also the overlap between the two (individuals receiving both an AEN alert and being contacted through MCT) reducing their overall effectiveness.

While the Google-Apple Exposure Notification service has emerged as a *de facto* world standard implementation, its risk estimation algorithm and Bluetooth sensor are only one possible and practical approach to estimating exposure "dosage" on smartphones. We decided that SimAEN would not encode the GAEN algorithm into its model; rather, SimAEN treats AEN's weights-and-thresholds configuration as an abstraction, and uses the BLEMUR [23] model's estimates for the probability of detection (*P(D)*) and false discovery rate (*FDR*) of two template configurations, the "narrow net" and "wide net" configurations (v1.0.1)[3] proposed at the November 2020 Risk Scoring Symposium Invitational.[26] This design decision for SimAEN means that it is simple to substitute another "close contact" detection and risk estimation system into the model if its *P(D)* and *FDR* are known.

Adaptation The SimAEN model assumes that there are two distinct levels of interaction: `NORMAL` and `RESTRICTED`. Agent in the `RESTRICTED` state have fewer close contacts, leading to lower levels of disease spread. Each day, agents are checked to see if they transition to `RESTRICTED` as a result of their condition. The probability of this transition is conditioned on several traits (such as receiving a positive test or being contacted by public health) that may also change during the course of the simulation.

Agents will also probabilistically don masks, conditioned on their level of interaction. It is assumed that as agents isolate themselves, they are also more likely to take other precautions. Once an agent starts wearing a mask, it will not stop wearing a mask.

The SimAEN model does not adapt agent behavior based on the progression of time or the prevalence of the disease. That is, agents do not respond to high disease rates by altering their levels of interaction or deciding to wear masks. Changes resulting from public health messaging such as deciding to promote mask wearing are also not modeled.

3.2.4 Objectives

Since the adaptation in SimAEN is purely driven by probabilities, there is no objective that is trying to be optimized. In the abstract sense, the agents are trying to minimize the amount of exposure to others, but this is projected on them by the selection of probabilities for behavior transition.

Sensing Agents are aware only of themselves with the exception that they are able to identify some fraction of the people who they have interacted with for the purposes of contact tracing by public health. Agents are able to identify if they are `SYMPTOMATIC`, as this is a trait that will affect their probability of getting tested or changing their behavior.

Interaction Agents are created along with the set of all individuals who they will ever interact with (apart from the agent that infected them). Each day a subset of this set is randomly selected for interaction, making them eligible for infection.

[3] The symposium participants reconvened in August 2021 due to mutual concern over the increased infectiousness of the Delta variant, and to review newly available data from the Exposure Notification Private Analytics[24] system. The RSSI-2 symposium produced a revised pair of template configurations.[25] The work shown in this paper used the original narrow and wide configurations, but it is a simple matter to adjust the inputs to BLEMUR and SimAEN for a jurisdiction's desired configuration.

During automated and manual contact tracing, individuals interact through an intermediary, either the AEN service or the public health system.

Stochasticity This model is driven significantly by stochasticity associated with the initial parameter settings. These probabilities are outlined in Appendix A, but a brief overview is given here.

How an agent experiences the disease are determined by the latent and incubation times, which are drawn randomly from a normal distribution on a per-agent basis.

Several agent traits have to be set during agent creation. First, the agent is probabilistically set to be wearing a mask. The mask-wearing state of the new agent is combined with the mask state of the infecting agent to determine whether the new agent is set as infected or uninfected. Finally, the new agent has some probability that they had previously downloaded and are running the AEN service. The agent is also assigned a collection of individuals who they have the potential to interact with over the course of their time in the simulation.

During transmission events, the number of individuals that an agent interacts with is probabilistic, conditioned on the behavior state of the transmitting individual. If the transmitting individual is running the AEN service there is also the probability that they will interact with some number of uninfected agents who were not close enough to have been infected but were identified by the AEN service. The false discovery rate models the inaccuracies inherent in the Bluetooth “close contact” detector. If the person transmitting the disease and the person being infected are both running the AEN service there is some probability that the AEN service on either end of the transmission will detect the signal.

Whether an agent gets a test is probabilistic, conditioned on whether they have been contacted by public health, received an notification from AEN, tested positive, or are feeling symptomatic.

The probability of a test coming back positive is based on the stage of the disease the agent is in at the time of the test. As noted in the assumptions, there are no false positive test results.

There is a probability that after receiving a positive test, an individual will contact public health for the purposes of contact tracing. During contact tracing there is a probability that any given call will successfully reach the agent. If the return call from public health is successful, then there is a probability that the traced person infected will be identified.

A positive test also comes with the probability that the individual will upload their key to the AEN server. When the key is subsequently downloaded by users of the AEN service, they are alerted of their potential exposure.

Finally, there are probabilities associated with an individual changing their interaction behavior, conditioned on the traits of the individual. Agents who transition to `RESTRICTED` will stay in that condition, except if they receive a negative test, in which case they will probabilistically transition to a behavior state based on the distribution of states currently occupied across the simulation.

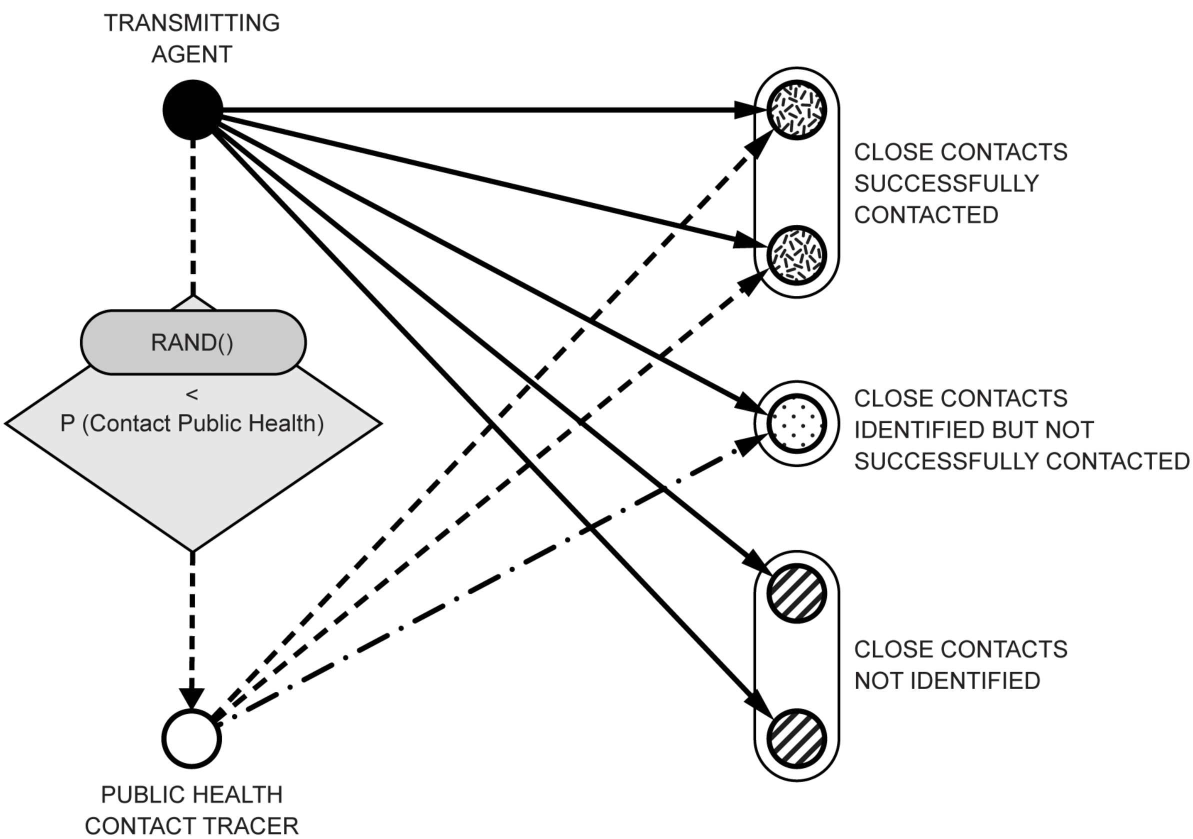

Figure 3. Agent generation (solid lines) showing the conditions that agents can obtain based on whether they are identified by the generator (dashed lines) according to a probability distribution.

Collectives There are no collectives, such as families or work groups, included in this model. All individuals are treated based only on their own traits and do not have any probabilities or parameters based on their generator or the agents they are generated alongside.

The closest aspect to a collective in this model is the set of individuals with which an agent is generated. This group represents the entire population of people with whom that agent has any chance of interacting. Individuals in this set are the only people that the agent can infect.

Observation Each agent keeps track of every transition that it makes while being simulated. Tracking this information supports forensic analysis of the simulation and how it progress. Examples of information that agents track include all of their traits (such as whether they have been tested) along with the days (since simulation start) on which those traits changed. They also keep track of all generation events including the status of all of the people involved in that event. Agents also track things for which they do not have direct knowledge. For example, agents keep track of how many times they have been unsuccessfully called by public health. This information is not used by the agent but supports deeper understanding of the simulation AEN process.

3.2.5 Details

Initialization The simulation begins with a collection of infected individuals. For simulations starting from the "initial outbreak" state, we chose 20 infected agents to start. This is a small enough number that it will not overwhelm the steady state yet is large enough that the disease will be able to take hold and propagate. It is assumed that there has been some low level of the disease circulating in the population prior to the exponential grown segment of the spread, and so these 20 individuals are initiated at a random point in the disease progression. This is accomplished by starting them with a "day in system" variable set to a random value drawn from a uniform distribution over the 17 days of the disease lifespan. Each run starts with a new random specification of these 20 agents. These individuals are also assumed to not have been tested, since we are starting from a time before widespread testing is available and AEN is in use.

For simulations of the disease in later states of spread, a multiplier was used to account for the potentially large number of individuals being infected. One simulation we performed looked at the state of Massachusetts in February 2021. At this point in time, there were an estimated 72270 active cases in the Commonwealth [27]. However, since the exact mix of individuals (what state of the disease they are in) was unknown, we had to adjust this number slightly in order to get day-to-day new case numbers which match the observed rates. We found that a starting cases count of 53000 produced this match. This lower number is likely because of lower infectiousness of individuals late in the disease progression. Instead of simulating all 53000 individuals, we started with 530 and multiplied all outputs by 100. Simulating only a fraction of the population also requires a change to the number of contact tracers and daily testing capacity since each agent now effectively represents 100. This method slightly decreases the fidelity of the simulation, but the gain in speed is fair compensation.

Input data All information necessary for progression of the model is contained within the simulation itself. The only outside input is through the selection of the parameters that drive the system. Once the model is started, it carries forward without any additional input.

3.2.6 Submodels

The SimAEN model includes submodels for testing, behavior, and public health interventions (AEN and MCT).

The testing submodel assumes that individuals will seek out a test contingent on their traits. Each day, a random draw occurs for each individual to determine if they will get a test performed. When an individual gets a test, it generates a `TEST` object. Each day, a number of tests is processed based on the testing capacity parameter. When a test is returned, a random draw is made to determine the result (positive or negative), conditioned on the state the agent was in when they were tested. A positive test will further prompt a random determination of whether they will upload their key to the AEN system and/or call public health. Contacting public health will add them to a list of index cases.

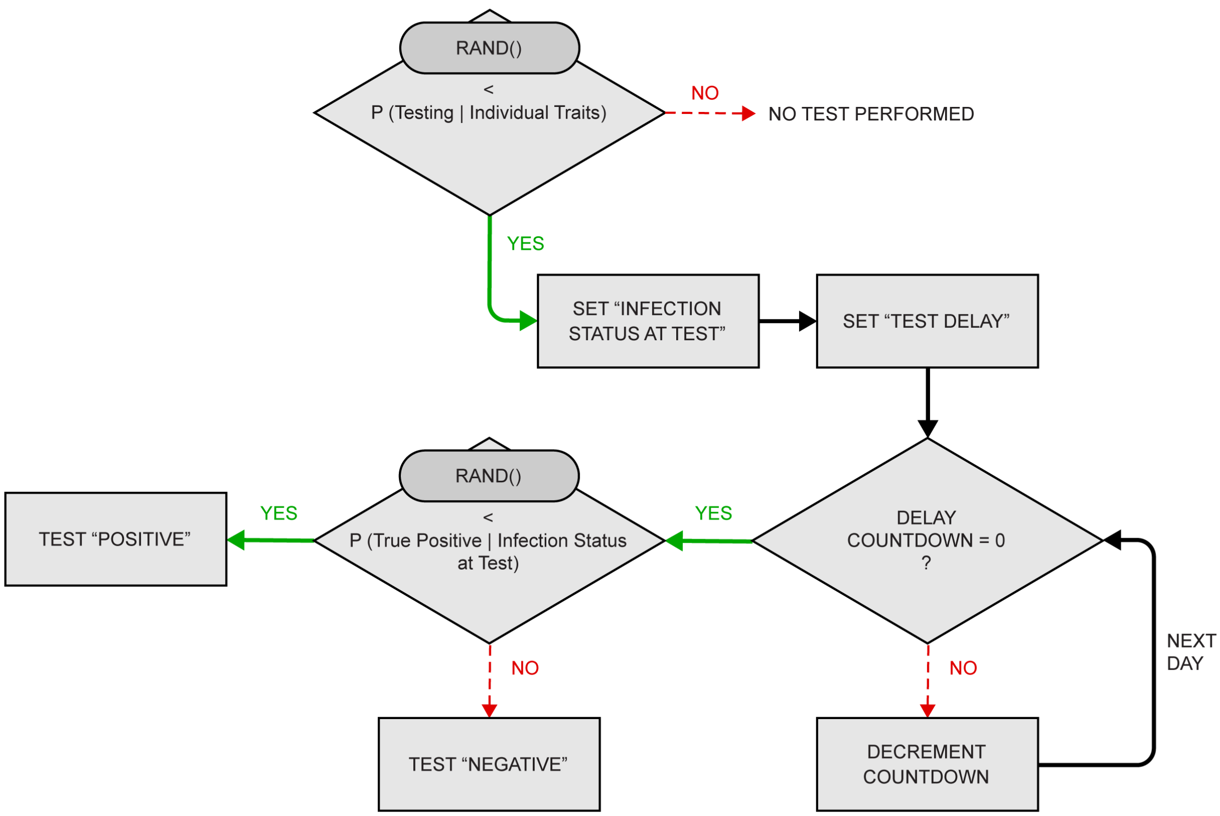

Figure 4. Testing methodology as implemented in SimAEN.

Behavioral changes of the individual are based on their traits. All agents start out in the normal or base state where they are interacting with society in the way that they would have were they not infected. Each day, a random draw is made to determine if the individual will transition to the `RESTRICTED` behavior. It is assumed that an individual in the `RESTRICTED` state will stay in that state unless they are not symptomatic and receive a negative test result. While people in the real world may have day-to-day-variation in their level of interaction, we assume that the probabilistic nature of transmission employed in this model is sufficient to describe this and that variation will not significantly affect the measures of interest being studied.

In the AEN submodel, individuals receiving a positive test have a random probability of uploading their keys to the AEN system. Some jurisdictions have selected a workflow that requires individuals talk to public health before uploading their key. In this case, the individual running the AEN service who received a positive test will also generate a `CALL` object. Once this call is processed, the agent will upload their key. Uploading a key triggers the AEN system to transmit an alert to all close contacts of the uploader who were running the AEN service and received a beacon message during the transmission event. All of these agents will update their object variables to note that a notification was received.

MCT encompasses both contact tracing and notification of identified contacts. It is assumed that there is a limited number of contact tracers and that each works some period of time each day. Each call they make takes some amount of time, which is based on whether it is a call for the purposes of contact tracing or just as a notification. There is also some time associated with missed calls. Our manual contact tracing submodel processes an ordered list of `CALL` objects (the "call list"). During the manual contact tracing portion of the daily simulation evaluation loop, calls go out to people in the order they were placed on the call list. When a call occurs the appropriate amount of time is subtracted from the available call time:

$$available\,call\,time \;=\; \#\,contact\,tracers\;\cdot\#\,work\;hours/day$$

If the call being made is to an index case, then a contact trace is performed. If any individuals are identified during the trace, they are added to the call list. Individuals who are not successfully called are added to the back of the contact list. Missed calls are logged and after an individual misses their allotment of calls, public health will assume that they are unreachable and remove them from the list. When there are large numbers of people on the call list, the call list may not be cleared on a single day. In that case, the contact tracing calls will pick up the next day at the point where they left off the day before.

3.3 VALIDATION

The parameters used at the time of model development and validation (i.e., before February 2021) were based on established research, where available. This includes the probability of transmission for presymptomatic, symptomatic, and asymptomatic individuals[28], the rate of asymptomatic cases[29], the effectiveness of masks[30][31], the lengths of the incubation [32], latent[21] and infectious periods[21], probability of a receiving a positive test [33], and daily personal interaction survey data [34].

Transmission also depends on the number of interactions that an individual has on any given day. Our model assumes that each day, an individual interacts with a number of people drawn from a log normal distribution parameterized by σ_{int}, μ_{int} as shown in Figure 5.

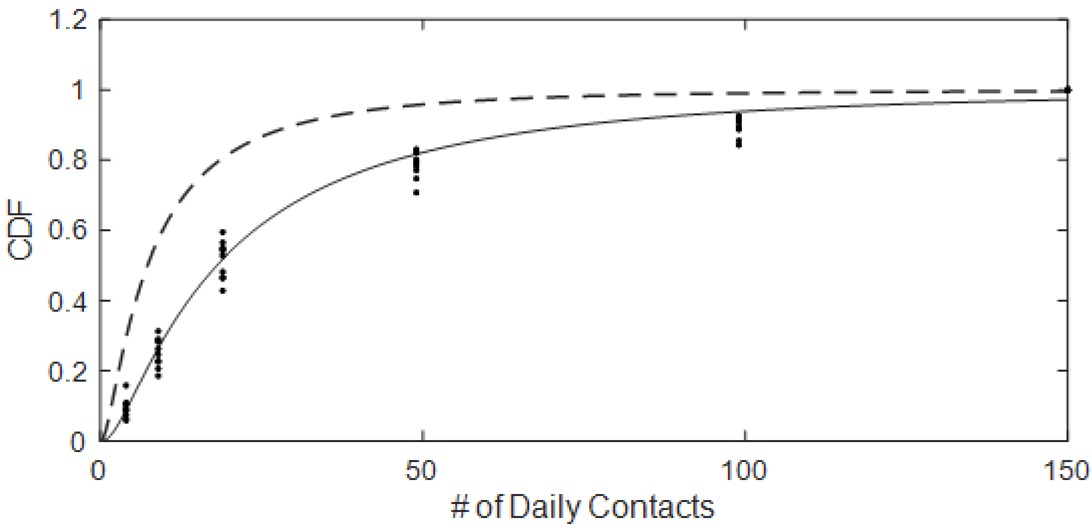

Figure 5. Cumulative distribution function (CDF) of daily interactions (dots), along with best fit log normal distribution (solid line; $\mu_{int} = 2.1, \sigma_{int} = 1.1$) and Pennsylvania fit (dashed line; $\mu_{int} = 1.9, \sigma_{int} = 1.1$).

Validation of the model was performed by showing that given these parameters, the model reproduces the rate of spread in real-world conditions. For our case, we compared the output of the model with the early days of the outbreak in Pennsylvania. Using the initial phase of the outbreak eliminates several confounding factors, such as the prevalence of mask usage and changes in interaction behavior.

As can be seen in Figure 6, the parameter set shown in Figure 5 produces infection curves that match the real-life data as observed in Pennsylvania. The use of a mean interaction parameter lower than the best fit curve for [34] ($\mu = 2.1$ vs. $\mu = 1.9$) is justified on the following bases:

- The counts from [34] are not unique interactions. Since most people interact with their family and coworkers regularly, we would anticipate a lower mean number of unique contacts.
- The Pennsylvania case counts are for the entire state, which includes many rural areas that may have significantly lower numbers of interactions than those found in the more urban study.
- Due to the high number of asymptomatic cases, the number of reported cases is likely much lower than the true number of cases.

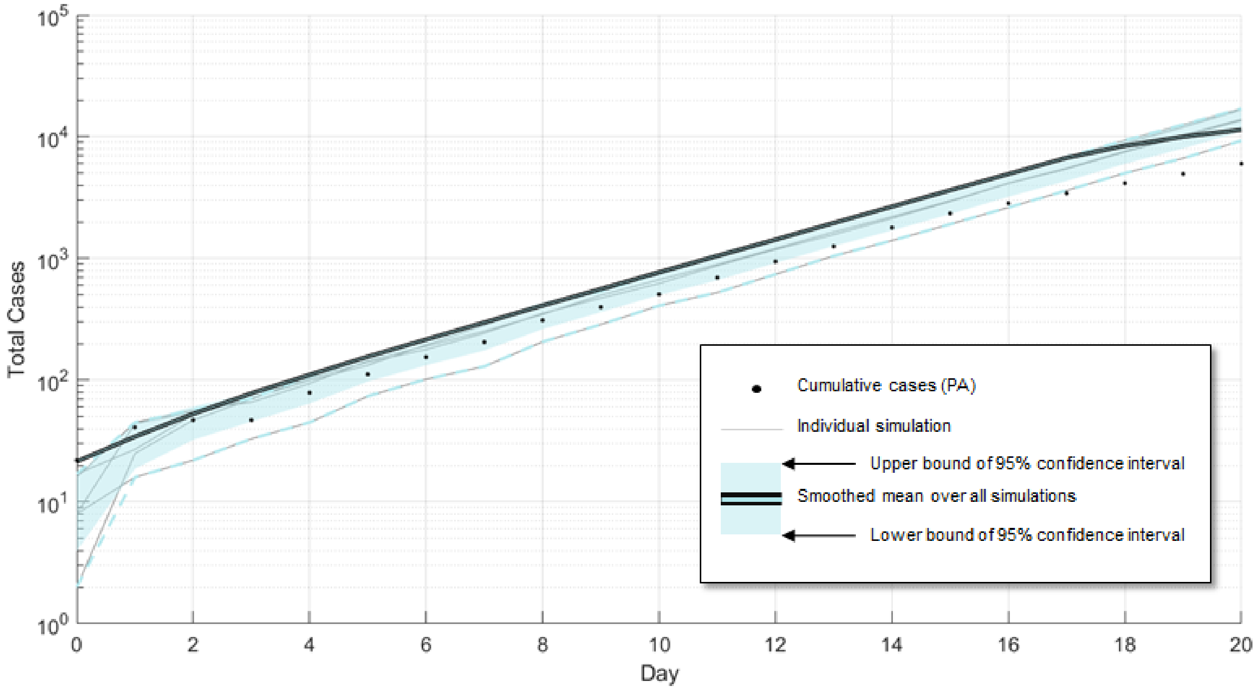

Figure 6. Infection growth for the first 20 days of the COVID-19 outbreak in Pennsylvania, alongside the results of simulation from SimAEN without any active interventions.

- The first cases in Pennsylvania (March 10, 2020) were several months after reports from Wuhan (January 2020), which may have prompted changes in behavior of the population.

The second stage of validation was against an analytic projection of Massachusetts metrics based on fractions derived from a variety of sources. The methodology for projecting the metrics is described in Appendix C. SimAEN's metrics are plotted against the projected mean metrics in Figure 7, showing that the behavior of the simulation across a wide range of measures is close to the analytic projections.

To create Figure 7, the simulation parameters were set to those listed in Appendix A and run 20 times, each with a different random seed. The results of these runs are shown in grey. The daily mean across these runs is shown as the solid black line. One standard deviation is denoted with dashed black lines. Averaging the daily means produces the overall mean, which is shown in cyan. Finally, the analytic result is the solid red line. We see good agreement between the cyan (average simulation) and red (analytic) lines, indicating that the implementation is consistent with what is observed from end to end.

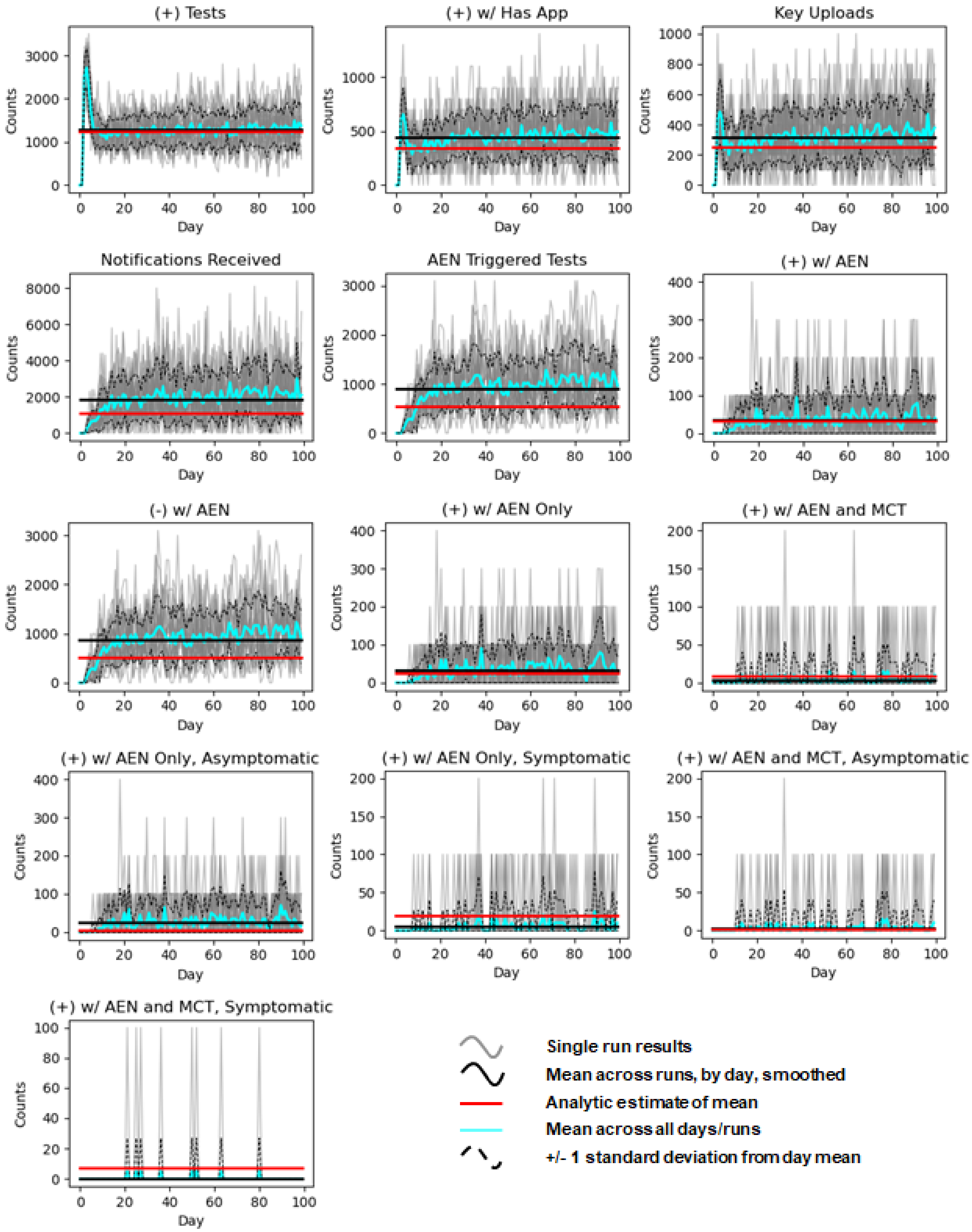

Figure 7. Comparison of SimAEN results and the analytic Massachusetts results.

4. RESULTS

4.1 EFFECTS OF AEN ADOPTION

The calculated effective reproduction number (the mean number of new infected per infected individual) is a useful metric for judging an intervention's effect on the *spread* of disease. The R_E predicted by SimAEN for the baseline parameters used here, and for a range of AEN adoption rates, are shown in Table 1. We see larger absolute effects on R_E for both higher interaction rates and higher AEN adoption rates, when varied independently. When using 0% AEN adoption (i.e., no AEN deployment) as a control, and increasing the AEN adoption rate, we also see larger percentage effects. Notably, at the interaction rate observed prior to the pandemic ($\mu = 2.9$), we see larger absolute and percentage changes in R_E than at the lower pandemic-level interaction rate.

The percentage reduction in R_E for each level of interaction suggests that increases in AEN adoption are generally equally effective. That is, a 50% AEN adoption rate will reduce effective reproduction number by ~8% regardless of the interaction rate between individuals. This is a result of the increased interaction rate affecting both the number of transmissions *and* the number of detections via the AEN service. For lower interaction levels, the percent change is essentially equivalent. At the highest interaction level, the percent change is slightly higher. This is due in part to the AEN-based notifications outpacing the effects of MCT.

TABLE 1

Effects of AEN adoption across a range of interaction rates. Percent change for each adoption rate was calculated with a control of 0% adoption.

R_E (% Change)		**AEN Adoption Rate**			
		0%	25%	50%	75%
Interaction Rate	2.1	1.02	0.98 (-3.8%)	0.94 (-7.6%)	0.89 (-12.8%)
	2.5	1.16	1.13 (-2.5%)	1.07 (-7.6%)	1.03 (-11.3%)
	2.7	1.21	1.18 (-2.8%)	1.12 (-8.1%)	1.07 (-11.7%)
	2.9	1.27	1.22 (-4.2%)	1.15 (-9.5%)	1.08 (-15.5%)

We also examined how SimAEN can help public health decision makers estimate the social and public health workload effects of varying intervention levels, in addition to the potential reduction in R_E. Figures 8-11 track the variations in the number of infected individuals, and in the number of uninfected individuals entering quarantine each day due to a close contact notification from AEN, for 20 runs of an 80-day simulation. (Please note that the scales of the y-axes are not uniform.)

With an interaction rate of 2.1 (Figure 8), as AEN adoption increases there is a ~50% decrease in the number of positive tests over time, but it comes at the cost of a fourfold increase in the number of people unnecessarily quarantined (as AEN adoption triples from 25% to 75%). As the interaction rate is increased, the suppression of positive cases increases to roughly 60% with increased AEN adoption. However, the cost of unnecessary quarantines continues to track upward

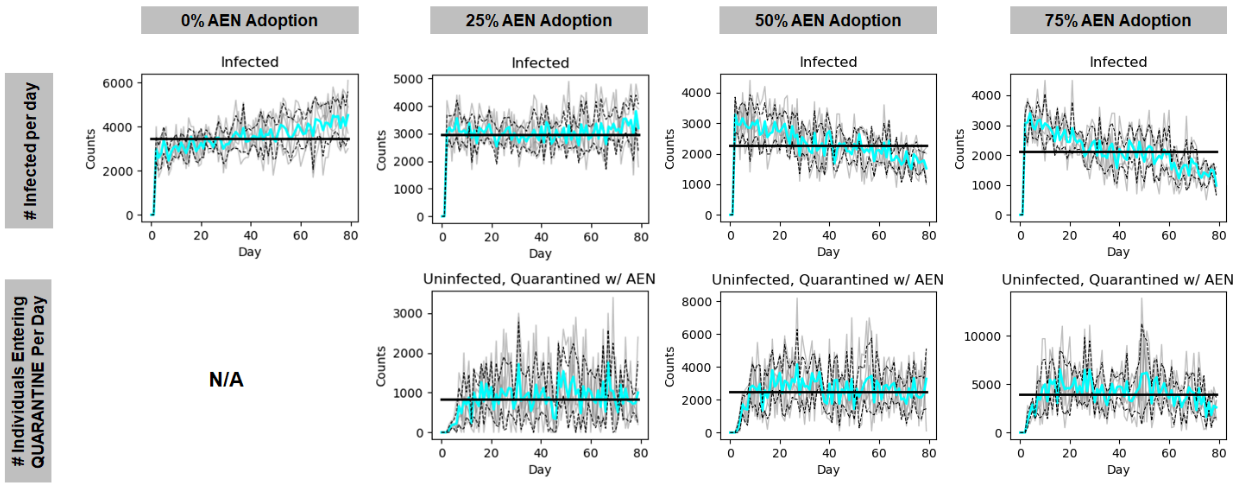

Figure 8. Costs and effects for the interaction rate $\mu = 2.1$.

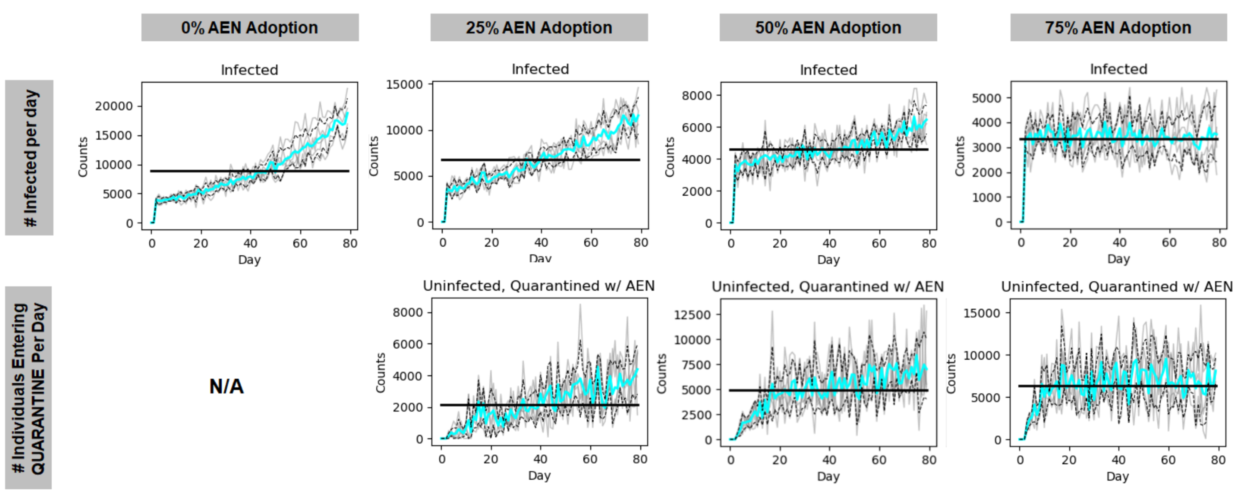

Figure 9. Costs and effects for the interaction rate $\mu = 2.5$.

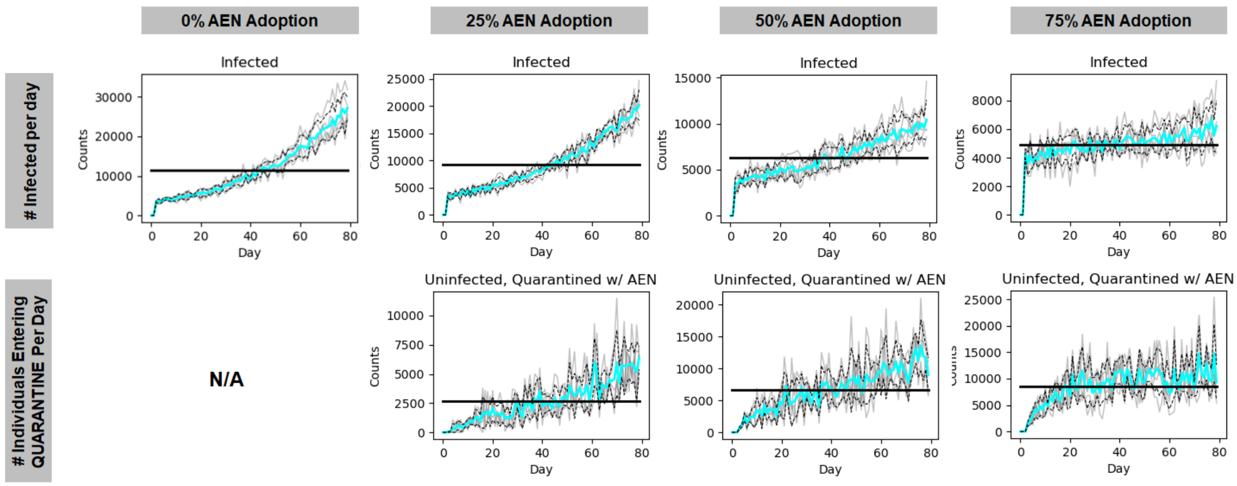

Figure 10. Costs and effects for the interaction rate $\mu = 2.7$.

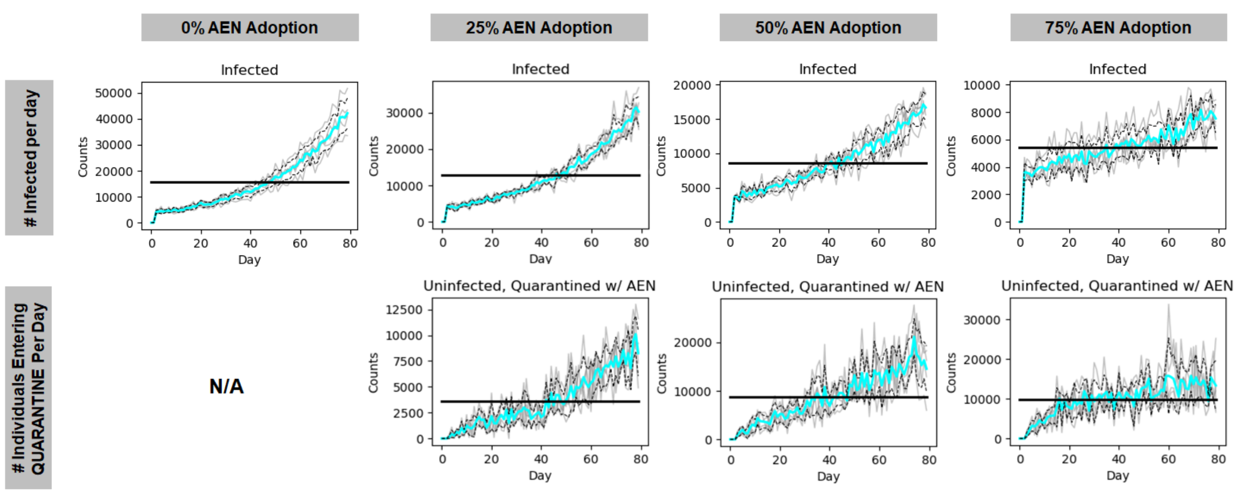

Figure 11. Costs and effects for the interaction rate $\mu = 2.9$.

with both the increased interaction rate and the increased AEN adoption. This suggests that AEN can prove to be generally more useful to a public health jurisdiction when the social and economic cost of quarantine is lower (e.g., when sufficient social supports exist for the quarantined, and/or when the inherent severity of quarantine is adjusted for vaccinated individuals who are less likely to infect others after an exposure). It also suggests that AEN is more generally useful when public health jurisdictions are faced with rising interaction rates ("re-opening") concurrent with greater testing capacity (both PCR and rapid antigen tests).

4.2 VARIATIONS IN AEN SENSITIVITY AND SPECIFICITY

One of the factors that public health has control over is the detector configuration for the AEN service. The Bluetooth attenuation weights and thresholds, in combination with the cumulative risk score thresholds for the detector, determine how likely it is that a person will receive an alert for an encounter at a certain distance and duration. The more restrictive the settings, the fewer people are identified, but as the aperture of the detector is widened, more people may be quarantined unnecessarily. Therefore, in SimAEN, the probability of detection and false discovery rate should not be varied independently, although they may appear to be independent parameters to the model. We examine here the effect on the total number of infections, and on the number of false quarantines due to AEN, as the sensitivity of the detector increases (from $P(D) = 0.36$ to 0.86) and the specificity of the detector decreases (from $FDR = 0.122$ to 0.3).

As can be seen in Figure 12, there is very little effect on the reproduction number (for the baseline configuration) as the weights are made less restrictive. On the least restrictive side, the reproduction number is 1.018, while at the most restrictive it is only reduced to 1.004. However, the number of daily false quarantined increased from 150 per day to 492 per day.

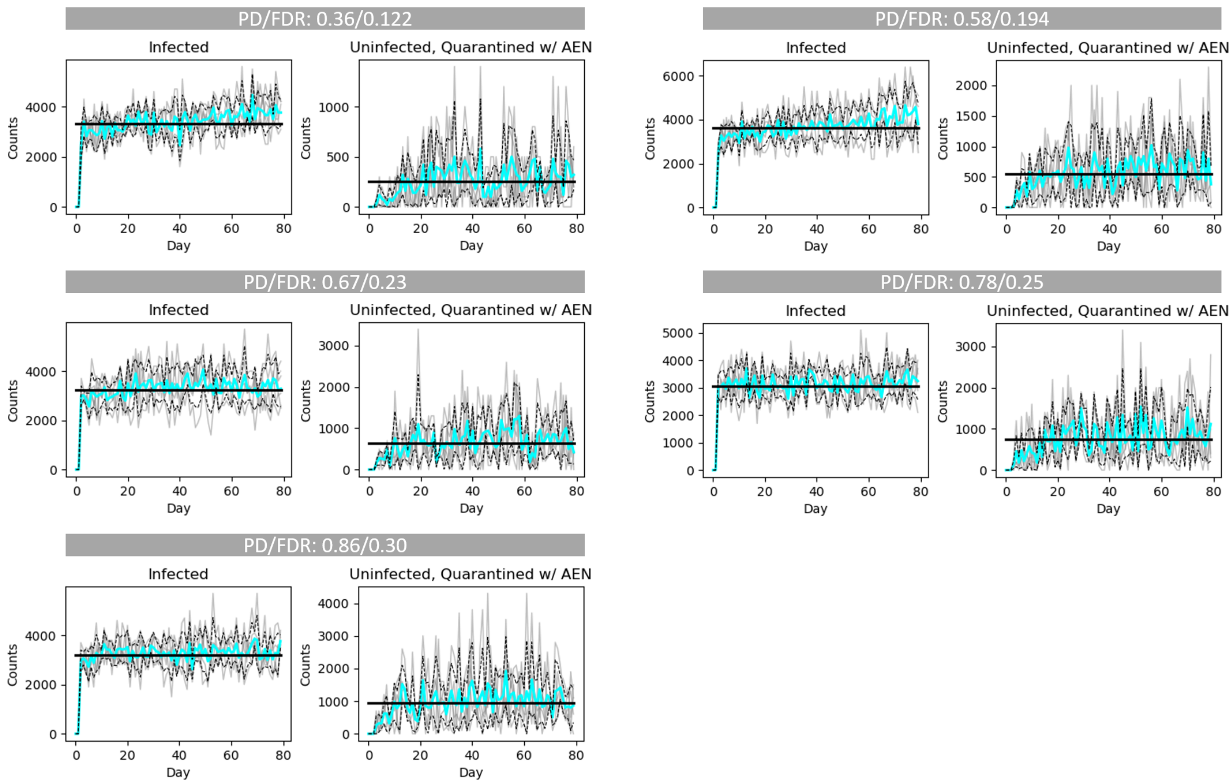

Figure 12. The number of new daily infections along with the number of false quarantines for a collection of probability of detection and false discovery rate, and AEN adoption rate of 25%.

5. DISCUSSION

5.1 SCOPE

The intention of this model is to provide public health officials with the ability to understand the effects of the actions that they have at their disposal. In order to accommodate a large range of potential input parameter values, it was critical to optimize the model for fast execution. To accomplish this, we sacrificed some fidelity of the actual spread of the disease, instead focusing on ensuring that the infection rate is appropriately affected by the public health controllable parameters. The speed gained by not incorporating factors such as mobility data or interaction topology allows the model to be executed numerous times to build up a statistical understanding of the range of potential outputs and to explore more of the parameter space.

This model is also intended for long-term understanding, not near-term high-accuracy estimates of cases or deaths. Assumptions about the predictability, or at least stability, of population mobility may be reasonable for a two-week horizon, but decisions made by public health are typically intended to be carried out on the order of months. Decisions about the hiring of contact tracing personnel or the setting of AEN weights are not going to change on short time scales in response to small fluctuations in infection rate.

5.2 LIMITATIONS

There are still many processes that are not addressed in the work. First, all of the actions are deterministic. While there is stochasticity, there is no feedback mechanism to alter behavior based on current pandemic conditions. For example, it is reasonable to expect that people are more likely to wear a mask when the number of cases is high, but in the current implementation, the rate of mask wearing is static for all time and set at the start of the run. Similarly, there are no changes as a result of seasonality. Interaction rate is likely dependent on the time of year, with riskier indoor interactions taking place more frequently in the cold winter months.

The model enjoys a computational speed boost from the simplifying assumption that the population is closed, but this is only reasonable for large populations with some amount of travel restriction. For smaller, mobile populations, this assumption becomes more questionable. Care needs to be taken when applying this model when these conditions are not met. Similarly, the assumption that only a small fraction of the population is infected at any time, and that the total number of infections is a low fraction of the overall population, may not be met. The model was created many months before the Omicron variant was detected and its effects on disease prevalence measured.

SimAEN does not attempt to model fluctuations in AEN adoption rates in response to either pandemic or individual conditions. Because there is no fixed population, we do not have any insight into the number of alerts an agent received from AEN. We have access to the number received for a given instantiation of that individual, but once they have passed out of relevance, all of their information is lost. Individuals in the real world may receive repeated AEN alerts if they interact with a lot of infected people over time, and they may choose to stop using AEN depending on this.

5.3 FUTURE WORK

As mentioned in the Limitations subsection, there is no implementation of adjustments to individual behavioral probabilities in response to changing pandemic conditions (e.g., trends in infection rates, testing availability, variants of concern). As richer data is gathered and shared about how people have responded to these trends, the model could be enhanced to incorporate a behavioral submodel. If the data is available, it would be simple to implement a Python function to generate the probabilities of wearing a mask, taking a COVID-19 test, and participating in contact tracing and quarantine from the environmental conditions, rather than using a fixed probability on each day of the model run.

Annual patterns should also be incorporated. This was not within the scope of the initial implementation of SimAEN, but as more information emerges about how COVID-19 is spread, and whether seasonal patterns obtain, it would enrich the model's predictions.

The assumptions that allow us to ignore changes in the susceptible population can be overcome by adding a parameter that adjusts the interaction rate in relation to the number of cumulative infections. This fix doesn't change the speed of the model execution and allows for more dynamic behavior. However, adding this into the model requires that the target population be specified. In the current model, there is no need to identify the size of the population, only that it is "sufficient". Once there is a factor changing susceptibility, the specific size of the population has to be specified.

Finally, the original public health workflow for AEN relied on lab-processed PCR testing to confirm an individual's infection of COVID-19. In the months since SimAEN was developed, Google and Apple developed a "self-report" feature for Exposure Notification, and many jurisdictions chose to empower their citizens to share AEN keys on the basis of a positive at-home rapid antigen test. This had the benefit of reducing public health workloads (at both PCR testing sites and labs as well as on the teams fielding citizen requests for key sharing codes), increasing participation in AEN key sharing, and reducing the time delay between testing and key sharing (by days). The AEN submodel of SimAEN could be enhanced to incorporate the new "path" to key sharing through rapid antigen testing, the regional availability of rapid antigen tests, and new data on participation rates in AEN key sharing after testing positive with a rapid antigen test.

6. CONCLUSION

We have presented an agent-based model of the effects of AEN and other public health interventions on the progression of the COVID-19 disease and the concomitant effects on public health resources and social costs (quarantines and isolation). As demonstrated in the Validation section, the proposed model produces a good fit for the early pandemic data, as measured in Pennsylvania, indicating that its most critical aspects are a reasonable reflection of real-world effects. Additional simulated effects resulting from varying mask usage, AEN adoption, and initial number of infections match expected behavioral trends and further support confidence in the model.

Some initial conclusions can be drawn from experiments with SimAEN. Because of the frequent asymptomatic or presymptomatic spread, *participation* is the most important aspect of any intervention relying on agent behavioral changes. Widespread masking is the most effective treatment because it reduces the transmission immediately, without the delay associated with AEN and MCT. These delays are a product of testing timeframes, but also the time that it takes for symptoms to develop in the generators. During this time there is ample opportunity for a large-scale spreading event to occur.

MCT is less effective than masking at reducing the number of cases. The low probability of an infectious agent identifying the agents they infected, coupled with the delays permitting transmission events, means that at real-world initial infection counts used in the model, manual contact tracing is not sufficient to control the spread of COVID-19.

This study also indicates that AEN may have a difficult time finding a role in disease mitigation. When the spread rate is approximately at replacement ($R_E \approx 1$), even a 75% adoption rate has very minimal effect on reducing spread further. Even the small reduction it does afford comes at the cost of a substantial number of false quarantines. For high spread rates, AEN can reduce the spread rate at moderate adoption rates, but it involves the quarantining of large number of individuals.

The best value to be achieved through a combination of AEN and MCT may come in the community messaging that they permit, in addition to their impact on the effective transmission rate. Public health could use this information to expand community testing efforts, direct messaging, and identify potential super-spreader events in these areas.

APPENDIX A PARAMETER DEFAULT VALUES

The tables below present default values for parameters of the SimAEN agent-based model. The values presented were derived from published literature, and from interviews with public health subject matter experts, when not available from literature.

STARTING CONDITIONS

The number of cases at the start of the simulation	50000	[27, 35]

STOPPING CONDITIONS

Length of the simulation	30 days	
Maximum number of current cases before program stops	1,500,000	

DISEASE PROPERTIES

The mean time between an individual being exposed and becoming infectious	2 days	[21, 36, 37]
The standard deviation of latent period	0.7	[21]
The mean time between an individual being exposed and becoming symptomatic	6 days	[21, 36, 38]
The standard deviation of incubation period	2.3 days	[21]
Infectious period	17 days	[21]
The likelihood an infected person will be asymptomatic	0.73	[39, 36]
The probability that a true contact event involving an infected person with no mask will result in infection...		
...if they are asymptomatic	0.03	
...if they are presymptomatic	0.03	
...if they are symptomatic	0.07	[40]

TESTING PARAMETERS		
The probability that a person who has been called by public health will get tested on any given day	0.5	
The probability that a person who has no symptoms and has not been notified in any way will get a test	0.01	
The probability that a person who has received a notification through the AEN service will get tested on any given day	0.5	
The probability that a person who is symptomatic will get tested on any given day	0.5	[33]
The mean and standard deviation of number of days that it takes for a test to get back (normal distribution)		
Mean $\mu_{testingdelay}$	2 days	[41]
Standard Deviation $\sigma_{testingdelay}$	1 day	
Daily testing capacity	∞	

PROBABILITY OF (+) TEST		
The probability that a person will test positive given they are...		
...exposed	0.5	
...presymptomatic	0.75	
...symptomatic	0.9	
...asymptomatic	0.9	

AEN PARAMETERS		
The probability that a person is running the AEN service	0.25	
The probably that the phone of an infected person will exchange information with the phone of a close contact through Bluetooth detector settings (narrow, wide)	0.67, 0.86	
The False Discovery Rate (FDR), used to create additional false positives picked up automatically by the system. At 0.5 the number of false discoveries will equal the number of true discoveries (narrow, wide)	0.23, 0.3	
The probability that a person who is running the AEN service who gets a positive test will upload their key to public health	0.72	[42]
Is an individual required to be contacted by public health before uploading their key?	No	

MANUAL CONTACT TRACING (MCT) PARAMETERS

The probability that a call from public health will reach a person identified through CT	0.5	
The probability that a call from public health will reach a person expecting the call	0.75	
The probability that an exposed individual will be found using traditional CT	0.1	
The maximum number of people an person can recall through traditional CT on a single phone call	10	
The number of contact tracers	500, 2,000	[43, 44]
How long each contact tracer can spend on calling in a day	8 hours	
The number of times contact tracers will try to contact an individual before giving up	3	
The length of time that a missed call takes	0.05 hours	
The length of time that a contact tracer takes to perform contact tracing on an index case	1 hour	
The length of time that a public health call takes	0.1 hours	
The length of time it takes for a call to upload key	0.1 hours	

STARTING BEHAVIOR

The probability that a newly initialized individual will start in quarantine	0.0	

MASK PARAMETERS

The probability that a person will wear a mask before receiving EN, being called by contact tracer, developing symptoms or receiving a positive test	0.25, 0.5	
The probability that a person will wear a mask while they are in the quarantine	0.9	[45]
How much maskless transmission rate is proportionally reduced for each person wearing a mask (higher numbers mean less transmission risk)	0.65	[31]

PERSONAL PARAMETERS		
The probability that a person will call public health after a positive test	0.75	
The probability that a person will call public health after receiving an EN notification	0.5	
Total number of people in a person's neighborhood (underlying Gaussian distribution for log-normal)		
Mean $\mu_{neighborhood}$	2.5	
Standard deviation $\sigma_{neighborhood}$	1.1	
The average number of contacts that an individual encounters each day if they take no precautions(underlying Gaussian distribution for log-normal)		
Mean $\mu_{int,standard}$	2.9	[34]
Standard deviation $\sigma_{int,standard}$	1.0	[34]
The average number of contacts that an individual encounters each day if they are in quarantine (underlying Gaussian distribution for log-normal)		
Mean $\mu_{int,quarantine}$	0.1	
Standard deviation $\sigma_{int,quarantine}$	0.1	
Probability of returning to starting behavior given negative test result and no symptoms	0.85	

PERSONAL BEHAVIOR	
Probability of entering isolation given the person is symptomatic	0.9
Probabilities of entering isolation given the person receives a positive test	0.9
Probability of entering quarantine given the person is successfully called by public health worker	0.75
Probability of entering quarantine given the person is notified by EN	0.5

VACCINATION PARAMETERS	
Probability that a person is vaccinated	0
Do vaccinated individuals spread disease asymptomatically?	No

APPENDIX B MODEL OUTPUTS

At the end of the simulation, three key summary metrics are calculated to estimate the impact of the selected input parameters.

DAILY OUTPUTS

Effective reproduction number, R_E	Mean number of new infected cases per infected individual
Cases prevented due to AEN	Calculated by subtracting the number of total cases for the given AEN adoption rate, from the number of total cases for AEN adoption = 0 (all other input paramters being the same)
Unnecessary quarantines due to AEN	Individuals identified as close contacts by the AEN detector, and notified by AEN to quarantine, but not actually infected

The following model outputs are tallied for each day of the simulation, and summed across the course of the simulation. If multiple runs are performed, they can be combined daywise to examine the minimum, maximum, mean, and standard deviation for each output over the course of the simulation, which smooths out some of the stochastic effects.

DAILY OUTPUTS	
New infections	total
	received an AEN alert
	called by MCT call
	Positive test only
	AEN alert + MCT call
	Not detected
Unnecessary Quarantines	
Positive cases ID'd and isolated	
Public Health calls	due to AEN
	due to MCT call
	(other)
Quarantines	due to AEN alert
	due to MCT call
	due to AEN alert + MCT call
	due to Symptoms
Tests	due to AEN alert
	due to MCT call
	due to symptoms
People who tested positive	& received AEN alert only
	& received MCT call only
	& received both
	& received neither
People who tested Negative	& received AEN alert only
	& received MCT call only
	& received both
	& received neither
Symptomatic positives	received AEN alert only
	received MCT call only
	received both
	received neither
Asymptomatic positives	received AEN alert only
	received MCT call only
	received both
	received neither

APPENDIX C AEN PROCESS-FLOW ANALYSIS FOR MASSACHUSETTS

At the request of the Massachusetts Department of Public Health, an analysis was conducted to assess the expected impact of a planned deployment of Automated Exposure Notification (AEN) in Massachusetts, based on SimAEN analysis and on AEN performance parameters obtained from other jurisdictions.

Specifically, the analysis focused on assessing certain cost and benefits parameters that were expected in a planned AEN deployment in Massachusetts. That included estimates of the expected levels of additional resources that would be required due to the AEN deployment on one hand (e.g., the expected increased level of testing capacity requirements), and the expected levels of public health benefits that would be attained from the AEN deployment, on the other hand (e.g., the expected level of earlier detected infected individuals).

In order to obtain key parameters from multiple jurisdictions in a consistent manner, we constructed a process model that represented a generic AEN flow, as depicted in Figure C.13.

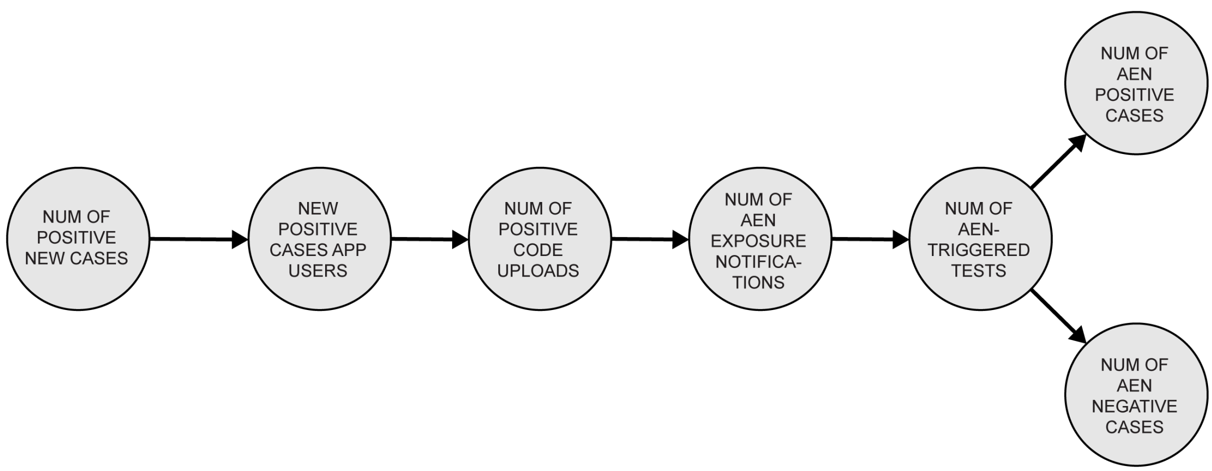

Figure C.13. Generic process flow model of an AEN system.

Given the total number of new positive cases in a particular jurisdiction on a particular day, one can derive the estimated number of positive cases who are AEN users, based on the AEN adoption rate at that jurisdiction.

From the number of positive AEN users one can derive the expected number of AEN key uploads, based on the average percentage upload rate at that jurisdiction.

Each upload results in a certain number of triggered exposure notifications to other AEN users, determined by the risk score setting of the AEN service at that jurisdiction.

The number of AEN users who receive AEN exposure notifications determines the number of expected AEN triggered tests, based on the percentage of users in that jurisdiction who are likely to go and get tested following the receipt of a notifications. Similarly one can derive the number of helpdesk sessions that are expected to be triggered by AEN exposure notifications.

The percentage of AEN triggered tests which end up with positive test results is the AEN secondary-attack-rate (SAR) and can be used to derive the number of AEN triggered users who are expected to test positive, as well as the number of these expected to test negative.

Figure C.14 represents an assessment of the above process flow values for the AEN deployment in the UK in a particular period. These values are derived from [42] and are based on data which can either be measured directly in the target environment or can be assessed indirectly from other sources.

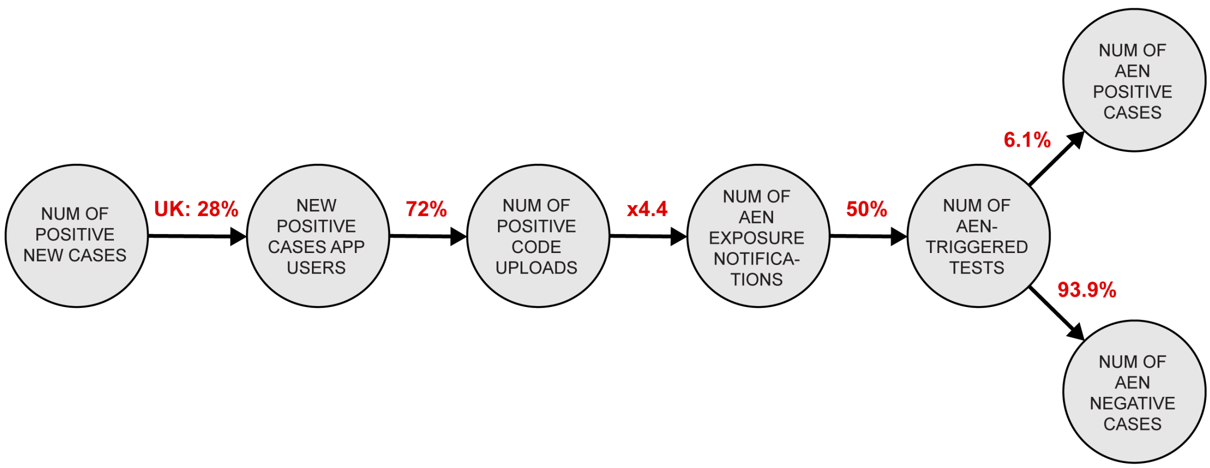

Figure C.14. AEN process flow with UK estimates of participation at each step.

At the particular target period of time, 28% of the UK population was using the AEN service and 72% of AEN users who tested positive actually went ahead and uploaded their recent AEN keys. Each upload resulted on average in 4.4 new exposure notifications to other AEN users. It was estimated that only 50% of the AEN users who received an exposure notification actually proceeded to get tested. Of these users who tested following a receipt of an exposure notification it was assessed that at least 6.1% tested positive. Similarly, corresponding values were assessed for other jurisdictions such as Switzerland and the Netherlands.

This type of process flow abstraction and high-level estimations served as a basis for initial rough estimations and what-if analysis which can then be further analyzed and refined. For example, applying the above UK derived parameters to the number of estimated new cases in Massachusetts on a given day, one can derive preliminary rough estimations as depicted in Figure C3.

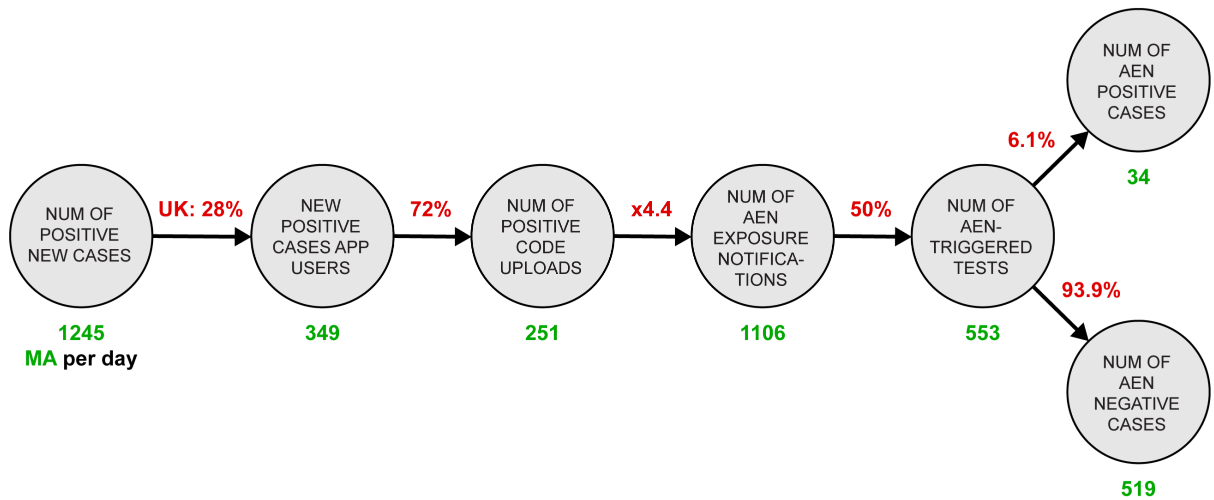

Figure C.15. AEN process flow, applying UK estimates to MA case rate estimate.

That is, given an estimated number of 1245 new cases in MA on a given day, based on the UK parameters we can expect an estimated 553 additional tests as a result of new AEN exposure notifications. Such process flow model values were used to calibrate certain SimAEN execution parameters and refine various insights.

The process flow model can serve as a basis for a more granular quantifiable cost-benefits analysis and optimization. For example, one can assess the number of cases prevented (and life saved) by the number of earlier detected of positive cases due to AEN exposure notifications. Similarly one can assess the extra resources associated with the additional notifications to individuals who are not infected.

Due to privacy-preservation concerns as well as to challenges of integration between AEN and public-health processes, assessing actual key AEN execution values (such as AEN's "secondary attack rate") has proven extremely challenging, and often impossible, in almost almost all jurisdictions. More effective and unified operational processes can and should be designed and implemented to enable better assessment and optimization of the AEN value and derive important new insights, without violating privacy objectives.

GLOSSARY

AEN	Automated Exposure Notification
BLEMUR	Bluetooth Low Energy Model of User Risk
COVID-19	The disease caused by the SARS-CoV-2 virus
CDF	Cumulative Distribution Function
CT	Contact tracing
EN	Exposure Notification
FDR	false discovery rate
GAEN	Google-Apple Exposure Notification
MCT	Manual (traditional) contact tracing
ODD	Overview, Design concepts, Details protocol
ODE	Ordinary Differential Equation
PACT	Private Automated Contact Tracing
PCR	Polymerase chain reaction
P(D)	Probability of detection
SARS-CoV-2	severe acute respiratory syndrome coronavirus 2
SimAEN	Simulation of Automated Exposure Notification
TC4TL	“Too close for too long” standard

NOTATION

μ	mean of the underlying Gaussian distribution
μ_{int}	mean number of daily interactions
$\mu_{neighborhood}$	mean number of people in a person's neighborhood
σ	standard deviation of the underlying Gaussian distribution
σ_{int}	standard deviation of number of daily interactions
$\sigma_{neighborhood}$	standard deviation of number of people in a person's neighborhood
A_i	i^{th} agent
C_i	potential interactions of the i^{th} agent with other agents $\{A_{i1}, ..., A_{in}\}$
R_E	effective reproduction number
t	time, i.e., day of simulation

REFERENCES

[1] S. M. Moghadas, M. C. Fitzpatrick, P. Sah, A. Pandey, A. Shoukat, B. H. Singer, and A. P. Galvani, “The implications of silent transmission for the control of COVID-19 outbreaks”, *Proceedings of the National Academy of Sciences*, vol. 117, no. 30, pp. 17 513–17 515, Jul. 2020. DOI: 10.1073/pnas.2008373117. [Online]. Available: https://doi.org/10.1073/pnas.2008373117.

[2] C. Watson, A. Cicero, J. Blumenstock, and M. Fraser. (Apr. 2020). A national plan to enable comprehensive COVID-19 case finding and contact tracing in the US. accessed 2022-04-05, [Online]. Available: https://www.centerforhealthsecurity.org/our-work/pubs_archive/pubs-pdfs/2020/200410-national-plan-tocontact-tracing.pdf.

[3] R. Canetti, Y. T. Kalai, A. Lysyanskaya, R. L. Rivest, A. Shamir, E. Shen, A. Trachtenberg, M. Varia, and D. J. Weitzner, “Privacy-preserving automated exposure notification”, *IACR Cryptol. ePrint Arch*, vol. 2020, no. 863, p. 31, 2020.

[4] H. Bair, J. D. Wanger, and N. R. Shah, “A brief history of exposure notification during the COVID-19 pandemic in the United States, 2020-2021”, *Public Health Reports*, vol. 137, no. 2_suppl, in press, 2022.

[5] J. Chan, S. Gollakota, E. Horvitz, J. Jaeger, S. Kakade, T. Kohno, J. Langford, J. Larson, S. Singanamalla, J. Sunshine, *et al.*, “PACT: Privacy sensitive protocols and mechanisms for mobile contact tracing”, *arXiv preprint arXiv:2004.03544*, 2020.

[6] L. Ferretti, C. Wymant, M. Kendall, L. Zhao, A. Nurtay, L. Abeler-Dörner, M. Parker, D. Bonsall, and C. Fraser, “Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing”, *Science*, vol. 368, no. 6491, 2020. DOI: 10.1126/science.abb6936.

[7] Apple, Inc. (2020). Apple Developer Documentation: ENExposureConfiguration. accessed 2020-12-16, [Online]. Available: https://developer.apple.com/documentation/exposurenotification/enexposureconfiguration.

[8] G. F. Hatke, M. Montanari, S. Appadwedula, M. Wentz, J. Meklenburg, L. Ivers, J. Watson, and P. Fiore, “Using Bluetooth Low Energy (BLE) signal strength estimation to facilitate contact tracing for COVID-19”, PACT, 2020. DOI: 10.48550/ARXIV.2006.15711. [Online]. Available: https://pact.mit.edu/wp-content/uploads/2020/11/TCFTL_paper_FINAL_V15_w_logos.pdf.

[9] D. J. Leith and S. Farrell, “Coronavirus contact tracing: Evaluating the potential of using Bluetooth received signal strength for proximity detection”, *ACM SIGCOMM Computer Communication Review*, vol. 50, no. 4, pp. 66–74, Oct. 2020. DOI: 10.1145/3431832.3431840. [Online]. Available: https://doi.org/10.1145/3431832.3431840.

[10] W. O. Kermack and A. G. McKendrick, “A contribution to the mathematical theory of epidemics”, *Proceedings of the Royal Society of London. Series A, Containing papers of a mathematical and physical character*, vol. 115, no. 772, pp. 700–721, 1927. DOI: 10.1098/rspa.1927.0118.

[11] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world' networks", *Nature*, vol. 393, no. 6684, pp. 440–442, 1998. DOI: 10.1038/30918.

[12] M. Wieczorek, J. Siłka, and M. Woźniak, "Neural network powered COVID-19 spread forecasting model", *Chaos, Solitons & Fractals*, vol. 140, p. 110 203, 2020. DOI: 10.1016/j.chaos.2020.110203.

[13] I. Traulsen and J. Krieter, "Assessing airborne transmission of foot and mouth disease using fuzzy logic", *Expert Systems with applications*, vol. 39, no. 5, pp. 5071–5077, 2012. DOI: 10.1016/j.eswa.2011.11.032.

[14] G. L. Werneck, C. H. Costa, A. M. Walker, J. R. David, M. Wand, and J. H. Maguire, "The urban spread of visceral leishmaniasis: Clues from spatial analysis", *Epidemiology*, pp. 364–367, 2002.

[15] A. Adamatzky, *Game of life cellular automata*. Springer, 2010, vol. 1. DOI: 10.1007/978-1-84996-217-9.

[16] S. H. White, A. M. Del Rey, and G. R. Sánchez, "Modeling epidemics using cellular automata", *Applied Mathematics and Computation*, vol. 186, no. 1, pp. 193–202, 2007. DOI: 10.1016/j.amc.2006.06.126.

[17] H. F. Gagliardi and D. Alves, "Small-world effect in epidemics using cellular automata", *Mathematical Population Studies*, vol. 17, no. 2, pp. 79–90, 2010. DOI: 10.1080/08898481003689486.

[18] D. Schuldt, T. Londner, J. Saunders, N. Hanson, G. Gianforcaro, M. C. Schiefelbein, J. Avellan, and B. Hallowell, *SimAEN source code*, Dec. 2021. DOI: 10.5281/zenodo.1234. [Online]. Available: https://github.com/informaticslab/SimAEN.

[19] V. Grimm, U. Berger, D. L. DeAngelis, J. G. Polhill, J. Giske, and S. F. Railsback, "The ODD protocol: A review and first update", *Ecological modelling*, vol. 221, no. 23, pp. 2760–2768, 2010. DOI: 10.1016/j.ecolmodel.2010.08.019.

[20] R. M. Anderson and R. M. May, "Population biology of infectious diseases: Part I", *Nature*, vol. 280, no. 5721, pp. 361–367, 1979.

[21] M. Peirlinck, K. Linka, F. S. Costabal, and E. Kuhl, "Outbreak dynamics of COVID-19 in China and the United States", *Biomechanics and Modeling in Mechanobiology*, p. 1, 2020. DOI: 10.1007/s10237-020-01332-5.

[22] A. T. Xiao, Y. X. Tong, and S. Zhang, "False-negative of rt-pcr and prolonged nucleic acid conversion in covid-19: Rather than recurrence", *Journal of Medical Virology*, 2020.

[23] G. Gettliffe, "Modeling probability of alert of Bluetooth Low Energy-based automatic exposure notifications", MIT Lincoln Laboratory, ACTA-4, Apr. 2022.

[24] Apple and Google. (Apr. 2021). Exposure Notification Privacy-preserving Analytics (ENPA) white paper. accessed 2022-04-05, [Online]. Available: https://covid19-static.cdn-apple.com/applications/covid19/current/static/contact-tracing/pdf/ENPA_White_Paper.pdf.

[25] E. Aronoff-Spencer, J. Benzler, T. Burner, C. Fraser, A. Higgins, B. T. Karras, A. King, W. B. Lober, R. Marsden, W. Pugh, D. Revere, C. Schiefelbein, J. Wanger, G. Zane, and S. Zimmerman. (Dec. 2021). Configuring Exposure Notification risk scores for COVID-19 v2.0. accessed 2021-03-30, [Online]. Available: https://github.com/lfph/gaen-risk-scoring/blob/main/risk-scoring.md.

[26] J. Benzler, M. Briers, M. Flowers, S. Halai, B. Harris, B. T. Karras, E. Kosten, R. Marsden, J. Masel, T. W. Menza, M. McConnell, K. Murphy, R. Palanki, M. Payer, R. Pillai, B. Pugh, R. Raskar, R. Rivest, T. Robb, K. Rough, G. Stringer, J. Wanger, S. Zimmerman, and M. Zissman. (Nov. 2020). Configuring Exposure Notification risk scores for COVID-19. accessed 2021-03-30, [Online]. Available: https://github.com/lfph/gaen-risk-scoring/blob/1.0.1/risk-scoring.md.

[27] E. Dong, H. Du, and L. Gardner, "An interactive web-based dashboard to track COVID-19 in real time", *The Lancet Infectious Diseases*, vol. 20, no. 5, pp. 533–534, 2020, ISSN: 1473-3099. DOI: https://doi.org/10.1016/S1473-3099(20)30120-1. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1473309920301201.

[28] A. A. Sayampanathan, C. S. Heng, P. H. Pin, J. Pang, T. Y. Leong, and V. J. Lee, "Infectivity of asymptomatic versus symptomatic COVID-19", *The Lancet*, 2020.

[29] K. Mizumoto, K. Kagaya, A. Zarebski, and G. Chowell, "Estimating the asymptomatic proportion of coronavirus disease 2019 (COVID-19) cases on board the Diamond Princess cruise ship, Yokohama, Japan, 2020", *Eurosurveillance*, vol. 25, no. 10, p. 2 000 180, 2020.

[30] O. Aydin, B. Emon, S. Cheng, L. Hong, L. P. Chamorro, and M. T. A. Saif, "Performance of fabrics for home-made masks against the spread of COVID-19 through droplets: A quantitative mechanistic study", *Extreme Mechanics Letters*, vol. 40, p. 100 924, 2020.

[31] H. Ueki, Y. Furusawa, K. Iwatsuki-Horimoto, M. Imai, H. Kabata, H. Nishimura, and Y. Kawaoka, "Effectiveness of face masks in preventing airborne transmission of SARS-CoV-2", *MSphere*, vol. 5, no. 5, 2020.

[32] J. A. Backer, D. Klinkenberg, and J. Wallinga, "Incubation period of 2019 novel coronavirus (2019-nCoV) infections among travellers from Wuhan, China, 20–28 January 2020", *Eurosurveillance*, vol. 25, no. 5, p. 2 000 062, 2020. DOI: doi:10.2807/1560-7917.ES.2020.25.5.2000062.

[33] J. Watson, P. F. Whiting, and J. E. Brush, "Interpreting a COVID-19 test result", *Bmj*, vol. 369, 2020. DOI: 10.1136/bmj.m1808.

[34] Y.-c. Fu, "Measuring personal networks with daily contacts: A single-item survey question and the contact diary", *Social Networks*, vol. 27, no. 3, pp. 169–186, 2005. DOI: 10.1016/j.socnet.2005.01.008.

[35] Centers for Disease Control and Prevention. (). United States COVID-19 cases and deaths by state over time. accessed 2021-06-18, [Online]. Available: https://data.cdc.gov/Case-Surveillance/United-States-COVID-19-Cases-and-Deaths-by-State-o/9mfq-cb36.

[36] Centers For Disease Control and Prevention. (Sep. 2020). COVID-19 pandemic planning scenario. accessed 2021-06-18, [Online]. Available: https://www.cdc.gov/coronavirus/2019-ncov/hcp/planning-scenarios.html.

[37] M. Johansson, H Wolford, and P Paul, "Reducing travel-related SARS-CoV-2 transmission with layered mitigation measures– symptom monitoring, quarantine, and testing", *BMC Med.*, vol. 19, no. 1, p. 94, 2021. [Online]. Available: https://https://doi.org/10.1186/s12916-021-01975-w, (Published 2021 Apr 14).

[38] C. McAloon, Á. Collins, K. Hunt, A. Barber, A. W. Byrne, F. Butler, M. Casey, J. Griffin, E. Lane, D. McEvoy, P. Wall, M. Green, L. O'Grady, and S. J. More, "Incubation period of COVID-19: A rapid systematic review and meta-analysis of observational research", *BMJ Open*, vol. 10, no. 8, 2020, ISSN: 2044-6055. eprint: https://bmjopen.bmj.com/content/10/8/e039652.full.pdf. [Online]. Available: https://doi.org/10.1136/bmjopen-2020-039652.

[39] P. Poletti, M. Tirani, D. Cereda, F. Trentini, G. Guzzetta, G. Sabatino, V. Marziano, A. Castrofino, F. Grosso, G. Del Castillo, *et al.*, "Probability of symptoms and critical disease after SARS-CoV-2 infection", *arXiv preprint arXiv:2006.08471*, 2020.

[40] R. Laxminarayan, B. Wahl, S. R. Dudala, K Gopal, S Neelima, K. J. Reddy, J Radhakrishnan, J. A. Lewnard, *et al.*, "Epidemiology and transmission dynamics of COVID-19 in two Indian states", *Science*, vol. 370, no. 6517, pp. 691–697, 2020. DOI: 10.1126/science.abd7672.

[41] D. Lazer, M. Santillana, R. H. Perlis, K. Ognyanova, M. Baum, A. Quintana, J. Druckman, J. Della Volpe, H. Chwe, and M. Simonson, "The state of the nation: A 50-state COVID-19 survey report #8: Failing the test", 2020. DOI: 10.31219/osf.io/gj9x8.

[42] C. Wymant, L. Ferretti, D. Tsallis, M. Charalambides, L. Abeler-Drner, D. Bonsall, R. Hinch, M. Kendall, L. Milsom, M. Ayres, C. Holmes, M. Briers, and C. Fraser, "The epidemiological impact of the NHS COVID-19 app", 2021. [Online]. Available: https://doi.org/10.1038/s41586-021-03606-z.

[43] Pennsylvania Pressroom. (Jun. 2020). Pennsylvania partners with staffing agency to hire 1,000 more COVID-19 contact tracers, [Online]. Available: https://www.media.pa.gov/Pages/Health-Details.aspx?newsid=938.

[44] D Lewis, "Why many countries failed at COVID contact-tracing but some got it right", 2021. [Online]. Available: https://www.nature.com/articles/d41586-020-03518-4.

[45] B. M. Gurbaxani, A. N. Hill, P. Paul, P. V. Prasad, and R. B. Slayton, "Evaluation of different types of face masks to limit the spread of SARS-CoV-2 – a modeling study", *medRxiv*, 2021. eprint: https://www.medrxiv.org/content/early/2021/04/27/2021.04.21.21255889.full.pdf. [Online]. Available: https://doi.org/10.1101/2021.04.21.21255889.

REPORT DOCUMENTATION PAGE	*Form Approved* *OMB No. 0704-0188*

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
15-04-2022	Project Report	

4. TITLE AND SUBTITLE

The Simulation of Automated Exposure Notification (SimAEN) Model

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Gwendolyn Gettliffe, Adam Norige, Ted Londner, Jonathan Saunders, Dieter Schuldt, Curran Schiefelbein, William Streilein

5d. PROJECT NUMBER
10383-2

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

MIT Lincoln Laboratory
244 Wood Street
Lexington, MA 02420-9108

8. PERFORMING ORGANIZATION REPORT NUMBER

ACTA-5

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Centers for Disease Control and Prevention

1600 Clifton Road, Atlanta, GA 30329 USA

10. SPONSOR/MONITOR'S ACRONYM(S)
CDC

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

DISTRIBUTION STATEMENT A. Approved for public release. Distribution is unlimited.
This material is based upon work supported by the Under Secretary of Defense for Research and Engineering under Air Force Contract No. FA8702-15-D-0001.

13. SUPPLEMENTARY NOTES

14. ABSTRACT
Automated Exposure Notication (AEN) was implemented in 2020 to supplement traditional contact tracing for COVID-19 by estimating “too close for too long" proximities of people using the service. AEN uses Bluetooth messages to privately label and recall proximity events, so that persons who were likely exposed to SARS-CoV-2 can take the appropriate steps recommended by their health care authority. This paper describes an agent-based model that estimates the effects of AEN deployment on COVID-19 caseloads and public health workloads in the context of other critical public health measures available during the COVID-19 pandemic. We selected simulation variables pertinent to AEN deployment options, varied them in accord with the system dynamics available in 2020-2021, and calculated the outcomes of key metrics across repeated runs of the stochastic multi-week simulation. SimAEN's parameters were set to ranges of observed values in consultation with public health professionals and the rapidly accumulating literature on COVID-19 transmission; the model was validated against available population-level disease metrics. Estimates from SimAEN can help public health ocials determine what AEN deployment decisions (e.g., conguration, workflow integration, and targeted adoption levels) can be most effective in their jurisdiction, in combination with other COVID-19 interventions (e.g., mask use, vaccination, quarantine and isolation periods).

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**		46	**19b. TELEPHONE NUMBER** *(include area code)*

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

This page intentionally left blank.